A Meta-Learning Approach for Hydrological Time Series Model Selection

Asma Slaimi

PhD

Supervised by

Prof. Noel O'Connor, Dr. Michael Scriney, Dr. Susan Hegarty and Prof. Fiona Regan

A thesis presented for the degree of Doctor of Philosophy

School of Electronic Engineering

Dublin City University

August 2024

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Asma Slaimi

ID No: 18215104 Date: 26/08/2024

Acknowledgements

I want to express my deep appreciation to my supervisors Prof. Noel O'Connor, Dr. Michael Scriney, Dr. Susan Hegarty, and Prof. Fiona Regan, for their unwavering support and invaluable guidance throughout my research journey.

In particular, I sincerely thank Dr. Michael Scriney for his exceptional dedication and tireless efforts, which have consistently provided me with his guidance, inspiration, and support.

I am also grateful to my friends, whose unwavering willingness to assist and listen during challenging times was crucial to my research.

Lastly, I extend my profound appreciation and love to my parents, Abed el Aziz and Salha, and my sisters and brothers for their unwavering support, deep understanding, and love over the years.

List of Publications

- Susan Hegarty, **Asma Slaimi**, Noel O'Connor, and Fiona Regan, "Using Citizen Science to meet the challenge of water monitoring: evidence from Dublin and beyond (a Dublin Case study)", January 2020.

- Chloe Richards, **Asma Slaimi**, Noel O'Connor, Alan Barrett, Sandra Kwiatkowska, and Fiona Regan, "Bio-inspired surface texture modification as a viable feature of future aquatic antifouling strategies; a review", International Journal of Molecular Sciences, July 2020.

- **Asma Slaimi**, Susan Hegarty, Fiona Regan, Michael Scriney, and Noel O'Connor, "Machine learning-based tools for water digitalisation," EGU General Assembly 2021, online, 19–30 Apr 2021, EGU21-13092, https://doi.org/10.5194/egusphere-egu21-13092, 2021.

- **Asma Slaimi**, Michael Scriney, Susan Hegarty, Fiona Regan, and Noel E. O'Connor, "Meta-learning for water level prediction," EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-16266, https://doi.org/10.5194/egusphere-egu23-16266, 2023.

Contents

List of Figures

List of Tables

List of Abbreviations

AICc	Akaike Information Criterion corrected
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
API	Application Programming Interfaces
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
BERT	Bidirectional Encoder Representations from Transformers
BI	Business Intelligence
CSV	Comma-Separated Values
DW	Data Warehouse
E	Eastern
EDA	Exploratory Data Analysis
ELT	Extract-Load-Transform
EMA	Exponential Moving Average
ETS	Exponential Smoothing State-Space Model
ETL	Extract-Transform-Load
EPA	Environmental Protection Agency
GDBT	Gradient Boosted Decision Trees
GIS	Geographical Information Systems
GSI	Geological Survey of Ireland
GRU	Gated Recurrent Unit
IoT	Internet of Things
IT	Information Technology
JSON	JavaScript Object Notation
KNN	K-Nearest Neighbors
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
MA	Moving Average
MAE	Mean Absolute Error
MD	Multidimensional Data
MDA	Multidimensional Analytical
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression
MOLAP	Multidimensional OLAP
MSE	Mean Squared Error
NB	Neagh Bann
NW	North Western
O&M	Observations and Measurements
OGC	Open Geospatial Consortium
OLAP	Online Analytical Processing
OPW	Office of Public Works
PACF	Partial Autocorrelation Function
RBD	River Basin District
RDB	Relational Database

RF	Random Forest
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
ROLAP	Relational OLAP
S	Shannon
SE	South Eastern
SMOTE	Synthetic Minority Oversampling Technique
SOM	Self-Organizing Map
SQL	Structured Query Language
STDI	Spatio-Temporal Environmental Data Integration
SVM	Support Vector Machine
SW	South Western
UTC	Coordinated Universal Time
VAE	Variational Autoencoder
W	Western
XML	Extensible Markup Language

Abstract

Asma Slaimi

A Meta-Learning Approach for Hydrological Time Series Model Selection

Time series forecasting is crucial in various fields, with significant socio-economic implications, as accurate predictions can aid in better resource management, disaster preparedness, and economic planning. However, selecting an appropriate forecasting model remains a labor-intensive task demanding expertise. This research introduces a novel meta-learning approach to automate and enhance the model selection process.

We curate extensive time series datasets specific to Ireland, spanning diverse temporal patterns and environmental attributes, including climate data, water level measurements, and landscape characteristics.

The initial part of the research focuses on developing a systematic architecture using Extract, Transform, and Load (ETL) technology to integrate heterogeneous data from various sources while ensuring data quality and consistency.

Then, this research concentrates on accurately predicting river water levels. Various Machine Learning (ML) models are employed, relying on previously observed river water level data. The research evaluates the predictive performance of these ML models across all hydrometric stations in Ireland and demonstrates the importance of careful model selection based on geographic and hydrological features. The results demonstrated that a universal 'one-model-fits-all' approach is not suitable for hydrological time series data.

Subsequently, this research explores the core contribution of applying meta-learning to context-aware model selection for river water-level prediction. The study demonstrates that meta-learning enhances the accuracy and reliability of hydrologic time series forecasting, addressing the complexities of this task and providing valuable insights into applying ML in this domain. The efficacy of our meta-learning approach is evaluated across various real-world time series datasets, consistently demonstrating its superiority over traditional model selection techniques. Importantly, our approach streamlines and expedites time series forecasting, making it more accessible to researchers.

In conclusion, this thesis significantly contributes to ML-based environmental time-series data prediction using a model-selection meta-learner approach and enhanced data integration techniques. The results show that our research aligns with the growing trend of automated machine learning and has the potential to revolutionise time series forecasting in diverse applications.

Chapter 1

Introduction

In a world grappling with unprecedented environmental changes, understanding, predicting, and mitigating their impacts has become increasingly urgent. Climate change, pollution, habitat loss, and other environmental factors pose significant challenges that reverberate through natural ecosystems, societies, and economies.

In this interdisciplinary research, we merge principles from environmental science, data integration methodologies, and machine learning techniques to address the pressing challenges posed by climate change and environmental dynamics. The necessity for precise predictions is paramount in both climate and environmental science, which span a broad spectrum of data—from ecosystems and land use to air and water quality, and biodiversity. However, grappling with the complexities of climate and environmental data, often nonlinear and influenced by various factors like geography and temporal dynamics, presents a formidable challenge.

This research seeks to bridge the gap between data-driven machine learning and the nuanced demands of climate and environmental science, particularly in predicting environmental features such as river water levels. We propose an innovative approach that integrates advanced meta-learning techniques with environmental data, empowering machine learning models to tackle the intricacies of climate-related and environmental phenomena.

The seamless integration of diverse environmental datasets is crucial for gaining a holistic understanding of the environment—a prerequisite for accurate predictions in this complex domain. This integration lays the foundation for leveraging the potential of meta-learning models, significantly improving the accuracy of time series predictions in climate and environmental forecasting.

This research centers around a meta-learning approach applied to model selection for time series prediction tasks, with a specific focus on climate and environmental time series data. Meta-learning, a methodology that involves learning how to learn, enables the construction of adaptable models capable of generalizing across various prediction tasks by leveraging knowledge from multiple datasets or tasks [198].

To effectively address this challenge, we emphasize the critical role of environmental data integration [97]. It unravels intricate connections between environmental factors, facilitating a deeper understanding of their ripple effects throughout ecosystems. By synthesizing data from multiple sources, we can gain a comprehensive view of environmental processes, unlocking insights to inform more effective conservation strategies and sustainable management practices.

In summary, this thesis is propelled by the urgent need to enhance climate science and environmental science predictions, focusing on water level prediction through the seamless integration of environmental data into the modelling process. We aim to provide more reliable and actionable predictions, ultimately bolstering the capacity to comprehend, adapt to, and address the multifaceted challenges of climate change and environmental dynamics.

This chapter serves as an introductory compass for this thesis, furnishing essential context for the research focus, delineating primary research inquiries, and highlighting anticipated contributions. The ensuing sections are structured as follows: Section 1.1 provides an overview of environmental data integration, emphasising its significance in addressing complex environmental issues; Section 1.2 explores the complexities of predicting environmental time series data, with a focus on hydrology; Section 1.3 outlines research objectives, particularly the application of meta-learning techniques to enhance prediction precision; and Section 1.4 presents the thesis structure, offering readers a roadmap to the upcoming chapters and their contents.

1.1 Integration of Catchment Datasets: Environmental Data Integration

A catchment is a region where the natural landscape collects water. It is the area of land from which water flows into a river, lake, reservoir or other body of water. Water management platforms and environmental models at the catchment scale have developed into essential tools for monitoring, managing, and evaluating catchment variables, such as climate and water [168].

The availability of hydrological datasets obtained from remotely sensed data has increased dramatically in the last decade, and a growing amount of research evaluating remotely sensed data for hydrological applications has emerged. Topographic data, precipitation, soil moisture, water flow, and variations in terrestrial water storage can now be measured or predicted at various spatial and temporal scales ([82, 188]). However, these data need to be centralised in one location to benefit from it the most.

Environmental data integration is crucial for understanding and addressing complex environmental issues and scenarios such as climate change, biodiversity loss, and natural resource management. With the increasing availability of diverse environmental data from sources such as satellite imagery, sensor networks, and citizen science initiatives, there are unprecedented opportunities to advance our understanding of environmental processes and interactions [129]. However, integrating these datasets poses significant challenges due to their heterogeneity in spatiotemporal resolution, data format, and attribute representation [185]. Accurate integration of environmental data is essential for informed decision-making, accurate analysis, and the development of sustainable policies [218].

Efficient data integration and management are crucial in environmental studies as they enable researchers to analyse and interpret diverse datasets from different sources [187]. Geographic Information Systems (GIS), a technology that captures, stores, analyzes, manages, and presents spatial or geographic data, are traditionally utilized in this regard. However, traditional GIS have limitations in handling the complexities of spatiotemporal data integration. They may struggle with complex temporal data, integrating non-geospatial data sources, and discrepancies in attribute values and data structures [225]. Additionally, computational power and storage capacity limitations can hinder the integration of large and complex datasets [208].

Spatiotemporal data integration is critical in environmental monitoring, where tracking changes in environmental conditions over time accurately requires data from multiple sources. Previous studies have demonstrated the potential of spatiotemporal data integration in monitoring water quality, land cover and land use changes, and air quality [215] [83] [126]. However, it is essential to recognise that the current methodologies in the literature for integrating environmental data frequently involve custom-tailored data engineering initiatives [3] [141]. These initiatives are meticulously crafted for particular datasets and goals, resulting in a disparate array of solutions [204]. The absence of an all-encompassing and standardised approach gives rise to inefficiencies, redundant work, and a lack of scalability [8].

By integrating heterogeneous environmental data sources and employing advanced machine learning techniques, the research seeks to improve the accuracy, reliability, and efficiency of hydrological time series forecasting, ultimately contributing to better resource management, disaster preparedness, and socio-economic planning.

1.2 Time Series Prediction

1.2.1 Time Series Definition

A time series can be defined as a sequential collection of data points that represent measurements of an object, phenomenon, or signal. Hence, a time series T is an ordered sequence of n real-valued variables $T = (t_1, \ldots, t_n)$, where $t_i \in \mathbb{R}$. [59]

These measurements are recorded at different time points, which may or may not be at regular intervals in time. At a high level, a time series dataset can be categorised as either equidistant or unequidistant.

An equidistant time series can be defined as a sequential collection of data points that represent measurements of an object, phenomenon, or signal, which are recorded at regular intervals in time [25]. An equidistant time series is a set of stochastic variables denoted as $x_1, x_2, \ldots, x_t, \ldots, x_T$, indexed by an integer value t. It is a specific instance of a time series T, which is an ordered sequence of n real-valued variables $T = (t_1, \ldots, t_n)$, where $t_i \in \mathbb{R}$.

The term *equidistant* refers to data points or observations that exhibit a uniform spacing or even distribution in relation to time intervals [27]. In simple terms, the time intervals between consecutive data points exhibit a constant and consistent pattern. The use of uniform spacing in the analysis and modelling of time series data facilitates the process of comparing and predicting values at specific time points. The consistency afforded by such datasets simplifies subsequent data cleaning, feature engineering and analysis stages.

An instance of an equidistant time series can be observed when daily temperature measurements are collected at a fixed time each day, with each data point representing the temperature recorded at the end of the day. In this scenario, the time intervals between measurements remain consistent, occurring at regular intervals of one day.

Another example is the the hydrological monitoring network managed by the Office of Public Works (OPW), responsible for water management and flood protection in Ireland. This network includes strategically positioned sensor-equipped stations along rivers. These sensors record water levels at consistent intervals, such as every 15 minutes. Table 1.1 provides an example of the water level data recorded for the Aclint Hydrometric station for the year 2022.

Table 1.1: Water-level time-series data for 2022 (Aclinet hydrometric station in Ireland)

Date	**Value**
2022-01-01 00:00:00	23.162
2022-01-01 00:15:00	23.163
...	...
2022-12-31 23:30:00	23.324
2022-12-31 23:45:00	23.325

The resulting dataset forms a time series, as illustrated in Figure 1.1, allowing for precise monitoring of changes in river water levels over time. This data is crucial for understanding river behaviour, forecasting potential floods, and conducting detailed hydrological analyses.

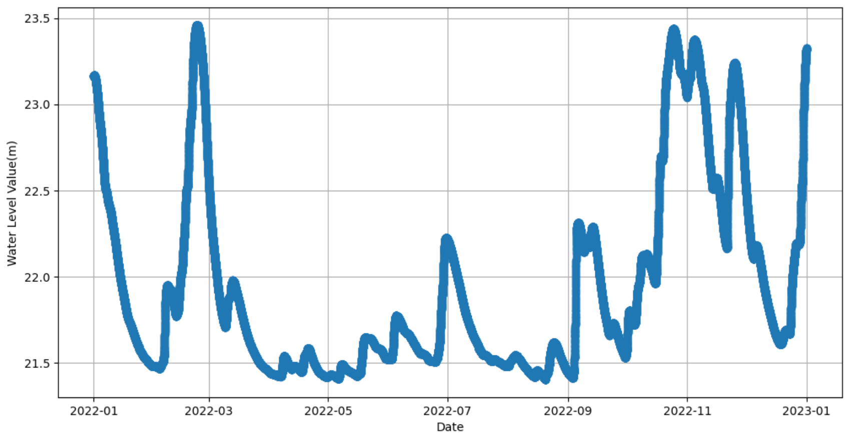

Figure 1.1: Plot of the water-level time-series data for 2022 (Aclinet hydrometric station in Ireland)

In contrast to equidistant time series data, "unequidistant" or "irregularly spaced" time series data refer to datasets where the observations are not uniformly distributed in terms of time intervals. However, it is essential to note that the time intervals between consecutive data points can exhibit variability, and it is possible for there to be irregular time points with gaps or missing data. Working with unequidistant time series data can pose greater complexity due to the presence of irregular time intervals, which can complicate various aspects such as time-based calculations, trend analysis, and forecasting. The analysis and modeling of such data may necessitate the utilization of specialized methodologies to address the irregular time grid and instances of missing data points. To manage non-equidistant time series data and missing points, we employed data aggregation to summarize and consolidate data over specified intervals. Additionally, we used interpolation to align data from different datasets to the same time intervals and fusion techniques to combine water level and weather condition data at the same time and location. These methods improved the consistency and usability of the dataset.

Instances of non-equidistant time series can be observed in water flow data provided by the OPW as displayed in Table 1.2. The OPW does not engage in continuous recording of river flow but, instead, relies on estimations derived from ratings (stage-discharge relationships) and recorded water levels. This estimation approach may lead to an inadequate time series dataset due to the absence of continuous and direct flow measurements.

Table 1.2: Water-flow time-series data for 2022 (Aclinet hydrometric station in Ireland)

Date	**Value**
1996-01-01 00:00:00	9.506
1996-01-01 04:15:00	9.788
...	...
1996-12-31 11:34:00	3.507
1996-12-31 20:10:00	3.000

In this thesis, we mainly focus on hydrology time series data sets; however, it's vital to acknowledge that the process of aggregating multiple datasets introduces an additional layer of complexity. The resulting time series may deviate from its equidistant nature, particularly when dealing with temporal irregularities in the original data. This potential transformation from equidistant to non-equidistant data necessitates careful consideration, as it can significantly influence the outcomes of subsequent analyses and the interpretations derived from them. Therefore, the choice of aggregation methods and the resulting data structure hold notable implications for this research.

1.2.2 Time Series Prediction

Time series prediction is a critical area of study with diverse applications in fields such as finance, climate science, and environmental monitoring. In this thesis, we embark on an in-depth exploration of methodologies and challenges associated with time series prediction, with a particular emphasis on environmental time series data, specifically in the domain of hydrology.

Time series prediction involves examining and modelling sequential data points arranged in a temporal order [28]. The aforementioned data points have the potential to serve as observations that are systematically recorded at consistent intervals, such as daily temperature readings, hourly fluctuations in stock prices, or monthly sales figures. Time series prediction mainly aims to make projections about forthcoming values, trends, or patterns by utilising past observations [36]. The importance of accurate time series prediction cannot be overstated in various fields, such as finance [133], economics [70], weather forecasting [212], and environmental science [77].

A key methodology in environmental science is the study of changing factors over time, either through repeated observations or autonomous sensor networks providing consistent data streams. As such, environmental science and time series analysis are intrinsically linked. Examples of applications of time series analysis within environmental science include: predicting ecological events using sensor measurements and identifying pollutant-related patterns in air quality monitoring. The field of time series prediction encompasses a range of methodologies, spanning from traditional statistical models to contemporary machine learning algorithms [89]. These methodologies frequently encompass identifying patterns, seasonality, and trends within the dataset and subsequently utilising this information to generate predictions. Furthermore, selecting suitable prediction models and interpreting results are significantly influenced by domain-specific knowledge and expertise [63].

1.2.3 Time Series Prediction Challenges

Predicting future values in time series data involves several challenges, particularly when dealing with temporal irregularities and complex seasonality or trends. These challenges can significantly complicate the modelling process and impact the accuracy of predictions. Time series prediction is a complex and multifaceted task, presenting several formidable challenges that demand careful consideration. These challenges extend across various domains and are fundamental considerations in numerous applications [140, 173]. These challenges are fundamental to time series prediction across various domains and underline the importance of addressing them for accurate and reliable predictive modelling. This thesis aims to investigate different methodologies and challenges related to time series prediction, with a specific focus on environmental time series data, particularly in the field of hydrology. In many environmental applications, time series data manifest complex patterns characterized by seasonality and trends. Environmental data, including river water levels, serves as a poignant example of this complexity due to the influence of various factors such as weather patterns, seasonal changes, and other environmental variables.

Temporal Irregularities

Time series data frequently exhibit temporal irregularities. Temporal irregularities occur when data points in a time series are not spaced at consistent time intervals. This can result from various factors such as missing data, variable sampling rates, or external disruptions. Irregularly spaced time series data introduces complications in applying standard analytical methods that assume uniform intervals. To handle temporal irregularities, advanced techniques such as imputation methods, resampling strategies, and time series-specific models like state-space models or Gaussian processes are often employed.

Consider, for instance, a river flow time series dataset. Although measurements are scheduled to be recorded hourly, real-world disruptions can occur. Factors such as equipment maintenance, adverse weather conditions, or unforeseen issues may result in missing data points during specific hours. Precisely handling these irregularities becomes indispensable for achieving robust predictions. Even a single gap or anomaly can exert a substantial impact on predictive model performance [62] [98] [128] [156].

In this research, we adopt a proactive approach by systematically checking for temporal irregularities and thoroughly examining the data for missing values, outliers, and inconsistencies in time intervals. By implementing rigorous data validation procedures, we aim to ensure the integrity and quality of our dataset, thereby bolstering the reliability and accuracy of our predictive models.

Complex Seasonality and Trends

Time series data often exhibit seasonality and trends that can vary in complexity. The complex seasonal and trend patterns observed in time series data, such as river water levels, pose challenges for machine learning because traditional modelling approaches may struggle to capture and account for these intricate patterns, potentially leading to inaccurate predictions [195]. Seasonal variations can introduce nonlinear, time-dependent relationships that standard machine learning algorithms might not adequately address without specialised techniques like seasonal decomposition or recurrent neural networks. Additionally, the presence of multiple influencing factors, such as weather patterns and environmental variables, further complicates the modelling process and requires sophisticated feature engineering and model selection to ensure dependable predictions. In many applications, time series data manifest complex patterns characterised by seasonality and trends. Environmental data, including river water levels, serves as a poignant example of this complexity due to the influence of various factors such as weather patterns, seasonal changes, temperature, The depth of the river, soil and other environmental variables [2].

For instance, the depth of the river and the amount of silt and soil deposits are crucial factors that significantly influence river water levels and patterns. These variables contribute to the overall complexity of the time series data and further complicate the modelling process. In river systems, the depth of the river directly affects the volume and flow rate of water, which in turn impacts water levels over time. Moreover, silt and soil deposits can alter the riverbed's topography, affecting the flow of water and potentially leading to changes in water levels.

Another example to mention is that during the spring months, river levels surge significantly due to snowmelt or increased rainfall, while in the dry summer months, levels recede. hence, the river water level dataset can exhibits distinct seasonal patterns.Accurately modelling and effectively incorporating these intricate seasonal and trend patterns into predictive models is paramount for attaining dependable predictions.

In summary, while time series prediction is a powerful tool for forecasting future values based on historical data, it requires careful consideration of temporal irregularities and complex seasonal/trend patterns. Advanced modelling techniques and robust data processing methods are essential for handling these challenges and achieving accurate predictions.

Multisource Heterogeneous Time-Series Fusion

Data collected for various applications often emanate from multiple sources, resulting in heterogeneity and diverse data structures [124]. Different types of sensors, devices, or data collection methods may lead to variations in sampling periods, frequencies, and information content. Temporal data from multiple datasets rarely occur with the same frequencies or granularity, leading to a bespoke integration methodology. Efficientl merging such multisource heterogeneous data is critical to achieving accurate predictions.

Low-Quality Time-Series Processing

The assurance of high-quality data is not always guaranteed, as environmental conditions, equipment failures, sensor malfunctions, and data transmission errors can introduce various issues [111]. These challenges may encompass missing data, unfilterable noise, or the presence of inaccurate and incomplete values. Effectively processing noisy and missing data is vital to enhance data quality for subsequent analysis and prediction.

Complex Time-Series Representation

In many applications, time series data become increasingly complex, characterised by high throughput, multidimensionality, nonlinearity, and non-Gaussian distributions [109]. Extracting meaningful insights from such complex data poses a formidable challenge. Traditional statistical and signal processing methods may struggle to capture the temporal and spatial correlations present in raw sensor data. The development of high-performance representation learning methods is essential to address these challenges [167]. While complexities in time series data are common, they are particularly challenging in river water level prediction. In this research, we emphasize the need for advanced representation learning methods to effectively handle these complexities, ultimately improving the accuracy of our predictions for better water resource management.

Time-Series Distribution Shift Alignment

Time-Series Distribution Shift Alignment" refers to the management and adjustment of variations or shifts in the distribution of time series data over time. For example, fluctuations in operating conditions, changes in data availability, or shifts in environmental factors can all lead to deviations from the assumption of data being independent and identically distributed. By addressing these shifts through appropriate methodologies, such as data normalization or adaptation techniques, we aim to ensure the robustness of predictive models in handling evolving data distributions. This concept is central to our investigation of time series prediction methodologies, particularly in the context of environmental time series data analysis, such as in hydrology. Operating conditions in various processes can exhibit high variability, challenging the assumption of data being independent and identically distributed [93]. Datasets availability may also vary, and security concerns, along with the cost of accessing data, can pose challenges. Furthermore, differences in operating environments and equipment conditions may lead to expensive, unlabeled, and imbalanced data. Managing distribution shifts in time series data is crucial to ensure the robustness of prediction models.

1.3 Proposed Solution: Metadata learning

In the preceding section, we discussed the challenges associated with model selection for time series data. We emphasised that time series data come in various forms and structures, and selecting a single, universal model to represent all types of time series data is often impractical. Each set of features within a time series dataset may necessitate a different type of machine learning model to effectively capture its underlying patterns and behaviors. As a result, the traditional approach of finding a one-size-fits-all model is not suitable for the complex and diverse world of time series data analysis. This limitation prompts the need for alternative approaches that can adapt to the unique characteristics of each dataset and select the most appropriate model accordingly. One promising avenue for addressing this challenge is meta-learning

1.3.1 Meta-Learning Definition

Meta-learning, often referred to as "learning to learn," is a subfield of machine learning and artificial intelligence that focuses on developing models or algorithms capable of learning and adapting quickly to new tasks or domains. A key issue in machine learning is in determining the optimal model for a given problem and dataset. In traditional data science a researcher will train and evaluate a number of models and hyperparameter configurations on a given dataset to evaluate which model provides the optimal performance. As datasets evolve or new datasets are presented to the system this process repeats to identify a "new" optimal model. This process is expensive in both time and resources utilised, as such, meta-learning has evolved as a subfield within machine learning to overcome these issues [200], [197], [175], [18], [171].

Unlike traditional machine learning, where models are trained for specific tasks, meta-learning aims to create models that can generalise their learning experiences across a wide range of tasks. In essence, meta-learners learn how to learn. Meta-learning can also be applied to the domain of model selection [99]. In this context, meta-learning focuses on developing techniques and algorithms that help automate and optimise the process of selecting the most suitable machine learning or statistical model for a given task or dataset.

In this context, meta-learning involves the following steps:

- **Meta-Feature Extraction:** The first step in meta-learning for model selection is the extraction of meta-features from the dataset. These meta-features characterise the dataset's properties, such as its size, dimensionality, skewness, and other statistical attributes. Meta-features serve as descriptors that help in understanding the dataset's characteristics.

- **Meta-Dataset Creation:** Meta-features from various datasets are collected to create a meta-dataset. Each instance in the meta-dataset corresponds to a specific dataset, and the associated label indicates the best-performing model or algorithm for that dataset. These labels are obtained through experimentation or benchmarking.

- **Meta-Learner Training:** The meta-dataset is used to train a meta-learner. The meta-learner's objective is to learn patterns or rules that relate the dataset's meta-features to the optimal model or algorithm choice. This training phase aims to capture the relationships between dataset characteristics and model performance.

- **Meta-Model Selection:** Once the meta-learner is trained, it can be employed to automatically select the most suitable model or algorithm for a new, unseen dataset. When presented with a new dataset, the meta-learner uses the extracted meta-features to predict which model is likely to perform best. This automates the model selection process.

- **Hyperparameter Tuning:** In addition to model selection, meta-learning can also be extended to optimise hyperparameters for the selected model. The meta-learner can predict both the model choice and the hyperparameters that are likely to yield the best results for a given dataset.

Benefits of Meta-Learning for Model Selection

The benefits of meta-learning for model selection offer valuable insights into the efficacy and efficiency of employing such techniques in predictive modeling.[163] Meta-learning reduces the need for manual experimentation and trial-and-error in selecting the appropriate model for a specific dataset. It automates the process and can save considerable time and computational resources. Meta-learning models can generalize from the patterns they have learned across various datasets. This means that they can make informed decisions for new datasets, even if those datasets differ significantly from those used during training. As the number of available machine learning models and algorithms continues to grow, meta-learning provides a scalable approach to model selection. It can adapt to a wide range of choices without the need for manual intervention. Meta-learning can help mitigate the risk of selecting suboptimal models, which can be especially important in critical applications where model performance is crucial.

In summary, meta-learning for model selection leverages machine learning techniques to automate the process of choosing the right model or algorithm for a given dataset. It enhances efficiency, generalisation, scalability, and robustness in the model selection process, making it a valuable tool in machine learning and data science.

Applications of Meta-Learning for Model Selection

Meta-learning for model selection has been applied in various domains and applications to automate and optimise the process of choosing the most appropriate machine learning or statistical model [112]. Here are some examples of how meta-learning has been applied elsewhere:

1. **Computer Vision:** In computer vision tasks, such as object recognition and image classification, meta-learning has been used to automatically select the most suitable convolutional neural network (CNN) architecture and hyperparameters for a given image dataset. This approach improves the efficiency and accuracy of image classification systems [139].

2. **Natural Language Processing (NLP):** In NLP applications, such as text classification and sentiment analysis, meta-learning helps in selecting the optimal text classification model, such as recurrent neural networks (RNNs) or transformer models like BERT. It also assists in fine-tuning hyperparameters to improve text classification accuracy [220].

3. **Recommendation Systems:** In recommendation systems, meta-learning is used to determine the most effective recommendation algorithm for a specific user or item dataset. It optimises the selection of collaborative filtering, content-based filtering, or hybrid recommendation models [49].

4. **Anomaly Detection:** In anomaly detection tasks, where the goal is to identify unusual patterns or outliers in data, meta-learning helps in selecting the most suitable anomaly detection algorithm and setting appropriate thresholds based on characteristics of the data [53].

5. **Healthcare:** Meta-learning is applied in healthcare for disease diagnosis and patient risk prediction [221]. It assists in selecting the best machine learning model for medical image analysis, clinical data interpretation, or patient outcome prediction based on different medical datasets.

6. **Finance:** In financial applications, meta-learning aids in selecting predictive models for stock price forecasting, credit risk assessment, and fraud detection [5]. It identifies the most effective algorithms for handling financial data with varying characteristics.

7. **Robotics and Autonomous Systems:** Meta-learning has been used in robotics and autonomous systems to select control policies or motion planning algorithms based on the robot's environment and task requirements. It helps robots adapt to different scenarios effectively [55].

8. **Energy Management:** In energy management applications, such as energy consumption prediction or load forecasting, meta-learning aids in selecting

regression models or time series forecasting algorithms to optimise energy resource allocation and consumption predictions [123].

9. **Environmental Monitoring:** Meta-learning can optimise the selection of models for environmental monitoring tasks, such as river water level prediction or weather forecasting. It considers the geographical location, historical data, and specific environmental factors to choose the most suitable forecasting model [58].

10. **Time Series Forecasting:** Meta-learning can be applied to time series forecasting, where it selects the best forecasting model (e.g., ARIMA, LSTM, or Prophet) based on the characteristics of the time series data. It also helps in setting appropriate hyperparameters for these models to achieve accurate predictions [121].

In each of these domains, meta-learning techniques automate and enhance the model selection process, making it possible to adapt to different datasets and tasks efficiently. This approach saves time and resources while improving the performance and accuracy of machine learning applications.

In the field of hydrology, with a specific focus on water level prediction, the challenge of model selection is prominent. River water level data are highly diverse, influenced by seasonal fluctuations, geographical factors, and meteorological conditions. These complexities result in intricate, nonlinear patterns, making it difficult to create a single, reliable predictive model for various river systems. Traditional modelling approaches, designed for specific datasets, often struggle to generalise effectively. They may fall short in providing accurate predictions due to the diverse nature of river water level data. The conventional practice of evaluating numerous models, each customised for a particular dataset or river system, is resource-intensive and impractical given the wide-ranging variations in the data. Here, meta-learning offers a promising solution. It introduces adaptability and efficiency to the model selection process. By employing meta-learning techniques, we can automate and optimise model selection. Meta-learners, adept at quick adaptation to new data, enhance their predictive performance by leveraging insights gained from diverse training datasets.

In summary, in the field of hydrology, especially in water level prediction, meta-learning stands out as an effective strategy for addressing the challenge of model selection. Its ability to enhance adaptability and selection aligns seamlessly with the nuances of water level prediction, promising more accurate and reliable results in this critical domain.

1.4 Thesis Contribution and Research Questions

This study embarks on a comprehensive exploration within the domain of enhancing time series prediction in environmental research, characterised by a multifaceted interplay of methodologies, data integration, and model selection. With a specific emphasis on the adoption of meta-learning methods, this research pursues a tripartite objective: Firstly, to dissect the intricacies associated with environmental data integration, a foundational component for informed decision-making; secondly, to examine the intricate challenges intrinsic to time series prediction, with a distinct focus on the hydrology domain; and lastly, to investigate applicability of the meta-learning techniques in model selection for time series predictions.

Question 1: How can we develop a robust and unique approach to integrating heterogeneous environmental datasets, considering variations in spatiotemporal resolution, data format, and attribute representation?

Robust environmental data integration is a critical component in improving the accuracy and reliability of time series predictions, especially when dealing with heterogeneous ecological datasets, including those in the field of hydrology. To address this, we propose a robust system adept at spatiotemporal integration, seamlessly merging multidimensional data from diverse sources to offer a comprehensive view that spans time and space. This multidisciplinary approach, which joins the sophistication of IT and the practicality of hydrology, is crucial for maintaining research integrity, guiding informed decisions, and assuring quality.

Question 2: What are the main challenges of developing accurate predictions for environmental time series data, particularly within the hydrology domain?

Whether traditional statistical models or modern machine learning algorithms, the choice of prediction models significantly impacts the accuracy of time series predictions. To address this challenge, this research explores machine learning algorithms for water level prediction in multiple locations within seven river basin districts in Ireland. Utilising historical water level data, we strive to establish reliable predictive models. Detailed methodologies and the results of the comparative analyses will be presented in subsequent chapters. The main research focus areas are developing and optimising deep learning methods, to improve river water level prediction accuracy. Through the innovative development of model architectures and data integration strategies tailored to the distinctive challenges inherent in river water level forecasting, this research endeavors to make substantive contributions to the continual evolution of predictive methodologies within this domain.

Question 3: Can meta-learning techniques be effectively applied to address the model selection problem for environmental time series data?

Environmental data differs from other sensor data due to its diverse variables, spatial and temporal variability, complex interactions, susceptibility to external factors, challenges in data quality and availability, and high stakes for stakeholders. These factors make predicting outcomes from environmental data particularly challenging and crucial for decision-making.

The application of meta-learning techniques can potentially improve the reliability of time series predictions, especially in the context of environmental data, with a particular focus on hydrology. We explore the potential of meta-learning as a valuable tool for addressing the complex challenges intertwined with the selection of an optimal prediction model. We conduct an in-depth analysis of the limitations and intricacies inherent in the process of identifying the most appropriate model for time series data. In response to these challenges, we introduce meta-learning as a compelling and advantageous solution. Through this exploration, we aim to shed light on the significant contributions that meta-learning can make in enhancing the model selection process.

In summary, by addressing these research questions and hypotheses and combining insights from environmental data integration, time series prediction, and meta-learning techniques, we aim to provide valuable insights into the utility and viability of meta-learning as a tool for enhancing the predictive capabilities of models in the complex and critical domain of river water level forecasting.

1.5 Thesis Outline

Chapter 2 offers a comprehensive exploration of previously discussed research related to data integration, machine learning for time series prediction and meta-learning. Our review covers environmental data integration, machine-learning techniques for hydrology predictions and meta-learning for time series model selection.

Chapter 3 provides the spatiotemporal methodologies for environmental data integration. We begin by introducing the chapter and then delve into the methodology and architecture required for this type of data integration. This includes data sourcing, extraction, transformation, mapping, and integration, focusing on spatial and temporal aspects. We also evaluate environmental case studies, showcasing real-world applications of the Spatio-Temporal Environmental Data Integration (STDI) approach. The chapter concludes with a summary of the findings.

Chapter 4 of this thesis is dedicated to evaluating machine learning approaches for river water level predictions. We introduce the chapter and then discuss the study area and the data used for the experiments. The Proposed methodology involves data collection, preprocessing, model development, and model evaluation.

In Chapter 5 we focus on meta-learning approaches for time series model selection. We start by addressing the challenges associated with model selection in hydrology, covering traditional approaches and highlighting the limitations in handling physical processes and uncertainties. We then delve into time series prediction, discussing time series features and model selection methods.

The thesis concludes with a summary in Chapter 6 that revisits the key points and contributions made by the research journey. We provided an overview of the entire thesis and its findings. Additionally, we discuss the implications of this work and how it advances the field of environmental data integration and hydrology predictions. Finally, we outline potential areas for future research.

Chapter 2

Literature Review

The primary problem addressed in this research is the prediction of hydrological time series data, specifically river water levels, to improve flood prediction, water resource management, and environmental monitoring. Accurate predictions are crucial for mitigating the impacts of floods, managing water resources efficiently, and ensuring environmental sustainability.

To develop a meta-learner capable of addressing this problem, we draw upon state-of-the-art research from multiple domains. This requires extensive knowledge of data engineering and data integration to construct a suitable dataset from heterogeneous sources. Subsequently, advanced machine learning techniques are employed to identify candidate models and develop a baseline dataset, which serves as the foundation for training a meta-learning system. Finally, the meta-learner is trained and evaluated to determine its efficiency and effectiveness in predicting hydrological time series data.

To conduct the literature review, we employed a systematic approach, ensuring comprehensive coverage of relevant studies and methodologies. We identified key databases and sources, including Google Scholar, IEEE Xplore, ScienceDirect, and SpringerLink, using specific keywords and phrases such as "environmental data integration," "machine learning for hydrology," "time series prediction," and "meta-learning for model selection." We screened articles based on their abstracts and titles, prioritizing recent publications (from the last 10 years) and including older seminal works where necessary. Detailed reviews focused on methodologies, data sources, model evaluation techniques, and key findings, critically analyzing each study's contributions, limitations, and applicability to our research.

The integration of multiple time series in this research is crucial for enhancing the prediction of hydrological data, specifically river water levels. The primary goal is to improve flood prediction, water resource management, and environmental monitoring. While predicting river water levels is inherently a time series problem, relying solely on historical river water levels may not capture the full complexity of the factors influencing these levels. Therefore, we integrate various time series from different sources to create a more robust and comprehensive predictive model. In addition to using previous river water levels, we incorporate other relevant variables such as temperature, humidity, precipitation, and real-time air and water quality measurements. These variables come from diverse data sources, including remote sensing platforms, weather stations, ground-based sensors, and citizen science initiatives. By combining these multiple time series, we can account for the various environmental factors that affect river water levels, thereby improving the accuracy and reliability of our predictions. This integrated approach allows us to address the complex and dynamic nature of hydrological systems more effectively than using a single time series alone.

Additionally, this chapter provides a comprehensive literature review structured to ensure a thorough understanding of the problem, the current state of research, and the innovative methodologies employed in this study to advance the field of hydrological predictions. Section 2.1: examines environmental data integration approaches, with a focus on spatiotemporal integration, to address the challenges of merging diverse environmental datasets. Section 2.2: delves into the application of machine learning techniques for hydrology predictions, highlighting various models and their performance in predicting river water levels. Furthermore, Section 2.3: discusses the application of meta-learning for model selection, exploring how meta-learning can enhance the accuracy and reliability of time series predictions by automating the model selection process. This structured approach ensures a thorough understanding of the problem, the current state of research, and the innovative methodologies employed in this study to advance the field of hydrological predictions.

2.1 Environmental Data Integration

Environmental management relies heavily on data collection, analysis, and integration. In this digital age, data-driven approaches, particularly in Machine Learning and Information Technology, are pivotal in addressing critical environmental concerns, assessing outcomes, and predicting future trends across various scientific domains [110] [209]. Environmental datasets are critical in understanding and addressing complex environmental phenomena and issues like climate change, floods, and natural resource management. The increasing availability of diverse environmental data from various sources, including sensor networks, field sampling and citizen science initiatives, offers unprecedented opportunities for advancing understanding of environmental processes and interactions [129]. However, these datasets often exhibit significant heterogeneity in terms of spatiotemporal resolution, data format, and attribute representation, posing significant challenges for effective data integration [185]. Effective integration of environmental data is crucial for accurate analysis, informed decision-making, and the development of sustainable policies [218]. While existing approaches for integrating environmental data, such as data harmonisation and fusion techniques, have made strides in addressing data heterogeneity [147], coping with the inherent complexity of environmental data remains challenging, especially when dealing with data represented by points or polygons [29]. This limitation impedes the development of comprehensive environmental models and analyses, highlighting the pressing need for innovative integration methods to navigate these complexities effectively [75]. Hence, the primary motivation for creating an environmental data integration system is to address the critical need to efficiently manage and analyse environmental data.

2.1.1 Data Sources

As we undertake a more thorough examination of the complexities involved in environmental data integration, it becomes evident that accessing reliable and diverse sources of information is paramount. To achieve a comprehensive understanding of complex environmental phenomena, we must turn this research attention to the critical data sources available from remote sensing platforms, as exemplified by the Earth Observing System by NASA and the Copernicus program by the European Space Agency, which provide indispensable datasets for monitoring land cover, vegetation health, and atmospheric conditions [127] [60].

Furthermore, weather stations are crucial in climate research as they provide essential data pertaining to temperature, humidity, precipitation, and other relevant variables [57]. Ground-based sensors are crucial in providing real-time air and water quality measurements, which are indispensable for conducting environmental assessments [199]. Geographic Information Systems (GIS) play a crucial role in spatial analysis and modelling by integrating geospatial data [72]. Government agencies, such as the Environmental Protection Agency (EPA), curate extensive datasets on air and water quality [199]. Including scientific research conducted by academic institutions enhances the data landscape. Citizen science initiatives offer significant contributions by collecting data from many individuals on diverse environmental parameters [47]. Continuous and high-resolution data can be obtained through remote sensors and Internet of Things (IoT) devices [22]. Numerical models generate simulated data for validation and forecasting [138]. Using historical records and archives contributes to comprehending enduring patterns over an extended period [157]. Social media and web sources can capture real-time public perceptions regarding environmental issues [177]. Environmental and biological surveys play a significant role in providing valuable data about species, population dynamics, and biodiversity [76].

Incorporating these various data sources is crucial in tackling urgent environmental issues comprehensively [75].

2.1.2 Data Types, Formats and Quality

Within the task of environmental data integration, the presence of various data types and formats offers both advantageous prospects and obstacles, thereby expanding upon the foundation of previously examined data sources. Researchers are faced with diverse data types in their research tasks. These include numerical measurements, such as temperature and pollutant concentrations; categorical data, such as land use classes and species presence; temporal data, such as time-series measurements and climate records; and geospatial data, such as GPS coordinates and remote sensing imagery. These data are collected from multiple sources, such as satellites, weather stations, and ground-based sensors.

The previously mentioned data types are commonly stored in many formats, including spreadsheets, databases, geospatial files (e.g., shapefiles), and text documents, which accurately represent the heterogeneous sources of the data. Moreover, it is common for environmental data to possess multidimensional features, requiring specialised data structures such as netCDF (Network Common Data Form) to manage intricate multidimensional datasets. The heterogeneity of data types and formats poses significant obstacles to achieving interoperability, demonstrating the importance of data harmonisation and standardisation in facilitating successful integration endeavours [213]. The utilisation of emerging standards, such as the Climate and Forecast (CF) metadata conventions for netCDF files [57] and open data protocols, such as those established by the Open Geospatial Consortium (OGC), play a crucial role in enabling the exchange and integration of data within the environmental research community. Furthermore, the utilisation of data modelling methodologies, such as the Observations and Measurements (O&M) model developed by the Open Geospatial Consortium (OGC), facilitates the organisation of environmental data to enable smooth integration and analysis [48].

Ensuring data quality is essential in environmental data integration. Attribute inconsistencies and data quality issues can lead to skewed or misleading conclusions. Solutions such as schema matching, data transformation, and data quality assessments have been proposed to maintain data consistency and reliability. However, these methods may require deep domain knowledge and scalability improvements for larger datasets. Addressing temporal and spatial variations in integrated datasets presents an ongoing research challenge [147] [52] [16].

2.1.3 Data Warehouse Concepts and Architecture

Nowadays, as mentioned in [17] research, datasets can be significant in some contexts. Therefore, they necessitate new technology solutions allowing storage, update, and efficient exploitation. Data Warehouses (DW) were developed in the 1980s to allow users to undertake Business Intelligence (BI) as an alternative to storing and organizing data in traditional databases.

2.1.3.1 Data Warehouses

When dealing with heterogeneous datasets, it is crucial to present a unified view of the data to ensure consistency and facilitate comprehensive analysis. This is where Data Warehouses (DWs) come into play, offering an effective solution to integrate diverse data sources and provide a coherent data environment.

Many DW definitions are available in the literature; according to [13], it is a collection of approaches and technologies that, when combined, give a systematic and pragmatic approach to solving the end-user problem of obtaining information that is dispersed across multiple systems inside an organisation. In [91], a Data Warehouse (DW) is defined as a data collection that is subject-oriented, integrated, evolving over time, non-volatile, and primarily utilised to facilitate decision-making processes. According to [106], a DW is a data source for an organisation formed by combining all appropriate data marts.

This study gives a generic view of DW phases and states that DW design should incorporate different approaches and solutions, such as data cubes. This study also defines six phases to build a DW (the DW Lifecycle): Data sources, staging area, integration patterns, DW, dimension construction, and data analysis tools.

The traditional operational database concept and the DW concept are different. Therefore, a brief comparison is needed to understand the DW concept and its applications better. Operational databases store the data required to run the organisation's daily activities. Users utilise them to register and execute predefined tasks. Therefore, their data may change when new organisational requirements emerge. This database does not require a vast storage capacity because there is no data redundancy, and no historical data is stored.

In contrast, a DW already holds the analytical data needed by management to make decisions. It is designed for vast amounts of data because it must support extensive analytical processes and long-term data storage. These capabilities are necessary to enable complex queries, generate comprehensive reports, and provide insights that drive strategic decision-making. As a result, the DW requires significant processing power and storage capacity to manage and analyze detailed and summarized data efficiently. This includes components such as ETL (Extract, Transform, Load) processes, metadata, data marts, and OLAP tools, all of which work together to facilitate efficient data management and analysis. Data in a DW are detailed and summarized, providing analytical views from operational datasets.

Additionally, DWs are multidimensional structures that integrate and consolidate information using different schemas and models [4]. Multidimensional schema was explicitly created for modelling data warehousing systems. The three primary types of DW schemas are the Star Schema, Snowflake Schema, and Galaxy Schema. The schemas are intended to meet the particular requirements of extensive databases used for analytical purposes or Online Analytical Processing (OLAP). OLAP is a subset of BI, and it is a technique used in computing that enables rapid response to multidimensional analytical (MDA) queries. OLAP is a data processing and control technique based on dimensional views. Multidimensional arrays are used as structures to store pre-calculated values in a given dimensional structure that takes a long time to build but reads quickly. The Snowflake model, for example, is utilised in Online Analytical Processing (OLAP) in a BI application [37] [106].

Multidimensional structures, such as data cubes, store information in the DW. Initially, the data cube was intended to use OLAP tools, making it easier to access multidimensional data and facilitate the interpretation of data [69]. It is particularly advantageous when expressing data with dimensions as specific measures of business requirements. Each cube dimension reflects a different database attribute, such as daily, monthly, or yearly precipitation. As a result, a data cube can aid in establishing trends and analyzing performance. This pre-computation allows users to pose previously unseen aggregate queries on a data warehouse efficiently. Since the data cube stores pre-calculated aggregate data, it significantly speeds up query response times for complex analytical queries. For example, management can quickly retrieve information on sales performance across different regions and time periods without having to perform complex calculations on the raw data. This enables faster decision-making and more effective analysis of large datasets.

Data cubes are primarily classified into two types [45]. The first is a multidimensional data cube. Most OLAP products are built on a multidimensional data cube structure. These multidimensional OLAP products (MOLAP) typically perform better than other approaches because they can be indexed directly into the data cube's design to gather large subsets of data. The second is Relational OLAP (ROLAP), which uses the relational database model. Compared to a multidimensional array, the ROLAP data cube is used as a collection of relational tables (roughly twice as many as the number of dimensions). Each of these tables is called a cuboid, representing a particular view. In addition, OLAP also has other types, such as HOLAP (Hierarchical-OLAP) [132]. The variations are related to the way data is stored for OLAP. It is true that the processing remains online analytical processing, yet the storage methodology is different.

Usually, A DW architecture comprises three main layers [174]. First, the data acquisition layer includes data sources (both internal and external). Data sources frequently utilise various systems and applications and a mix of network and relational data models to store the information they collect. Then, the data storage layer extracts the data from the various source systems. The information is then supplied to the staging area, where it is processed. Last, the data delivery layer prepares data that can be accessed via dynamic queries while maintaining a high efficiency level.

Data and information are always available in the DW, allowing strategic plans to be tested, evaluated, assessed, and monitored confidently. A DW is part of an overall data environment that serves as a single integrated data source for information processing that is subject-oriented, integrated, time-variant, and non-volatile [91]:

- Subject-oriented means that information in a DW is categorised around specific subjects, which differs from transactional systems that organise data according to business processes. The DW stores crucial information following its future use-case scenarios. The vital point is that defining what information is critical when building DWs is essential.

- Integration is a crucial feature of DWs. Integration defines a unique representation of data from many sources stored in the DW.

- Time variant means that DWs store historical data and include date-time as an important variant. Historical data is required to discern patterns and long-term relationships in a specified time frame [91]. Like the data, metadata has temporal characteristics. Hence, future changes will not alter the current meta-data status.

- Non-volatile means that data within the DW is not updated directly by the users; this ensures that all users work on the same data version. Therefore, users can only query data after the loading phases.

Data quality is a critical aspect of DWs, as it directly impacts the reliability and usability of the information stored. References to frameworks like OGC (Open Geospatial Consortium) [101] and INSPIRE (Infrastructure for Spatial Information in the European Community) [14] highlight the importance of standardization and data quality in geospatial datasets. A key consideration in maintaining data quality within a DW is deciding whether to address issues such as null values and outliers at the data integration stage or defer these adjustments until the analytics phase. This decision impacts both the design and the operational efficiency of the DW.

2.1.3.2 Data Integration Architectures: ETL and ELT

Data integration involves transferring data from source systems to a Data Warehouse (DW) using various architectural patterns. One critical distinction among these patterns lies in the timing of data transformations within the data-processing pipeline. Additionally, the technologies, algorithms, and optimisation methods employed in the process depend on the chosen pattern [92].

ETL (Extract-Transform-Load) [201] is the conventional data integration pattern involving transforming data before loading it into a DW. This method follows a structured three-step procedure: extraction, transformation, and loading. Initially, ETL collects vast amounts of raw data from diverse sources, transforms it into a unified predefined schema, and then loads it into the data warehouse. Typically, ETL serves as the primary method for importing data from one source to another. In a typical ETL process, data is gathered from various sources, staged, and subsequently loaded into the warehouse. Transformations occur before the data is loaded into a predefined schema. This results in a data repository highly optimised for analytical queries; however, the bespoke integration transformation steps coupled with a predefined schema mean that subsequent changes to the source dataset or the addition of new data sources require significant investment to change.

In contrast, the ELT (Extract-Load-Transform) [136] pattern involves loading data into a data lake first and then performing transformations. ELT offers a notable advantage regarding rapid data ingestion, making it suitable for use with NoSQL databases and data lakes when immediate data collection is necessary, regardless of raw data format or value. In a typical ELT process, data integration technologies push multiple data sources into the data lake, which is then transformed into a cube or data mart. Unlike ETL, ELT allows real-time data loading from source systems, and it does not require coding changes in response to alterations in the source column or data structure.

The choice between ETL and ELT depends on various factors, including data volume, system compatibility, and the desired speed of data ingestion. ETL focuses on transformation before loading, while ELT prioritises rapid data ingestion and post-load transformations. Understanding the distinctions between ETL and ELT is important for effective data management, as they influence how quickly data can be processed and made available for analysis. ETL is often preferred when complex data transformations are required before storage, ensuring data quality and consistency. In contrast, ELT is suitable for environments where quick data loading is essential, with the processing power of the destination system being leveraged for transformation tasks. ELT is also advantageous when requirements are unclear and subject to change, making it an ideal mechanism for data science applications.

2.1.4 Environmental Data Integration Challenges

Environmental data integration is essential for gaining insights into complex environmental and climatic phenomena. However, this process has its challenges. Outlined below are some common limitations associated with environmental data integration, as well as considerations related to ETL (Extract, Transform, Load) and ELT (Extract, Load, Transform) data integration processes:

1. Data Variety: Environmental data comes in various forms, including remote sensing imagery, climate models, field measurements, and textual reports. Integrating these diverse data types can be a complex task.

2. Data Quality: Ensuring the accuracy and reliability of environmental data is crucial for informed decision-making. Errors or inconsistencies in data can lead to flawed analyses.

3. Real-Time Data Needs: Environmental monitoring often requires real-time data integration to respond promptly to changing conditions, such as natural

disasters or pollution events.

4. Data Volume: Environmental data, especially high-resolution imagery and continuous sensor data, can result in massive data volumes. Managing, processing, and storing such data can be resource-intensive.

5. Data Source Diversity: Environmental data originates from numerous sources, including government agencies, research institutions, satellites, and local sensors. Integrating data from these diverse sources can be a significant challenge.

6. Spatial and Temporal Considerations: Environmental data frequently involves complex spatial and temporal dimensions. Integrating data with varying resolutions and timeframes requires specialised techniques.

7. Data Governance: Environmental data integration must adhere to data governance practices, including security and compliance with environmental regulations, which can be intricate due to the diverse sources and sensitive nature of the data.

8. Metadata Management: Maintaining metadata for environmental data is essential for understanding data lineage, context, and quality. Proper metadata management is critical for meaningful interpretation.

9. Interoperability: Achieving interoperability between various environmental data systems and standards is essential for seamless data integration and stakeholder collaboration.

10. Complex Transformations: Environmental data often requires complex spatial and statistical transformations to derive meaningful insights, adding complexity to integration processes.

11. Sustainability: Environmental data integration solutions should consider environmental sustainability, as large-scale data processing can have an ecological footprint.

Addressing these challenges in environmental data integration requires a holistic approach that combines advanced technologies, domain expertise, data governance practices, and ongoing monitoring. The ability to harness and integrate environmental data effectively is pivotal for making informed decisions related to climate change, natural resource management, and environmental conservation.

In response to data integration limitations, researchers have explored alternative strategies for data integration in environmental studies. These alternatives include data harmonisation, data fusion, and interoperability solutions. These approaches aim to align, transform, and amalgamate datasets, reducing the risk of information loss in the integration process [147].

Spatiotemporal data integration (STDI) is an emerging research domain that addresses the challenges of integrating data with temporal and spatial variations, particularly in the context of environmental data. STDI holds the promise of improving the understanding of environmental systems and enhancing the ability to predict human-induced environmental impacts. The integration of data from diverse sources is particularly critical in environmental monitoring to track temporal shifts in environmental conditions.

Several studies have utilised spatiotemporal data integration (STDI) for diverse applications. Some of these include monitoring water quality in lakes [215], tracking land cover and usage changes [83], and air quality monitoring [126]. These studies underscore STDI's potential for accurately detecting environmental changes and demand efficient data integration and management to facilitate the analysis and interpretation of diverse datasets from various sources [187]. Geographic Information Systems (GIS) have traditionally served as a primary tool for integrating such datasets [186]. However, these conventional methods have inherent limitations that pose significant challenges to the integration process. GIS-based integration techniques, while valuable, often require adjustments when dealing with complex temporal data or data from non-geospatial sources. Additionally, they may need help with seamless integration due to inconsistencies in attribute values and varying data structures [225]. Moreover, the computational and storage constraints of GIS systems can become bottlenecks when handling extensive and intricate datasets [208].

While STDI offers promising solutions for integrating environmental data with temporal and spatial variations, it is essential to acknowledge that challenges persist in achieving comprehensive and reliable data integration. Attribute inconsistencies and data quality issues remain significant hurdles that can lead to skewed or misleading conclusions in environmental studies, particularly within the context of STDI [52]. Various solutions have been proposed to address these challenges, including schema matching, data transformation, and data quality assessments[16]. However, implementing these solutions effectively within the realm of STDI can be particularly complex due to the unique characteristics of environmental data. Additionally, scaling these methods for handling more extensive and diverse datasets remains an ongoing research focus within the field of environmental data integration.

In the context of environmental research, the integration of diverse data sources and the challenges it presents lay the foundation for more advanced predictive modelling techniques. As the previous section explored the complexities of environmental data integration, the following section move into the realm of predictive modelling, explicitly focusing on machine learning algorithms and their applications in hydrology predictions. This section illustrates how overcoming data integration challenges can lead to more informed and accurate predictions, ultimately contributing to a better understanding of complex environmental systems.

Advancements in spatiotemporal data integration for environmental research have enabled a more comprehensive understanding of complex environmental processes and interactions. These advancements are crucial for addressing critical environmental concerns, such as climate change, floods, and natural resource management. Diverse data sources, including remote sensing platforms, weather stations, ground-based sensors, GIS, government agencies, citizen science initiatives, IoT devices, numerical models, historical records, social media, and biological surveys, have provided an extensive data landscape for researchers to work with. Furthermore, the development and adoption of data standards, such as netCDF and OGC protocols and data modelling methodologies like Observations and Measurements (O&M), have improved data harmonisation and standardisation.

However, several limitations persist in spatiotemporal data integration for environmental research. These limitations include data quality issues, data volume and scalability challenges, heterogeneity of data sources, the modelling of complex spatial and temporal relationships, interoperability difficulties, data privacy and security concerns, and the need for substantial computational resources. These challenges can impede the efficient integration and analysis of environmental data, especially when dealing with large and complex datasets. Nevertheless, researchers and practitioners in this field continue to innovate and develop solutions to address these limitations, as spatiotemporal data integration remains pivotal for informed decision-making and sustainable environmental policies.

2.2 Contemporary Approaches in Hydrological Predictive Modeling: Traditional and Machine Learning Techniques

In this section, we will thoroughly investigate hydrological predictive modeling, focusing on both traditional and machine-learning algorithms for hydrology predictions. This analysis will address prediction challenges and the complexities involved in model selection. Hydrology, the scientific study of water movement and distribution on Earth's surface, plays a crucial role in water resource management and natural disaster mitigation, such as flood prevention. Selecting the appropriate models is essential to achieving these objectives. However, the field of hydrology lacks a universally applicable modeling framework due to the complex and dynamic nature of hydrological systems. Consequently, no single solution can be universally applied [80], [21], [21].

This section will explore the challenges of model selection in hydrology and emphasize the importance of adopting diverse methodological approaches.

2.2.1 Traditional Hydrological Predictive Models

Hydrology predictions involve forecasting various aspects of the water cycle, including river discharge, rainfall, and water quality. Accurate predictions in hydrology are essential for various applications, such as flood forecasting, drought monitoring, and water resource management [125] [33] [181].

Time series data in hydrology often exhibit complex patterns and dependencies, making traditional statistical methods less effective. Machine learning algorithms, such as artificial neural networks (ANNs), support vector machines (SVMs), and decision trees, have shown promise in capturing these intricate relationships [134] [43] [50]. ANNs, in particular, have been extensively used in hydrology due to their ability to model non-linear relationships and handle high-dimensional data [34] [108]. SVMs have also gained popularity for handling non-linear data and providing robust predictions [15] [67] [134]. While less complex, decision trees can still offer valuable insights into hydrological processes [108] [44].

Feature selection and engineering play a crucial role in improving the performance of machine learning models in hydrology predictions. Selecting relevant hydrological attributes and transforming data to highlight meaningful patterns are essential steps in model development [219] [11]. Additionally, ensemble learning techniques, such as random forests and gradient boosting, have enhanced prediction accuracy by combining the outputs of multiple base models [54] [191] [183].

Various research efforts have been dedicated to developing water level prediction systems across different types of river basins. Factors such as the contextual needs of the forecast, availability and relevance of historical data, required level of accuracy, and time constraints play a crucial role in selecting an appropriate forecasting model [115] [151].

Traditional hydrodynamic methods have been extensively utilised, focusing on differential equations to describe the physical processes involved in water movement accurately [104] [154]. Physical models, primarily of the distributed type, aim to capture the intricacies of the hydrological cycle. However, their application on a larger scale often proves to be computationally expensive and time-consuming [146].

Table 2.1 captures the summary of some traditional hydrological predictive models, including their descriptions, advantages, disadvantages, and references.

Table 2.1: Summary of Traditional Hydrological Predictive Models

Model	**Description**	**Advantages**	**Disadvantages**	**References**
Artificial Neural Networks (ANNs)	Machine learning algorithms that model non-linear relationships and handle high-dimensional data	Ability to model non-linear relationships, effective with high-dimensional data	Requires large datasets, risk of overfitting	Chang (2012), Kisi (2008)
Support Vector Machines (SVMs)	Algorithms that handle non-linear data and provide robust predictions	Robust with non-linear data, good generalization ability	Computationally intensive, requires careful parameter tuning	Basheer (2000), Govindaraju (2000), Maier (2003)
Decision Trees	Simple algorithms that can offer valuable insights into hydrological processes	Simple to interpret, fast computation	Prone to overfitting, less effective with complex data	Kisi (2008), Chui (2007)

Table 2.1 continued from previous page

Model	Description	Advantages	Disadvantages	References
Ensemble Learning (e.g., Random Forests, Gradient Boosting)	Combines outputs of multiple base models to enhance prediction accuracy	Improved accuracy, reduces overfitting	Computationally expensive, complex to implement	Duan (2012), Sun (2015), Solomatine (2008)
Traditional Hydrodynamic Methods	Focus on differential equations to describe physical processes in water movement	Accurate physical process modeling	Computationally expensive, time-consuming for large-scale applications	Kaya (2019), Paiva (2011)
Physical Models (Distributed Type)	Aim to capture the intricacies of the hydrological cycle	Detailed process representation	High computational cost, time-consuming	Nagatani (2012)
Feature Selection and Engineering	Involves selecting relevant attributes and transforming data to highlight meaningful patterns	Enhances model performance, highlights important patterns	Requires expert knowledge, time-consuming	Yaseen (2018), Azarnivand (2020)

Table 2.1 continued from previous page

Model	**Description**	**Advantages**	**Disadvantages**	**References**
Water Level Prediction Systems	Systems developed for predicting water levels across different river basins	Contextual and specific to river basins	Varies with context, dependent on historical data	Kure (2009), Nourani (2015)

2.2.2 Challenges in Hydrological Model Predictions

The variability of hydrological data is well-known and is influenced by factors such as precipitation patterns, geographical characteristics, and climate conditions [203] [196]. These seminal studies underscore the multifaceted nature of hydrological variability and its profound sensitivity to a plethora of determinants, ranging from climatic variations and precipitation patterns to anthropogenic activities, such as irrigation practices and groundwater management. Recognising and comprehending these multifarious complexities assumes paramount importance in the development of precise and adaptable hydrological models equipped to address the multifarious challenges posed by this intrinsic variability.

Drawing from the work of [203], an essential concern emerges pertaining to the sustainability of global water resources. This research highlights that one-third of the world's lowland regions, equipped for irrigation purposes, presently reside in locales heavily reliant on runoff contributions originating from mountainous regions. Simultaneously, these regions exhibit unsustainable utilisation of local blue water resources. Projections indicate that this precarious scenario is poised to escalate, with anticipated figures surpassing the 50% threshold in forthcoming decades. These revelations emphatically underscore the exigency of effective water resource management, particularly in mountainous terrains, and advocate vehemently for the preservation of such regions as an integral facet of the broader tapestry of sustainable development. Within the context of hydrological data variability, this research illuminates the pivotal role played by mountainous regions in the realm of water resource management and sustainable development, thereby accentuating the intricate dynamics encompassing global hydrology.

Furthermore, [196] work augments the comprehension of hydrological variability. This research introduces a large-scale hydrological modelling approach with specific attention to evaluating the impact of irrigation practices on hydrological processes, encompassing facets such as evapotranspiration and groundwater recharge across both irrigated and non-irrigated regions. A salient lesson derived from this study lies in the imperative necessity of accounting for localised water management practices, including groundwater pumping, as instrumental contributors to the variegated tapestry of hydrological data variability.

2.2.2.1 Time Series Features

[166] observed that the performance of prediction methods varies depending on the characteristics of the data, which has spurred further investigation into these variations. Building on this idea, [85], [118], and [10] proposed that understanding the characteristics of a time series can provide valuable insights for selecting appropriate prediction methods. Instead of directly working with individual time series observations, we advocate for the analysis of time series through the lens of features that define an "instance space."

Time series features represent quantifiable attributes of time series data. For instance, we can compute the strength of seasonality and trend using metrics introduced by [208]. These features encompass measures such as autocorrelation, spectral entropy, as well as indicators of self-similarity and nonlinearity. The work by [61] has identified various techniques for extracting features from time series data.

2.2.2.2 Time Series Predictions

Time series prediction is a crucial method for predicting future values based on historical data, widely applicable across various domains, including finance and industry [24] [144]. However, there is no one-size-fits-all approach; the choice of a prediction model relies on the specific characteristics of the time series and the constraints of the application. This selection process often presents significant challenges.

Traditionally, model selection has involved a laborious process of testing all candidate models exhaustively with available data, which is impractical for large datasets containing numerous potential models [144]. In response to this challenge, researchers have explored alternative methods, including knowledge-based approaches and empirical techniques. In particular, a comparison of analytical and algorithmic approaches for multistep-ahead prediction has shown that different methods can offer varying levels of accuracy depending on the dataset and prediction horizon [12]. Knowledge-based approaches, such as expert systems [46], rely on human expertise to develop rule-based systems that assign weights to candidate models based on predefined rules. While effective, these systems are limited by the availability and cost of human experts [9]. Empirical methods involve conducting competitions among models using various datasets and analysing the results. For instance, the M3-Competition employed this approach, but it relies on human experts for analysis, leading to imprecise insights [135]. However, these competitions have not clearly defined when to choose a more complex model over a simpler one.

The ETS (Exponential smoothing state space model) algorithm developed by [88] and the automatic ARIMA (Autoregressive integrated moving average) algorithm proposed by [87] are two widely utilised automated algorithms. Both algorithms are implemented in the forecast package in R, as documented by the [164, 86]. A predetermined set of models is chosen in this particular framework, and multiple models from that set are estimated for each time series. The model exhibiting the lowest Akaike Information Criterion corrected (AICc) value is selected and employed for the purpose of making forecasts. The methodology employed in this approach is dependent on the expertise of the forecaster in initially determining the most suitable category of models to utilise. Comparing AICc values between different model categories is typically not feasible due to variations in the computation of the likelihood and the treatment of initial conditions.

The proposed potential alternative method, which circumvents the need to preselect a specific class of models, involves utilising a straightforward "hold-out" test set. However, this approach often suffers from a lack of sufficient data, thereby limiting the ability to draw reliable conclusions. In order to address this limitation, one can employ time series cross-validation as suggested by [165] and [86]. This approach allows for the application of models from various classes, ultimately selecting the model with the lowest cross-validated mean squared error (MSE). Nevertheless, this leads to a significant increase in the computational time required, at least on the order of $N \times M$, where N represents the total number of series that need to be forecasted and M represents the number of the models.

Evidently, there is a necessity for a rapid and adaptable algorithm to mechanise the procedure of model selection precisely to predict future outcomes. This process is commonly referred to as forecast-model selection. Table 2.2 summarizing time series predictions, it encapsulates the various approaches, their descriptions, advantages, disadvantages, and relevant references for further reading.

Table 2.2: Summary of Time Series Predictions

Approach	**Description**	**Advantages**	**Disadvantages**	**References**
Traditional Model Selection	Involves testing all candidate models exhaustively with available data, impractical for large datasets	Potentially thorough model evaluation	Impractical for large datasets	Montgomery (1990)
Knowledge-Based Approaches	Utilizes human expertise to develop rule-based systems assigning weights to candidate models	Effective with expert knowledge	Limited by availability and cost of human experts	Collopy (1992), ARINZE (1994)
Empirical Methods	Conducts competitions among models using various datasets and analyzing results	Provides insights from real-world competitions	Imprecise insights and reliance on human analysis	Makridakis (2000)

Table 2.2 continued from previous page

Approach	**Description**	**Advantages**	**Disadvantages**	**References**
Automated Algorithms	ETS and automatic ARIMA algorithms implemented in the forecast package in R	Automated and widely used	Dependent on forecaster's expertise for model category selection, AICc comparison issues	Hyndman (2002, 2008, 2021), R (2022)
Proposed Alternative Method	Utilizes a hold-out test set or time series cross-validation to select the best model based on cross-validated MSE	Does not require preselecting a specific class of models	High computational time requirement, limited by data availability	Racine (2000), Hyndman (2021)

2.2.3 Machine Learning Hydrological Predictive Models

To address the unique challenges posed by climate science data, researchers have turned to more advanced Data-driven algorithms, particularly those based on machine learning algorithms.

Artificial neural networks (ANNs) and Support Vector Machines (SVMs), for instance, have shown great promise in the domain of climate forecasting. ANNs have proven valuable in predicting a wide range of climate variables, including but not limited to temperature trends, rainfall patterns, and even more complex phenomena like climate system behaviour. Similarly, SVMs have been applied effectively in climate science to model and forecast climate-related processes, such as long-term temperature changes or regional climate variations [180] [56].

Recent years have also witnessed the rise of deep learning models, including Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, which have gained attention for their ability to model complex time-series data. While these models have found substantial success in fields like natural language processing and image recognition, their application in climate science, including water level prediction, still needs to be explored. Their potential to capture intricate climate patterns is yet to be fully harnessed in climate science research.[79] [149]

These models offer improved accuracy by capturing complex non-linear relationships in the data, adaptability to changing conditions, efficient handling of large datasets, and automatic feature engineering, enabling the integration of diverse data sources [56] [181]. Furthermore, they can provide real-time updates by leveraging advances in sensor technologies, ensuring timely responses to changing river conditions. Machine learning models are robust to non-stationarity in river systems, quantifying uncertainty effectively and capturing localised variations, which is crucial for understanding specific river regions [56] [180] [223] [151]. Their seamless integration with remote sensing technologies enhances the accuracy of predictions. At the same time, continuous learning ensures model performance improvement over time, making them indispensable tools for modern river system management, particularly in the context of climate change and urban development [183] [151].

The selection of modelling methods in climate science should consider the specific challenges and datasets unique to this field. ANNs, SVMs, and deep learning models like RNNs and LSTMs have demonstrated their capabilities in forecasting various climate variables, including temperature, precipitation, and meteorological phenomena. These models provide valuable tools for understanding climate patterns at different spatial and temporal scales, ranging from urban microclimates to regional climate trends, with the potential to significantly advance research in climate science and environmental studies.

Hence, the application of these machine learning algorithms is not confined to climate science and environmental research alone. They hold significant utility in hydrology predictions, where accurate forecasts of water-related variables such as river flow, precipitation, and groundwater levels are crucial for effective water resource management, flood control, and environmental preservation. Moreover, their versatility extends to diverse fields beyond environmental sciences, including healthcare, finance, and urban planning, where predictive analytics play a vital role in decision-making processes [130] [151] [183].

The hydrology domain faces the need for a universally applicable modeling framework to better predict and understand hydrological processes and address various water-related challenges. This need arises from the complexity and variability of hydrologic systems, which require models that can integrate multiple physical processes across different spatial and temporal scales.

However, it is imperative to underscore that the hydrology domain confronts a need for a universally applicable modelling framework. Recent advancements have been made in developing such frameworks. For example, the GLOFRIM (Global Flood Risk with Integrated Model) framework is designed to couple hydrologic and hydrodynamic models, allowing for improved simulation of flood wave propagation and inundation extents. This framework integrates coarse-resolution global hydrologic models with fine-resolution hydrodynamic models, enhancing the accuracy of

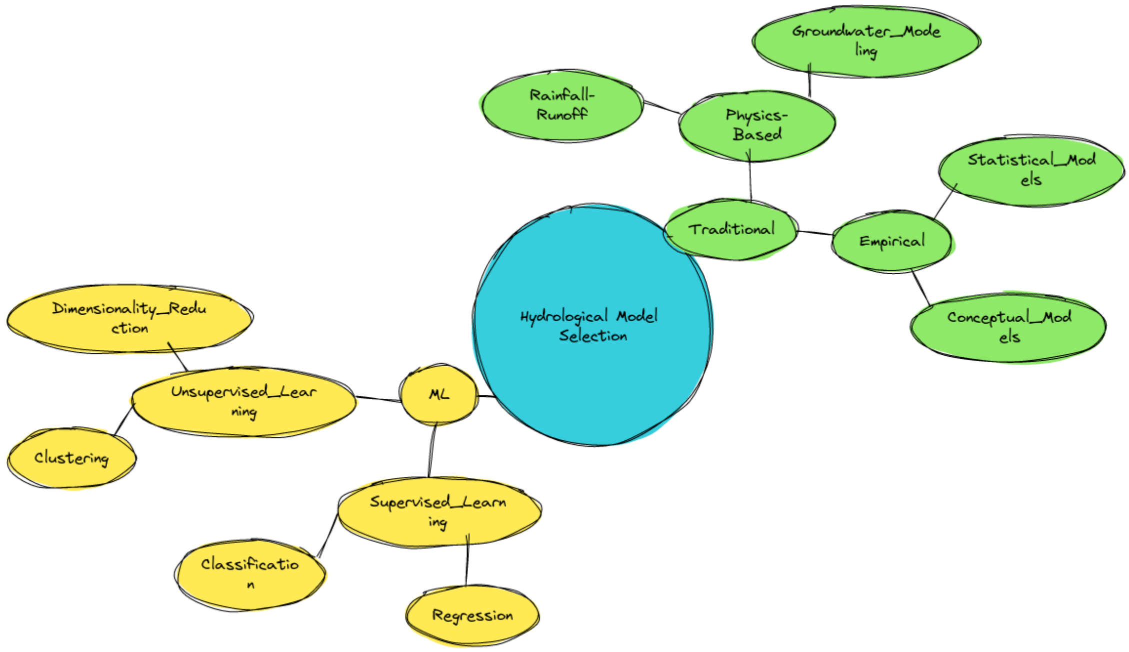

Figure 2.1: hydrological Models

flood modeling by addressing different physical processes at appropriate scales [80] [20]. Another significant development is the Python Modeling Tool (pyMT), which provides a flexible "plug-and-play" approach to coupling models. This tool allows for the creation of fit-for-purpose models by integrating different hydrologic and hydrodynamic processes, tailored to specific applications [21].

These advancements highlight the ongoing efforts to create more comprehensive and adaptable modeling frameworks in hydrology, addressing the need for tools that can manage the complexities of water systems and provide reliable predictions for water resource management and risk mitigation.

2.2.4 Hydrological Model Selection

Hydrological models, whether they adhere to traditional methodologies or employ machine learning techniques, 2.1 often grapple with the formidable challenge of faithfully representing the intricate physical processes inherent to watersheds, as noted in studies by [103] [137].

The process of selecting a suitable model within the hydrology domain entails a critical consideration: the inherent uncertainties accompanying model predictions.

Hydrological models are plagued by inherent uncertainties arising from the inherent variability in data, the structural intricacies of the models themselves, and the determination of model parameters, as underscored by [142]. To comprehensively address and communicate these uncertainties, it becomes not just important but imperative to employ methodological tools such as Bayesian modelling, ensemble modelling, and sensitivity analysis.

These methodologies are essential in hydrology as they allow for the quantification of uncertainties, improving the reliability of hydrological predictions. By enhancing model validation and calibration, they refine the accuracy of model simulations, supporting better-informed decision-making in water resource management and disaster preparedness. Effective communication of uncertainty through these tools fosters transparency and trust among stakeholders, facilitating adaptive strategies for managing hydrological systems in the face of changing conditions.

Traditional hydrologists have employed conventional model selection approaches, wherein a singular model is customised to suit a particular dataset or river system [96]. Although this approach may produce satisfactory outcomes within the parameters of a specific dataset, its effectiveness is often limited when extrapolated to diverse river systems or novel data [95]. The inflexibility of traditional models constrains their capacity to adapt to changing hydrological conditions, thereby impeding their efficacy in tackling broader hydrological challenges.

Traditional model selection methods, such as Multiple Linear Regression (MLR), have a long history of application in climate science. However, their performance can be limited, mainly when dealing with datasets that are inherently challenging due to their scarcity and diversity. In the context of climate science, MLR has often been used to explore the relationships between climate variables. For example, it has been employed to understand how changes in meteorological factors like temperature, humidity, and atmospheric pressure relate to phenomena such as precipitation patterns or extreme weather events. Nevertheless, MLR's simplicity and linearity can hinder its effectiveness when dealing with the intricacies of climate data. [105].

The inherent inflexibility of traditional hydrological models poses a significant challenge. These models are often characterised by rigid structures and fixed parameterisations, making them ill-suited to adapt to the dynamic and ever-changing nature of hydrological conditions. As a consequence, their effectiveness is constrained when confronted with broader hydrological challenges that transcend the boundaries of a specific dataset or geographical region. Consider, for instance, the implications of climate change on hydrological processes. Climate change introduces new patterns of temperature and precipitation, leading to shifts in the hydrological regime. Traditional models, which were calibrated based on historical data, may struggle to accurately capture and predict these emerging patterns. Their inability to adapt to changing conditions hampers their efficacy in addressing the evolving hydrological challenges posed by climate change. Moreover, the limitations of traditional models become particularly evident when dealing with diverse river systems. Each river system is inherently unique, influenced by a combination of geographical, geological, climatic, and anthropogenic factors. Attempting to shoehorn a single model, no matter how sophisticated, into such heterogeneous environments can lead to significant inaccuracies and unreliable predictions. In essence, the conventional approach of customising a single model to specific datasets or river systems, while suitable for certain scenarios, falls short when confronted with the complexities and uncertainties inherent in hydrology. To navigate the intricate landscape of hydrological modelling effectively, it is imperative to embrace more adaptable, data-driven and flexible modelling approaches that can accommodate the diversity of hydrological systems and respond to changing conditions. This shift toward adaptability and versatility is essential to meet the growing challenges posed by a dynamic and uncertain hydrological future.

2.3 Solution 'meta-learning' for time-series prediction model selection

Meta-learning has emerged as a powerful approach for automating the time series model selection process. Instead of manually choosing a specific model or algorithm for a given time series dataset, meta-learning leverages higher-level learning algorithms to make data-driven model selection decisions [26] [122]. In this section, we conduct a comprehensive literature review, particularly emphasising model selection in hydrology using meta-learning techniques. We explore the complexities surrounding the selection of models in hydrology and how meta-learning provides a transformative strategy for addressing these obstacles.

Meta-learning has acquired prominence in machine learning, particularly in time series prediction, as it offers automation for model selection. In this section, we will explore meta-learning techniques, explicitly focusing on model selection and how these techniques can be applied to adapt models to new datasets.

The advantage of meta-learning in time series model selection is its adaptability and ability to handle a wide range of time series datasets without manual intervention. It leverages historical performance data and dataset characteristics to make informed decisions about which forecasting model is likely to work best for a given dataset [122].

Recent studies have explored the application of meta-learning techniques in this domain, emphasising its potential to enhance the precision and effectiveness of model selection. A foundational study by [179] proposed using meta-learning to automate forecast model selection based on time series data features. This research laid the groundwork for subsequent investigations in the field. [162] made significant contributions by examining various meta-learning algorithms for selecting prediction models. Their work expanded the knowledge base in this area. In 2010, [120] conducted a study that delved deeper into the application of meta-learning in time series prediction. They explored the impact of different meta-features and algorithms on model selection, contributing valuable insights. They observed that certain meta-features, such as statistical characteristics of time series data or characteristics of the underlying time series models, could significantly influence the choice of the best predictive model. For example, their research showed that the presence of strong seasonality in a time series, as indicated by certain meta-features, could lead to the selection of specific algorithms that are well-suited for handling seasonal patterns. This finding provided valuable insights for practitioners, highlighting that the choice of a predictive model should consider not only the available data but also the inherent characteristics of the time series being analyzed. This holistic approach ensures that predictive models are tailored to effectively capture and utilize the unique temporal patterns, such as seasonality or trends, present in the data. [114] built upon previous research by exploring the practical integration of meta-learning methodologies into prediction systems, highlighting its potential to enhance accuracy and efficiency across diverse fields. These studies collectively illustrate the growing importance of meta-learning in time series prediction, where researchers leverage time series data features and meta-learning algorithms to optimise forecast accuracy. The field continues to evolve, likely leading to further advancements in automated prediction systems. In 1992, [46] developed 99 rules based on 18-time series features for economic and demographic prediction. [10] extended this work to reduce human involvement. [178] study categorised time series based on features like observations, turning point ratio, step change ratio, skewness, kurtosis, and autocorrelations. [161] introduced the term "meta-learning" and evaluated it with two case studies, using features like length, autocorrelation coefficients, skewness, and kurtosis, employing decision trees to acquire knowledge. [100] utilised the NOEMON system (NOEMON measures model performance for a collection of datasets) methodology to establish a hierarchy among time series predicting algorithms. [120] investigated the effectiveness of meta-learning methods, considering ARIMA, exponential smoothing, neural network models, various statistical measures, and machine learning algorithms. [205] proposed a meta-learning framework with a novel metric called simple percentage better (SPB). They used nine features and eight prediction models, employing decision trees and SOM clustering. [211] reduced dimensionality with PCA using features introduced by [205] in their meta-learning framework. [113] developed a neural network-based meta-learning framework using 78 time series data from the NN3 competition. They utilised error prediction and average symmetric mean absolute percentage error for model selection, considering various prediction techniques.

2.4 Summary

In summary, the model selection procedure in the context of time series analysis is crucial for ensuring precise and dependable predictions. This literature review has adopted a comprehensive methodology to investigate the current advancements in this field, delineating it into three principal segments.

Section 2.1 detailed advancements in integrating environmental data, emphasizing spatiotemporal integration in particular. This aspect of time series analysis highlights the significance of integrating heterogeneous data sources and comprehending the intricate interaction of environmental variables when modeling time series data, with a particular focus on DW/ELT processes as essential components.

In Section 2.2, the application of machine learning methods in hydrology predictions was explored. This section emphasised the significance of utilising sophisticated algorithms to model and forecast hydrological phenomena with precision. Trends in approaches include the use of artificial neural networks (ANNs), support vector machines (SVMs), and deep learning models like Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. Outstanding challenges in the domain include handling the variability of hydrological data, capturing complex non-linear relationships, and ensuring model adaptability to changing conditions. The integration of diverse data sources and continuous learning were identified as critical for improving prediction accuracy and reliability in water resource management and environmental sustainability.

Section 2.3 examined the emerging field of meta-learning in the context of time series model selection. This section highlighted the potential of meta-learning to automate the selection and optimisation of time series models using advanced learning algorithms. Key takeaways include the adaptability of meta-learning to various datasets, the ability to leverage historical performance data and dataset characteristics, and the reduction of manual intervention in model selection. Challenges include the need for effective feature extraction, handling the diversity of time series data, and ensuring the scalability of meta-learning frameworks.

Overall, this literature review underscores the complexity of time series analysis, where machine learning, data integration, and meta-learning converge to enhance the capacity to predict and model time-dependent phenomena. As progress continues in these domains, there is a promising prospect of achieving higher precision, resilience, and automation in time series model selection. This advancement holds significant potential for improving decision-making processes across various disciplines, particularly in hydrology, environmental monitoring, and resource management.

Chapter 3

Spatiotemporal Environmental Data Integration

3.1 Introduction

Environmental management is intrinsically tied to data collection, analysis, and integration. In the digital age, data-driven approaches, especially in Machine Learning and Information Technology, have proven pivotal in addressing pertinent concerns, evaluating outcomes, and forecasting future trends across various scientific disciplines. However, the complexities of the data landscape, characterized by issues such as data unavailability, incompatibility of formats, and the challenge of integrating disparate systems, often make the task daunting. Hydrological models emerge as a beacon in this intricate domain. They serve a dual purpose: aiding researchers in simulating the effects of landscape structures and climatic changes on the water while generating intermediate results that require careful processing and conversion into actionable insights. These models represent, comprehend, and hypothesize the operations of environmental systems, making their integration into IT frameworks indispensable.

As spatiotemporal data analysis continues to revolutionize source location classification, climate impact forecasting, and water quality assessments, the storage, visualization, and integration hurdles remain prominent. The rapid advancements in Geographical Information Systems (GIS), driven by the evolution of information and communication technologies, have ushered in an era of data abundance. However, this wealth of data often remains underutilised, as it is frequently confined within isolated platforms or rendered incompatible across disparate systems. In hydrology, the situation is further exacerbated by the scarcity of specific measurements capturing spatial and temporal variations. Datasets are often heterogeneous, appearing in many different formats (.shp, .json, .csv, etc..) at varying levels of granularity with respect to both location and time and may (or may not) contain overlapping spatiotemporal windows. These issues compound the traditional issues associated with data integration and provide additional complexities driven by the times and location components offered by such data. Responding to these challenges, this chapter proposes a robust system adept at spatiotemporal integration. Seamlessly merging multidimensional data from diverse sources aims to offer a comprehensive view that spans both time and space. This multidisciplinary approach, which marries the sophistication of IT and the practicality of hydrology, is crucial for maintaining research integrity, guiding informed decisions, and assuring quality. As we delve deeper into this chapter, we will dissect the components of this promising data integration system, spotlighting its potential to revolutionize sustainable water resource management for the digital era.

3.2 Methodology and Architecture

Integrating heterogeneous data collection from multiple sources demands a robust and versatile methodological architecture capable of efficiently managing various data types and formats while maintaining data quality and consistency. The complexity of today's data landscapes, marked by myriad sources and formats, necessitates a comprehensive approach to integration that transcends traditional methods.

Many ML models developed and used in the literature utilised multiple input features [30]. However, due to the lack of a robust environment for analysis and an inability to offload the work to more extensive, scalable computing environments, datasets are often limited in scope geographically (there is not much national modelling). Researchers often spend a significant portion of their time on data collection and maintenance, with efforts extending over several weeks to months. This extensive time investment is due to the need for independent collection, wrangling, and management of datasets from various providers. [84] discuss how data acquisition technologies can streamline maintenance management systems, offering real-time retrieval of information and improving field operations. [148] highlight the challenges and methodologies for analyzing maintenance work orders, underscoring the importance of transforming data into a more analyzable format.

This research discusses the challenges of sourcing environmental data in Ireland. Although it does focus on specific types of data, such as river nutrients and water flow, it highlights a broader issue that could apply to various environmental datasets. Therefore, while the examples are specific, the problem addressed could be relevant to a wider range of environmental data collection and analysis efforts. After the initial data-sourcing phase, it became evident that many datasets were not available for all areas in Ireland. For some locations, environmental data, such as river nutrients, were almost nonexistent. This could be due to either the data not being measured in those areas or not being publicly accessible. However, nutrient datasets are crucial for creating AI models to predict water quality. In other instances, like water flow data, availability depends on specific conditions, such as the deployment of a hydrometric station during the queried time period. Additionally, obtaining data can be complicated by each provider's unique access methods and protocols. While some datasets can be downloaded from websites, others are accessed via the Data Access API (Application Programming Interface) provided by the data supplier.

Although using a Data Access API is arguably the best method for obtaining data, it can be quite sophisticated for researchers without an IT background. As a result, many researchers prefer to manually download data from data provider websites, despite the tedious nature of this task. For example, the Office of Public Works (OPW)(see Appendix A) provides water flow data for up to five weeks on one website, while historical data must be accessed through a different site (see Appendix A). Similarly, Met Éireann's (see Appendix A) website requires users to select each station, time period, and time resolution separately for each download operation. Consequently, the data collection process was a challenging and time-consuming aspect of this project. This highlights the need for a unified and easily accessible platform for researchers to access and analyze environmental data efficiently.

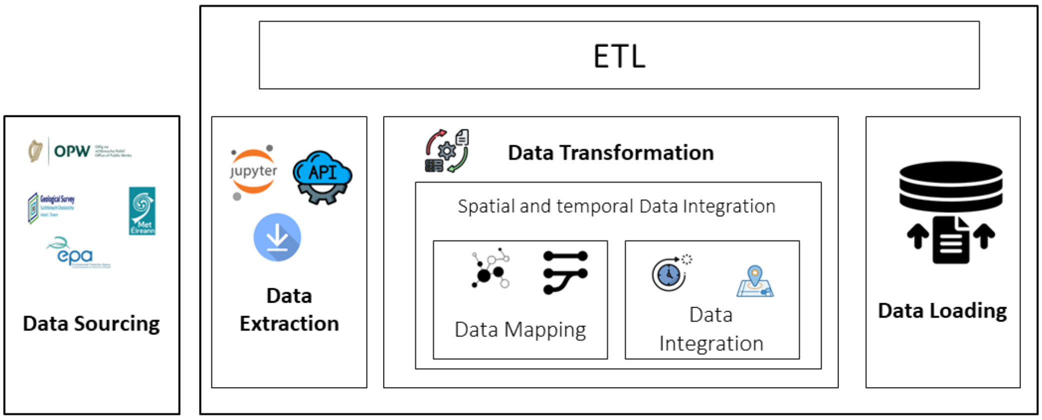

Figure 3.1: Spatiotemporal data integration system architecture

The key components of the proposed integration system for merging heterogeneous environmental data architecture are:

- Data Sources: These are the various heterogeneous sources from which data is collected.

- Data Extraction: This component extracts the data from different sources.

- Data Transformation & Mapping: Data mapping is an essential part of the ETL process that establishes relationships between data elements from different sources.

- Data Integration: This component consolidates and integrates data from multiple sources into a unified view.

- Data Loading: This component loads the transformed data into the target system.

3.2.1 Data Sourcing

This phase includes identifying the various heterogeneous sources from which data is collected. Data sources include databases, warehouses, APIs, web services, flat files, spreadsheets, or other data storage systems. Identifying and understanding these sources is crucial for a successful ETL process. In the context of Environmental spatiotemporal Data Integration (STDI) in Ireland, various datasets can be collected from multiple sources, such as the Office of Public Works (OPW), the Environmental Protection Agency (EPA), Met Éireann and others. Table 3.1 includes some examples of the data these organizations provide. When integrating these datasets, it is crucial to consider the spatiotemporal resolution, data formats, and any potential discrepancies or gaps in the data.

3.2.2 Data Extraction

The data extraction component is responsible for extracting the data from the different sources identified in Ireland's Environmental Spatiotemporal Data Integration (STDI). The extraction process involves connecting to the data sources, selecting the required data, and retrieving it in a format that can be processed further. Data extraction involves handling different data formats, structures, and access protocols. Extracting data from these sources may involve connecting to APIs, web services or downloading files in various formats, such as CSV, GeoJSON, NetCDF, or raster data. The extraction process should be designed to handle the specific data access protocols and formats of the environmental data sources. Given the variety of sources and data types, using appropriate extraction methods and tools for each source is crucial. Table 3.1 details some of the access and extraction methods from the sources mentioned in the previous section.

During the data extraction phase, it is essential to consider factors such as data

Table 3.1: Data Sources, Types and Access Methods

Provider	**Data Type**	**Access**
Office of Public Works (OPW)	River flow and water level data	Access data file downloads (e.g., Excel formats)
Office of Public Works (OPW)	Flood risk mapping and management plans	Access data through web services, or file downloads (e.g., GIS formats)
Environmental Protection Agency (EPA)	Water quality data (e.g., lakes, rivers, groundwater, coastal)	Use web services to retrieve data or download data files (e.g., CSV, JSON, XML, or GIS formats)
Environmental Protection Agency (EPA)	Air quality data (e.g., pollutants, emissions, air quality)	Use web services to retrieve data or download data files (e.g., CSV, JSON, XML, or GIS formats)
Environmental Protection Agency (EPA)	Waste and resource management data	Use APIs or web services to retrieve data or download data files (e.g., CSV, JSON, XML, or GIS formats)
Environmental Protection Agency (EPA)	Biodiversity and habitat data	Use APIs or web services to retrieve data or download data files (e.g., CSV, JSON, XML, or GIS formats)
Met Éireann	Meteorological data (e.g., temperature, precipitation)	Connect to APIs or web services for real-time and historical data or extract data in formats such as CSV, JSON or XML
Met Éireann	Climate data (e.g., historical records, climate models)	download data files (Zip files)
Met Éireann	Weather forecasts and warnings (e.g., storms, flooding)	Connect to APIs or web services for real-time and historical data or extract data in formats such as CSV, JSON or XML
National Earth Science Knowledge Centre	Geological data (e.g., bedrock, soils, mineral resources)	Access data through APIs, web services, or file downloads (e.g., CSV, Excel, or GIS formats)
National Parks and Wildlife Service	Data on protected areas, species distribution, conservation	Access data through web services or file downloads (e.g., CSV, Excel, or GIS formats)
Central Statistics Office (CSO)	Socio-economic and demographic data	Use web services to retrieve data or download data files (e.g., CSV, Excel, or GIS formats)

access permissions, API rate limits, data refresh rates, and data versioning. Hence it was necessary in some cases to create bespoke application code to extract these datasets.

3.2.3 Data Transformation and Mapping

Data transformation is the process of converting the extracted data into a format that can be easily integrated, analysed, and utilised. During this phase, various data manipulation techniques are applied to clean, normalise, aggregate, and enrich the data. Given the heterogeneous nature of the data sources, ensuring that the data is consistent, accurate, and compatible before integration is crucial.

3.2.3.1 Data Transformation

The data transformation process involves four key steps: data cleaning, validation, normalization and standardization.

1. Data cleaning and validation:
    - Identify and handle missing values, errors, and outliers.
    - Validate data accuracy and integrity by comparing it with reference datasets or data quality rules.
2. Data normalization and standardization:
    - Convert data to a common measurement unit, scale, or coordinate system.
    - Standardize data formats, such as date and time formats, and categorical values.

Throughout the data transformation process, it was essential to maintain data quality, ensure data consistency, and track data lineage. This was achieved by analysing the data for inconsistencies, errors, or missing values, checking the transformed data to ensure it meets formatting requirements and tracking changes to the data throughout the transformation process. To support these steps, a comprehensive specification for data extraction and integration was developed. This specification outlined the procedures for sourcing data from various providers, detailing the methods for accessing and retrieving data, whether through manual downloads or using Data Access APIs. It also included guidelines for integrating data from different sources, ensuring that the transformed data adhered to a unified schema compatible with subsequent analytical processes.

3.2.3.2 Data Mapping

Data mapping is the process of defining relationships between fields or attributes in different datasets, often originating from different sources, to enable data integration, transformation, and analysis. Data mapping is a crucial component of data integration as it informs researchers *how* data should be integrated and acts as the start of defining a common data model. It involves identifying corresponding fields, determining data types, and defining transformations or conversions needed to harmonize the data. A simple data mapping can be represented using a table or a diagram that links the source and target fields.

Table 3.2 presents an example of data mapping between two datasets related to the river water level in Ireland, one from the Office of Public Works (OPW) and the other from the Environmental Protection Agency (EPA). The table details a sample of the mapping rules derived to integrate the dataset provided by the OPW and the EPA. The Station_ID and Gauge_ID fields are mapped, representing the same information. The Station_Name and Gauge_Name fields are also mapped, indicating that they represent the same information. The Timestamp_UTC and Date_Time_UTC fields are linked, indicating that they represent the same information, and the Water_Level and Level fields, as well as the Water_Flow and Flow fields, are mapped accordingly.

Table 3.2: Example of data mapping rules for mapping OPW and EPA data

Dataset A (OPW water-level data)	**Transformation/Conversion**	**Dataset B (EPA water-level data)**
Station_ID	Station_ID = Gauge_ID	Gauge_ID
Station_Name	Station_Name = Gauge_Name	Gauge_Name
Timestamp_UTC	Date_Time_UTC	Date_Time_UTC
Water_Level	Level	Level
Water_Flow	Flow	Flow

In this example presented in Table 3.2 , the data mapping rules are applied to map attributes from two datasets, "Dataset A" (OPW water-level data) and "Dataset B" (EPA water-level data). The mapping involves the following rules:

- **Station_ID:** This attribute in "Dataset A" is mapped to "Gauge_ID" in "Dataset B."
- **Station_Name:** This attribute in "Dataset A" is mapped to "Gauge_Name" in "Dataset B."
- **Timestamp_UTC:** This attribute in "Dataset A" is directly used as "Date_Time _UTC" in "Dataset B."
- **Water_Level:** This attribute in "Dataset A" is directly used as "Level" in "Dataset B."
- **Water_Flow:** This attribute in "Dataset A" is directly used as "Flow" in "Dataset B."

These mapping rules ensure that corresponding attributes between the two datasets are properly aligned, and the data can be integrated effectively. The rules also specify cases where direct attribute usage without any transformation is appropriate, simplifying the integration process. A full set of the data mapping rules are provided in Appendix D. These rules identify common points of integration between sources and act as a guide when developing a data integration plan and common data model (schema). By adhering to these mapping rules, we ensured that corresponding attributes between datasets were properly aligned, facilitating effective data integration without the need for a formal ontology.

3.2.4 Data Integration

In traditional ETL integration rules usually define common sets of attributes and identifiers (e.g. CustomerID) these rules, while taking time to discover are easily handled by modern ETL systems. Spatiotemporal data provides its own unique set of challenges for integration, time points may be irregular, present in overlapping windows or contain differing granularities, providing bespoke application logic to accomplish integration. Similarly location components pose additional challenges, locations may represent a point or a polygon, points present at differing levels of granularity and polygons may have overlapping areas of intersection, similarly to the issues presented by temporal data, spatial data requires bespoke logic to facilitate integration. Spatiotemporal data integration combines and harmonises data from multiple sources with spatiotemporal attributes to enable analysis, visualisation, and decision-making in a spatiotemporal context. This type of integration is particularly relevant for environmental data, where spatiotemporal factors play a significant role in understanding and managing natural resources and ecosystems. This component consolidates and integrates data from multiple sources into a unified view, making it easier to analyse and gain insights. It consists of two primary steps: spatial and temporal data integration. Spatial data integration involves merging data from multiple sources with a common geographic location or region. On the other hand, temporal data integration involves combining data from various sources related to the same time period or time series.

3.2.4.1 Spatial Data Integration

Spatial data integration is crucial for combining diverse environmental datasets into a unified framework, enabling comprehensive analysis and decision-making. This research focuses on developing an algorithm that integrates preprocessed environmental datasets, represented as points and polygons, by considering their spatial and attribute relationships. The integration process aims to minimize information loss and ensure optimal data alignment, which is vital for accurate environmental analysis [131] [51] [35] [210] [189].

3.2.4.1.1 Distance Calculation Using the Haversine Formula To integrate spatial data, calculating the distance between points and polygons is essential. This research employs the Haversine formula, a well-known method ,the Haversine Formula 3.1, for computing the shortest distance over the Earth's surface, which is defined as :

$$\operatorname{hav}\left(\frac{d}{r}\right) = \operatorname{hav}(\varphi_2 - \varphi_1) + \cos(\varphi_1)\cos(\varphi_2)\operatorname{hav}(\lambda_2 - \lambda_1) \tag{3.1}$$

Where:

- d is the distance between two points,
- r is the Earth's radius,
- ϕ_1 and ϕ_2 are the latitudes of the two points,
- $\Delta\phi$ is the difference in latitudes,
- $\Delta\lambda$ is the difference in longitudes.

This standardized distance measurement ensures consistency and reliability in selecting appropriate data for integration, facilitating accurate environmental analysis.

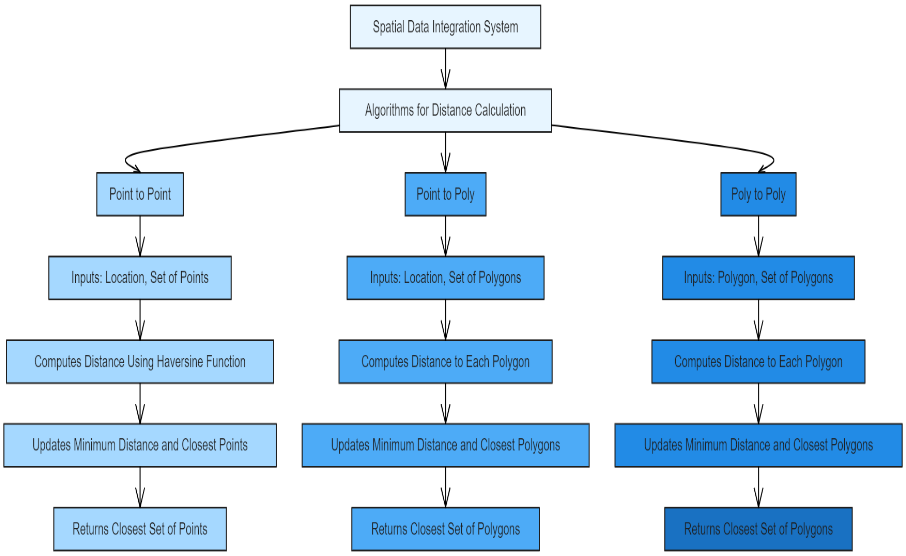

Figure 3.2: Flowchart of Distance Calculation Algorithms in a Spatial Data Integration

3.2.4.1.2 Vector Data Structures To store and represent spatial data, vector data structures are used. Vectors consist of vertices, which are discrete geometric points (e.g., x, y values) that define the shape of spatial entities. The organized vertices determine the vector type:

- **Points:** Single x, y coordinates define each point. These are used for sampling locations and discrete geographic features.
- **Lines:** Multiple connected points form lines, representing linear features like roads or streams.
- **Polygons:** Three or more connected vertices create closed shapes, representing areas such as lakes, catchments, and administrative boundaries.

3.2.4.1.3 Algorithms This research introduces three algorithm 3.2 designed to calculate the closest set of points, polygons, or neighboring polygons to a given location or polygon. These algorithms are crucial for integrating spatial data from various sources, ensuring accurate and efficient data alignment. Each algorithm calculates the distance between the location or polygon of interest and the other locations or polygons in the dataset to identify the closest set.

These algorithms have applications in various fields, such as finding the nearest station or boundary, spatial analysis, and planning. They can be implemented using various methods, such as the minimum distance between a point and a line segment or the minimum distance between two vertices.

Point to Point Algorithm

This section presents an algorithm in Algorithm 1 for finding the closest set of points to a given location in a dataset. The algorithm takes as input a location and a set of points and returns the closest set of points to the given location. To accomplish this, the algorithm computes the distance between the location and each point in the set using the Haversine function, which measures the shortest surface-level distance between two points on a sphere. The minimum distance and closest set of points are then updated accordingly. If multiple points have the same minimum distance, the algorithm adds all of them to the closest set of points. This algorithm can be used in various applications, such as finding the nearest station or any other location of interest.

The input of this algorithm is a location and a set of points, and the output is the closest set of points to the given location. To use the algorithm, initialize an empty set and set the minimum distance to infinity. For each point in the set of points, compute the distance between the given location and the point using the Haversine function. If the computed distance is less than the current minimum distance, update the minimum distance to the computed distance and set the closest point to that point. If the computed distance equals the current minimum distance, add the point to the closest point set. Finally, return the closest point set as the output of the algorithm.

Consider a is a location in dataset A and b is location in dataset B, then this method returns the closest location set B_{min} to location a, within all locations in dataset B.

$$B_{min} = \underset{\forall b \in B}{\operatorname{argmin}}(\operatorname{hav}(a, b)) \tag{3.2}$$

Algorithm 1 Point to point

Input: *location* and *points*
Return: Closest point set P_{min}

$P_{min} \leftarrow \varnothing$
$minDist \leftarrow \infty$
for p **in** $points$ **do**
    $dist \leftarrow \text{hav}(location, p)$
    if $dist < minDist$ **then**
        $minDist \leftarrow dist$
        $P_{min} \leftarrow \{p\}$
    else if $dist = minDist$ **then**
        $P_{min} \leftarrow P_{min} \cup \{p\}$
    end if
end for

Point to Poly Algorithm

The Point to Poly algorithm defined in Algorithm 2 is used to find the closest polygon set to a given location. The polygon set is defined by a set of line segments that make up the polygon's borders. The algorithm takes a location and a set of polygons as input and calculates the distance between the location and each polygon using the distPointToPoly function. If the point is inside the polygon, it returns a distance of 0. Otherwise, it iterates over the polygon's edges to find the closest edge to the point. If the intersection is on the line segment, the distance between the point and the edge is calculated as the length of the perpendicular line from the point to the line segment. Otherwise, it is the minimum distance to the vertices of the line segment. The algorithm updates the minimum distance and closest polygons accordingly and returns the set of closest polygons to the given location. This algorithm can be used in applications such as finding the nearest boundary or region of interest or in spatial analysis and planning. The distPointToPoly function can be implemented using various methods, such as the minimum distance between a point and a line segment or the minimum distance between a point and the polygon's vertices.

Consider a is a location in dataset A, and polygon *pol* is dataset POL, where a polygon is defined by the line segments that make up a polygon's borders, then this method returns the closest polygon set POL_{min} to location a.

$$POL_{min} = \underset{\forall pol \in POL}{\operatorname{argmin}} (\text{distPointToPoly}(p, pol)) \tag{3.3}$$

Here, distPointToPoly function calculates the distance from a point to a polygon. If the point is inside the polygon, it returns 0 as the distance. Otherwise, it iterates over the polygon's edges (i.e., line segments) to find the closest polygon edge to the point. The distance between the point and the edge is calculated as the length of the perpendicular line from the point to the line segment if the intersection is on the line segment. Otherwise, it is the minimum distance to the vertices of the line segment.

Algorithm 2 Point to Poly

Input: *location* and *polygons*
Return: Closest polygon set POL_{min}

$POL_{min} \leftarrow \varnothing$
$minDist \leftarrow \infty$
for *pol* **in** *polygons* **do**
  $dist \leftarrow \text{distPointToPoly}(location, pol)$
  if $dist < minDist$ **then**
    $minDist \leftarrow dist$
    $POL_{min} \leftarrow \{pol\}$
  else if $dist = minDist$ **then**
    $POL_{min} \leftarrow POL_{min} \cup \{pol\}$
  end if
end for

Poly to Poly Algorithm

This section describes an algorithm in Algorithm 3 for finding the closest polygon set Dmin to a given polygon c within a set of polygons D. The algorithm takes advantage of the distPolyToPoly function, which calculates the distance between two polygons. If the two polygons intersect or if one is inside the other, the function returns 0 as the distance. Otherwise, the algorithm iterates over the edges of the polygons to find the closest edge of the other polygon and returns the distance to that edge. The algorithm takes as input a polygon and a set of polygons and returns the closest set of polygons to the given polygon. It uses the distPolyToPoly function to compute the distance between the given polygon and each polygon in the set and updates the minimum distance and closest set of polygons accordingly. If multiple polygons have the same minimum distance, the algorithm adds all of them to the closest set of polygons. This algorithm can be used in various applications, such as finding the nearest neighbouring polygon or in spatial analysis and planning. The distPolyToPoly function can be implemented using various methods, such as finding the minimum distance between two line segments or between two vertices.

Consider c is a polygon in dataset C and d is polygon in dataset D, then this method returns the closest polygon set D_{min} to polygon c, within all polygons in dataset D.

$$D_{min} = \underset{\forall d \in D}{\operatorname{argmin}}(\text{distPolyToPoly}(c, d)) \tag{3.4}$$

Here, distPolyToPoly function calculates the distance between two polygons. If the first polygon is inside or intersects with the other polygon, it returns 0 as the distance. Otherwise, it iterates over the polygon's edges (i.e., line segments) to find the closest edge of the other polygon. It returns the distance to that edge, where the distance between two line segments is the closest distance between the vertices of the two line segments.

Algorithm 3 Polygon to polygon

Input: $polyC$ and $polygons$
Return: Closest polygon set POL_{min}

$POL_{min} \leftarrow \varnothing$
$minDist \leftarrow \infty$
for pol **in** $polygons$ **do**
  $dist \leftarrow \text{distPolyToPoly}(polyC, pol)$
  if $dist < minDist$ **then**
    $minDist \leftarrow dist$
    $POL_{min} \leftarrow \{pol\}$
  else if $dist = minDist$ **then**
    $POL_{min} \leftarrow POL_{min} \cup \{pol\}$
  end if
end for

3.2.4.2 Temporal Data Integration

Temporal data integration involves combining data that have temporal attributes, such as timestamps or intervals. Temporal data can be represented using various formats, such as time series, event, or interval data. Temporal data integration involves aligning data from different sources based on their temporal attributes and ensuring that they use consistent time zones, calendars, and date formats.

An example of how temporal data integration can be performed based on the data sources provided in the previous section for Environmental SpatioTemporal Data Integration (STDI) in Ireland:

Suppose we aim to integrate temporal data from the Office of Public Works (OPW) and Met Eireann related to Ireland's water level and weather conditions (see Table 3.3).

To perform temporal data integration, we would first need to align the two

Table 3.3: Water and weather data tables

Dataset A (OPW)
Station ID (integer)
Station Name (char)
Timestamp UTC(timestamp)
Water Level (float)

Dataset B (Met Eireann)
Station ID (integer)
Station Name (char)
Timestamp UTC(timestamp)
Temperature (float)
Precipitation (float)
Wind Speed (float)
Humidity (float)

datasets based on their Timestamp_UTC field. We would then need to determine the appropriate time interval for the analysis, which could be daily, weekly, monthly, or any other relevant interval. Once the time interval is determined, we can perform the integration using various techniques, such as:

- Aggregation: Aggregate the data from both datasets based on the chosen time interval (e.g., daily) by computing summary statistics (e.g., mean, sum, max, min) for each variable within the interval. This would result in a new dataset summarising each time interval's water level and weather conditions.

- Interpolation: Interpolate the data from one dataset to align with the time intervals of the other dataset (e.g., interpolate water level data from the OPW dataset to align with the daily intervals of the Met Eireann dataset). This would allow us to create a new dataset that combines the interpolated data with the original data from the other dataset.

- Fusion: Fuse the data from both datasets to create a new dataset that combines the water level and weather condition data at the exact same time and location. This would require aligning the spatial attributes of the datasets, such as the station IDs and names, and the temporal attributes.

By performing temporal data integration, we can create a new dataset that combines the water level and weather condition data from multiple sources, enabling us to analyze and visualize the data in a spatiotemporal context. After performing temporal data integration, the resulting dataset will depend on the specific integration technique used.

3.2.5 Data Loading

Data loading is the last phase of the ETL process, where the transformed data is loaded into the target system or storage for further analysis, visualization, or reporting. The loading process must ensure that the data is stored efficiently, securely, and in a format that is compatible with the intended applications or systems. In the context of Environmental spatiotemporal Data Integration (STDI) in Ireland, the target system includes a data warehouse and a GIS platform. Each chosen data set was analysed to identify the relevant information. Therefore, the DW design is dependent on the harvested data. Appendix C provides a complete overview of the entire data table schema. Each box describes a table, and each line represents a connection between tables. In addition to loading the transformed data, metadata and data cataloging are crucial components of the data loading process. Metadata provides descriptive information about the data, such as its source, format, and transformation history, ensuring that users can understand and utilize the data effectively. A data catalog is maintained to document all datasets, including details about their structure, relationships, and access protocols. This catalog facilitates efficient data management, retrieval, and integration, ensuring that the data warehouse and GIS platform are well-organized and user-friendly.

3.3 Environmental Case Studies Demonstrating Data Integration and Query Capabilities

Building on the detailed methodologies and architecture discussed, this section presents various environmental case studies to demonstrate the practical application and effectiveness of the spatial and temporal data integration techniques. To validate the proposed integration function's effectiveness and efficiency, three experiments were conducted using real-world environmental datasets with varying spatiotemporal resolutions and different attribute representations. In each experiment, the integration function was applied to the selected datasets, and the results were evaluated using quantitative and qualitative measures. These measures included the degree of information loss, the reduction in data inconsistencies, and the overall improvement in data quality. Additionally, the performance of the proposed function was compared with traditional integration methods, highlighting the advantages and limitations of each approach.

3.3.1 River Water Levels and Weather Data Integration

We applied a merging algorithm that combined point-based rainfall data with point-based water level data based on their spatial proximity. Specifically, we utilized the "Point to Point" algorithm to achieve this integration. This algorithm was selected for its ability to match and link the data based on their precise geographic locations. This enabled us to conduct a spatial analysis of the correlation between rainfall and water levels. The two datasets we aimed to use in this study have distinct properties. The river water level data is a continuous, time-series dataset that records the water levels of various rivers and streams in Ireland.

Meanwhile, the rainfall data is also a time-series dataset, recording the amount of rainfall in a given location over a specific time period, along with other features such as temperature and wind. Integrating these two datasets was essential to the analysis, as it allowed us to investigate the relationship between water levels and precipitation events comprehensively. By linking the data based on their geographic locations, we could identify areas where rainfall significantly impacted water levels and vice versa. While the datasets are correlated, each provides unique and complementary information crucial for a thorough analysis. Rainfall data indicates precipitation patterns, but without water level data, it does not show how these patterns affect water bodies. Conversely, water level data without rainfall information cannot reveal the potential causes of changes in water levels. By integrating both datasets, we can better understand the dynamics between rainfall and water levels, identify causative relationships, and make more informed predictions and decisions regarding water resource management.

The "Point to Point" algorithm was specifically chosen for its ability to merge the datasets based on their precise locations. This allowed us to conduct a detailed spatial analysis of the relationship between rainfall and water levels, which would have been difficult to achieve using other methods. The insights gained from this analysis have the potential to inform flood management and water resource planning efforts. They may help mitigate the impact of extreme weather events on water levels in Ireland.

3.4 presents a sample output of the "Point to Point" algorithm, which compares water level monitoring stations to the closest available weather station. This table is crucial as it demonstrates the practical application of our integration methodology, linking hydrological data with meteorological data based on spatial proximity.

Table 3.4: Sample output of the "Point to Point" algorithm comparing water level monitoring stations to the closest available weather station

Water_N	**Water_ID**	**Weather_N**	**Weather_ID**	**Dist**
Broadmeadow	8008	DUBLIN AIRPORT	532	5.276 (Km)
Ballincolly	19056	CORK MONTENOTTE	5404	2.123 (Km)
Riverstown	23001	LIMERICK CITY	6205	4.789 (Km)
Kilcurry	34002	GALWAY SALTHILL	7112	3.450 (Km)
Dundalk Bay	45003	BELFAST CITY	8123	6.512 (Km)

Where:

- Water_N: the water station name.
- Water_ID: Identifier for the water station.
- Weather_N: the nearest weather station name.
- Weather_ID: Identifier for the nearest weather station.
- Dist: the distance between the two stations represented in x (Km)

Now that we have identified the closest weather station to a given water station, 3.5 illustrates an example of the integrated data set; each row represents a combined record of water data and climate data for a 'Broadmeadow' water station and its corresponding closest weather station 'DUBLIN AIRPORT'. The integrated dataset includes the following:

Table 3.5: Example of integrated water level and climate data for Broadmeadow and Dublin Airport

Tstamp	**WL**	**P**	**T**	**vappr**	**H**	**msl**	**wdsp**	**wddir**	**vis**
1/1/2023 0:00	9.241	0.3	6.7	9.2	94	995.1	6	220	9000
12/31/2022 23:00	9.244	0	6	8.9	95	994.4	4	220	7000

Where:

- Tstamp: Temporal attribute representing the date and time of the measurements (hourly in this case).
- WL (m): Water level measurement in meters.
- T(°C): Temperature measurement in Celsius.
- P (mm): Precipitation measurement in millimetres.
- H (%): Humidity measurement in percentage.
- msl(hPa): Mean Sea Level Pressure
- wdsp(m/s): Wind speed measurement in meters per second.
- wddir (°): Wind direction measurement in degrees.
- vis(m): Visibility

3.3.2 Weather Station Proximity to Geological Features

In this scenario, we aimed to integrate weather station data with geological feature data to investigate the influence of geological features on local weather patterns and microclimates. To achieve this, we must utilise a merging algorithm that combines weather station data with geological data based on their spatial proximity. The algorithm should be able to match and link the data based on their precise geographic locations and their proximity to the geological features of interest. This enabled us to conduct a spatial analysis of the correlation between weather patterns and geological features.

The two datasets we aimed to use in this study have distinct properties. The weather station data is a continuous, time-series dataset that records various meteorological variables such as temperature, humidity, wind speed, and precipitation. Meanwhile, the geological data is a spatial dataset that captures the physical characteristics of the terrain, such as elevation, slope, and land cover. The integration of these two datasets was essential to the analysis, as it allowed us to investigate the impact of geological features on weather patterns in a comprehensive manner. By linking the data based on their geographic locations, we can identify areas where certain geological features significantly influence weather patterns and how weather stations located near these features are affected. This information is crucial for various applications such as climate modelling, weather forecasting, and natural disaster preparedness. we used the `Point to Poly` algorithm, which assigned each weather station to the nearest geological feature polygon based on their proximity. The weather station data used in this study was collected from a total of 2084 weather stations. The data encompassed various meteorological parameters, such as temperature, precipitation, wind speed, humidity, and additional measurements contingent upon the specific monitoring station and the region of interest. Meanwhile, the geological feature data from the Geological Survey of Ireland included information on topography, geology, and other relevant terrain features surrounding each weather station. In this instance, the `Point to Poly` algorithm was necessary to associate each weather station with the nearest geological feature. This was important because it provides the ability to analyze the influence of geological features on local weather patterns. Using this algorithm, we can better understand how different geological features could impact temperature and precipitation patterns in their surrounding areas, which was crucial for climate research and forecasting.

Table 3.6 provides an example of the integrated dataset that includes weather station data alongside geological features such as bedrock type, national soil classification, and Corine land cover. The purpose of including this table is to illustrate the type of integrated data used in our meta-learning approach. This integration allows us to comprehensively analyze the relationships between different environmental factors and improve the accuracy of our predictive models.

Table 3.6: Example of integrated weather stations and geological features

Tstamp	**Bedrock**	**soil**	**corine**	**P**	**T**	**vappr**	**H**	**msl**	**wdsp**	**wddir**	**vis**
1/1/2023 1:00	Limestone	urban	124	0	6.2	8.2	86	995.1	8	230	20000
1/1/2023 2:00	Limestone	urban	124	0.2	6.1	8.3	87	994.8	8.2	235	19900
1/1/2023 3:00	Limestone	urban	124	0	6.0	8.4	88	994.5	8.5	240	19800
1/1/2023 4:00	Limestone	urban	124	0	5.9	8.5	85	994.2	8.8	245	19700
1/1/2023 5:00	Limestone	urban	124	0.3	5.8	8.6	84	993.9	9	250	19600
1/1/2023 6:00	Limestone	urban	124	0	5.7	8.7	83	993.7	9.2	255	19500

For instance, our exploration into how mountain ranges influence precipitation patterns or whether urban areas experience higher temperatures compared to their rural neighbors. The integrated dataset has empowered us to conduct a spatial analysis of the intricate interplay between weather phenomena and geological characteristics, providing invaluable insights into the intricate relationship between climate and our surroundings.

Likewise, let's consider soil moisture datasets. They play a pivotal role in curbing nutrient losses and greenhouse gas emissions in the agricultural and forestry sectors. The moisture content in the soil plays a vital role in shaping water levels. When soil moisture is abundant, it aids in retaining more water in the soil, potentially causing groundwater levels to rise and elevating water levels in nearby water bodies like rivers, lakes, and aquifers. Conversely, when soil moisture is scant, it curtails water runoff into these water bodies, possibly causing a reduction in their water levels. As such, soil moisture levels wield a direct influence on water level [224].

This analysis could be helpful in various applications, from climate modelling and forecasting to urban planning and natural resource management. By understanding how geological features impact local weather patterns, we can better predict and adapt to climate change's impacts and make informed land use and resource management decisions.

3.3.3 Integration of Water Sensor Data with Geological Features for Assessing Water Quality Patterns

A key component for meta-learning is annotating the datasets with domain-specific knowledge such as geological characteristics. As such, further spatiotemporal integration was required to integrate data obtained from water level sensors with the geological features surrounding the sensor's placement and features of the wider river basin. In merging these datasets, we aimed to learn more about the regional patterns and relationships between geological characteristics and water quality assessments. The water sensor data included readings from various stations placed strategically throughout the river basin. These stations continuously measured turbidity, pH, and other water quality indicators. In contrast, the geological feature data was acquired from the Geological Survey, Ireland and included details on the kind of bedrock, the soil, and the land cover.

Using the `Point to Poly` algorithm and considering the water sensor and the geological features' spatial proximity, we linked each water sensor station to the closest geological feature polygon.

Table 3.7: Example detailing integration of water level sensor data and geological features

Water_N	**Water_ID**	**Lat**	**Lon**	**Unit**	...	**Area**	Formation
Aclint	06026	53.92	-6.64	GLYDE	...	Deep marine turbidite sequence	Palaeozoic, Silurian
Brewery Park	06015	53.99	-6.42	RAMPARTS MAIGUE	...	Deep marine turbidite sequence	Palaeozoic, Silurian
Riverstown	06033	54.01	-6.54	BOYNE	...	Continental sedimentary deposits	Mesozoic, Triassic
Dundalk Bay	06047	53.95	-6.37	CASTLETOWN	...	Shallow marine limestone formation	Cenozoic, Triassic
Kilcurry	06052	54.03	-6.48	FANE	...	Deltaic sandstone deposits	Palaeozoic, Devonian

The example presented in Table 3.7 showcases the integration of individual water level sensor data with wider geological features highlighting a part of the bedrock and water sensor data integration. It effectively demonstrates how the integration has been carried out, with information on the station number, location (latitude and longitude), geological unit, area, and formation.

3.3.4 Geology and Topology and Climate

In this section, we aim to investigate the influence of geological and topological features on climate patterns in Ireland. To achieve this, we have constructed a map by aggregating data from various sources. This comprehensive dataset enables us to explore and gain a visual understanding of the geological and topological features in Ireland and their relationship with climate variables.

Objectives and Methodology

Our primary objective is to understand how geological features, such as bedrock composition and topography, interact with climatic factors like temperature, precipitation, and wind patterns. This understanding is crucial for improving climate models and predicting local climate variations.

To accomplish this, we utilized several data layers including:

- Mineral Deposits and Occurrences
- Public Drillholes
- Bedrock Dykes
- NRFA Stations (accessible to the public)
- Water Regions Surface
- Water Quality
- Groundwater Quality
- National Soils Database
- Hydrometric Stations
- Weather Stations

These layers were integrated to create a detailed map that pinpoints the locations of different monitoring stations across Ireland. By analyzing the spatial distribution of these features, we can examine how geological and topological characteristics influence local climate conditions.

Significance and Applications

The integration of these data layers allows us to conduct a comprehensive spatial analysis, providing insights into:

- How bedrock composition affects soil moisture and, consequently, local weather patterns.
- The impact of topographical features such as mountains and valleys on wind and precipitation distribution.
- The relationship between groundwater quality and surface water dynamics.

For instance, Figure 3.2 displays the bedrock information and the location of different hydrometric stations within the Neghan RDB. This visualization helps in identifying patterns and correlations that are not immediately obvious through numerical data alone.

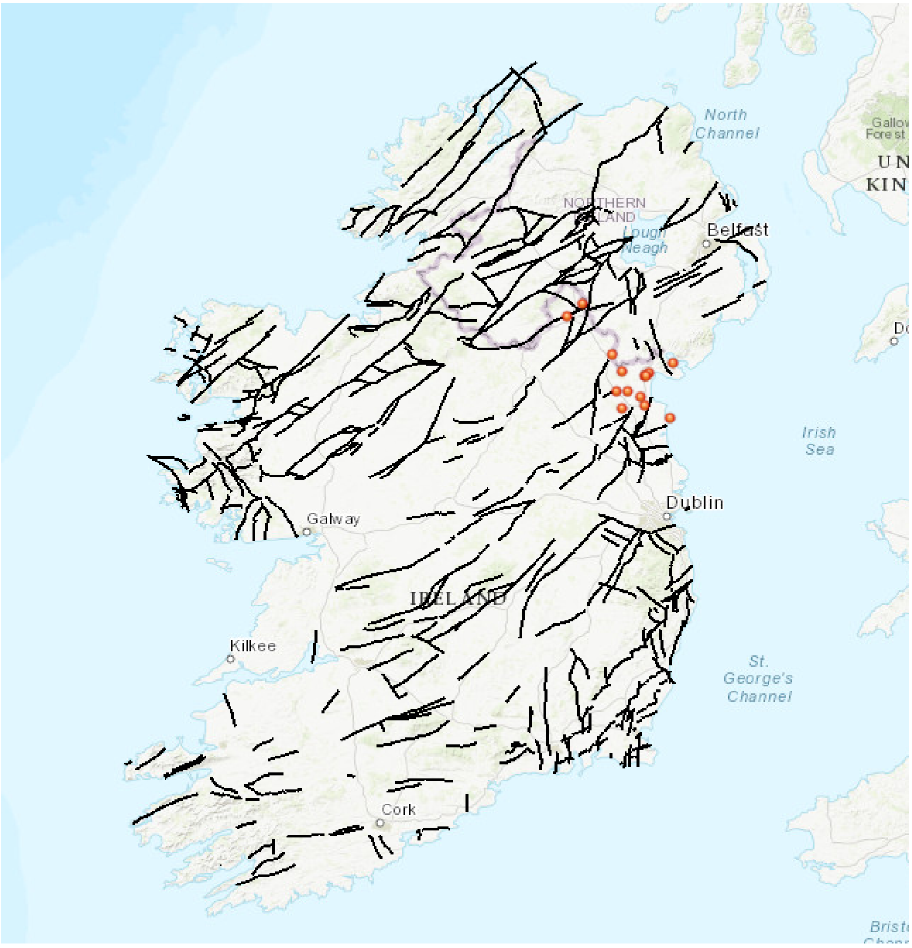

Figure 3.3: Geological Bedrock Data and Distribution of Hydrometric Stations on the Neaghan RDB

By understanding these interactions, we can enhance climate models, improve weather forecasting accuracy, and inform land use and resource management decisions. This research is particularly relevant for climate adaptation strategies, helping policymakers and scientists to develop more effective measures to mitigate and adapt to climate change impacts.

3.4 Summary

This chapter presents a methodological architecture leveraging ETL technology to integrate heterogeneous environmental data from multiple sources. The approach addresses the challenges of data sourcing, extraction, transformation, loading, mapping, spatiotemporal integration, and data quality validation, providing a solution for Environmental Spatiotemporal Data Integration (STDI). The case studies of STDI in Ireland demonstrate the applicability and effectiveness of the proposed architecture in handling and integrating diverse data types, formats, and sources. The distance-based algorithms for spatial data integration and the temporal data integration techniques used in this approach ensure seamless data merging, optimal alignment, and minimal information loss. The algorithms implemented have been tailored to address specific spatiotemporal integration challenges within environmental data contexts. Their design aims to balance efficiency and accuracy, facilitating the integration of diverse datasets crucial for comprehensive environmental analysis. The successful integration of real-world datasets, as illustrated in the case studies, provides practical evidence of the architecture's capabilities. This methodological architecture can be adapted and applied to various fields where integrating heterogeneous data from multiple sources is essential for practical analysis, visualization, and decision-making. The customizable rules and parameters within the algorithms allow users to tailor the integration process to their specific needs, making it a versatile solution for diverse integration scenarios.

Future work in this area may focus on exploring the application of machine learning and artificial intelligence techniques to enhance the automation and accuracy of the data integration process. Additionally, the development of more advanced algorithms and tools to handle complex data types and relationships, as well as the integration of real-time streaming data, can further improve the capabilities of the proposed architecture.

In conclusion, the proposed methodological architecture provides a practical and efficient solution for integrating heterogeneous environmental data from multiple sources, enabling comprehensive analysis, visualization, and decision-making in spatiotemporal contexts. The case studies have demonstrated its effectiveness, proving its value in real-world applications and offering a foundation for further enhancements.

Chapter 4

Evaluating Machine Learning Models for River Water Level Predictions

The preceding chapter introduced a system designed to efficiently integrate spatiotemporal data into a unified dataset. Leveraging this dataset as our foundational resource, this chapter focuses on a critical endeavour: evaluating and selecting the optimal machine learning model for predicting river water levels. Through rigorous analysis, we seek to elucidate the strengths and limitations of each model, thereby constructing a dataset of experimental results that will underpin the development of the meta-learner.

4.1 Introduction

Building upon the previous chapter's emphasis on spatiotemporal data integration, this chapter explores the practical applications of the integrated dataset in addressing real-world challenges. One such application is the forecasting of river water levels, which demands a combination of precision and adaptability. Accurate river water-level predictions are crucial for managing water resources, mitigating the risks associated with floods and droughts, and enhancing decision-making processes. Traditional methodologies often employ hydrodynamics and differential equations to model the complex physical processes involved in runoff, river flow, and confluence [104] [154]. Despite their ability to accurately model natural phenomena, these traditional approaches can be resource-intensive, requiring substantial computational power, detailed data, and specialized expertise. Therefore, there is a need for improvement, particularly in providing real-time forecasts that are both timely and actionable for decision-makers.

The field of hydrology has significantly benefited from advancements in machine learning (ML). These novel approaches have garnered considerable attention, driven by the increased availability of data and the pressing need to address the impacts of climate change and human activities on the environment [117] [216]. Machine learning models, such as artificial neural networks (ANNs), are now considered promising tools that have the potential to address the limitations of traditional models. These models have significantly expanded the toolkit available for hydrological forecasting, offering a range of options that provide high accuracy and adaptability.

4.2 Study Area and Data

The study area encompasses the river basin districts (RBDs) in Ireland, which are divided into seven districts:

- Shannon (S)
- South Eastern (SE)
- Western (W)
- Eastern (E)
- South Western (SW)
- North Western (NW)
- Neagh Bann (NB)

These districts, as shown in Figure 4.1 (source: OPW RBD), are responsible for monitoring the water-level status of rivers in Ireland, including the shared water courses with Northern Ireland.

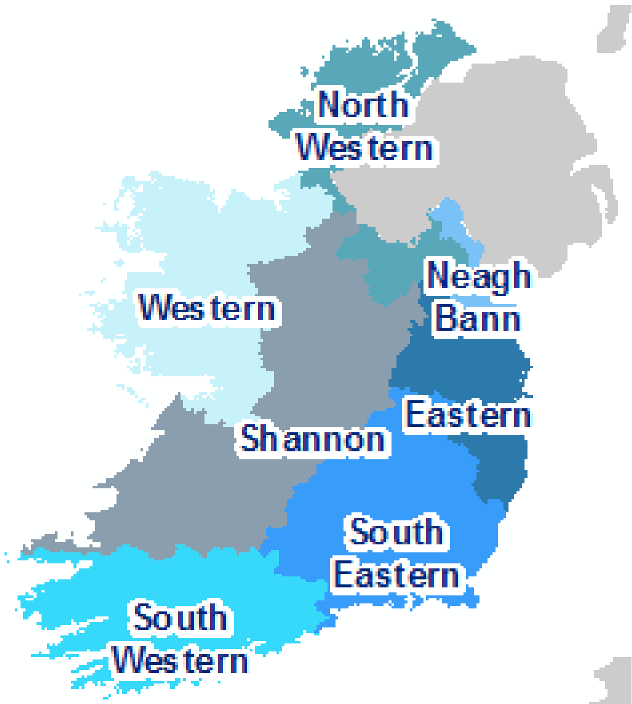

Figure 4.1: River Basin District from Office of Public Works [152]

The water-level time series (datasets) were collected from the Office of Public Works (OPW). OPW's hydrometric surface water network currently has 386 stations.Figure 4.2, Figure 4.3, Figure 4.4, Figure 4.5, Figure 4.6, Figure 4.7 and Figure 4.8 below display the hydrometric station distribution within each RBD:

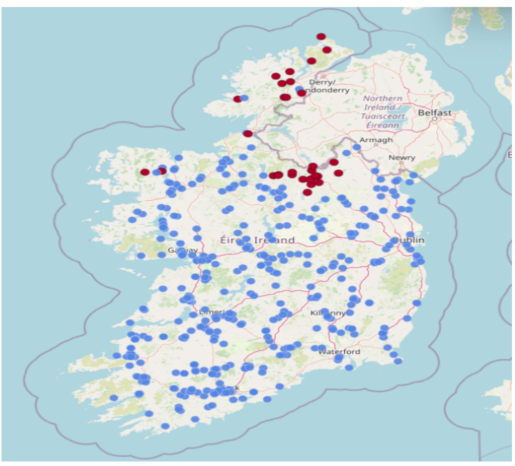

Figure 4.2: North Western (NW)

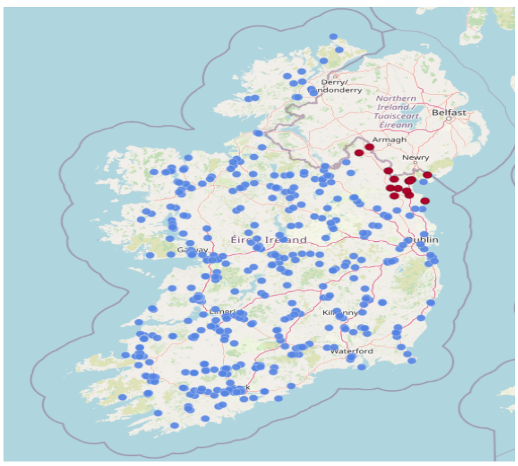

Figure 4.3: Neagh Bann (NB)

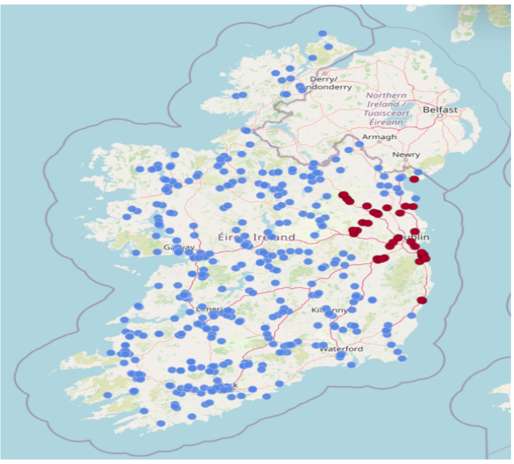

Figure 4.4: Eastern (E)

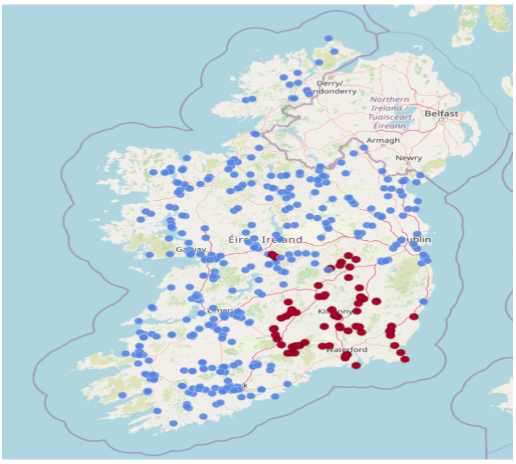

Figure 4.5: South Eastern (SE)

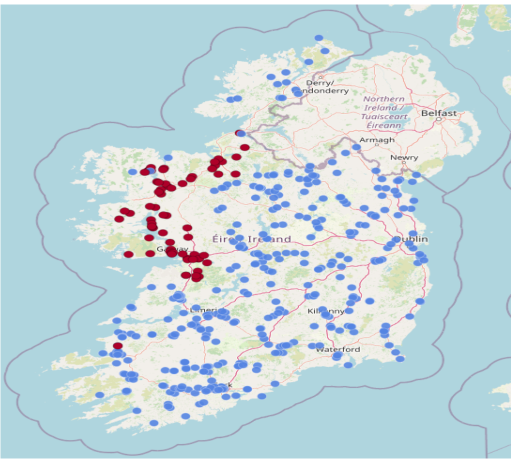

Figure 4.6: Western (W)

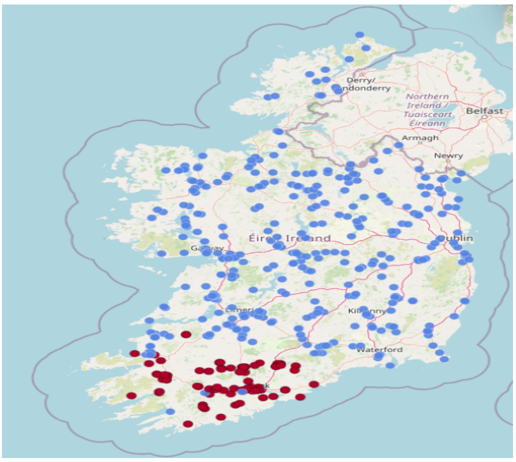

Figure 4.7: South Western (SW)

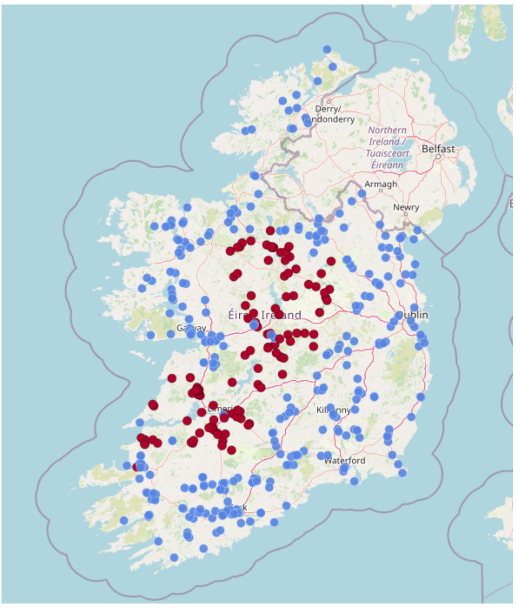

Figure 4.8: Shannon (S)

4.3 Exploratory Analysis

When starting any analysis related to time series (and most data in general), one of the first steps is visualising the data. This aims to inspect the data in order to gain insights and extract meaningful information from it, for example:

- Does the data exhibit any patterns?
- Are there any atypical observations (outliers) found in the data set?
- Is there evidence of non-stationarity in the properties of the series of observations over time?
- Are there any correlations or associations observed between the variables? If multiple variables are presented, it is essential to consider each variable and its implications carefully.
- Are there identifiable clusters of flow patterns?

For this initial assessment of our approach, we will use the Aclint station as our illustrative example. The Aclint station is located within the GLYDE catchment area. It is identified by station number 6026 and belongs to the Neagh Bann (NB) RBD. The station is situated near the LAGAN river, specifically in the GLYDE section of the river. Its geographic coordinates are approximately 53.92 degrees latitude and -6.64 degrees longitude. The station serves as an illustrative example in our analysis or study.

4.3.1 Trend Analysis of Water-Level Data

Trend analysis is crucial for understanding the behavior and patterns within hourly water-level data recorded at a hydrometric station. It helps uncover underlying trends, seasonal variations, and short-term fluctuations, providing valuable insights for purposes such as flood prediction, water resource management, and environmental monitoring.

In this section, we explore the selection of appropriate window sizes for calculating moving averages. Different window sizes capture short-term, mid-term, and long-term trends, offering a comprehensive view of water-level patterns at Aclint Station.

4.3.1.1 Selection of Window Sizes

Choosing an appropriate window size for calculating a moving average depends on the data's characteristics and the analysis objectives:

- **Trend Detection:** Larger window sizes, such as a 24-hour window, capture long-term trends and smooth out noise, aiding in the identification of daily trends.
- **Short-Term Trends:** Smaller window sizes, like a 12-hour window, reveal short-term patterns and diurnal fluctuations.
- **Data Granularity:** High-volatility data may require larger window sizes to effectively detect meaningful trends.
- **Experimentation:** Testing various window sizes helps determine the most insightful for the specific analysis.

4.3.1.2 Applying Different Window Sizes

We applied various window sizes to the water-level data to capture different trend durations:

- **12-hour window:** Captures short-term diurnal patterns (Figure 4.9).
- **24-hour window:** Highlights daily trends (Figure 4.10).
- **7-day window:** Reveals weekly cycles (Figure 4.11).
- **30-day window:** Shows monthly variations (Figure 4.12).
- **90-day window:** Identifies quarterly patterns (Figure 4.13).

- **365-day window:** Displays annual trends (Figure 4.14).

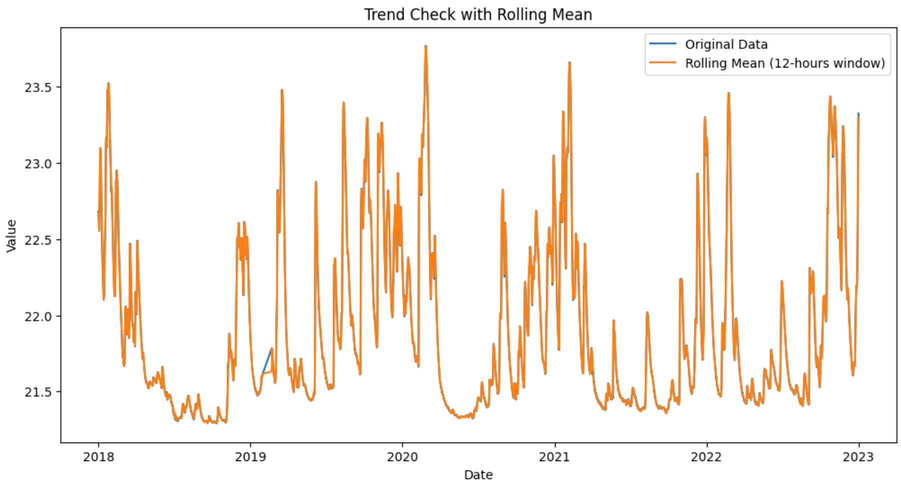

Figure 4.9: Short-term trend (12 hours) for water-level data (Aclint Station)

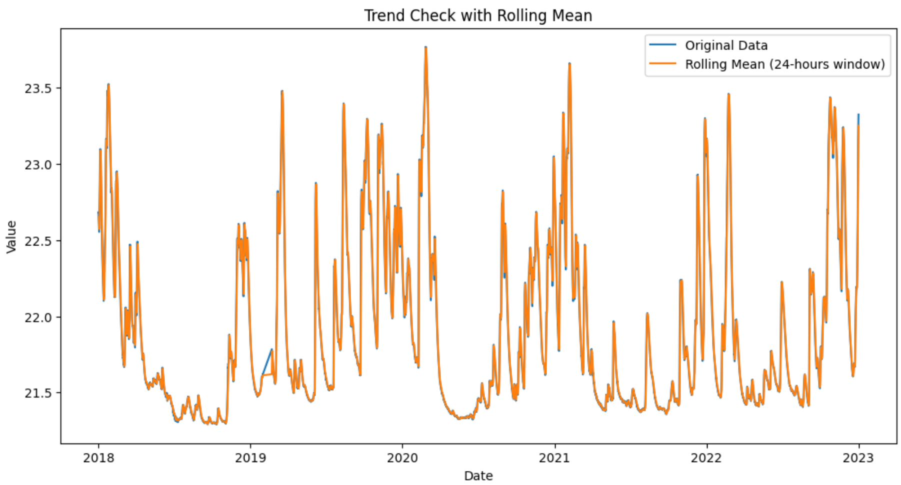

Figure 4.10: Daily trend (24 hours) for water-level data (Aclint Station)

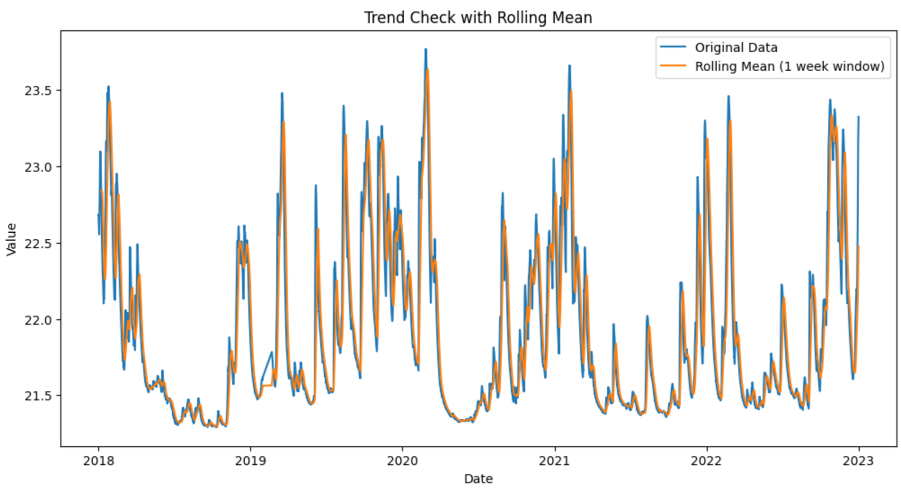

Figure 4.11: Weekly trend (7 days) for water-level data (Aclint Station)

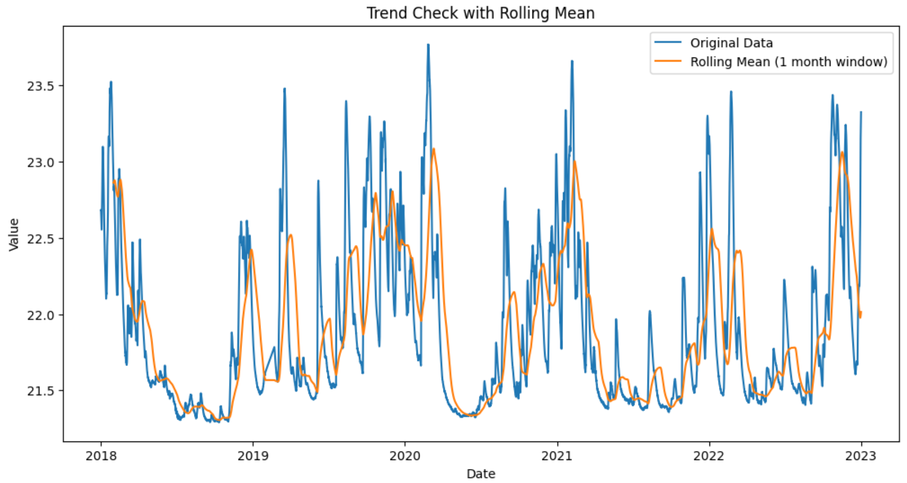

Figure 4.12: Monthly trend (30 days) for water-level data (Aclint Station)

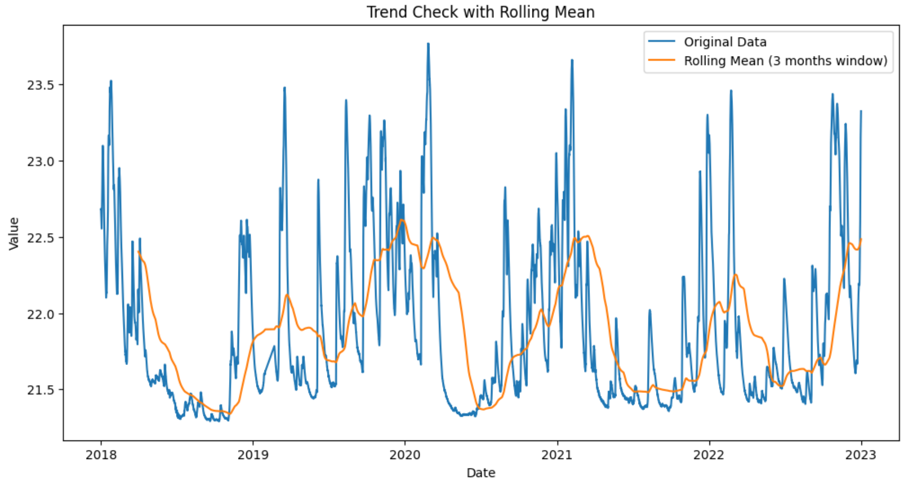

Figure 4.13: Quarterly trend (90 days) for water-level data (Aclint Station)

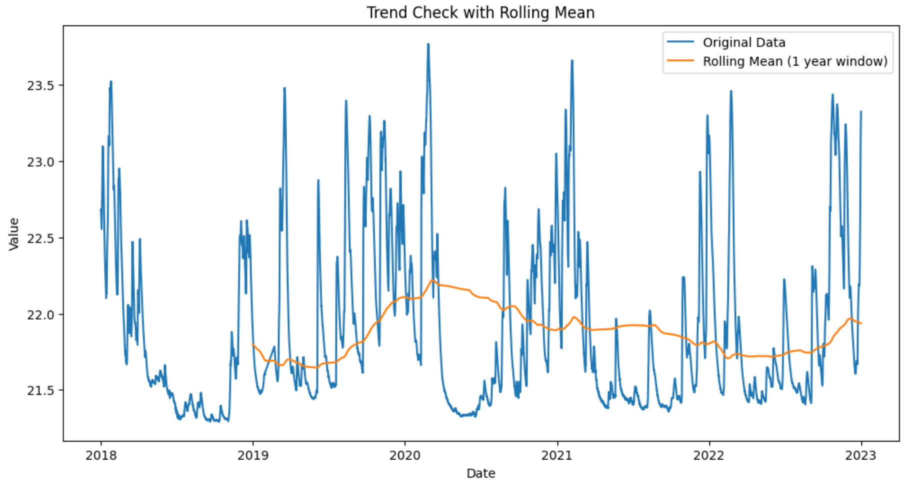

Figure 4.14: Annual trend (365 days) for water-level data (Aclint Station)

In conclusion, the initial selection of a 24-hour window (equivalent to one full day) for the computation of the rolling mean is advantageous for discerning daily trends, which often hold significance in analyzing such temporal datasets. Trend analysis using moving averages provides a foundation for model selection and further time series analysis, helping to identify the most suitable models for predicting water levels.

Figure 4.15 illustrates the comprehensive application of different window sizes for capturing trends in water-level data at Aclint Station.

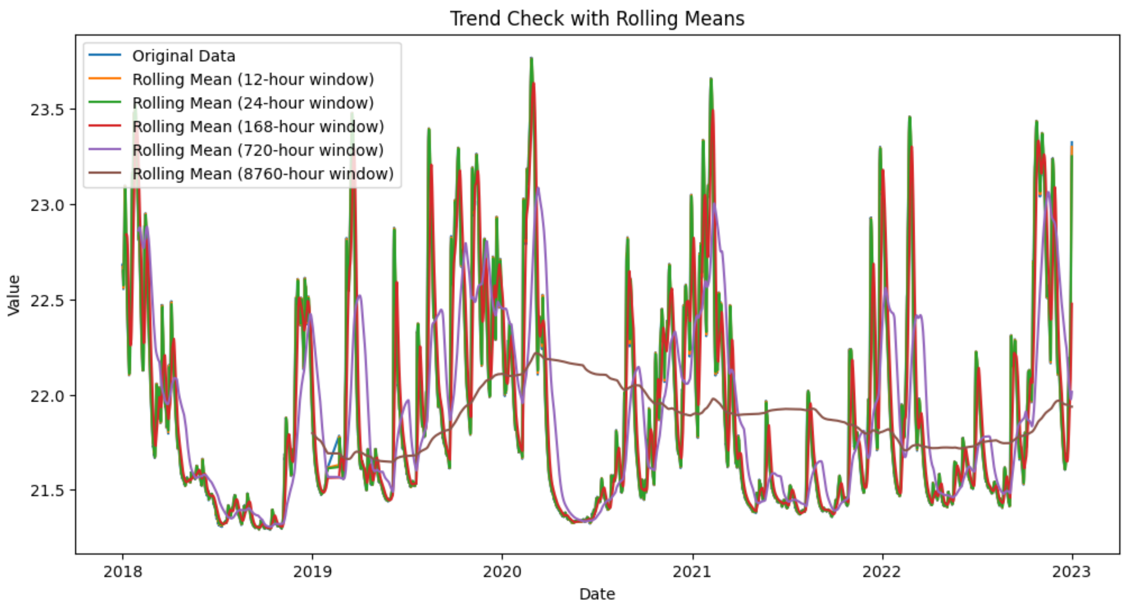

Figure 4.15: Trends for water-level data (Aclint Station)

4.3.2 Autocorrelation and Partial Autocorrelation Analysis

This section explores the autocorrelation and partial autocorrelation patterns within our time series data. Autocorrelation, also known as serial correlation, is the measure of the correlation between observations of a time series separated by various lags. Essentially, it quantifies the relationship between a variable's current value and its past values, providing insights into the degree of dependency within the data.

These analyses provide valuable insights into the temporal dependencies and potential seasonality present in the dataset, which are crucial for selecting an appropriate time series model for forecasting or further analysis. Autocorrelation func-

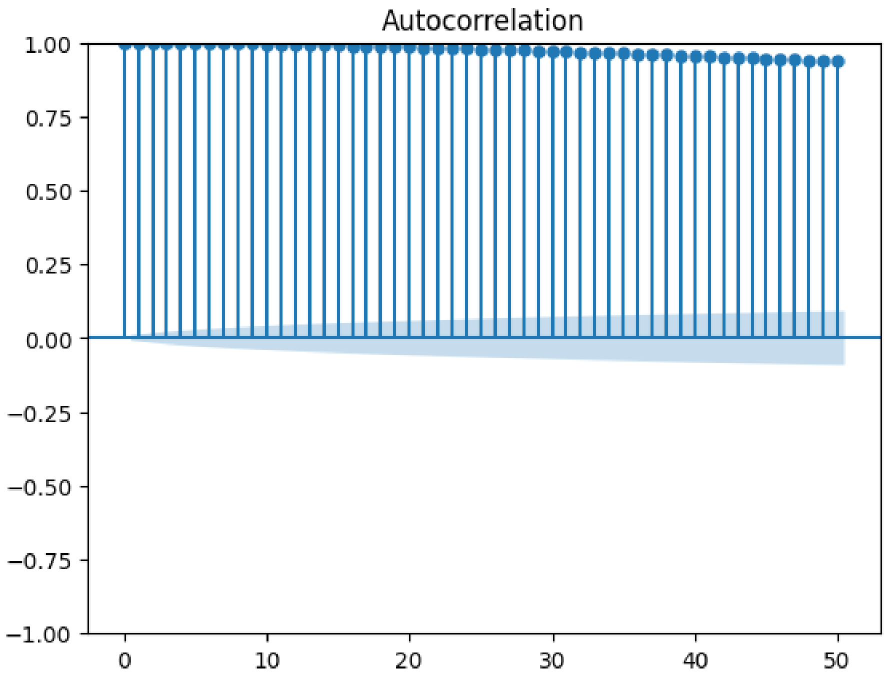

Figure 4.16: Autocorrelation for water-level data (Aclint Station)

tions (ACF) and partial autocorrelation functions (PACF) help identify the extent and nature of these dependencies, enabling the identification of significant lags that contribute to the series' behavior.

The ACF measures the correlation between the series and its lagged values, capturing both direct and indirect effects. In contrast, the PACF isolates the direct effect of a particular lag by removing the influence of intermediate lags. Together, these tools allow for a comprehensive understanding of the internal structure of the time series, guiding the model selection process to enhance forecasting accuracy.

The trend, autocorrelation, and partial autocorrelation analyses performed earlier help to reveal the structural characteristics of the water-level time series data, which will guide the modeling approach in subsequent sections.

ACF (AutoCorrelation Function): The ACF value at lag 1 being suggests a strong positive autocorrelation at lag 1 as shown in Figure 4.16. The ACF starts dropping slightly around lag 20 and reaches almost 0.90 around lag 50. This suggests a positive autocorrelation that persists for a relatively long period (lags 20 to 50). There might be a seasonal or cyclic pattern with a periodicity of approximately 50-time units.

The drop in ACF values as we move further from lag 1 is expected. If we notice a spike at lag one and a gradual decrease, it indicates a first-order autoregressive (AR(1)) component in the data. This typically means:

1. AR(1) Process: The time series data may follow an autoregressive process of order 1. In other words, the current value depends linearly on the previous value with a lag of 1. This is a typical pattern in many time series data.

2. No Clear Seasonality: The gradual decrease in ACF values without any recurring peaks at specific lags suggests that there might not be a seasonality in the data, at least not one that dominates the autocorrelation structure.clear

To further investigate the extent of the lags and identify the order of an autoregressive model, we now look at the partial autocorrelation function (PACF).

PACF (Partial AutoCorrelation Function): The initial two lags at 1 in the PACF indicate a strong correlation with the first two lags as shown in Figure 4.17. The gradual ascent in the PACF from -0.25, stabilizing at 0 at around lag 20, suggests a possible moving average (MA) component in the time series model. The initial negative values followed by stabilization at 0 indicate that there might be some seasonality or cyclic behaviour captured by the MA component.

Based on the ACF and PACF plots (Figure 4.16 and Figure 4.17), the data might follow an AR(1) process.

4.3.3 Stationarity Tests: ADF and KPS

In this section, we perform two fundamental stationarity tests, the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, to evaluate the stationarity properties of our time series data.

Augmented Dickey-Fuller (ADF) Test: The ADF test is used to test for the presence of a unit root in a time series sample. The null hypothesis of the ADF test is that the time series has a unit root (i.e., it is non-stationary). The test statistic is computed as follows:

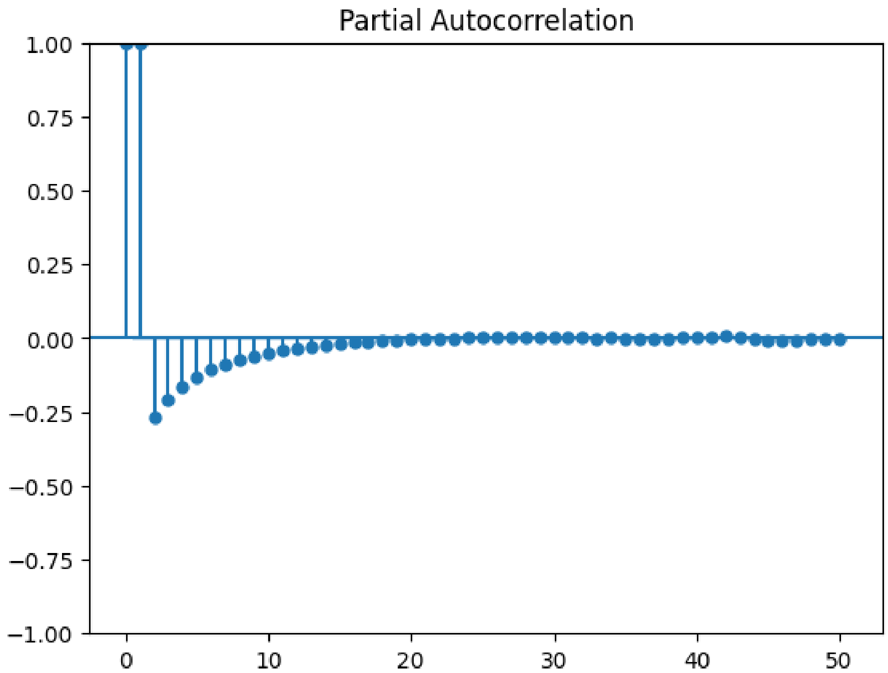

Figure 4.17: Partial autocorrelation for water-level data (Aclint Station)

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{k} \delta_i \Delta Y_{t-i} + \epsilon_t \tag{4.1}$$

where:

- ΔY_t is the first difference of the series.
- α is a constant.
- βt is the coefficient on a time trend.
- γY_{t-1} is the lagged value of the series.
- $\delta_i \Delta Y_{t-i}$ represents lagged differences of the series.
- ϵ_t is the error term.

A negative test statistic indicates stronger evidence against the null hypothesis, suggesting that the series is stationary.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test: The KPSS test tests the null hypothesis that a time series is stationary around a deterministic trend (trend-stationary). The test statistic is calculated as:

$$\text{KPSS} = \frac{1}{T^2}\sum_{t=1}^{T} S_t^2/\sigma^2 \tag{4.2}$$

where:

- T is the number of observations.
- S_t is the partial sum of residuals.
- σ^2 is the variance of the residuals.

A smaller test statistic suggests stronger evidence against the null hypothesis of stationarity.

Table 4.1: Stationarity test results (Aclint Station)

	Test	ADF	KPSS
Test Statistic		-6.4315	0.4782
p-value		1.6932e-08	0.0466
Critical Values	*1%*	-3.4305	0.7390
	2.5%	-	0.5740
	5%	-2.8616	0.4630
	10%	-2.5668	0.3470

Table 4.1 presents the results of the ADF and KPSS tests conducted to assess the stationarity properties of the time series data collected from Aclint Station.

Augmented Dickey-Fuller (ADF) Test Results:

- **Test Statistic:** The ADF test statistic is -6.4315, which is quite negative, indicating strong evidence against the null hypothesis of non-stationarity.
- **p-value:** The p-value is 1.6932e-08, which is extremely small, suggesting we can reject the null hypothesis.
- **Critical Values:** The test statistic is more negative than all critical values at the 1%, 5%, and 10% levels, further supporting the rejection of the null hypothesis.

The ADF test indicates evidence for stationarity, with a highly negative test statistic and an extremely small p-value.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test Results:

- **Test Statistic:** The KPSS test statistic is 0.4782.
- **p-value:** The p-value is 0.0466, which is less than the typical significance level of 0.05, indicating we can reject the null hypothesis of stationarity.
- **Critical Values:** The test statistic is less than the critical values at the 10% and 5% levels but greater than those at the 2.5% and 1% levels.

Based on the KPSS test results, the evidence suggests that the Aclint hydrometric station time series data may not be strictly stationary but might be trend-stationary. Trend-stationarity means that the data has a constant mean but may have a changing variance over time.

In summary, both tests provide evidence regarding stationarity, with the ADF test suggesting stationarity and the KPSS test indicating potential trend-stationarity.

4.4 Methodology

4.4.1 Methodology Overview

Our study aims to build accurate predictive models based on historical data for river water-level forecasting. To achieve our research goals, this section provides a detailed explanation of the methodology used to create and assess accurate predictive models for forecasting river water levels using machine learning algorithms. The subsequent subsections outline our approach in depth.

The flowchart of the proposed method is illustrated in Figure 4.18.

The flowchart details the sequential steps involved in the process, starting from defining the model's purpose and context, followed by inputting raw data. This data undergoes feature engineering, which includes normalization, handling outliers, managing missing data, and extracting relevant features. A machine learning model is then constructed using the engineered features. Subsequently, the top k features are selected to optimize the model, and the best-performing model is chosen based on a set of performance metrics. The final step involves validating, testing, and verifying the model on unseen data to ensure its reliability and generalizability. Throughout the process, reassessment and iteration are conducted as necessary to refine the model and improve its performance.

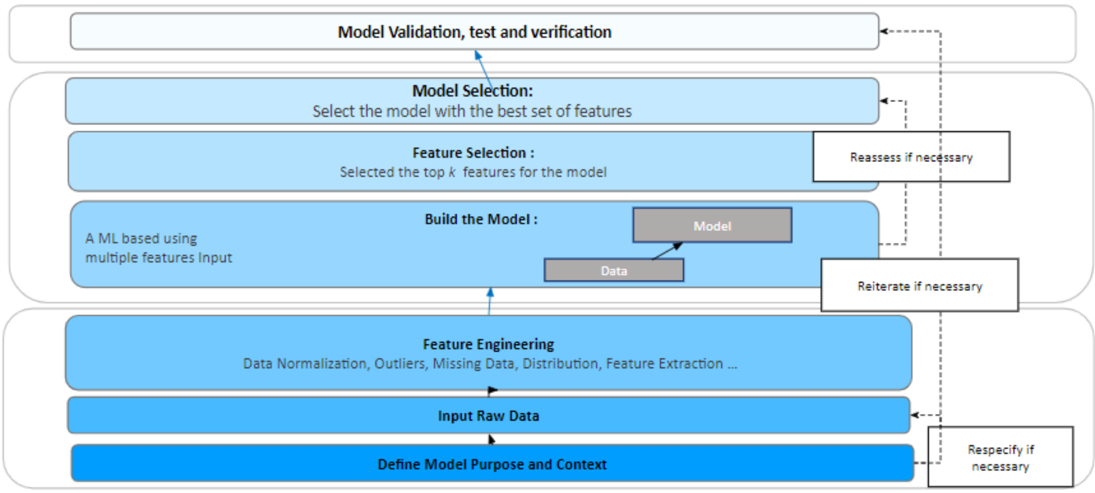

Figure 4.18: Flowchart of the proposed method in this study.

4.4.2 Data Preprocessing

Environmental data, including water level, temperature, wind, and rainfall, is collected as time series data. The data undergoes preprocessing to ensure its integrity and suitability for machine learning testing.

4.4.3 Machine Learning Models

In this research, a diverse set of machine learning models is employed to predict water levels at multiple locations. The objective is to build models that can forecast water levels based on historical data, leveraging the relationships between past water-level observations and future data. Each model offers unique capabilities, making it suitable for specific prediction scenarios.

1. **Baseline Model**: The baseline model serves as a fundamental reference point for comparing the performance of more complex models. Its primary function is to provide a simple yet meaningful prediction. In this case, the baseline model predicts "no change" in the water level by merely returning the current water-level as its prediction. This decision is based on the understanding that water levels typically change gradually over time, making the current level a reasonable approximation for the near future. However, it's important to note that as we project further into the future, the accuracy of this baseline model is likely to diminish. Additionally, the figure displaying the model's results may show a noticeable shift to the right by one hour when compared to the actual labels. This shift is due to factors such as data processing or inherent characteristics of the baseline model.

2. **Linear Model**: The Linear Model builds upon the baseline approach by incorporating a linear transformation between the input data and the predicted output[176]. In this model, predictions are generated independently based on consecutive time steps. The primary objective of the linear model is to identify and capture straightforward linear relationships within the data. It is essentially a single-layer neural network with a single neuron, using a linear activation function and a learning rate of 0.001. The model undergoes a specified number of training cycles, or epochs, set at 20 in this case.

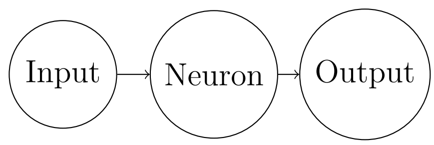

Figure 4.19: Structure of the Linear Model (Single Perceptron).

3. **Dense Model**: A dense model, also known as a "fully connected" model, is a type of neural network where all neurons in one layer are connected to every neuron in the next layer. This means that each neuron in a dense layer receives input from all neurons in the previous layer. The Dense Model in this study consists of a single fully connected layer with ReLU activation functions and a

learning rate of 0.001. Dense models are versatile and well-suited for various applications, including river water-level predictions, where intricate patterns may exist in the data[190].

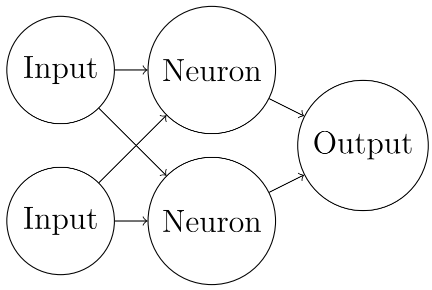

Figure 4.20: Structure of the Dense Model (Single Fully Connected Layer).

4. **MultiDense Model**: The MultiDense Model, similar to the dense model, introduces additional dense layers positioned between the input and output layers. This architectural depth enhances the model's capability to discern intricate patterns and relationships inherent in time series data. The MultiDense Model used in this study comprises three fully connected layers, each with ReLU activation functions and a learning rate of 0.001. This augmented depth results in a more intricate model architecture, affording the potential to capture nonlinear dependencies that a simple linear model might overlook [159].

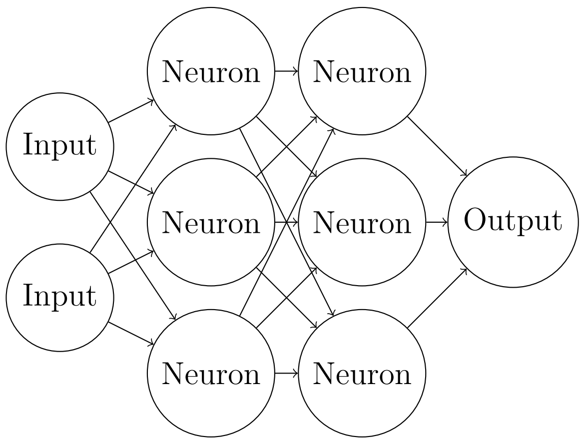

Figure 4.21: Structure of the MultiDense Model (Multiple Fully Connected Layers).

5. **CNN Model**: Convolutional Neural Networks (CNNs) are renowned for their exceptional feature extraction capabilities. In this context, a CNN model is

specifically designed for water-level prediction. To ensure a consistent training and plotting process, the output sequence length is adjusted to match the labels. CNNs excel at identifying spatial and temporal patterns within data, making them a valuable tool for time series forecasting, especially when there are underlying patterns and features that can be extracted.

6. **RNN Model**: Recurrent Neural Networks (RNNs) are a specialized class of neural networks tailored for modeling sequential data. They are highly suitable for time series forecasting tasks because they can capture the sequential dependencies and temporal patterns present in the data. RNNs work by maintaining an internal state that evolves as new information is processed, allowing them to exhibit behavior akin to the human brain when processing sequential data [155].

7. **LSTM Model**: Long Short-Term Memory (LSTM) models are a specific type of RNN that addresses the vanishing gradient problem, which is a common issue in training deep neural networks. LSTMs are equipped with memory cells that can store and retrieve information over longer sequences, making them particularly effective for modeling long-range dependencies in time series data. Fine-tuning the LSTM model involves selecting appropriate hyperparameters such as the training window size, the number of batches, and the number of training epochs [42].

8. **GRU Model**: Gated Recurrent Unit (GRU) models are a more advanced variant of standard RNNs and are designed to mitigate the vanishing gradient problem while retaining some of the efficiency of RNNs. GRUs are especially adept at capturing long-term dependencies in time series data, which makes them a valuable choice for tasks like water-level prediction. Their "gated" architecture allows them to control the flow of information through the network, enabling them to capture relevant information over extended sequences without losing context [64, 155].

9. **LSTM with Attention**: LSTM with attention combines the capabilities of LSTM (Long Short-Term Memory) and attention mechanisms. This fusion allows the model to focus on specific parts of the input sequence, enabling it to capture fine-grained dependencies within water-level data. This model proves advantageous when certain time steps or patterns in water-level data demand more attention for accurate predictions. It excels at handling situations where localized details are critical for forecasting [150].

10. **Autoencoder**: Autoencoders are adept at learning compressed representations of input data. In the context of water-level predictions, they can capture the most pertinent features within the data and provide a concise representation suitable for forecasting. Autoencoders find utility when there's a need to reduce dimensionality or extract meaningful features from water-level data. They are particularly valuable for preprocessing data for subsequent prediction models [1]. Autoencoders are not used directly for prediction in this study. Instead, they are employed for feature extraction and dimensionality reduction. By compressing the input data into a lower-dimensional representation and then reconstructing it, autoencoders can capture essential patterns and structures in the data. These extracted features are then fed into other predictive models, such as dense models or CNNs, to enhance their performance. This approach leverages the strengths of autoencoders in uncovering complex data patterns, which are subsequently used to improve the accuracy and robustness of our water-level prediction models.

11. **Transformer**: Transformers represent formidable models for capturing global dependencies and long-term patterns in water-level data. They employ self-attention mechanisms to process sequences, making them adept at capturing intricate relationships. Transformers excel when dealing with water-level data that exhibits global dependencies and non-local interactions. They are well-suited for applications where a holistic understanding of the data is crucial [214].

12. **Variational Autoencoders (VAEs)**: Variational Autoencoders (VAEs) possess the capability to model the underlying distribution of data. In the context of water-level predictions, they can capture the uncertainty and variability within the data, offering insights into a range of possible outcomes. VAEs find their strength in situations requiring probabilistic modeling and the generation of diverse samples. They are valuable for understanding the probabilistic nature of water-level data and assessing potential scenarios.

Each of these models offers unique capabilities, making them suitable for specific prediction scenarios, providing valuable insights into water-level forecasting.

Table 4.2: Summary of Machine Learning Models

Model	**Type**	**Layers**	**Activation Function**	**Learning Rate**
Baseline	-	-	-	-
Linear Model	Single Layer Perceptron	1	Linear	0.001
Dense Model	Fully Connected NN	1	ReLU	0.001
MultiDense Model	Fully Connected NN	3	ReLU	0.001
CNN Model	Convolutional NN	Multiple	ReLU	0.001
RNN Model	Recurrent NN	Multiple	Tanh	0.001
LSTM Model	LSTM NN	Multiple	Tanh	0.001
GRU Model	GRU NN	Multiple	Tanh	0.001
LSTM with Attention	LSTM + Attention	Multiple	Tanh	0.001
Autoencoder	Encoder-Decoder	Multiple	ReLU	0.001
Transformer	Self-Attention	Multiple	ReLU	0.001
Variational Autoencoder	Probabilistic Encoder-Decoder	Multiple	ReLU	0.001

4.4.4 Model Training and Validation

The process of training a machine learning model aims to develop a predictive model that effectively captures patterns within the dataset. However, it is essential to acknowledge the possibility of the model encountering a phenomenon known as overfitting. Overfitting occurs when the model, instead of grasping the underlying structural elements of the data, tends to memorize specific idiosyncrasies present solely in the training dataset. The implication of this issue becomes particularly significant when the model is applied to entirely new, previously unseen data. In order to mitigate the potential consequences of overfitting and to assess the generalization capabilities of a model, various validation techniques are employed. These techniques involve evaluating the model's performance on distinct data subsets, thereby providing a more comprehensive understanding of its predictive capabilities and highlighting any potential limitations or biases.

The models were developed using 70% of the water level (WL) data. The rest of the data were used for model validation and testing (20% and 10%, respectively). The data was not randomly shuffled before splitting for two reasons. First, to ensure that splitting the data into windows of consecutive samples is still possible. Moreover, to ensure that the validation and test results were evaluated after the model was trained.

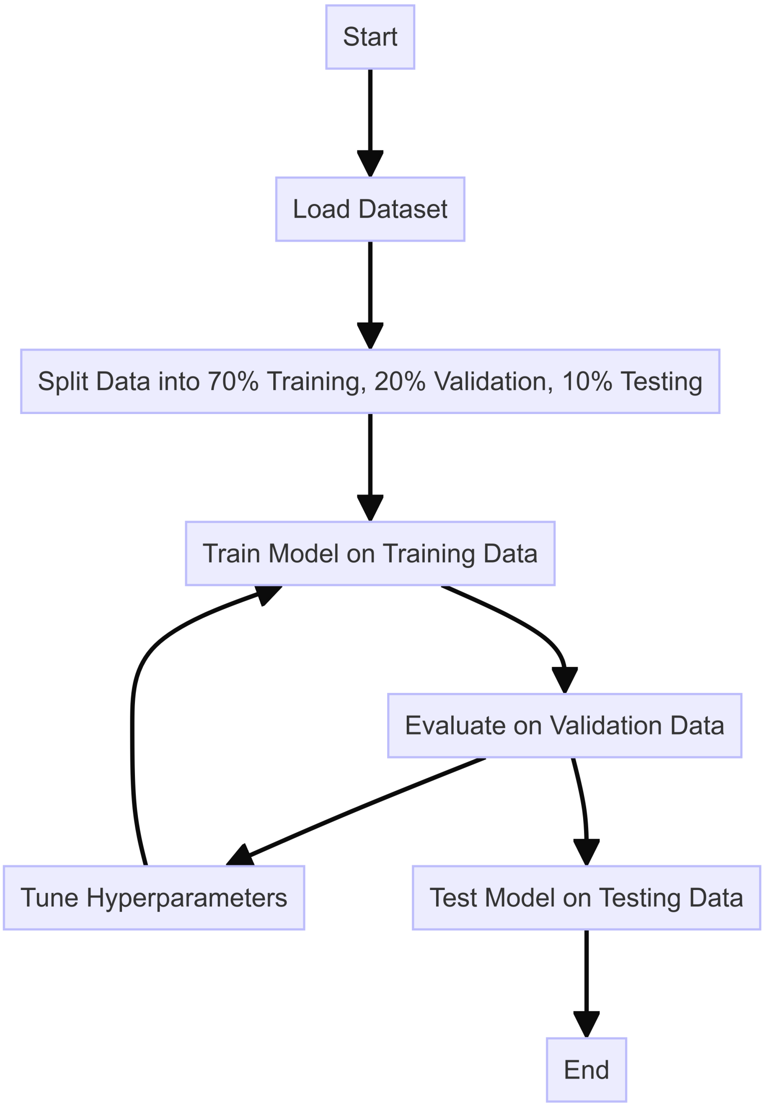

Figure 4.22: Schematic overview of the Model Training process.

Training Process:

1. **Data Splitting:** The dataset is divided into three parts: training (70%), validation (20%), and testing (10%).

2. **Model Training:** The model is trained on the training dataset. During this phase, the model learns to recognize patterns within the data.

3. **Validation:** The model's performance is evaluated on the validation dataset. This step helps in tuning the model's hyperparameters and assessing its ability to generalize to unseen data.

4. **Testing:** Finally, the model's performance is tested on the testing dataset to evaluate its predictive accuracy and generalization capabilities on completely unseen data.

By employing this structured approach, we ensure that the model is not only trained effectively but also evaluated rigorously to prevent overfitting and ensure robust performance on new data.

4.4.4.1 Normalization

The next step before training a neural network is to scale the data. Normalization is a typical way of scaling: subtracting the mean and dividing it by the standard deviation of the feature. The equation for normalizing a feature is as follows:

$$X' = \frac{X - \bar{X}}{\sigma} \tag{4.3}$$

where X is the original feature vector, $\bar{X}$ and σ are the mean and standard deviation of that feature vector, respectively.

Advantages of Normalization for Water Level Time Series:

1. Improved Model Convergence: Normalization helps in speeding up the training process by ensuring that the features have similar scales. This makes the gradient descent optimization more stable and efficient, reducing the time needed for the model to converge.

2. Prevention of Numerical Instability: Normalization prevents numerical instability during training, especially in neural networks where large values can cause the model weights to grow excessively and lead to overflow errors.

3. Enhanced Model Performance: By scaling the data, normalization helps in achieving better model performance. It allows the model to learn the weights more effectively, leading to improved accuracy and generalization on unseen data.

While normalization can sometimes lead to the loss of absolute scale information, this is not a significant issue for predictive modeling. The relationships and patterns in the data, which are crucial for making predictions, are preserved. The main goal of normalization is to facilitate the learning process of the model by standardizing the input features, thereby improving its predictive capabilities without losing essential information.

In summary, normalization is a fundamental preprocessing step in training neural networks for time series prediction. It ensures that the model training is efficient and robust, leading to better overall performance. However, it is important to choose the normalization method carefully, considering the specific characteristics of the dataset. The normalization technique that centers data around zero and scales it to have a standard deviation of one, known as mean-std normalization, is widely used for its simplicity and effectiveness, especially with datasets featuring normally distributed features. This method is crucial in machine learning, particularly for neural networks, as it promotes faster convergence during training.

Rationale for Employing Normalization in Water Level Predictions:

- Scale Consistency: Water level data, like many other types of measurements, can vary widely in scale. Normalization ensures that all features are on a uniform scale, preventing features with larger numeric ranges from unduly influencing the learning process.

- Enhanced Convergence: Normalization standardizes the data, typically centering it around zero with a standard deviation of one. This standardization facilitates the training of machine learning models, particularly neural networks, by expediting the convergence process. It mitigates issues such as vanishing or exploding gradients, which often affect unscaled data.

- Feature Balance: Normalization ensures that each feature contributes equally to the model's learning process by removing the influence of scale differences. This balance is essential for algorithms that assume a standard distribution of data.

- Robustness to Outliers: While normalization is sensitive to outliers, it generally improves the handling of these values compared to other scaling methods. Additionally, specialized normalization techniques like robust scaling are designed explicitly to manage extreme values.
- Interpretability: Normalized data tends to be more interpretable because the coefficients or weights of a model become more comparable and understandable when features share a common scale. This aids in understanding the impact of each feature on the model's predictions.
- Algorithm Compatibility: Many machine learning algorithms, including neural networks, assume data is normally distributed and centered around zero with a standard deviation of one. Normalization adheres to this foundational assumption, enabling the optimal utilization of these algorithms.
- Reduced Computational Load: In some cases, normalization can reduce the computational complexity of machine learning models, leading to faster training times and lower memory utilization.

However, it is equally important to acknowledge that alternative normalization methods are available, each tailored to address distinct requirements. For instance, the Z-score normalization method, which mirrors mean-std normalization, involves subtracting the mean and dividing by the standard deviation. In scenarios where datasets exhibit the presence of outliers, it becomes imperative to explore other scaling techniques. Robust scaling and Min-Max scaling, for instance, are specifically engineered to cope with data containing extreme values while maintaining effective normalization [71].

4.4.4.2 Data Windowing

In this research, our models generate predictions using a window of consecutive data samples. These windows are characterized by several key attributes, including the width of the input and label windows, the time offset between them, and the specific

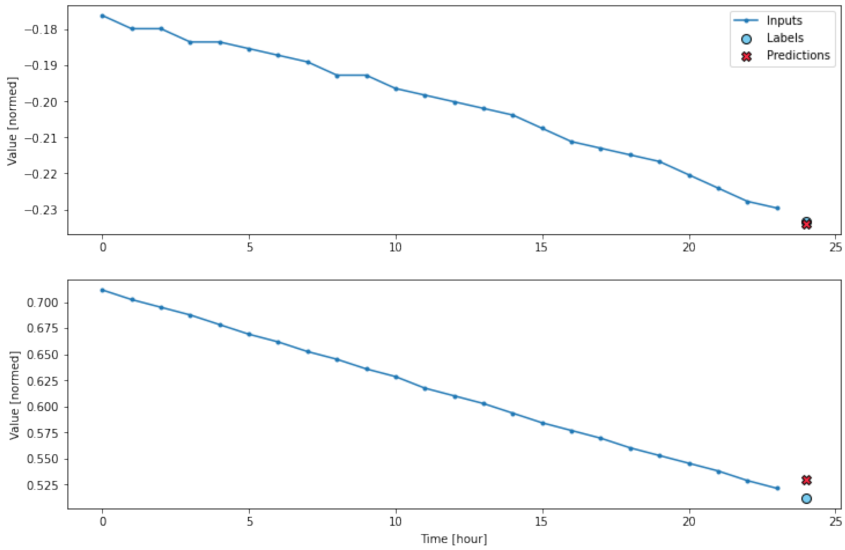

Figure 4.23: Aclint Station CNN Model Predictions (1-hour ahead)

selection of features utilized within these windows. Consequently, the training, evaluation, and test datasets undergo transformation into windowed datasets to facilitate this modeling approach. Depending on the experiment and the type of model employed, various data window configurations were created.

For example, when working with a water level dataset with hourly measurements, the configuration denoted as {`input_width=24, label_width=1, shift=1`} defines the data window setup used for training the models. Here, each input window comprises data from the past 24 time steps (e.g., the past 24 hours), the model predicts one time step into the future (e.g., the water level for the next hour), and the label window starts immediately after the input window with no gaps. This configuration is particularly relevant for scenarios necessitating hourly predictions grounded in the most recent 24-hour data. Figure 4.23 illustrates 1-hour ahead water level predictions generated by the Convolutional Neural Network (CNN) model, reflecting the model's performance in making hourly forecasts based on the most recent 24 hours of data.

On the other hand, the second window configuration, represented as {`input_width =24, label_width=24, shift=1`}, encompasses different characteristics. In this setup, the model uses the preceding 24 hours of data to predict the water levels for the next 24 hours. This configuration is apt for making daily predictions, as depicted in Figure 4.24, which presents 48-hour ahead forecasts using the LSTM model.

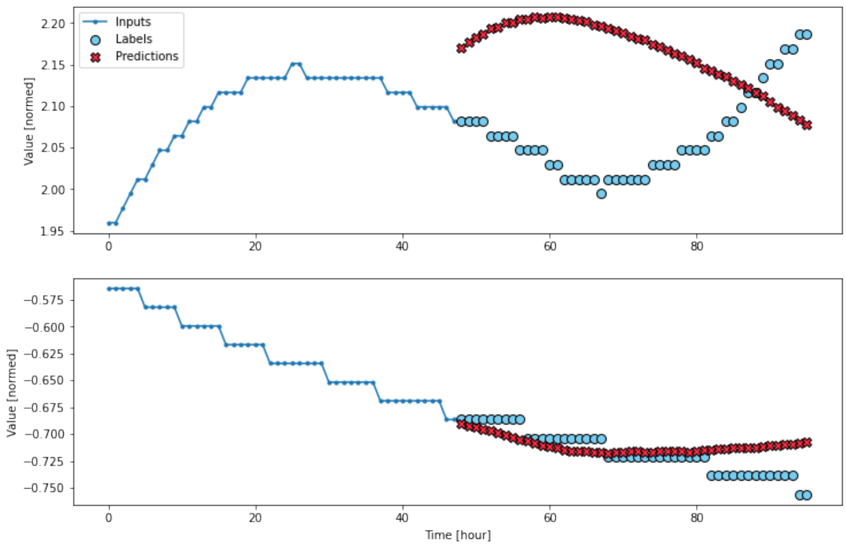

Figure 4.24: Aclint station LSTM Model Predictions (48 hours ahead)

However, while the idea of predicting beyond one hour is introduced, it was not thoroughly analyzed in the results. Increasing the prediction horizon often impacts the accuracy of the results due to the accumulation of errors in multi-step ahead predictions. This phenomenon is well-documented in the literature, where errors compound over each prediction step, leading to a significant decline in accuracy for longer prediction horizons. For instance, research by [19] highlights that multi-step predictions can suffer from what is known as "error accumulation," where small errors in each prediction step propagate and magnify, resulting in larger errors as the prediction horizon extends .

Moreover, [65] discuss the challenges in training models for multi-step predictions due to this compounding effect, which often requires sophisticated techniques such as sequence-to-sequence models or the incorporation of temporal dependencies to mitigate . These challenges are particularly relevant in the context of water level forecasting, where accurate long-term predictions are crucial for effective water resource management and flood prevention.

Studies by [192] and [68] have further demonstrated that traditional models struggle with multi-step forecasting because they do not adequately capture the long-term dependencies required for accurate predictions. This issue is particularly pronounced in recurrent neural networks (RNNs), which, despite their ability to handle sequential data, often exhibit vanishing gradient problems that exacerbate error propagation over longer sequences. The vanishing gradient problem, as described by [81], refers to the difficulty in training deep neural networks due to the gradients becoming exceedingly small, effectively preventing the network from learning long-term dependencies .

Addressing these challenges typically involves model enhancements or hybrid approaches, integrating techniques like attention mechanisms or ensemble methods. For example, [41] introduced the use of attention mechanisms in neural networks, which allow the model to focus on different parts of the input sequence when making predictions. This approach has been shown to significantly improve the performance of multi-step predictions by reducing the impact of error accumulation. Similarly, [68] demonstrated that utilizing Long Short-Term Memory (LSTM) networks, which are designed to better capture long-term dependencies, can also mitigate some of the issues associated with traditional RNNs .

Despite these advancements, the issue of error accumulation in multi-step predictions remains a significant challenge. [193] explored the bias-variance tradeoff in multi-step forecasting and concluded that as the forecast horizon increases, the model's bias and variance also increase, leading to less accurate predictions . This finding underscores the importance of carefully selecting and designing models for long-term forecasting tasks.

To address these concerns and provide a more focused analysis, we decided to limit our experiments in this chapter to 1-hour ahead predictions. This decision ensures a more reliable evaluation of model performance, which is detailed in Section 4.6. By focusing on short-term predictions, we can provide a clearer assessment of the model's capabilities without the confounding effects of error accumulation.

4.4.4.3 Hyperparameter configuration

The number of units and epochs in the ML models is essential in learning. Thus, multiple experiments were conducted to find the optimal number of units and epochs to train the model. To find the optimal number of units, we compared the model performance where the unit number was equal to 16, 32, 64, 128, and 256. The comparison results showed that:

- The best results in training were recorded when the number of units was 64, followed by the case where the units were 32.
- When the number of units was 128 and 256, the training loss value was almost the same. There was no significant difference in training loss values for the 32, 128, and 256 cases.
- When the unit number was set to 16, the models delivered the worst recorded performance.
- The validation loss value also obtained the best performance when the number of units was 64, followed by 32, 128, 256, and 16 units.

It was confirmed that the learning time was proportional to the number of units. When the number of units was 64, it was confirmed that both learning and validation showed the best performance, and when the number of units was 16, it was confirmed that the worst performance was shown. Thus, according to the experimental results, the number of units was finally determined to be 64. In order to discover the optimal epoch number, the models were trained up to 50 epochs, and the training loss values and validation loss values were compared against several epochs. The result shows that 20 epochs delivered the highest performance; hence, the epoch's max number was set to 20.

To conclude, Adam [107] was used as an optimisation function, the number of units used to train the models was 64, and the number of epochs (training iterations) was 20. Mean Squared Error (MSE), was used as the loss function for model training, and Mean Absolute Error (MAE) was used as the indicator for the validation and test to compare the observed and predicted values. The relative performance of different models developed in this study was evaluated for both the validation and test periods.

While Bayesian Optimization was not employed in this study, it represents a promising approach for hyperparameter tuning. Bayesian Optimization is a sequential model-based optimization technique that constructs a probabilistic model to map hyperparameters to the objective function. This method could help in efficiently navigating the hyperparameter space to identify optimal values by balancing exploration and exploitation, potentially leading to further improvements in model accuracy and training efficiency.

4.4.5 Models Performance Evaluation

The data is divided into training, testing, and validation subsets, where the selection of subsets can significantly affect the model performance [23]. The unseen dataset (validation) becomes a crucial element at this stage. It allows us to compare our model against data that has never been used for preparation. This test will let us see how the model can respond to the knowledge it has not seen before. It is meant to show how the model could work in the real world.

To evaluate the execution of our model, we quantified errors using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) [90]. These metrics provide a comprehensive view of the model's accuracy and reliability in predicting water levels. In addition, these metrics measure the differences between the predicted values and the actual values, enabling an evaluation of prediction quality.

The Mean Squared Error (MSE):

The MSE computes the average squared difference between the predicted values, Y', by a model or estimator and the actual values, Y, as expressed below.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2 \tag{4.4}$$

The Root Mean Squared Error (RMSE):

The root-mean-square error (RMSE) (also known as root-mean-square deviation (RMSD)) is a commonly used calculation of the variations between predicted (Y') and actual values (Y) by a model or estimator, i.e., it is the square root of MSE, as shown below.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\left(\frac{Y_i - Y_i'}{\sigma_i}\right)^2} \tag{4.5}$$

The Mean Absolute Error (MAE):

The mean-absolute-error (MAE) can be defined as the average of the entire absolute errors, as shown below.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_i'| \tag{4.6}$$

Each of these metrics captures different aspects of prediction error. MSE calculates the average squared difference between predictions and actual values, giving more weight to larger errors. This metric is sensitive to outliers and penalizes them more than the other two metrics. RMSE is the square root of MSE and provides a measure of the standard deviation of the errors. It is in the same units as the target variable, making it easier to interpret. RMSE is also sensitive to outliers but is often used when the magnitude of errors is significant. MAE computes the average of the absolute differences between predictions and actual values, treating all errors equally and being less sensitive to outliers.

These metrics are widely accepted and used across different regression problems. Having a common set of evaluation metrics (like MSE, RMSE, and MAE) makes it easier to assess and compare the performance of different models [32]. The choice of these metrics can also depend on the problem and the trade-off between bias and variance. MSE and RMSE tend to emphasize larger errors more, which might be suitable when small errors are more tolerable. MAE treats all errors equally and can be preferred when a more balanced view of error is needed. These metrics are particularly useful in the validation phase with a separate dataset (validation set) that the model has not seen during training. This enables the evaluation of the model's performance when confronted with novel, unobserved data.

While MSE, RMSE, and MAE provide robust metrics for evaluating model performance, additional evaluation metrics and analyses can offer additional insights and improve model assessment: **R-Squared (R^2)** is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. It provides an indication of how well the predicted values approximate the actual data points. Incorporating R^2 could offer more insight into the model's explanatory power [184].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - Y_i')^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

Analyzing the distribution of errors (residuals) is crucial for understanding model performance. Ideally, residuals should be normally distributed with a mean of zero, indicating unbiased and consistent predictions across different data points. Plotting residuals can help identify patterns, trends, or outliers that the model might not have captured, providing deeper insights into model performance and areas for improvement [116].

In the context of hydrology, these metrics help to assess the quality of the predictions and ensure that the model is performing well, particularly when predicting a critical variable like water levels, which can have significant real-world implications [145].

4.5 Computational Resources and Tools

This study utilized a combination of Python and R programming languages to implement and evaluate the machine learning models. Specifically, the following tools and frameworks were employed:

- **Programming Languages:** Python and R were used for data processing, model implementation, and evaluation. Python was primarily utilized for developing machine learning models, while R was employed for initial data exploration and visualization.
- **Frameworks and Libraries:** The implementation of machine learning models was carried out using the following frameworks and libraries:
  - **TensorFlow:** Used for building and training deep learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Unit (GRU) networks.
  - **Scikit-learn:** Utilized for traditional machine learning models and preprocessing tasks.
  - **Keras:** A high-level neural networks API, written in Python and capable of running on top of TensorFlow.
  - **Seaborn and Matplotlib:** Used for data visualization and plotting.
- **Development Environments:** The code development and model training were conducted primarily on Google Colab and Anaconda environments.
  - **Google Colab:** An online platform that provides free access to GPUs, used extensively for model training and experimentation.
  - **Anaconda:** A local development environment that was used for data preprocessing and initial exploratory data analysis in R.
  - **PyCharm:** An integrated development environment (IDE) used for coding, debugging, and testing the machine learning models in Python.

Computational Resources: The computational experiments were performed using the following resources:

- **Google Colab:** Utilized for leveraging free GPU resources to speed up the training of deep learning models. The specifications include:
  - **GPU:** NVIDIA Tesla K80 or T4, depending on availability.
  - **CPU:** 2.3 GHz Intel(R) Xeon(R) CPU.
  - **RAM:** 100 GB.
- **Local Machine (Anaconda Environment):**
  - **Processor:** Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz 2.11 GHz
  - **RAM:** 16 GB.
  - **Storage:** 512 GB SSD.
  - **Operating System:** Windows 11.

These computational resources and tools provided the necessary infrastructure to conduct extensive experiments, ensuring efficient processing and model training.

4.6 Experiments

In this study, we will undertake two separate experiments to evaluate the performance of machine learning models in predicting hydrometric data. The first experiment focuses on 70 hydrometric stations and eight different models, while the second experiment extends the scope to 349 stations and introduces 12 models.

4.6.1 Experiment 1: Eight Machine Learning Models on 70 Hydrometric Stations

4.6.1.1 Experimental Setup

The first experiment was designed to provide an initial evaluation of the model performance using a smaller, yet representative sample of the overall dataset. We randomly selected 70 hydrometric stations, with 10 stations chosen from each of the seven River Basin Districts (RBDs) 4.25). This approach ensures that the sample is representative of the diverse hydrological conditions across the RBDs, while keeping the dataset manageable for an initial analysis. In this experiment, we tested eight different machine learning models, which include:

- Baseline Model
- Linear Model
- Dense Model
- MultiDense Model
- CNN Model
- RNN Model
- LSTM Model
- GRU Model

The rationale for selecting this smaller subset of stations and models is to establish a foundational understanding of model performance across different hydrometric conditions. By starting with a manageable dataset and a limited number of models, we can identify initial trends and insights that can inform the design of larger-scale experiments.

As mentioned above, we randomly selected ten stations from each river basin district. Randomly selecting ten stations from each river basin district is a crucial methodology in this research, serving multiple purposes. Firstly, it ensures representative sampling, minimizing potential biases and providing an unbiased overview of the entire river basin district. Secondly, it upholds statistical validity by allowing for valid inferences about the entire population based on the characteristics of the randomly selected sample. By ensuring that each station has an equal chance of being selected, the randomness of the selection process helps to ensure that the sample accurately reflects the diversity and variability of the entire river basin district.

Moreover, this approach is efficient in terms of time and resources, striking a balance between data comprehensiveness and practicality. It also facilitates generalizability, enabling insights from the selected stations to be applied more broadly to understand hydrological characteristics and trends in the entire river basin district. Finally, it reduces the risk of bias and outliers, contributing to more reliable and robust research outcomes. In essence, random station selection enhances data quality and the applicability of research findings to the broader context, such as providing insights into hydrological patterns and trends across all of Ireland.

Exemplary location coordinates and comprehensive information pertaining to a single hydrometric station selected from each of the 7 River Basin Districts (RBDs) are meticulously documented in Table 4.3. For a comprehensive listing encompassing all chosen stations, please refer to Appendix B.

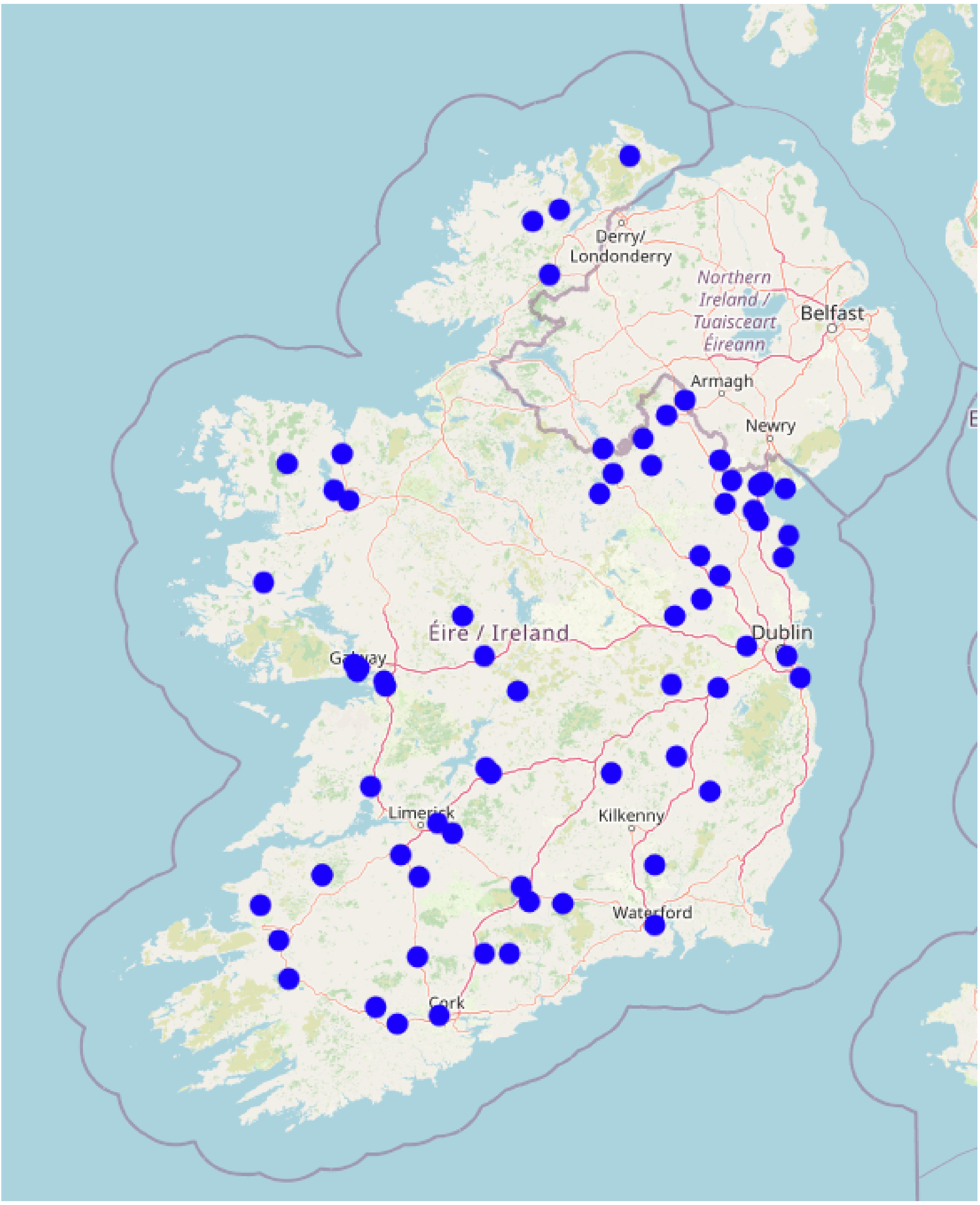

Figure 4.25: Geographic distribution of 70 randomly selected stations, with 10 stations from each River Basin District (RBD).

Table 4.3: Hydrometric stations details

Station name	**Number**	**RBD**	**River name**	**Catchment**	**Area**	**Latitude**	**Longitude**
Aclint	6026	N	LAGAN(GLYDE)	GLYDE	144.00 km^2	53.92476528	-6.640019444
Aghawoney	39009	NW	LEANNAN	LEANNAN	207.00 km^2	55.04378556	-7.720692778
Abington	25003	S	MULKEAR	Shannon	397.00 km^2	52.63186778	-8.421220833
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
Aasleagh Bridge	32060	W	ERRIFF	ERRIFF	166.32 km^2	53.61774917	-9.671156667
Anglesea Road	9369	E	Dodder	LIFFEY	106.22 km^2	53.327878	-6.230943
Adelphi Quay	16160	SE	JOHN'S RIVER	SUIR	3508 km^2	52.25966639	-7.102433056
Athea D/S	23051	SW	GALEY	FEALE	36.30 km^2	52.461378	-9.286944

The data collection period for this experiment spans from 2017 to 2022, with measurements recorded at 15-minute intervals. Specifically, water-level datasets were used as the primary input data for training and testing the water-level prediction models. These datasets were also resampled at 1-hour intervals from January 1, 2017, to January 1, 2022 as we did in Experiment 1. In addition to water-level data, the OPW provides information on flow estimation and water temperature. Flow estimation involves deriving flow values based on ratings and observed water-levels determined through a series of flow gaugings at the monitoring stations. Water temperature is recorded as a by-product of the water-level measurement, but caution should be exercised when interpreting or using this data, as it has not undergone calibration or quality assurance procedures. Therefore, this research's analysis and prediction models will be based exclusively on the water-level datasets. To comprehensively evaluate the models across the 70 stations, we will conduct a total of $70 \times 8 = 560$ model runs. For each station, all eight models will be trained, validated, and tested.

This experiment will provide valuable insights into how different machine-learning models perform across various hydrometric stations within distinct RBDs.

4.6.1.2 Results

The purpose of this experiment was to determine the optimal model for predicting river water levels predicting hour ahead using a time window of 48. 70 hydrometric stations were selected across Ireland and each was trained on eight different models with hyperparameter configurations, totalling 560 model runs. The following section details the results of our analyses, first presenting the results from validation and subsequently testing. Finally a discussion of the overall performance and main tkeaways of one-hour ahead prediction of hydometric data. The table in Appendix E shows all models' prediction performance index values for the 70 hydrometric stations.

Validation

Figure 4.26 illustrates the frequency of highest performing machine learning model selections for water level prediction in various River Basin Districts (RBDs).

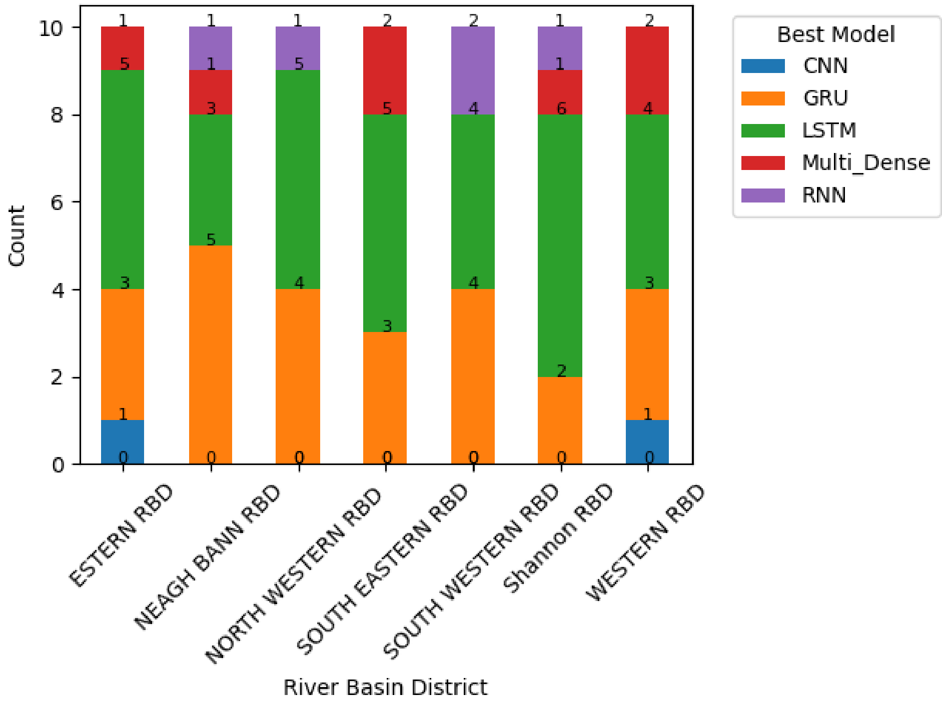

Figure 4.26: The distribution of top-performing models for each River Basin Districts (RBDs) based on the MAE values

The Table 4.4, summarizes the frequency of top-performing model selections for each RBD during validation. The top-performing model was selected based on its Mean Absolute Error (MAE). MAE was chosen as the primary metric for model selection due to its simplicity, interpretability, and robustness. MAE measures the average magnitude of errors in a set of predictions, providing a straightforward indication of model accuracy by averaging the absolute differences between predicted and actual values. Unlike other metrics, such as Mean Squared Error (MSE) or Root Mean Squared Error (RMSE), MAE is less sensitive to outliers, ensuring that the evaluation is not disproportionately affected by extreme values. This makes MAE a reliable measure for assessing model performance in hydrometric data, where variations and anomalies can occur naturally. By focusing on MAE, we aim to identify models that consistently perform well across different datasets and conditions, en-

Table 4.4: Frequency of the Highest Performing Model for Water Level Prediction in Different River Basin Districts (RBDs) During Validation (baased on the MAE values)

RBD	CNN	GRU	LSTM	MultiDense	RNN
EASTERN RBD	1	3	**5**	1	0
NEAGH BANN RBD	0	**5**	3	1	1
NORTH WESTERN RBD	0	4	**5**	0	1
SOUTH EASTERN RBD	0	3	**5**	2	0
SOUTH WESTERN RBD	0	**4**	**4**	0	2
SHANNON RBD	0	2	**6**	1	1
WESTERN RBD	1	3	**4**	2	0

suring both accuracy and robustness in water-level predictions.

Eastern RBD: In the Eastern RBD, the most frequently selected model is the LSTM model, chosen 5 times out of the total 10 station. The GRU model follows with 3 selections. The Multi_Dense model is preferred once. 0 instances of RNN models being chosen as the highest performing model are observed in this RBD.

Neagh Bann RBD: For the Neagh Bann RBD, the GRU model is the most frequently selected highest performing model, chosen 5 times out of 10 station. The LSTM model follows with 3 selections. The MultiDense and RNN models each secure one instance as the highest performing model, while the MultiDense model is also preferred once. No CNN models are chosen in this RBD.

North Western RBD: Within the North Western RBD, the LSTM model is the most prevalent choice, being selected 5 times out of 10 station. The GRU model closely follows with 4 selections. The RNN model is chosen once. No instances of the CNN or MultiDense models being chosen as the highest performing model are noted in this RBD.

South Eastern RBD: In the South Eastern RBD, the LSTM model is the dominant choice, selected 5 times out of 10 station. The GRU model follows with 3 selections. The MultiDense model is preferred twice. No CNN or RNN models are chosen in this RBD.

South Western RBD: Within the South Western RBD, the GRU model stands out as the most frequently chosen highest performing model, selected 4 times out of

10 station. The LSTM model closely follows with 4 selections. The RNN model is the only other model chosen as the highest performing model, securing 2 instances. The MultiDense model is also preferred twice. No CNN models are chosen in this RBD.

Shannon RBD: The Shannon RBD presents the LSTM model as the dominant choice, being selected 6 times out of 10 station. The GRU is chosen 1 time. The MultiDense and RNN models are each chosen once, and the CNN model is preferred once.

Western RBD: Finally, the Western RBD showcases the LSTM models as the most frequently selected highest performing models 4 times out of the 10 station. Followed by the GRU model 3 times. The MultiDense model is preferred twice. The CNN model is chosen once, and the RNN model is not favoured within this RBD.

These results provide insights into the choice of machine learning models for specific geographic regions (RBDs) in the context of the given application. The predominance of LSTM and GRU models in multiple RBDs suggests their efficacy in capturing temporal patterns or dependencies within the datasets associated with these regions. However, model selection may also depend on specific dataset characteristics, and further analysis may be required to determine the reasons behind these preferences.

The highest LSTM result shown in Figure 4.27 was recorded for the Banagher hydrometric station, which belongs to the SHANNON RBD. It achieved an impressive MAE equal to 0.0033, indicating the model's ability to accurately track water level trends, as shown in Figure 4.28.

Among the evaluated models, the worst MAE (Mean Absolute Error) for the LSTM model was observed at the Anglesea Road station in the EASTERN RBD region, with a value of 0.0536. This signifies that LSTM had difficulty accurately predicting water levels at this specific station.

The GRU (Gated Recurrent Unit) model emerged as the second-highest performer, being selected as the optimal model for 24 out of the 70 experimental

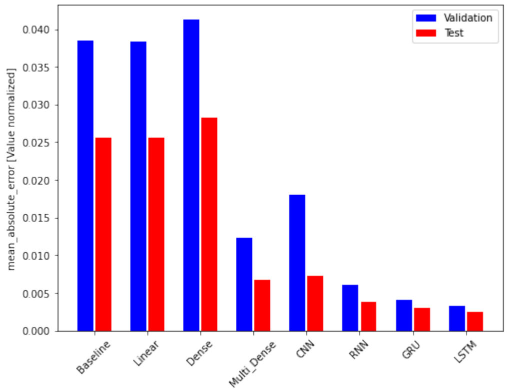

Figure 4.27: Performance in eight prediction models: MAE values for the Banagher hydrometric station

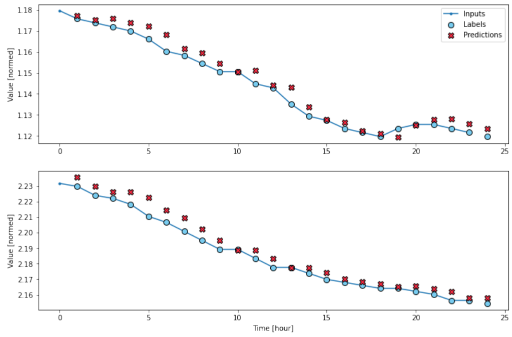

Figure 4.28: LSTM Model validation performance: Banagher hydrometric station

configurations. GRU is closely related to LSTM, as both are designed to handle sequential data effectively. In this experiment, GRU demonstrated comparable performance to LSTM. For instance, at stations like Brewery Park and Aclint, both LSTM and GRU achieved an MAE of 0.0077, highlighting their similar predictive capabilities. The GRU model displayed a range of MAE values, from its highest performance at 0.0035 to its worst at 0.113. This variance indicates that while GRU can excel in some scenarios, it may struggle in others, emphasizing the importance of considering different models based on station-specific characteristics.

MultiDense, RNN, and CNN models consistently delivered the top-performance, being selected 7, 5, and 2 times, respectively, across different station and regional contexts. These models showcased their versatility and effectiveness in capturing the complexities of water level prediction. Nevertheless, it is noteworthy that these models are highly customizable. This implies that researchers have the flexibility to introduce additional layers or fine-tune various parameters to enhance their performance. Hence, it would be unjust to broadly categorize all CNNs as unsuitable. The underperformance observed is specific to these particular configurations. Such malleability allows for meticulous adjustment and optimization. In different scenarios, CNNs can certainly deliver commendable results.

On the contrary, the Baseline, Linear, and Dense models consistently yielded the lowest performance throughout the experiment. These models struggled to provide accurate predictions across a range of conditions and locations.

On the contrary, the Baseline, Linear, and Dense models consistently yielded the lowest performance throughout the experiment. These models struggled to provide accurate predictions across a range of conditions and locations. This performance disparity highlights a fundamental insight: when simpler models like Baseline and Linear fail to deliver accurate results, it underscores the complexity of the underlying data. Such findings strongly motivate the need for more complex modeling approaches. In the face of intricate and multifaceted data patterns, employing advanced techniques becomes imperative.

This comprehensive analysis sheds light on the suitability of various machine learning models for specific geographic regions and provides valuable guidance for future model selection in similar applications.

Test

Figure 4.29 below illustrates the frequency of highest performing machine learning model selections for water level prediction in various River Basin Districts (RBDs) during the test phase.

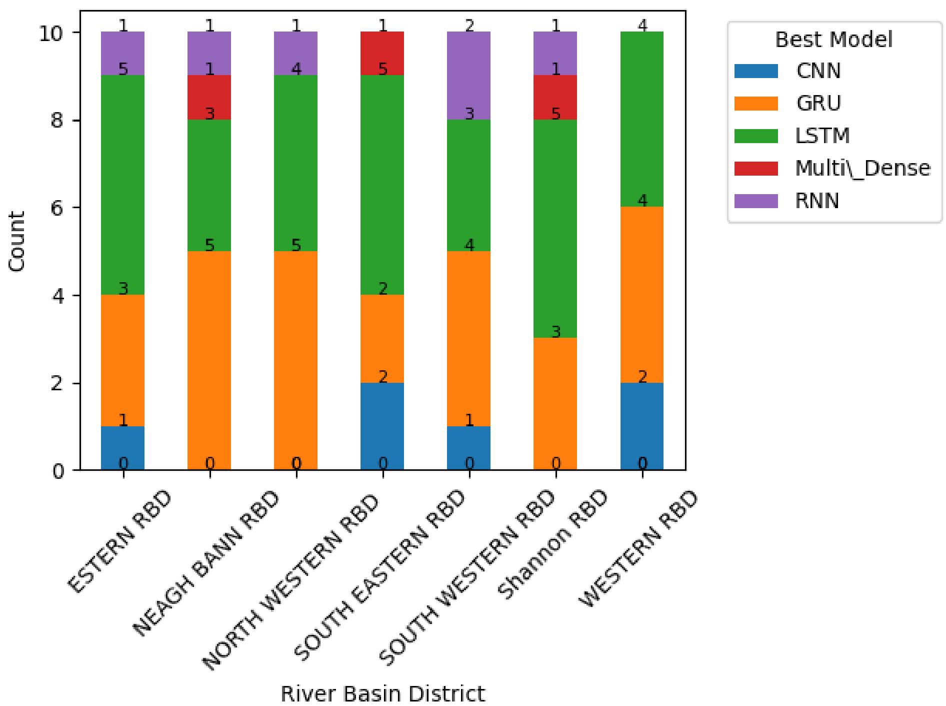

Figure 4.29: The distribution of highest performing models for each River Basin Districts (RBDs)

The Table 4.5 below provides a summary of the frequency of model selections for each RBD during the testing phase:

Eastern RBD: In the Eastern RBD, the most frequently selected highest performing model is the LSTM model, chosen 5 times out of the total 10 stations. The GRU model follows with 3 selections. The RNN and CNN models are each preferred once. No instances of the MultiDense models being chosen as the highest performing model are observed in this RBD.

Neagh Bann RBD: For the Neagh Bann RBD, the GRU model is the most frequently selected highest performing model, chosen 5 times out of 10 stations. The LSTM model follows with 3 selections. The MultiDense and RNN models each secure one instance as highest performing model. The CNN model was not chosen in this RBD.

North Western RBD: Within the North Western RBD, the LSTM model is the most prevalent choice, being selected 5 times out of 10 stations. The GRU model closely follows with 4 selections. The RNN model is chosen once. No instances of the CNN or MultiDense models being chosen as the highest performing model are noted in this RBD.

South Eastern RBD: In the South Eastern RBD, the LSTM model is the dominant choice, selected 5 times out of 10 stations. The GRU and CNN models follows both with 2 selections. The MultiDense model is preferred once. No RNN model is chosen in this RBD.

South Western RBD: Within the South Western RBD, the GRU model stands out as the most frequently chosen highest performing model, selected 4 times out of 10 stations. The LSTM model closely follows with 3 selections. The RNN model secures 2 instances. The RNN model is preferred once. No MultiDense model is chosen in this RBD.

Shannon RBD: The Shannon RBD presents the LSTM model as the dominant choice, being selected 5 times out of 10 stations. The GRU model is chosen 3 times, respectively. Both the CNN and the Multi_Dense models are preferred once. No

Table 4.5: Frequency of Highest Performing Machine Learning Model Selections for Water Level Prediction in Different River Basin Districts (RBDs) During the Test Stage

RBD	CNN	GRU	LSTM	MultiDense	RNN
EASTERN RBD	1	3	**5**	0	1
NEAGH BANN RBD	0	**5**	3	1	1
NORTH WESTERN RBD	0	**5**	4	0	1
SOUTH EASTERN RBD	2	2	**5**	1	0
SOUTH WESTERN RBD	1	**4**	3	0	2
SHANNON RBD	0	3	**5**	1	1
WESTERN RBD	2	**4**	**4**	0	0

CNN model is chosen in this RBD.

Western RBD: Finally, the Western RBD showcases a mix of the GRU and LSTM models as the most frequently selected highest performing models, each chosen 4 times out of 10 stations. The MultiDense model is preferred twice. The CNN model is chosen once, and the RNN model is not favored within this RBD.

The choice of the highest-performing model varies among locations within the same RBD, indicating that the suitability of a model depends on the specific location and its characteristics. Some locations consistently show lower MAE values (seeAppendix E for MAE results), suggesting more accurate predictions, while others have higher MAE values, indicating the need for model improvement. The choice of machine learning model (GRU, LSTM, CNN, MultiDense, RNN) also varies across RBDs, emphasizing the importance of model selection based on local data patterns. To improve model accuracy in certain locations with high MAE, further model tuning and data preprocessing may be necessary. In summary, the analysis highlights the importance of considering location-specific factors and model performance when deploying machine learning models for water level prediction in different River Basin Districts. The choice of the most suitable model should be based on a combination of model performance metrics and domain knowledge of the specific region.

Discussion

According to Table 4.6 for the test phase, we observe that LSTM dominates the experiments and outperforms all competing models with 29 times lower MAE values for predicting the water level. Similarly, the GRU model gave the second-highest performance, with 26 times better prediction than all the others. In addition, RNN and CNN performed well six times each. Lastly, the MultiDense model was preferred only three times out of the 70 experiments.

According to Table 4.6 when comparing the highest performance for each model during the validation and test phases, we observe a significant difference in the MAE values. For instance, the linear model's lowest MAE is 0.0283, while the LSTM achieves 0.0025. This substantial difference underscores variations in model performance.

The MAE (Mean Absolute Error) values represent the average magnitude of the errors between the predicted and actual values, without considering their direction. A lower MAE indicates a model's predictions are closer to the actual values, thus providing more accurate and reliable forecasts. The figures suggest that the LSTM model significantly outperforms the linear model in predicting water levels, with the LSTM achieving an MAE of 0.0025 compared to the linear model's 0.0283. This implies that on average, the LSTM's predictions are much closer to the actual water levels, highlighting its superior performance in capturing the underlying patterns in the data.

Table 4.6: Comparison of highest and lowest performing models for validation and test phases

Validation								
Model	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM
min	0.0385	0.0207	0.0223	0.0067	0.0082	0.0033	0.0035	0.0033
max	1.1466	1.0170	0.9601	0.1395	0.1964	0.2260	0.2132	0.2062
Test								
Model	Baseline	Linear	Dense	Multi_Dense	CNN	RNN	GRU	LSTM
min	0.0256	0.0256	0.0283	0.0055	0.0067	0.0028	0.0026	0.0025
max	1.1624	1.0608	1.1149	0.2652	0.3113	0.2475	0.2760	0.2820

The performance of all models improved slightly in the test compared to the validation. Figures 4.30, 4.31, 4.32, and 4.33 illustrate examples of the baseline model's highest and lowest performance during both validation and test phases. The baseline model serves as a fundamental reference point for comparing the performance of more complex models. Its primary function is to provide a simple yet meaningful prediction by assuming that the water level does not change from one time step to the next. This approach is based on the understanding that water levels typically exhibit gradual changes over short periods, making the current level a reasonable predictor for the near future. This simplistic model helps to set a performance benchmark that more sophisticated models aim to surpass.

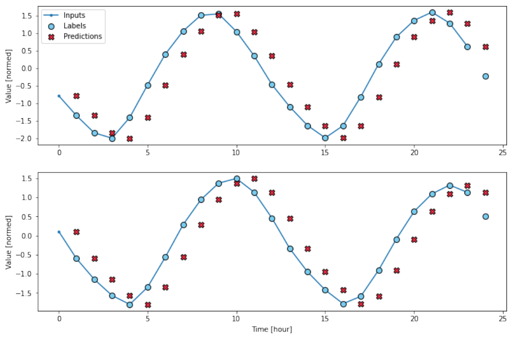

Figure 4.30: Baseline model lowest validation performance

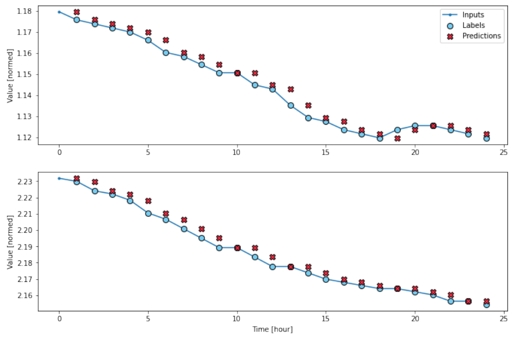

Figure 4.31: Baseline model highest validation performance

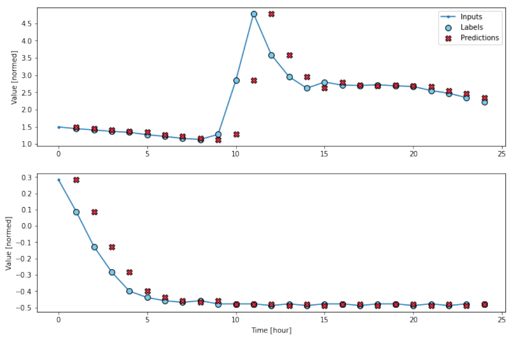

Figure 4.32: Baseline model lowest test performance

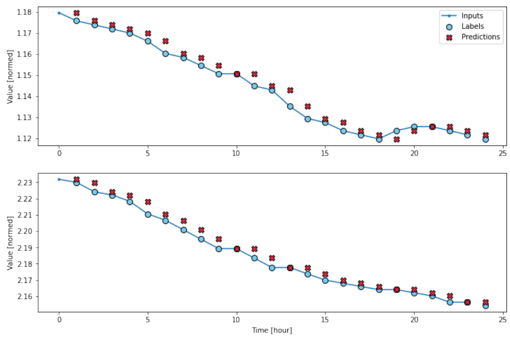

Figure 4.33: Baseline model highest test performance

As demonstrated in Figures 4.34 and 4.35, the distribution of each model's performance showcases distinct characteristics, albeit some similarities are discernible among specific models. The LSTM model emerges as the frontrunner, closely followed by the GRU model in terms of effectiveness.

Intriguingly, upon a meticulous examination of both the validation and test results, a noticeable divergence in model performance becomes apparent. This underscores the critical significance of methodical model selection, as choosing an inappropriate model can result in significantly inferior prediction outcomes. So, we

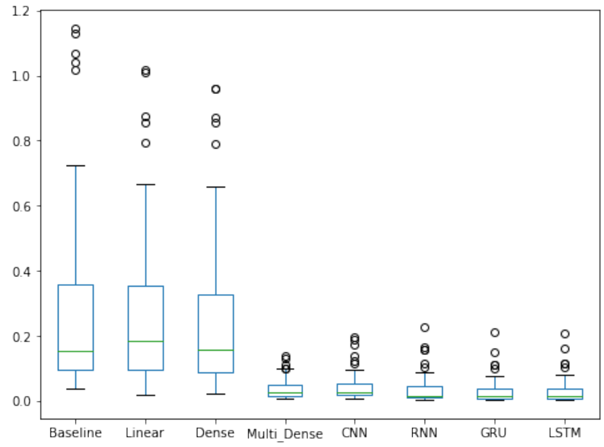

Figure 4.34: Boxplot of performance in eight prediction models: Validation MAE values

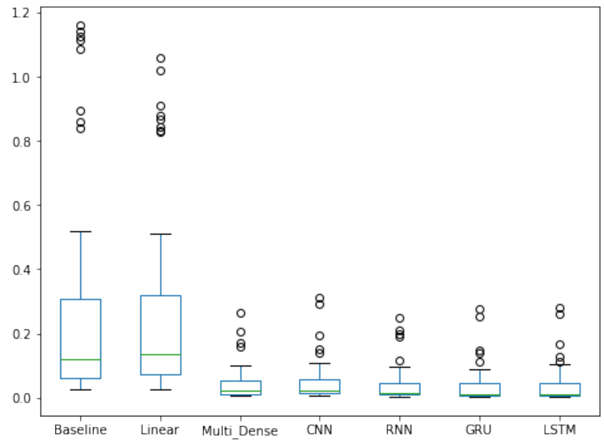

Figure 4.35: Boxplot of performance in eight prediction models: Test MAE values

can devise a formula to signify the error of choosing a wrong model in term of percentage, as follows:

$$\text{error\%}_{\text{model}} = \frac{\text{MAE}_{\text{model}} - \text{MAE}_{\text{highest-performing-model}}}{\text{MAE}_{\text{model}}} = \frac{\Delta\text{MAE}}{\text{MAE}_{\text{model}}}. \tag{4.7}$$

Note that highest performing model will always have a lower MAE than the other models. Thus, if the highest performing model is chosen error percentage would be zero, i.e., error% $= 0$.

The case of the "Aclinet" hydrometric station provides a compelling illustration of the need to gauge the performance disparities among various models in relation to the highest-performing model, designated as GRU. Thus, Table 4.7 below displays the calculated error percentages for each model with respect to the highest performing model (i.e., GRU) for the Aclinet Station dataset.

Table 4.7: Model error percentages with respect to the highest performing model: Validation and Test (Aclinet Station)

Validation								
Model	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM
MAE	0.0665	0.0667	0.0716	0.0136	0.0209	0.0103	**0.0039**	0.0039
ΔMAE	0.0626	0.0628	0.0677	0.0170	0.0170	0.0064	0	0.0000
error%	94.14%	94.15%	94.55%	81.34%	81.34%	62.14%	0%	0.00%
Test								
Model	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM
MAE	0.0455	0.0463	0.048	0.0098	0.028	0.0085	**0.0026**	0.0077
ΔMAE	0.0429	0.0437	0.0454	0.0072	0.0254	0.0059	0	0.0051
error%	94.29%	94.38%	94.58%	73.47%	90.71%	69.41%	0%	66.23%

These findings shed light on the degree to which each model deviates from the highest performing model (in this case, GRU) in terms of MAE, expressed as a percentage. Smaller MAE differences imply a closer alignment in accuracy with the highest performing model, while larger disparities signify substantial performance variations. Consequently, this comprehensive analysis serves as a valuable tool for discerning the relative performance of different models and, thereby, facilitates well-informed decisions regarding model selection tailored to specific tasks.

Nevertheless, a pivotal takeaway from these results emphasizes the absence of a universally optimal model for water level prediction. This is due to the inherent variability and complexity of hydrological data, which can vary significantly across different geographical locations and temporal scales. Models may excel under distinct circumstances or datasets, underscoring the paramount importance of considering contextual factors and data specificity in the process of model selection.

4.6.2 Experiment 2: 349 Hydrometric Stations on 12 Models

4.6.2.1 Experimental Setup

Building on the insights gained from the first experiment, the second experiment expands the scope significantly. This experiment involves 349 hydrometric stations, providing a more comprehensive evaluation of model performance across a broader range of hydrological conditions. Additionally, we introduce four more models, bringing the total to 12 models tested in this experiment:

- Baseline Model
- Linear Model
- Dense Model
- MultiDense Model
- CNN Model
- RNN Model
- LSTM Model
- GRU Model
- LSTM with Attention
- Autoencoder
- Transformer
- Variational Autoencoder (VAE)

The purpose of increasing the number of stations and models in the second experiment is to validate the robustness and scalability of the models identified as promising in the first experiment. By using a larger dataset, we can better assess the generalizability of the models and their performance across varied hydrological settings.

Extending the study to a larger dataset and incorporating more models aims to provide a comprehensive understanding of the performance of various machine learning models in the context of hydrometric data prediction. This approach allows us to explore the impact of model complexity on prediction accuracy, ensuring that the selected models are not only effective but also adaptable to diverse hydrological conditions.

For this experiment, the data collection period spans from 2017 to 2022, with measurements recorded at 15-minute intervals. This five-year timeframe ensures a comprehensive examination of machine learning model performance across diverse temporal nuances, seasonal variations, weather conditions, and the intricate dynamics of hydrological systems. With 349 stations and 12 models, this experiment involves a total of $349 \times 12 = 4188$ model runs. Each model will be trained, validated, and tested for all 349 stations.

This augmentation allows us to scrutinize the performance of established models in tandem with newer counterparts, elucidating whether contemporary techniques exhibit superior predictive capabilities or complement the insights of their predecessors. Moreover, we significantly elevate the number of hydrometric stations under consideration, amplifying the diversity of data sources to explore how variations in data origin influence the observed performance variations among different models.

The choice of the most appropriate model remains contingent upon the specific characteristics of the water-level data, including the presence of temporal dependencies, the intricacy of underlying patterns, manifestations of localized or global interactions, and the prevalence of uncertainty. Our expanded approach underscores the importance of experimenting with a variety of models, enabling the selection of the most fitting model tailored to the precise prerequisites and performance objectives governing water-level prediction in the context of Irish rivers.

To facilitate a comprehensive assessment of model performance, we have established a detailed set of evaluation metrics, including loss values, mean squared error (MSE), and mean absolute error (MAE), calculated both during training on the validation set and post-training on the test set. Additionally, we catalog essential architectural details and parameters of each model, such as model complexity, learning rate, number of layers, number of dense layers, and the dimensionality of input features.

In essence, this experiment represents a pivotal step forward in our research to unravel the intricacies of hydrological prediction. Through this extensive investigation, we aim to provide valuable insights that empower decision-makers in selecting the most suitable models for water-level forecasting, ultimately contributing to more effective water resource management strategies on a national scale.

4.6.2.2 Results

The table in Appendix E shows all models' prediction performance index values for the 349 hydrometric stations. Table 4.8 below provides an illustrative snapshot of the evaluation metrics derived from previous forecasting models applied to various time series datasets:

Table 4.8: Time-series dataset evaluation metrics

Station	Model	Sample Size	Execution Time	Val Loss	Val MSE	Val MAE	Test Loss	Test MSE	Test MAE	Model Complexity	Learning Rate	Number of Layers	Number of Dense Layers	Number of Features
Aclint	GRU	43,305	190.7900	0.0000	0.0000	0.0027	0.0000	0.0000	0.0031	12,929	0.0010	2	1	1
Brewery Park	GRU	29,704	94.2700	0.0005	0.0005	0.0075	0.0005	0.0005	0.0103	12,929	0.0010	2	1	1
Burley	LSTM Attention	43,535	191.1800	0.0002	0.0002	0.0065	0.0004	0.0004	0.0089	16,961	0.0010	3	1	1
Cappoge Bridge	Transformer	42,492	383.6000	0.0020	0.0020	0.0230	0.0043	0.0043	0.0249	141,825	0.0010	2	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Ladyswell	LSTM	34,695	104.2000	0.0003	0.0003	0.0066	0.0005	0.0005	0.0078	16,961	0.0010	2	1	1
Mansfieldstown	GRU	43,539	177.1400	0.0001	0.0001	0.0071	0.0002	0.0002	0.0046	12,929	0.0010	2	1	1
Moyles Mill	GRU	43,585	158.6100	0.0004	0.0004	0.0096	0.0006	0.0006	0.0131	12,929	0.0010	2	1	1
Port Oriel	Transformer	42,027	393.0900	0.0020	0.0020	0.0224	0.0012	0.0012	0.0216	141,825	0.0010	2	1	1
Tallanstown	LSTM	43,390	173.5900	0.0001	0.0001	0.0073	0.0006	0.0006	0.0099	16,961	0.0010	2	1	1

The objective was to evaluate the performance of various models in predicting water levels and to identify the best-performing models across different stations.

Figure 4.36 shows the highest-performing models for all 349 stations during the validation stage. According to these results, the highest-performing model was LSTM-Attention with 124 stations (35.53%), followed by GRU with 75 stations (21.49%). The Baseline model, Transformer, and LSTM also performed well, covering 62, 43, and 38 stations respectively.

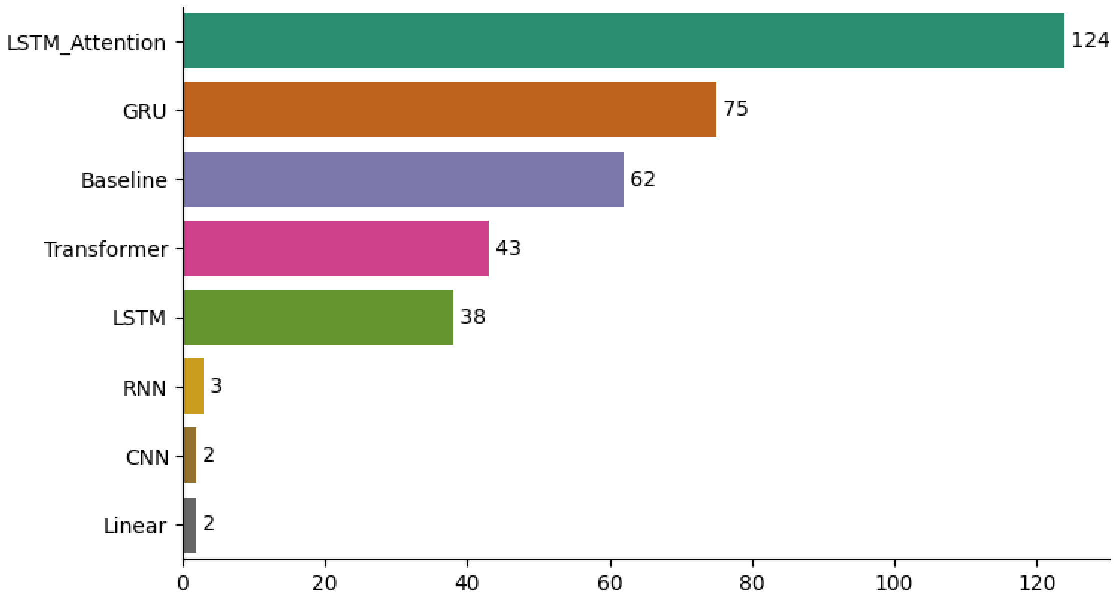

Figure 4.36: Highest performing prediction models during validation, based on MAE values

The dominance of LSTM-Attention and GRU models during the validation phase suggests that these models are particularly effective at capturing the temporal dependencies in the hydrometric data. The Baseline model's strong performance indicates that, for some stations, simple predictions based on recent observations can be highly effective.

Additionally, it's important to highlight that during the test phase, the performance metrics remained consistent, as shown in Figure 4.37. The highest-performing model remained LSTM-Attention, covering 129 stations (36.96%), followed by GRU at 71 stations (20.34%). The performance of other models, such as Baseline, Transformer, and LSTM, was consistent with their validation phase results.

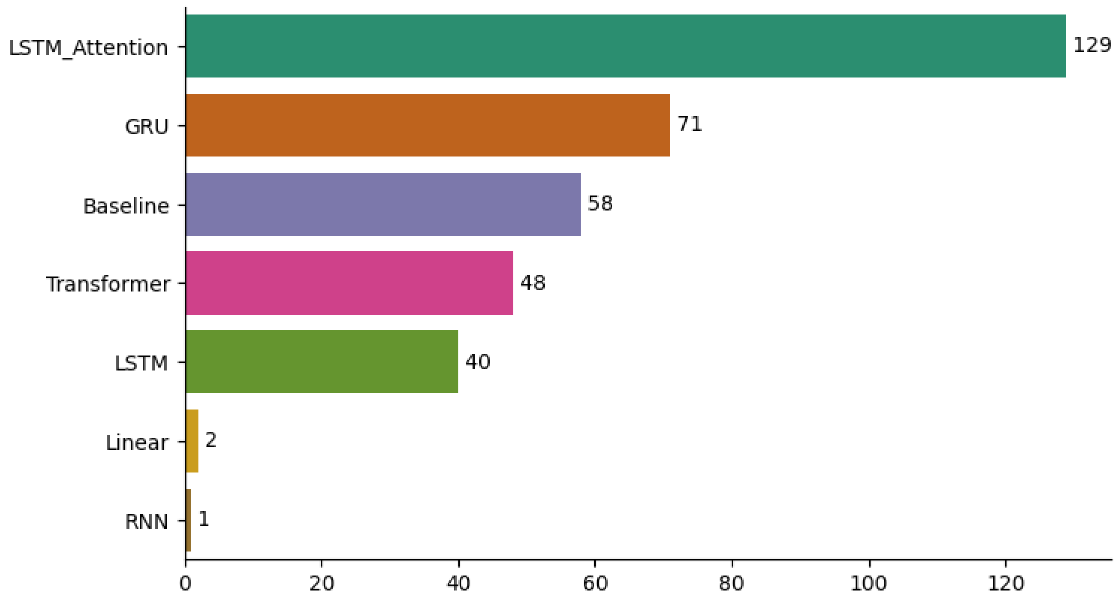

Figure 4.37: Highest performing prediction models during test, based on MAE values

The consistency in performance between the validation and test phases reinforces the reliability of LSTM-Attention and GRU models. These models maintained their effectiveness across different data splits, indicating their robustness and generalizability. The analysis reveals that no single model consistently outperforms others across all stations. This variability underscores the importance of model selection tailored to the specific characteristics of each station. For instance, the LSTM-Attention model's ability to handle long-term dependencies made it particularly suitable for stations with complex temporal patterns.

The distribution of test MAE values across models (see Figure 4.38) further confirms the superior performance of LSTM-Attention and GRU models.This distribution shows a narrower spread for these models, indicating more consistent performance.

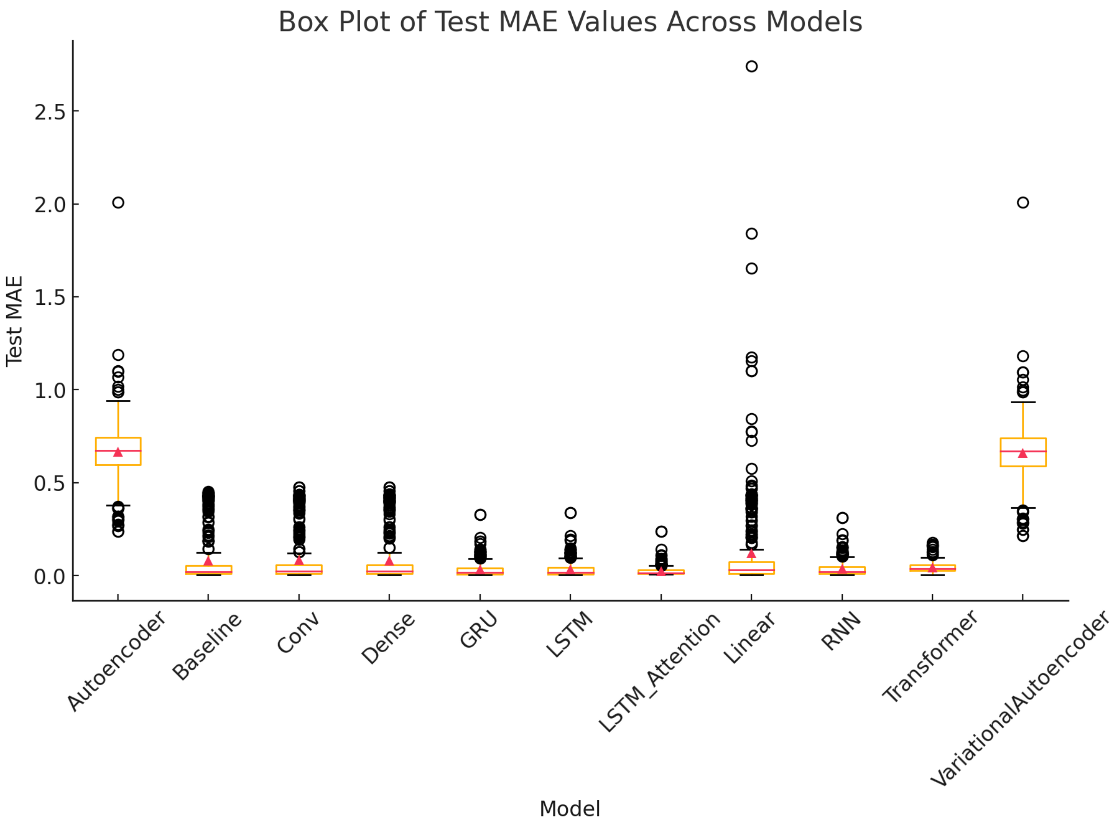

Figure 4.38: Box Plot of Test MAE Values Across Models

Furthermore, the Baseline model's performance highlights the simplicity and effectiveness of recent observation-based predictions for certain stations. This finding suggests that, in some cases, more complex models may not necessarily offer significant improvements over simpler approaches. These findings have important implications for hydrological forecasting. The superior performance of LSTM-Attention and GRU models suggests that advanced machine learning techniques can enhance predictive accuracy in hydrology. However, the need for tailored model selection indicates that a one-size-fits-all approach is not feasible. Future research should focus on developing adaptive frameworks that can dynamically select and tune models based on real-time data characteristics.

In terms of computational efficiency, while advanced models like Transformer and LSTM-Attention offer high accuracy, they require significantly more computational resources. For real-time applications, a balance between accuracy and computational efficiency must be considered.**

The results also highlight the importance of considering the specific properties of each hydrometric station when choosing a model. Factors such as the station's location, climate, and historical data patterns should inform the model selection process to achieve optimal forecasting accuracy.

Experiment 2 demonstrates the effectiveness of LSTM-Attention and GRU models in hydrological forecasting. The findings underscore the necessity for tailored model selection and highlight the potential for advanced machine learning techniques to improve water level predictions. These insights provide a foundation for future research aimed at refining model selection processes and enhancing the accuracy of hydrological forecasts.

LSTM-Attention and GRU models consistently showed superior performance in both validation and test phases, with low MAE values and reliable results across different data splits. However, no single model was consistently the best across all stations, indicating the need for station-specific model selection. Advanced models like Transformer and LSTM-Attention, while accurate, are computationally intensive. In contrast, simpler models like Baseline and Linear can be effective and more efficient in some cases.

Future research should focus on developing adaptive frameworks for model selection and tuning based on real-time data, improving the balance between accuracy and computational efficiency. These detailed analyses and additional insights are essential for understanding the performance dynamics of various prediction models and formulating strategies for effective hydrological forecasting.

4.7 Summary

In summary, this chapter investigated the prediction of water levels based on longitudinal observational data through two experimental configurations. In order to understand the properties of the RBDs, Experiment 1 performed one-hour ahead predictions on randomly selected 10 hydrometric stations from each of the seven RBDs (i.e., 70 hydrometric stations in total) across eight models. The second experiment increased the scale of our research (using 349 stations and 12 models) to further demonstrate the necessity of model selection in time series analyses based on the properties of the dataset, as it was not possible to pick a single model for all stations.

In the field of hydrology, where traditional physical modelling has long been the norm, selecting the right forecasting technique or combining multiple methods effectively has posed a persistent challenge. This decision often hinges on the availability of domain-specific expertise, which can be scarce due to various constraints. However, our research seeks to bridge this gap by establishing a comprehensive experimental foundation for time series forecasting while simultaneously introducing a diverse spectrum of forecasting methodologies.

A noteworthy feature of our approach is its potential applicability across a wide range of hydrological forecasting scenarios, even in contexts where domain-specific expertise is elusive. Moreover, as the machine learning field experiences a surge of interest in meta-learning, our research trajectory is now aimed at harnessing meta-learning methodologies. The paradigms of few-shot and zero-shot learning are particularly promising, as they seek to incorporate domain knowledge into the feature and model selection processes. Importantly, while these methodologies have found success in climate data forecasting, their application to water prediction remains relatively unexplored. Our forthcoming endeavors are geared toward filling this gap, specifically targeting the challenges inherent in spatial-temporal water prediction in catchments characterized by sparse datasets.

From a practical standpoint, the proposed methodology has substantial potential for application by various entities, including water management bodies, urban planning commissions, and environmental conservation agencies. The study demonstrated the ability to make accurate water-level predictions through rigorous experimentation and validation across multiple hydrometric stations. This methodology is particularly useful in mitigating the impact of floods and ensuring the efficient allocation of water resources, especially in regions with complex and varied hydrological dynamics. The experiments showed that the predictive models, when correctly selected and tuned, can improve forecasting accuracy.

In essence, this chapter contributes to the advancement of hydrological forecasting by introducing a data-driven approach. The study offers versatile and effective tools for addressing the challenges of water-level prediction, as demonstrated by the improved accuracy of the predictive models across different hydrometric stations. These findings are supported by extensive experimentation and validation, highlighting the practical applications of the methodology.

However, the results also demonstrate the extensive time investment required for proper model selection and the consequences of an incorrect selection. This poses further issues for hydrology as datasets are continuous and new previously "unseen" datasets may become available. In the face of such a volatile landscape, it is impractical to undertake a full model selection process in a continuous manner. The following chapter discusses meta-learning approaches to provide an automated approach to model selection to overcome these issues.

Chapter 5

Meta-Learning Approaches for Time Series Model Selection

In the previous chapter we demonstrated that there is no 'one-model-fits-all' approach for hydrological time series data and that exhaustive model testing is impractical, our focus now shifts to implementing a meta-learning process to identify the selected optimal model from a time series.

This chapter begins with an introduction in Section 5.1, where we explore the limitations encountered when identifying the most suitable model for time series data. We also introduce meta-learning as a valuable solution to these challenges.

In Section 5.2, we present our methodology, which aims to formalize the process of selecting an appropriate time-series prediction model. In addition, we provide an overview of the Meta-Dataset, where we discuss and explain the input data for the meta-learner (Section 5.2.1). Additionally, we consider the potential of an automated machine learning (AutoML) approach for model selection and hyperparameter optimization, which can streamline the process and improve efficiency. AutoML involves automating various stages of the machine learning workflow, including data preprocessing, feature selection, model selection, and hyperparameter optimization, thereby making machine learning more accessible and efficient. The implementation and Experimental Setup are presented in Section 5.5 The results for this chapter are presented in the following Section 5.6. Finally, we conclude this chapter by summarising the findings in Section 5.7.

5.1 Introduction

Precise predictions are critical for making informed decisions and mitigating potential catastrophic impacts, particularly in the hydrology domain. Accurate forecasts are necessary for effective flood management, water resource allocation, and environmental conservation. In the context of hydrology, precise predictions enable timely and effective responses to extreme weather events, which can significantly reduce the risk of property damage and loss of life. However, achieving accurate forecasts in this complex field is hampered by the inherent variability in river datasets. This variability stems from a multitude of factors, including seasonal fluctuations, geographical attributes, and meteorological conditions. These diverse influences manifest as non-linear and dynamic patterns, making it challenging to develop a single, universally applicable model capable of reliably predicting water levels across different river systems. Traditional machine learning approaches often fall short in achieving this due to their tendency to be dataset-specific, hindering effective generalization.

Addressing this issue typically involves creating and evaluating numerous models tailored to specific datasets or river systems. However, this approach has its limitations. It is computationally expensive and time-consuming to comprehensively test multiple models on unfamiliar datasets. Furthermore, the sheer number of models required to accommodate the variability in river water level data is impractical. As such, a more adaptive and efficient solution is needed to overcome the challenges of model selection.

This chapter centres on meta-learning, a subfield within machine learning that offers a promising methodology for tackling the challenge of model selection in river water level prediction. Meta-learning algorithms empower models to adapt and generalize effectively across diverse datasets. Our focus lies in harnessing the power of meta-learning to enhance model adaptability and selection.

Challenges

The growing availability of time series prediction methods has intensified the challenge of selecting an appropriate time series prediction model. Typically, the decision process involved in selecting a method typically entails a comprehensive evaluation of the time series, utilizing metrics as described in Chapter 4.

In hydrology, as discussed in Chapter 2 expert systems are an appropriate potential solution for model selection. However, their effectiveness often depends on the presence and expertise of hydrologists, which may not always be practical or scalable. This raises a fundamental question within the hydrological community: Can a more adaptable and efficient approach be developed for selecting the optimal model in hydrological scenarios or time series data?

The challenge of developing models that can generalize well across diverse datasets is a significant concern in the field of machine learning. This issue is particularly pertinent when dealing with highly variable and non-linear data patterns, as seen in domains such as hydrology. Several recent studies have highlighted various approaches to enhance the generalization capabilities of models.

- According to [202], while prior studies have explored diverse approaches for applying meta-learning in time series prediction, the unique requirements of hydrology call for further innovation. The field of meta-learning has shown remarkable promise in various domains, but it still faces significant challenges that require attention to progress further.

- [207] focus on domain generalization (DG). The authors emphasize the importance of domain-invariant representations and propose methods to disentangle spurious correlations and enhance meaningful ones through both sample and feature perspectives, achieving significant improvements in generalization performance across multiple datasets .

- [226] investigate the factors that influence model generalization, particularly in compositional generalization tasks. Their empirical analysis shows that increased dataset complexity and a balanced mixture of simple and hard examples can improve generalization performance. This study underscores the importance of dataset diversity in training robust models .

- [40] explore leveraging the diversity of pretrained models from a model zoo to improve out-of-distribution generalization. The authors argue that even weaker models contain valuable knowledge that can enhance generalization when integrated effectively, demonstrating state-of-the-art results across various datasets .

- [217] propose a method to improve the generalization of deep metric learning models by augmenting diverse out-of-distribution samples and regularizing feature distributions. Their approach effectively enhances generalization to both unseen categories and domains .

These studies collectively highlight the importance of model adaptability and the integration of diverse data and knowledge sources to enhance generalization capabilities.

To address the pressing need for adaptive model selection in the field of hydrology and to advance our understanding of this domain, this study endeavours to conduct a comprehensive exploration of the relationship between time series features and the selection of forecasting models in hydrology. Our specific focus is on the prediction of water levels, and to accomplish our research objectives, we will employ interpretability techniques within the realm of machine learning.

This study advocates for the application of meta-learning techniques to autonomously acquire knowledge in the context of selecting suitable models for water level prediction in hydrology. By leveraging machine learning methods for predicting water levels in hydrology, we can harness the common practices of meta-learning used in algorithm selection for classification problems. This approach streamlines the model selection process, akin to identifying the most suitable learning algorithm for precise water level predictions. Here, hydrology models serve as the foundational learners, and meta-features denote the attributes characterizing the time series of water levels under analysis.

Our primary objective is to enhance water level prediction in hydrology using a meta-learning approach for time series model selection. This involves identifying key similarities among prediction models, understanding how hydrological features relate to water level variations, and characterizing their interactions to improve model selection.

Contribution

Our motivation for this study primarily arises from recognizing the inherent limitations of employing a fixed, universally applicable methodology for model selection in the field of hydrology. Hydrological systems are marked by intricate characteristics and substantial variability, which preclude the existence of a singular superior model. Instead, our aim is to explore how meta-learning can autonomously select and adapt models based on the unique attributes of each river system, ultimately enhancing the precision of predictions regarding river water levels. The identification of an appropriate prediction method for a given dataset remains an ongoing challenge in the domain of time series prediction problems. In general, researchers are required to align the characteristics of time series data with the most suitable models. The area of inquiry referred to as Meta-Learning involves the task of characterizing a time series using various metrics, which are subsequently employed in the process of selecting a suitable algorithm. A meta-learner is commonly constructed through the utilization of a classifier such as Decision Tree, Random Forest (RF) and Support Vector Machine (SVM) [153]

Our research aims to comprehensively assess the effectiveness of meta-learning in enhancing the precision of river water level predictions. This assessment is grounded in the autonomous selection and adaptation of models tailored to the unique attributes of each river system. This realization forms the basis for implementing meta-learning as a promising solution in this vital field.

Initially, relevant datasets are gathered and integrated through spatiotemporal integration of heterogeneous environmental data, ensuring their proper formatting and cleanliness. These datasets are combined into a unified dataset for subsequent analysis, as detailed in Chapter 3.Next, various candidate models, including linear models and dense neural networks, are trained on a designated training dataset. The performance of each model is assessed using a separate validation dataset. This comprehensive training and evaluation process is elaborated upon in Chapter 4. The meta-learner utilizes the performance results from the validation set to identify and recommend the best-performing model.Furthermore, the process involves the creation of meta-datasets, which consist of sets of input features provided to the meta-learners. Leveraging these meta-datasets, the meta-learner makes informed model selection decisions. Drawing upon the insights gained from the prior stages, the meta-learner identifies the best-performing model for a given dataset. The details of this model selection process are outlined in the evaluation section, Section 5.6, of this chapter.The foundation of this assessment lies in the autonomous selection and adaptation of models tailored to the unique attributes of each river system. This realization forms the basis for implementing meta-learning as a promising solution in this critical area. The research aims to enhance understanding and enable more informed and dependable adaptive model selection for water level prediction in hydrology. The general meta-learning architecture for model selection investigates the decision-making processes in complex hydrological systems, specifically focusing on water level prediction.

5.2 Meta-Learning Methodology

The underlying concept of meta-learning involves identifying a suitable prediction technique, given a collection of characteristic metrics obtained from a particular dataset. The use of these metrics within a pre-existing model yields a recommendation. Our approach is designed to effectively choose the appropriate model for a given dataset based on the meta-features of that dataset. By enhancing our understanding and enabling more informed and dependable adaptive model selection for water level prediction in hydrology, our general meta-learning architecture for model selection, presented in Figure 5.1, investigates the decision-making processes in complex hydrological systems, specifically focusing on water level prediction.

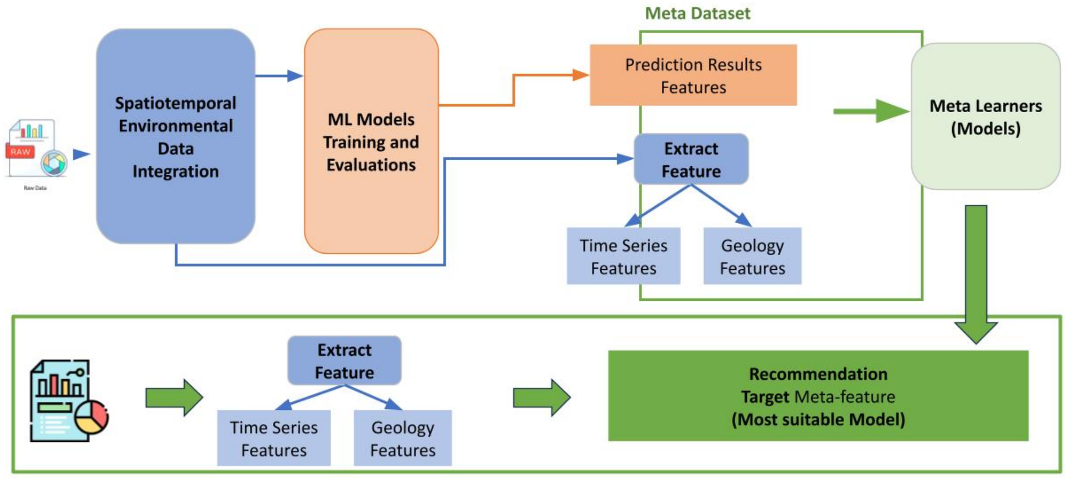

Figure 5.1: Decision-making process diagram based on the meta-learning architecture

This systematic process can be deconstructed into various components that form an integral part of the general architecture presented in Figure 5.1.

- Meta-Datasets: A collection of time series features as described in section Section 5.2.1.
- Candidate models: The ensemble of time series prediction models detailed in Section 5.2.3
- Meta-learner: The models employed for the implementation of the meta-learner, as outlined in Section 5.2.3
- Evaluation: The evaluation metrics explained in Section 5.6

5.2.1 Meta-Learner Input: Meta-Dataset

For our meta-learner, we have categorised our input features into three distinct categories, each offering a unique perspective on the problem 5.2.

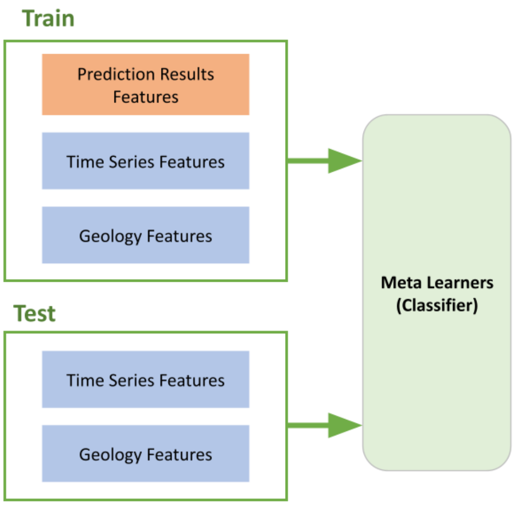

Figure 5.2: EInput Features for the Train and test Phases

1. Time Series Input Features

2. Geology Features

3. ML Prediction Results Features

5.2.1.1 Time Series Input Features

This category encompasses a range of temporal data related to the water levels time series. Time series features are crucial for capturing the historical patterns and trends in water level data. The selection of some of these features for this research was guided by their demonstrated utility in prior studies conducted by [102, 143]. The features mentioned below encompass a comprehensive range of statistical and time series metrics, offering valuable insights into the dataset's characteristics. Starting with the central tendency, the mean and median provide a glimpse into the dataset's central values, while the variance and standard deviation quantify its dispersion and spread. The range indicates the span between the highest and lowest values. The skewness and kurtosis shed light on the data's shape and asymmetry, with the former measuring skew and the latter assessing tail heaviness. Seasonal strength gauges the presence of recurring patterns.

The input feature set also includes series-specific metrics, the exponential moving averages (EMAs) at different smoothing levels offer dynamic insights into trends. Rolling statistics with various window sizes, such as rolling mean, standard deviation, skewness, and kurtosis, help capture evolving patterns over time. Autocorrelation and partial autocorrelation functions at multiple lags (ACF and PACF) reveal the presence of temporal dependencies. Cross-correlation values at different lags highlight relationships with lagged variables.Additionally, unit root tests (Phillips-Perron and KPSS) provide information about stationarity, a crucial concept in time series analysis. Finally, the autocorrelation of residuals from a linear model at lag 1 (LMResiduals_ACF1) indicates any remaining temporal dependencies in the data. These features collectively equip analysts with a robust toolkit for understanding the underlying patterns and dynamics within the dataset, facilitating informed decision-making and predictive modelling.

The choice of input meta-dataset is influenced by the need to capture various dimensions of the problem being addressed. Time-series features provide insights into historical patterns and trends, geological features offer contextual information about the physical environment, and ML prediction results features allow for an understanding of the performance of different models. This comprehensive approach ensures that the meta-learner can make informed decisions based on a rich set of data attributes.

5.2.1.2 Geology Features

The geology data category provides valuable contextual information about the geographical and geological characteristics of the area where water level measurements are taken. This category includes:

- Geographical Coordinates: Latitude and longitude coordinates pinpointing the measurement locations. Latitude and longitude are geographical coordinates that can be autocorrelated with variables such as elevation and precipitation. These spatial dependencies can influence local climate conditions and, subsequently, the water levels. Incorporating these variables helps in capturing the underlying geographical influences on the data.

- Geological Features: Information about the type of soil, bedrock, or geological formations in the vicinity, which can influence water behaviour.

5.2.1.3 ML Prediction Results Features

In our pursuit of identifying the most appropriate forecasting model for a given time series, It encompasses both hyperparameters and evaluation metrics.

Within this critical category, we introduce an innovative approach wherein we harness the outcomes of prior forecasting models as additional input features for training a meta-learner. These metrics are instrumental in empowering the meta-learner to make informed decisions regarding model selection.

This additional data provides important information about various performance metrics and characteristics of different machine learning models used for predicting water levels at different stations. This data is essential for the meta-learner to learn for multiple, for instance:

1. **Model Performance Metrics (Validation Loss, Validation MSE, Validation MAE, Test Loss, Test MSE, Test MAE):** These metrics quantify the quality of predictions made by each model. A meta-learner can use this information to assess how well a model generalizes to unseen data. Models with lower validation and test metrics indicate better predictive performance, which is valuable for model selection [73, 65]. For instance, [182] introduced the MetaSieve method, which optimizes the trade-off between performance and complexity, highlighting the importance of these metrics in model evaluation.

2. **Sample Size:** Understanding the sample size used for training and testing models is crucial. It provides insights into the amount of data available for each station, and a meta-learner can consider this when assessing model reliability. Larger sample sizes generally lead to more robust models. For example, [7]

demonstrated the effectiveness of the N-BEATS model for production forecasting with limited historical data, emphasizing the challenges of small sample sizes .

3. **Execution Time:** Execution time indicates how long it takes each model to make predictions. This information is vital for resource allocation and operational considerations. A meta-learner can decide whether to prioritize models with shorter execution times for real-time applications or choose more accurate models with longer execution times when computational resources are available. Studies such as those by [182] highlight methods that balance execution time and performance, reducing computational effort without significant loss in forecasting quality .

4. **Model Complexity (Number of Layers, Number of Dense Layers):** These parameters provide insights into the architectural complexity of the models. A meta-learner can use this information to assess the trade-off between model complexity and predictive performance. Some applications may favor simpler models to reduce overfitting. [194] introduced a feature-based approach for forecast model performance prediction, emphasizing the impact of model complexity on accuracy .

5. **Learning Rate:** The learning rate is a hyperparameter that affects the training process. A meta-learner can consider learning rates when comparing models, as different values may lead to varying training behaviors and convergence speeds. For example, [172] analyzed different error metrics, including learning rates, and their impact on meta-learning performance, underscoring the importance of choosing appropriate hyperparameters .

6. **Number of Features:** This parameter indicates the dimensionality of the input features used by each model. It's essential to assess whether models with higher or lower feature dimensions perform better for a particular station, as this can guide feature engineering decisions. [194] discussed how feature-

based meta-learning approaches could predict model performance effectively, considering the number of features used.

Table 5.1 below provides an illustrative snapshot of the evaluation metrics derived from previous forecasting models applied to various time series datasets from the Neagh Bann (NB) Aclinet Station:

Table 5.1: Aclint station ML prediction results

Feature	**Value**
Station	Aclint
Model	GRU
Sample Size	43,305
Execution Time	190.79
Validation Loss	0.00003
Val MSE	0.00003
Val MAE	0.00267
Test Loss	0.00004
Test MSE	0.00004
Test MAE	0.00314
Model Complexity	12.929
Learning Rate	0.001
Number of Layers	2
Number of Dense Layers	1
Number of Features	1

Table E.2 in Appendix E details the prediction results the stations tested in Neagh Bann RBD. These evaluation metrics constitute a comprehensive evaluation of each forecasting model's performance, serving as a valuable addition to our meta-learning framework. They encompass diverse metrics, including sample size, execution time, validation loss, validation mean squared error (MSE), validation mean absolute error (MAE), test loss, test MSE, test MAE, model complexity, learning rate, number of layers, number of dense layers and number of features.

By analyzing this comprehensive dataset, a meta-learner can identify patterns, trends, and trade-offs among different models and their associated characteristics. This knowledge empowers the meta-learner to make informed decisions about model selection, hyperparameter tuning, and resource allocation. The meta-learner leverages these evaluation metrics as input features to make data-informed decisions regarding model selection. By learning from the outcomes of prior models applied to diverse time series data, the meta-learner enhances its capacity to recommend the most appropriate forecasting method for each unique dataset. Incorporating these comprehensive evaluation metrics derived from prior model predictions, we bolster the capabilities of our meta-learner, enabling it to make well-informed choices when selecting the most fitting forecasting method for each unique time series dataset. To achieve this, the following steps are followed:

1. **Data Collection and Preprocessing:** Gather a diverse set of time series datasets and preprocess them to ensure consistency and quality. This includes handling missing values, normalizing data, and extracting relevant features.

2. **Model Training and Evaluation:** Train multiple candidate models on these datasets, evaluating their performance using metrics such as validation loss, mean squared error (MSE), and mean absolute error (MAE). Record these metrics along with model characteristics such as complexity, execution time, and sample size.

3. Meta-Dataset Creation: Compile the evaluation metrics and model characteristics into a meta-dataset. This meta-dataset serves as the input for training the meta-learner.

4. **Meta-Learner Training:** Train the meta-learner using the meta-dataset. The meta-learner learns to map the evaluation metrics and model characteristics to the best-performing model for each dataset.

5. **Model Selection and Recommendation:** For a new dataset, input its features into the trained meta-learner. The meta-learner will analyze these features and recommend the most suitable forecasting model based on its learned knowledge from previous datasets.

6. **Continuous Learning and Improvement:** Continuously update the meta-learner with new data and model evaluations to improve its recommendation

accuracy over time.

By following this structured approach, the meta-learner can effectively enhance model selection, leading to more accurate and efficient forecasting for diverse time series datasets.

5.2.2 Meta-Learner Output

In this subsection, we provide a detailed insight into the output produced by our meta-learner, focusing on its application in the context of model selection for time series prediction. The meta-learner, as previously described, plays a pivotal role in the process of identifying the most suitable forecasting model for a given time series. However, the pivotal question arises: what exactly is this output, and how does it contribute to the broader goal of model selection?

To address this, it's essential to understand that the output from the meta-learner essentially represents a well-informed decision regarding the choice of the forecasting model. This decision is guided by an intricate evaluation process that scrutinizes various aspects of each potential model's performance.

These evaluation metrics encompass a range of factors, including not only the predictive accuracy of the models but also factors such as execution time, model complexity, and other essential characteristics. The goal is to ensure that the chosen model not only provides accurate predictions but also does so efficiently and without unnecessary complexity.

The meta-dataset used in this process comprises a substantial amount of data, with numerous rows and columns. However, it's important to note that not all of this data is used in the actual model selection. The reason is that some of these features or metrics pertain to the evaluation of model predictions and are not available during the testing phase.

In essence, the meta-learner synthesizes a wealth of information to make a well-informed decision on the most suitable forecasting model for a specific time series. By analyzing diverse performance metrics and model characteristics, the meta-learner

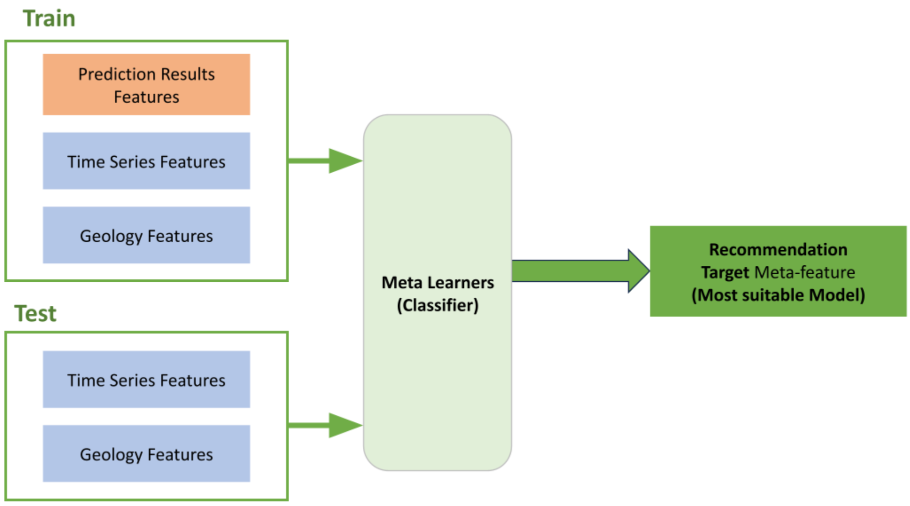

Figure 5.3: EInput Features for the Train and test Phases

ensures the selected model provides accurate predictions efficiently, balancing complexity and practicality. The output of the meta-learner thus represents the optimal model choice for time series prediction, underscoring the significant enhancement in the model selection process achieved through meta-learning (see 5.3)

The output function of the meta-learner can be mathematically expressed as follows: Given a new metadataset D_{new} and a set of metadatasets $D = \{D_1, D_2, \ldots, D_n\}$, the function selects the best time series prediction model for D_{new} based on the similarity between the features of D_{new} and the metadatasets in D.

1. **Calculate Similarity**: Compute the similarity $S(D_i, D_{\text{new}})$ between D_{new} and each metadataset D_i in D.

2. **Select Most Similar Metadataset**: Identify the metadataset D_{selected} that is most similar to D_{new}:

$$D_{\text{selected}} = \arg\max_{D_i \in D} S(D_i, D_{\text{new}})$$

3. **Retrieve Associated Model**: Retrieve the time series prediction model M_i

associated with D_{selected}. This model is the selected output:

$$M_{\text{selected}} = M_i \quad \text{for the most similar } D_i$$

By following this process, the meta-learner adeptly identifies the best model for new datasets, thereby optimizing predictive performance and enhancing the overall model selection framework in time series forecasting.

Objective Function

The objective function of the meta-learner is designed to optimize the model selection process. The goal is to minimize the prediction error while balancing the complexity and computational efficiency of the selected model. This is achieved through the following optimization problem:

$$\theta^* = \arg\min_\theta \sum_{i=1}^{n} L(M(D_i;\theta), y_i)$$

where L is the loss function, $M(D_i;\theta)$ represents the model trained on dataset D_i with parameters θ, and y_i is the actual outcome.

Training Phase

The training phase of the meta-learner involves updating its parameters to improve its model selection capability. This phase can be defined as follows:

1. **Sample a Metadataset D_i from D**:
   - Randomly select one of the existing metadatasets D_i from the collection D. Each metadataset consists of time series data and its associated features, performance metrics, and model characteristics.
2. **Evaluate Similarity**:
   - Compute the similarity $S(D_i, D_{\text{new}})$ between the sampled metadataset D_i and a new time series dataset D_{new}. This similarity measure helps

determine how closely D_i matches the new dataset in terms of relevant features.

3. **Update Meta-Learner's Parameters**:

   - Adjust the meta-learner's parameters θ based on the performance of the best model for the sampled metadataset D_i. The goal is to minimize the loss function L, which measures the prediction error of the model $M(D_i; \theta)$:
   $$\theta \leftarrow \theta - \nabla_\theta L(M(D_i; \theta), y_i)$$
   - Here, $\nabla_\theta L$ represents the gradient of the loss function with respect to the parameters θ. This gradient descent step helps the meta-learner improve its predictions by reducing the error iteratively.

Testing Phase

The testing phase of the meta-learner involves applying the learned parameters to new data to select the most appropriate model. This phase can be defined as follows:

1. **Calculate Similarity for New Metadataset**:

   - For the new metadataset D_{new}, compute the similarity $S(D_i, D_{\text{new}})$ with all existing metadatasets D_i in the collection D. This step helps identify which existing dataset is most similar to the new one.

2. **Choose the Most Similar Metadataset**:

   - Identify the metadataset D_i that has the highest similarity score with D_{new}. The most similar metadataset D_{selected} is chosen as:
   $$D_{\text{selected}} = \arg\max_{D_i \in D} S(D_i, D_{\text{new}})$$

3. **Select the Associated Model**:

- Retrieve the time series prediction model M_i associated with the selected metadataset D_{selected}. This model M_{selected} is deemed the best choice for the new dataset D_{new}:

$$M_{\text{selected}} = M_i \quad \text{for the most similar } D_i$$

- By selecting the model associated with the most similar metadataset, the meta-learner ensures that the chosen model is well-suited for the new dataset based on past performance and characteristics.

The meta-learner learns to identify the best model for a new dataset based on the similarity between the new time series dataset's features and the metadatasets' features. The specific similarity metric and model selection strategy will depend on your problem and available data.

5.2.3 Meta-Learner Candidate Models

Why Use a Classifier Instead of a Regression Model?

In the context of meta-learning for time series model selection, the choice between using a classifier or a regression model is critical. Here are the key reasons for preferring a classifier:

1. **Nature of the Output**:

   - **Classification Task**: Our objective is to categorize or select the most suitable forecasting model from a predefined set based on the features of the time series data. This makes the problem a classification task, where the output is a discrete label indicating the chosen model.
   - **Discrete Model Selection**: Given that the output involves selecting from a finite set of models, it aligns more naturally with classification rather than regression.

2. **Performance Metrics**:

- **Precision and Recall**: Classifiers provide essential performance metrics such as precision, recall, and F1-score. These metrics are crucial for evaluating the effectiveness of the meta-learner in accurately selecting the best model.
- **Handling Imbalanced Data**: Techniques such as oversampling, undersampling, and Synthetic Minority Oversampling Technique (SMOTE) are specifically designed for classifiers to handle imbalanced datasets, enhancing the meta-learner's performance.

3. **Interpretability**:
   - **Model Interpretability**: Classifiers, particularly tree-based methods such as Decision Trees and Random Forests, offer interpretability through feature importance metrics. This helps in understanding the rationale behind the model selection, making the process transparent and explainable.
4. **Optimization**:
   - **Hyperparameter Tuning**: Classifiers allow extensive hyperparameter tuning, which can be leveraged to optimize the model selection process effectively, ensuring the meta-learner performs at its best.

Using a classifier, therefore, provides a structured and efficient approach to solving the model selection problem, making it the preferred choice over regression models for this specific task.

We employ a set of classifiers to assess their performance using various resampling techniques. Evaluating these models across different strategies allows us to identify which models and techniques perform best for specific time series datasets and problems. Meta-learning aids in the selection of the most appropriate models and preprocessing strategies, ultimately improving predictive performance for time series analysis.

1. Random Forest: Random Forest is an ensemble method that combines multiple decision trees to make predictions. It's known for its robustness and the ability to handle a variety of data types. It's often used as a baseline model for many tasks [78].

2. Logistic Regression: Logistic Regression is a simple yet effective model for binary classification. It's interpretable and widely used for tasks where understanding feature importance is essential.

3. Naive Bayes: Naive Bayes is a probabilistic classifier based on Bayes' theorem. It's particularly suited for text classification and other tasks where feature independence assumptions are reasonable.

4. K-Nearest Neighbors: K-Nearest Neighbors is a non-parametric classification method that classifies a data point based on the majority class among its k-nearest neighbours. It's effective for both classification and regression tasks.

5. Support Vector Machine (SVM): Support Vector Machine is a powerful classification algorithm known for finding optimal decision boundaries in high-dimensional spaces. It's particularly useful when dealing with complex datasets [74].

6. Gradient Boosted Decision Trees (GDBT): GDBT is an ensemble method that builds decision trees sequentially, with each tree correcting the errors of the previous one. It's known for high accuracy and is used in many machine learning competitions.

7. MLP (Multilayer Perceptron):MLP is a type of neural network that consists of multiple layers of interconnected neurons. It's a versatile model used in a wide range of applications, from image classification to natural language processing.

8. XGBoost: XGBoost is an optimized and efficient gradient-boosting algorithm known for its speed and performance. It's often used in machine learning competitions and for regression and classification tasks [39].

9. LightGBM: LightGBM is another gradient-boosting framework known for its speed and efficiency. It's suitable for large datasets and offers distributed computing support.

10. CatBoost: CatBoost is a gradient-boosting library designed for categorical feature support and high performance. It's particularly useful when dealing with structured data [160].

11. Isolation Forest: Isolation Forest is an anomaly detection method that isolates anomalies in data using decision trees. It's often used for identifying rare events or anomalies.

12. OneClassSVM: One-Class SVM is a support vector machine variant used for one-class classification or outlier detection. It's helpful when you have limited positive examples and want to identify anomalies.

13. Dummy Classifier: Dummy Classifier is used as a baseline for comparison. It generates predictions using simple rules (e.g., random or most frequent class) and helps assess the performance of other models.

14. LDA (Linear Discriminant Analysis): LDA is a dimensionality reduction and classification method. It's particularly useful for feature extraction and data visualization.

5.3 Meta-Learner AutoML Automated Machine Learning (AutoML)

Automated Machine Learning (AutoML) refers to the process of automating the end-to-end process of applying machine learning to real-world problems [207] [170] [66] [6] [119]. This includes automating tasks such as data preprocessing, feature selection, model selection, and hyperparameter optimization. By leveraging AutoML techniques, we can streamline the machine learning workflow, making it more efficient and accessible.

- **Data Preprocessing:** AutoML tools can automatically handle missing values, encode categorical variables, and scale features, ensuring that the data is ready for modeling.
- **Feature Selection and Engineering:** AutoML can identify the most relevant features and create new features from existing data, improving model performance.
- **Model Selection:** AutoML can evaluate multiple machine learning algorithms and select the best-performing model for the given dataset.
- **Hyperparameter Optimization:** AutoML tools can tune hyperparameters to achieve optimal model performance, often using techniques such as grid search, random search, or Bayesian optimization.
- **Model Evaluation:** AutoML uses cross-validation and other techniques to evaluate model performance, ensuring robust and generalizable results.
- **Deployment:** AutoML can automate the deployment process, allowing models to be easily integrated into production environments.

In this study, we explore the potential of AutoML to enhance the meta-learning approach for time series model selection and prediction. While AutoML can automate various aspects of the machine learning workflow, our primary goal is to avoid the traditional model selection problem through meta-learning. We hypothesize that there are correlations between datasets based on wider hydrological features present across datasets.

This assumption necessitates a meta-learner capable of capturing these correlations, thereby improving model selection and prediction performance. Therefore, although AutoML provides valuable tools for automation, our focus is on leveraging meta-learning to address the specific challenges of time series forecasting in hydrological contexts.

5.4 Evaluation Metrics

In this section, we present the key evaluation metrics used for assessing the performance of the meta-learner. The equations below are used to calculate accuracy, precision, recall, and F1-score, which are essential metrics for assessing the performance of the meta-learner.

5.4.0.1 Accuracy Calculation

Accuracy is calculated by comparing the predicted model rankings against the actual rankings of the models. It is defined as the number of correct predictions divided by the total number of predictions. The formula is:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}}$$

5.4.0.2 Precision, Recall, and F1-Score

Precision is a measure of the accuracy of a binary classification model. It quantifies the ability of the model to correctly identify positive instances out of all instances it predicted as positive. In other words, precision answers the question: "Of all the instances the model classified as positive, how many were actually positive?" [158]:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall (Sensitivity or True Positive Rate) measures the model's ability to correctly identify all relevant positive instances. It answers the question: "Of all actual positive instances, how many did the model identify as positive?" [158]

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1-Score The F1-Score is the harmonic mean of Precision and Recall. It provides a balanced evaluation of a classification model's performance. The F1-Score combines both Precision and Recall into a single value, making it particularly useful when there is a need to balance the trade-off between false positives and false negatives. Higher F1-Scores indicate better model performance, with values closer to 1 being more desirable [158]:

$$\text{F1-Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

5.4.0.3 Success and Failure

These metrics are commonly used in the evaluation of classification models, such as those used in machine learning and data analysis. They help in assessing the model's ability to correctly classify instances into different categories.

We will be using the F1 score as it is a measure of a meta-learner model's accuracy that considers both precision and recall. It's particularly useful when you have imbalanced classes in your dataset [169].

While an F1 score may indicate overfitting, differentiating between a truly accurate model and one that overfits can be achieved by comparing performance on training and validation datasets, and by using cross-validation techniques to ensure generalization.

- **Success**: Success is defined as the meta-learner correctly predicting the rank of the best model for a given dataset. This means that the top-ranked model predicted by the meta-learner matches the top-ranked model based on actual performance metrics.

- **Failure**: Failure occurs when the meta-learner's predicted top-ranked model does not match the actual top-ranked model. This can be due to incorrect evaluation metrics or misinterpretation of the features.

5.5 Experimental Setup

In order to identify the most appropriate forecasting model for a given time series, We recognize that the process is far from arbitrary; rather, it relies on a foundation of meticulous evaluation metrics. Within this critical category, we introduce an innovative approach wherein we harness the outcomes of prior forecasting models as additional input features for training a meta-learner. These metrics are instrumental in empowering the meta-learner to make informed decisions regarding model selection.

In this experiment, the dataset was divided into training and testing sets to ensure robust evaluation of the model's performance. The training set comprised 661 samples, while the testing set included 166 samples. This distribution resulted in 80% of the data being allocated for training purposes, ensuring the model has sufficient information to learn from, and 20% reserved for testing, allowing for a reliable assessment of the model's generalization capability. The division of the data was carefully designed to strike a balance between training the model effectively and evaluating its performance accurately.

5.5.1 Meta-Feature

In this section, we discuss the use of various meta-features in the model. These meta-features include sample size, execution time, validation metrics (such as validation loss, validation MSE, and validation MAE), test metrics (such as test loss, test MSE, and test MAE), model complexity (including the number of layers, number of dense layers, and learning rate), and statistical measures (such as mean, median, variance, standard deviation, range, skewness, and kurtosis). Additionally, geographical and geological attributes such as latitude, longitude, and other geological features are used. These features are crucial in training the meta-learner to make informed decisions regarding model selection. While training the meta-learner, all these comprehensive features are considered. However, during testing, the meta-learner relies more on statistical and geographical features, ensuring it can generalize well to new data and make accurate predictions even when some model-specific metrics are not available.

Overall, the meta-dataset comprises a total of 154 rows and 58 columns. However, we will use distinct features for training and testing purposes when constructing our machine-learning models. This distinction arises because the evaluation metric for model predictions will not be available during the testing phase. It is that the training and testing samples are independent to avoid overfitting. If the same data points are used in both sets, the model may perform well on the test set but fail to generalize to new data. This independence ensures that the evaluation metrics accurately reflect the model's performance.

Why Use Separate Features for the Train/Test Set?

To ensure that our meta-learner model generalizes well to new data, we utilize distinct sets of features during the training and testing phases. This approach helps our model make accurate predictions or classifications based on the specific attributes provided for each dataset.

1. **Training Features:**
    - Includes a comprehensive set of attributes such as station information, model details, sample size, execution time, validation metrics, test metrics, model complexity, and more.
    - These features provide detailed insights into each model's performance, which helps in effectively training the meta-learner.
2. **Testing Features:**
    - Focuses on statistical measures and geographical attributes such as mean, median, variance, standard deviation, range, skewness, kurtosis, latitude, and longitude.
    - These features are more readily available and do not depend on the performance of individual models, making them practical for use in the testing phase.
    - Ensures that the model can generalize well to new data and make accurate predictions or classifications based on the specific attributes provided for each dataset.

applications.

For training, we include a comprehensive set of features such as model parameters, sample size, execution time, and various validation metrics (loss, MSE, MAE). In contrast, during testing, we focus on a subset of these features, primarily statistical attributes and specific domain-related characteristics. For neural networks, the number of layers is a measure of complexity. For other models such as Random Forests (RF), complexity can be measured using parameters such as the number of trees, depth of trees, and the number of features considered for splitting at each node. These parameters determine the capacity of the model to capture patterns in the data.

For a complete list of the features used in both training and testing phases, please refer to Appendix H. By using distinct sets of features for training and testing, we aim to ensure that our learners can generalize well to new data and make accurate predictions or classifications based on the specific attributes provided for each dataset.

5.5.1.1 Imbalanced Data

The imbalance in the hydrological time series datasets used in this thesis arises due to the inherent characteristics of the data collected from various hydrometric stations. Imbalanced data is a common challenge in machine learning, particularly in classification tasks, where certain classes (minority classes) have significantly fewer instances compared to others (majority classes). In the context of this research, the class imbalance issue is evident as presented in the 5.4, where the distribution of the classes is skewed. This class imbalance issue can give rise to biased models that exhibit a preference for the majority class, resulting in subpar performance when predicting the minority class.

Handling class imbalance is crucial for improving the predictive performance of machine learning models. When a dataset is imbalanced, models tend to be biased towards the majority class, leading to poor performance in predicting the minority

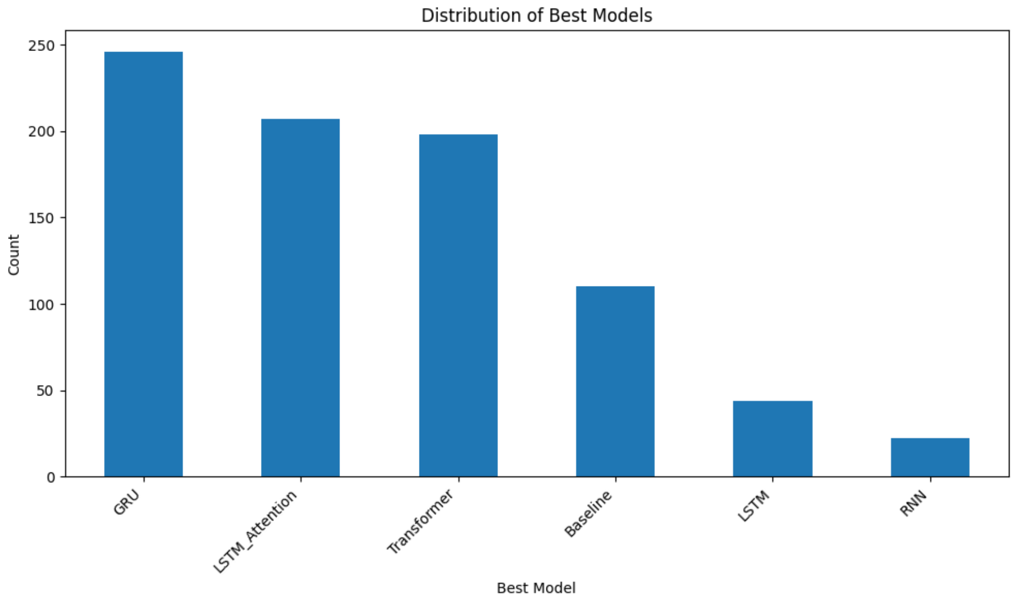

Figure 5.4: Class Distribution (Model)

class. This bias can result in misleading evaluation metrics, where high accuracy might not necessarily translate to good predictive power for the minority class. To address this, various resampling techniques such as oversampling, undersampling, and the Synthetic Minority Oversampling Technique (SMOTE) are employed.

Methods to Mitigate Class Imbalance

1. **Original Data No Resampling:** Using the dataset as initially collected without any modifications often leads to biased models favoring the majority class, resulting in poor performance on the minority class [94].

2. **Undersampling:** This technique reduces the number of instances in the majority class by randomly removing some of them, aiming to balance the class distribution. While it can effectively address class imbalance, it may lead to a loss of important information from the majority class [222].

3. **Oversampling:** This method increases the number of instances in the minority class by duplicating existing instances or generating synthetic ones using algorithms such as SMOTE. It helps balance the class distribution and improve model performance by increasing the representation of the minority class [206].

4. **SMOTE (Synthetic Minority Over-sampling Technique):** SMOTE creates synthetic instances by interpolating between existing minority class instances and their nearest neighbors. This technique enhances the diversity of the minority class and reduces the risk of overfitting [38].

The presence of imbalanced data in the datasets used for this thesis significantly impacts the performance of the meta-learner. By addressing class imbalance through the aforementioned techniques, the thesis aims to enhance the meta-learner's capability to select the most appropriate models for time series forecasting, leading to improved predictive performance and more reliable results

5.5.2 Meta-Learners

During the development Meta-Learners , we evaluated the performance of 14 different Classifiers: Random Forest, Logistic Regression, Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Gradient Boosted Decision Trees (GDBT), Multilayer Perceptron (MLP), XGBoost, LightGBM, CatBoost, Isolation Forest, OneClassSVM, Dummy Classifier, and Linear Discriminant Analysis (LDA). These models were trained and tested on a specific dataset, and their performance was measured using standard classification metrics, including Accuracy, Precision, Recall, and F1-score.

The next step in assessing the meta-learner is to evaluate its ability to rank the top models for a given time series accurately. The choice of the most suitable model should depend on the specific problem and the trade-offs between accuracy and model complexity. In this evaluation, Random Forest and K-Nearest Neighbors stand out as top contenders due to their high accuracy and well-balanced precision and recall scores. However, it's essential to consider factors such as computational resources, interpretability, and the nature of the problem when selecting the final model for deployment. Additionally, further analysis and experimentation, such as hyperparameter tuning, cross-validation, and feature engineering, may be necessary to fine-tune the models and potentially improve their performance. Ultimately, the selection of the best model should align with the project's objectives and requirements, ensuring that it delivers the desired level of performance and meets specific constraints.

Drawing from the specific outcomes discussed in the Chapter 4, our research findings reveal that certain best models dominate while others exhibit comparatively poorer performance. Consequently, the datasets utilized as input for the meta-learner exhibit class imbalance concerns.

5.6 Evaluation of Meta-Learners for Optimal Time Series Model Selection at Hydrometric Stations Across Ireland

This section provides a comprehensive evaluation of meta-learners aimed at selecting optimal time series models for hydrometric stations across Ireland. By leveraging data from 249 distinct geographical locations, we ensure a robust and diverse analysis, capturing the variability inherent in different regions.

The experiment utilized a variety of classifiers to select the most suitable time series models from a set of potential models (such as GRU, LSTM, and Transformer). Each classifier's performance was evaluated based on its ability to generalize and accurately predict the appropriate model for new data points.

A detailed analysis of performance metrics, summarized in Table 5.2, provides deeper insights into the classifiers' effectiveness in model selection.

Table 5.2

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**
		Original		
Random Forest	1.0000	1.0000	1.0000	1.0000
Logistic Regression	0.2952	0.1527	0.2952	0.1996

Table 5.2 continued from previous page

Model	Accuracy	Precision	Recall	F1-score
Naive Bayes	0.3675	0.5234	0.3675	0.2483
K-Nearest Neighbors	0.9819	0.9828	0.9819	0.9820
Support Vector Machine	0.3193	0.3096	0.3193	0.1618
Gradient Boosting	1.0000	1.0000	1.0000	1.0000
MLP	0.3795	0.4616	0.3795	0.3490
XGBoost	1.0000	1.0000	1.0000	1.0000
LightGBM	1.0000	1.0000	1.0000	1.0000
CatBoost	1.0000	1.0000	1.0000	1.0000
Isolation Forest	0.3072	0.1125	0.3072	0.1647
OneClassSVM	0.1566	0.1116	0.1566	0.1303
Dummy Classifier	0.3133	0.0981	0.3133	0.1494
LDA	0.7952	0.8155	0.7952	0.7990
		Random Oversampling		
Random Forest	1.0000	1.0000	1.0000	1.0000
Logistic Regression	0.1325	0.3722	0.1325	0.1217
Naive Bayes	0.3554	0.5116	0.3554	0.2360
K-Nearest Neighbors	0.9819	0.9836	0.9819	0.9822
Support Vector Machine	0.2470	0.4334	0.2470	0.1960
Gradient Boosting	1.0000	1.0000	1.0000	1.0000
MLP	0.3916	0.4347	0.3916	0.3428
XGBoost	1.0000	1.0000	1.0000	1.0000
LightGBM	1.0000	1.0000	1.0000	1.0000
CatBoost	1.0000	1.0000	1.0000	1.0000
Isolation Forest	0.3012	0.1160	0.3012	0.1675
OneClassSVM	0.1205	0.0964	0.1205	0.1071
Dummy Classifier	0.1386	0.0192	0.1386	0.0337

Table 5.2 continued from previous page

Model	Accuracy	Precision	Recall	F1-score
LDA	0.6928	0.7387	0.6928	0.7024
		Random Undersampling		
Random Forest	0.8494	0.9311	0.8494	0.8736
Logistic Regression	0.2470	0.2719	0.2470	0.2373
Naive Bayes	0.3434	0.4107	0.3434	0.2512
K-Nearest Neighbors	0.4578	0.5248	0.4578	0.4583
Support Vector Machine	0.2410	0.4455	0.2410	0.1861
Gradient Boosting	0.8313	0.8544	0.8313	0.8363
MLP	0.2470	0.3092	0.2470	0.2119
XGBoost	0.8855	0.8998	0.8855	0.8859
LightGBM	0.8855	0.8983	0.8855	0.8854
CatBoost	0.8976	0.9267	0.8976	0.9065
Isolation Forest	0.2952	0.1154	0.2952	0.1659
OneClassSVM	0.1325	0.0999	0.1325	0.1139
Dummy Classifier	0.1386	0.0192	0.1386	0.0337
LDA	0.7108	0.7935	0.7108	0.7170
		SMOTE		
Random Forest	1.0000	1.0000	1.0000	1.0000
Logistic Regression	0.4096	0.4850	0.4096	0.3895
Naive Bayes	0.3554	0.5098	0.3554	0.2338
K-Nearest Neighbors	0.9819	0.9828	0.9819	0.9820
Support Vector Machine	0.2470	0.4334	0.2470	0.1960
Gradient Boosting	1.0000	1.0000	1.0000	1.0000
MLP	0.3193	0.4737	0.3193	0.3097
XGBoost	1.0000	1.0000	1.0000	1.0000
LightGBM	1.0000	1.0000	1.0000	1.0000

Table 5.2 continued from previous page

Model	Accuracy	Precision	Recall	F1-score
CatBoost	1.0000	1.0000	1.0000	1.0000
Isolation Forest	0.2952	0.1129	0.2952	0.1633
OneClassSVM	0.1205	0.1062	0.1205	0.1129
Dummy Classifier	0.1386	0.0192	0.1386	0.0337
LDA	0.7048	0.7481	0.7048	0.7146

Classifier Analysis

Random Forest consistently achieved perfect scores across all metrics, demonstrating its robustness and reliability in model selection. It achieved an accuracy, precision, recall, and F1-score of 1.0000, indicating flawless performance in predicting the optimal time series model.

Logistic Regression performed significantly lower, with an accuracy of 0.2952 and an F1-score of 0.1996. This indicates that while it may occasionally predict correctly, its overall reliability is limited.

Naive Bayes showed moderate performance with an accuracy of 0.3675 and an F1-score of 0.2483, suggesting some capability in model selection, but with considerable room for improvement.

K-Nearest Neighbors (KNN) demonstrated high effectiveness with an accuracy of 0.9819 and an F1-score of 0.9820, closely trailing the top-performing classifiers.

Support Vector Machine (SVM) achieved an accuracy of 0.3193 and an F1-score of 0.1618, indicating moderate performance.

Gradient Boosting classifiers, including **XGBoost**, **LightGBM**, and **CatBoost**, matched the perfect scores of Random Forest, each achieving an accuracy, precision, recall, and F1-score of 1.0000.

Multi-Layer Perceptron (MLP) showed varied performance across different sampling methods, with its highest performance under original sampling (accuracy of 0.3795 and F1-score of 0.3490) and the lowest under random undersampling (accuracy of 0.2470 and F1-score of 0.2119).

Isolation Forest and **OneClassSVM** classifiers performed poorly, with low accuracies (0.3072 and 0.1566, respectively) and corresponding F1-scores (0.1647 and 0.1303), highlighting their unsuitability for this task.

Dummy Classifier, used as a baseline, predictably performed the worst, with an accuracy of 0.3133 and an F1-score of 0.1494, confirming that the predictive power of other classifiers is not by chance.

Linear Discriminant Analysis (LDA) showed good performance, with an accuracy of 0.7952 and an F1-score of 0.7990 under original sampling, and moderate performance under different sampling methods.

Influence of Resampling Techniques

The study also explored the impact of various sampling methods on classifier performance, summarised in Table 5.3.

Table 5.3: Impact of Resampling Techniques on Classifier Performance

Classifier	**Original**	**Random Oversampling**	**Random Undersampling**	**SMOTE**
CatBoost	166	166	149	166
Dummy Classifier	52	23	23	23
Gradient Boosting	166	166	138	166
Isolation Forest	0	0	0	0
K-Nearest Neighbors	163	163	76	163
LDA	132	115	118	117
LightGBM	166	166	147	166
Logistic Regression	49	22	41	68
MLP	63	65	41	53
Naive Bayes	61	59	57	59
OneClassSVM	0	0	0	0

Table 5.3 continued from previous page

Classifier	Original	Random Oversampling	Random Undersampling	SMOTE
Random Forest	166	166	141	166
Support Vector Machine	53	41	40	41
XGBoost	166	166	147	166

The results indicate distinct patterns of performance for each resampling technique as shown in Figures 5.5 and 5.6.

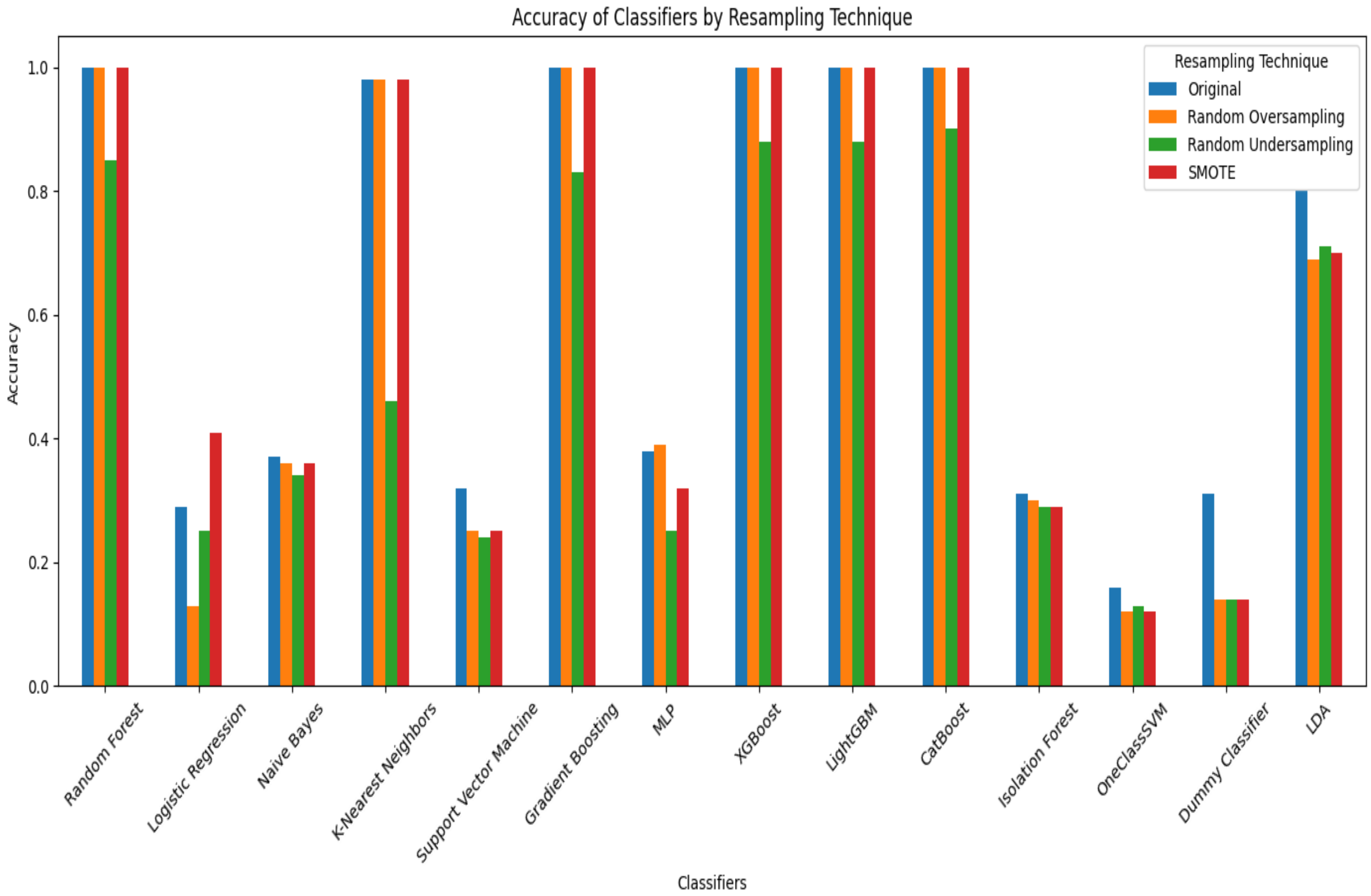

Figure 5.5: Accuracy of Classifiers by Resampling Technique

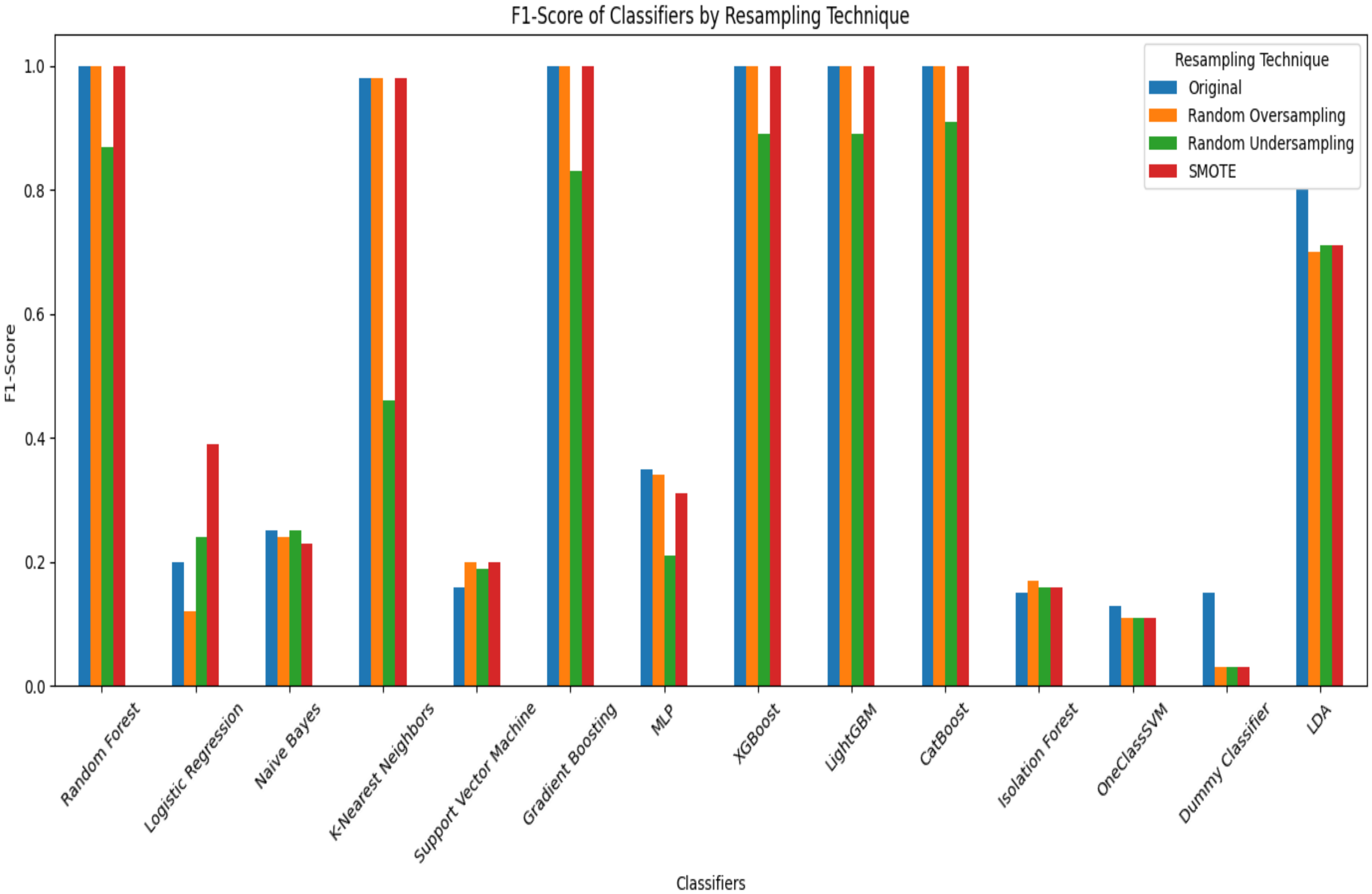

Figure 5.6: F1-Score of Classifiers by Resampling Technique

- Random oversampling shows mixed results across different classifiers. Classifiers such as Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost maintain high accuracy and F1-scores with this technique, indicating their robustness and ability to handle oversampled data effectively. These ensemble methods leverage the additional data points to improve their decision boundaries without overfitting significantly. However, simpler classifiers like Logistic Regression and Support Vector Machine (SVM) exhibit lower performance. This suggests that these models might be more prone to overfitting with duplicated examples, as they do not have the inherent complexity to distinguish between genuine and replicated instances effectively. Naive Bayes also shows moderate performance, but its probabilistic nature helps it manage the oversampled data better than logistic regression and SVM. From Figure 5.5, we observe that the accuracy of Logistic Regression drops significantly when using Random Oversampling, possibly due to the model overfitting on the duplicated samples. Similarly, SVM shows a substantial decrease in performance, indicating that it struggles with the redundancy introduced by oversampling.

- Random undersampling generally results in lower performance metrics across most classifiers, as shown in Figures 5.5 and 5.6. The reduction in the majority class can lead to loss of valuable information, affecting the classifiers' ability to generalize. This is particularly evident in ensemble methods where the performance drops significantly compared to the original or oversampled datasets. However, classifiers such as K-Nearest Neighbors (KNN) and Naive Bayes are less affected by random undersampling. KNN benefits from a more balanced class distribution as it relies on the neighborhood of data points, while Naive Bayes, with its assumption of feature independence, can still perform reasonably well with fewer samples from the majority class. The performance of KNN, as seen in both accuracy and F1-score metrics, remains relatively stable with Random Undersampling. This stability suggests that KNN can effectively utilize the smaller, balanced dataset to maintain its classification ability. Naive Bayes also performs consistently, benefiting from its probabilistic approach that is less reliant on the volume of data.

- SMOTE appears to provide a balanced approach, improving performance metrics for several classifiers. This technique generates synthetic examples for the minority class, thus creating a more diverse and representative dataset. For example, Logistic Regression and Multi-Layer Perceptron (MLP) show significant improvements in both accuracy and F1-scores when using SMOTE. This highlights the benefits of synthetic data generation in helping these classifiers generalize better. SMOTE also enhances the performance of ensemble methods, though not as dramatically as random oversampling. The generated synthetic instances provide a better spread of the minority class without simply duplicating existing examples. Figure 5.6 clearly shows that the F1-scores of Logistic Regression and MLP improve notably with SMOTE. This improvement suggests that SMOTE's synthetic examples help these classifiers better capture the underlying patterns of the minority class, leading to enhanced performance.

The detailed analysis of the performance of each classifier under different resampling techniques reveals several insights. Ensemble methods such as Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost consistently show high performance across different resampling techniques, particularly excelling with Random Oversampling and SMOTE. These methods benefit from the additional diversity provided by these techniques, improving both their accuracy and F1-scores. The robustness of these ensemble methods is evident from their ability to maintain high performance even with resampling, which adds complexity to the dataset. Simple classifiers like Logistic Regression and SVM struggle with oversampling and undersampling but show noticeable improvements with SMOTE. This indicates that the introduction of synthetic samples helps these classifiers to better capture the decision boundaries without overfitting. The significant improvement in both accuracy and F1-scores for these classifiers when using SMOTE highlights the effectiveness of synthetic data in mitigating the issues of class imbalance. KNN and Naive Bayes handle undersampling better than others, suggesting their robustness to smaller datasets. However, they also benefit from SMOTE, showing that synthetic data can enhance their performance without introducing significant bias. The consistent performance of KNN and Naive Bayes across different resampling techniques underscores their flexibility and adaptability to various dataset distributions. In summary, the experiment results demonstrates that the choice of resampling technique can have a profound impact on the performance of different classifiers. Ensemble methods and sophisticated techniques like SMOTE generally provide better results, highlighting their importance in handling class-imbalanced datasets. The detailed analysis of classifier performance under various resampling techniques offers valuable insights for selecting appropriate methods to address class imbalance in time series data.

The overall accuracy of the predictions stands at 55.86%, indicating that just over half of the time, the predicted labels align with the true labels (see F). Utilizing the Random Forest classifier for prediction, a detailed analysis of the classifier's performance across different labels reveals distinct patterns of accuracy and misclassification.

5.7 helps to visualize these patterns, where diagonal elements indicate correct predictions and off-diagonal elements show misclassifications.

The diagonal elements, which represent correct predictions, show that the classifier performs best on the GRU and Transformer labels, with high counts of accurate predictions at 1546 and 1687 respectively. However, the off-diagonal elements reveal significant misclassifications. For instance, the Baseline label is frequently misclassified as GRU and Transformer, with 142 and 126 occurrences respectively. Similarly, the GRU label is often misclassified as Inlier (287 times) and Transformer (294 times). The LSTM label, though correctly predicted 181 times, is notably misclassified as Inlier (35 times). The LSTM_Attention label also shows a considerable number of correct predictions (978), but is misclassified as GRU (239 times) and Inlier (181 times). The RNN label has a relatively lower count of correct predictions

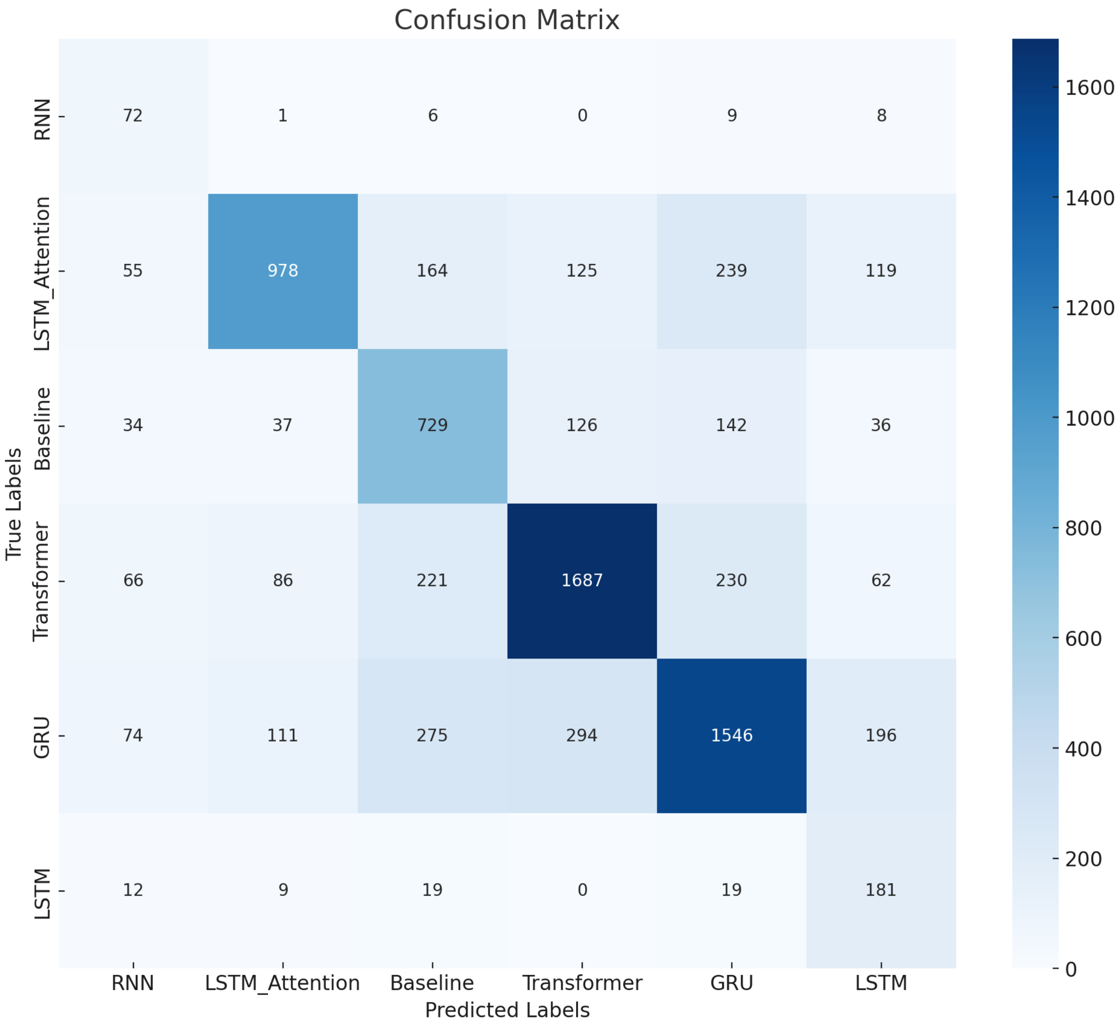

Figure 5.7: Confusion Matrix of Classifier Predictions

(72) and is frequently misclassified as Inlier (16 times). These patterns indicate specific areas where the classifier excels and other areas where it has difficulties, suggesting potential directions for improving its precision, such as focusing on reducing misclassifications among similar labels and enhancing the distinction between frequently confused classes.

5.7 Summary

In conclusion, this chapter has offered an evaluation of the meta-learner's performance in the context of model selection time series forecasting. We assessed the effectiveness of 14 machine learning models and explored various resampling techniques to address class imbalance within the dataset. While some models, such as Random Forest and K-Nearest Neighbors, demonstrated strong performance and may be considered as promising candidates for model selection in specific scenarios, others, such as the Support Vector Machine, struggled consistently. Additionally, we highlighted the critical role of choosing appropriate resampling strategies to enhance model performance, with oversampling techniques often proving effective.

Furthermore, this research was part of a multidisciplinary project, and our industrial partner, ARUP, provided valuable domain expertise. They found the findings highly significant and applicable, emphasizing the practical relevance and potential impact of our results in real-world applications.

These findings underscore the importance of carefully tailoring the choice of machine learning models and class imbalance mitigation techniques to the specific needs of time series forecasting projects. Further evaluation of diverse datasets will be essential to ascertain the meta-learner's robustness and its applicability to different forecasting tasks.

1. **Model Performance Variability:** The evaluation of different classifiers for time series prediction revealed a substantial variability in their performance. Some models, such as Random Forest and K-Nearest Neighbors, consistently

demonstrated high accuracy and balanced precision and recall, indicating their suitability for the task. Others, such as Logistic Regression and Naive Bayes, showed competitive performance depending on the specific resampling strategy. Meanwhile, the Support Vector Machine consistently underperformed, suggesting it may not be the best choice for this task.

2. **Overfitting Concerns:** The Random Forest model, while achieving near-perfect scores, raised concerns about potential overfitting. Overfitting occurs when a model becomes overly tuned to the training data, capturing noise and idiosyncrasies that do not generalize well to unseen data. This highlights the importance of model selection and the need for further analysis and experimentation to fine-tune hyperparameters and prevent overfitting.

3. **Class Imbalance Challenge:** The real-world time series dataset exhibited class imbalance, a common challenge in machine learning. Class imbalance can lead to biased models that favor the majority class and perform poorly on the minority class. Different resampling techniques, including oversampling and undersampling, were explored to address this issue. The choice of resampling technique significantly impacted model performance, with oversampling often yielding better results.

4. **Meta-Learner Benefits:** The use of a meta-learner for model selection proved valuable in this context. By leveraging a meta-learner, it was possible to combine the strengths of different base classifiers and mitigate their weaknesses. This approach allowed for more robust model selection, especially when dealing with highly variable performance across different classifiers.

5. **Resampling Strategy Importance:** The selection of the appropriate resampling strategy played a crucial role in enhancing the performance of the meta-learner. Oversampling techniques, such as Random Oversampling and SMOTE, were particularly effective in improving model performance, addressing class imbalance, and preventing models from being biased toward the ma-

jority class.

6. **Recommendations for Model Selection:** Based on the evaluation results, Random Forest and K-Nearest Neighbors emerged as top contenders for the time series prediction task. However, the choice of the best model should align with the project's objectives, constraints, and specific problem requirements. Considerations such as computational resources, model interpretability, and the nature of the problem should guide the final model selection.

In conclusion, using a meta-learner for model selection in time series prediction offers a powerful approach to leverage the strengths of different classifiers and address challenges such as class imbalance. However, carefully considering the base classifiers and resampling strategies is crucial to achieve the best results. Further fine-tuning and experimentation may be necessary to optimize model performance and ensure that the selected model aligns with the desired objectives.

Chapter 6

Conclusion

In this concluding chapter, Section 6.1 provides a comprehensive summary of the work conducted throughout this study. In Section 6.2, potential prospects for future research are discussed to expand upon the work presented in this thesis.

6.1 Thesis Overview

This thesis is dedicated to the systematic selection of predictive models using meta-learning methodologies, specifically within the domain of time series data analysis. The central focus of this research addresses the complex challenges associated with selecting optimal models suited to various typologies of time series data. By employing meta-learning techniques, this study aims to improve the accuracy and efficiency of predictive model selection, thereby making a significant contribution to the fields of machine learning and data science.

In Chapter 1, we commenced with an introduction to environmental data integration. Given that time series data constitute a substantial subset of environmental datasets, this was followed by a discussion on time series predictive algorithms, eventually leading to the exploration of the application of meta-learning for time series predictive models. Subsequently, we engaged in an analysis of open research issues, which played a pivotal role in crystallizing our overarching hypothesis and, consequently, in the formulation of three specific research questions: The first research question focuses on the challenges and complexities of integrating heterogeneous environmental datasets, considering variations in spatiotemporal resolution, data format, and attribute representation. The second question addresses the primary challenges in developing accurate predictions for environmental time series data, specifically within the hydrology domain. The final question explores the feasibility of constructing a meta-learner to aid time series researchers in selecting the most suitable predictive model for the dataset they are analyzing.

Chapter 2 This chapter lays the foundation for meta-learning in model selection by providing a comprehensive literature review covering several critical aspects of our research. It starts with environmental data integration, discussing various data sources, types, formats, and quality issues. Significant advancements have been made in integrating diverse environmental data, but challenges remain in maintaining data quality and consistency, especially with spatiotemporal resolution and data format variations. Data warehousing concepts, architectures, and integration patterns (ETL and ELT) are examined, highlighting the strengths and limitations of current approaches. Data warehouses offer robust solutions for organizing and analyzing large volumes of data but often require substantial computational resources and lack flexibility to adapt to changing data environments. The review then delves into hydrological predictive models, focusing on traditional and machine learning-based models. Traditional models are robust and reliable, while machine learning models excel in handling complex, non-linear relationships in data. However, machine learning models often demand significant computational resources and specialized expertise, and traditional models may not effectively capture the dynamic nature of hydrological systems. Time series analysis is also scrutinized, emphasizing the characteristics of time series data, prediction techniques, and the complexities of model selection. Automated algorithms like ETS and ARIMA have facilitated more efficient model selection, but the task remains challenging due to the diverse characteristics of time series data. Meta-learning is introduced as a promising approach for automating the time series model selection process. Meta-learning leverages historical performance data and dataset characteristics to make informed model selection decisions, reducing the need for manual intervention and improving prediction accuracy. Despite its potential, the field of meta-learning is still in its early stages, with many open research questions and opportunities for further development. In summary, this chapter synthesizes critical insights from the domains of environmental data integration, data warehousing, hydrological predictive modeling, time series analysis, and meta-learning. It identifies strengths and weaknesses in contemporary work, highlighting gaps that guide the formulation of the following research questions: What are the challenges of integrating heterogeneous environmental datasets, considering variations in spatiotemporal resolution, data format, and attribute representation? What are the primary challenges in developing accurate predictions for environmental time series data, specifically within the hydrology domain? Is it feasible to construct a meta-learner to aid time series researchers in selecting the most suitable predictive model for the dataset they are analyzing?

In Chapter 3, we introduce a methodological architecture that leverages ETL (Extract, Transform, Load) technology to address the integration of heterogeneous environmental data from diverse sources. This Spatiotemporal Data Integration (STDI) approach effectively handles challenges encompassing data sourcing, extraction, transformation, loading, mapping, spatiotemporal integration, and data quality validation, providing a holistic solution for Environmental Spatiotemporal Data Integration. Our methodological architecture is adaptable and versatile, making it suitable for various domains where integrating diverse data from multiple sources is essential for practical analysis, visualization, and decision-making. Users can customize algorithms' rules and parameters to tailor the integration process to their needs, offering flexibility for various integration scenarios. To illustrate this architecture's practical application and effectiveness, we present a case study centred on STDI in Ireland. We employ distance-based algorithms for spatial data integration and implement temporal data integration techniques to seamlessly merge data, ensuring optimal alignment and minimal information loss. The case study includes integrating river water levels and weather data, proximity analysis of weather stations to geological features, and merging water sensor data with geological features for assessing water quality patterns. Additionally, we explore the interplay between geology, topology, and climate.

Chapter 4 built upon Chapter 3's emphasis on spatiotemporal data integration and highlights the pressing need to apply this integrated dataset to real-world applications. Predicting river water levels is vital, demanding precision and adaptability. Accurate predictions are significant in managing water resources, mitigating flood and drought risks, and improving decision-making processes. Traditional methodologies, although accurate, often require substantial computational resources and specialized expertise, leaving room for improvement. This chapter evaluates and selects the optimal machine learning model for predicting river water levels. The objective is to provide insights into the advantages and drawbacks of each model, ultimately creating an experimental dataset that will serve as the basis for the meta-learner.

Chapter 5 addresses the formidable challenge of model selection for time series data encountered, marked by inherent complexity and the impracticality of exhaustive testing. Within this context, a groundbreaking meta-learning process is introduced, capable of identifying the optimal model for time series data. This chapter elaborates on the methodology that underpins selecting a suitable time-series prediction model, reinforced by the Meta-Dataset, which serves as the foundational input for the meta-learner. Furthermore, the practical implementation of the methodology is presented, and the research outcomes are shared. This chapter provides a comprehensive understanding of how the challenges associated with model selection for time series data can be surmounted, along with an appreciation of the advantages of applying meta-learning in this context.

In summary, the main contribution of the research is the meta-learning approach for selecting the appropriate model for water level time series prediction using the previously proposed Meta-Dataset. This approach integrates distinct categories of input features, including Time Series Input Features, Geology Features, and ML Prediction Results Features, to provide a comprehensive understanding of the intricate dynamics underlying water time series. The research empowers hydrologists and researchers to make informed decisions regarding model selection, contributing to more accurate and timely water level predictions with far-reaching implications for water resource management, environmental conservation, and decision-making processes.

6.2 Future Research

This thesis has established a strong foundation for advancing the field of time series data analysis, specifically in selecting predictive models through innovative meta-learning methodologies. The journey so far has illuminated several avenues for future research and exploration.

Future work can encompass the following directions:

Enhanced Meta-Learning Framework: The Meta-Dataset and the meta-learning process introduced in this research exhibit immense promise in facilitating the selection of optimal models for time series data. Enhancing this framework could involve exploring advanced techniques in few-shot and zero-shot learning to augment further the meta-learner's capacity to generalize across different datasets and model categories.

Incorporating Additional Features: While the current research incorporates Time Series Input Features, Geology Features, and ML Prediction Results Features, broadening the scope of features used in the model selection process can offer more comprehensive insights into the behaviour of water time series, For instance, the use of the uses of satellite data to analyze the temporal and spatial variability of surface waters around Ireland [31].

Real-Time Predictions: Extending the research to encompass real-time predictions involving data streaming and the capacity to update models on the fly could prove valuable, particularly for applications that necessitate immediate decision-making, such as flood management.

Multimodal Data Integration: Incorporating multiple types of data sources, including other sensor data, satellite imagery, and climate models, can provide a more holistic understanding of environmental dynamics. Subsequent research may explore the challenges and opportunities of integrating multimodal data into the predictive modelling framework.

Environmental Impact Assessment: Expanding the applications of accurate water level predictions, future research can delve into the environmental impact assessment of water resource management decisions. This could encompass evaluating the ecological consequences of water level changes and their effects on aquatic ecosystems.

Hydrological Model Integration: Integrating machine learning-based predictive models with traditional hydrological models can create hybrid systems that leverage the strengths of both approaches. Such collaborative modelling has the potential to enhance the accuracy and reliability of predictions, particularly in regions with complex hydrological dynamics.

Counterfactuals: A key underpinning of meta-learning is that of explainable AI (XAI). Within this field there exists the topic of counterfactuals where a user can post "what-if" questions to a model to gain actionable insights. The additional of counterfactuals and more investment into the explainability of the meta-learner could provide valuable insights into the characteristics of each dataset further broadening the models predictive power.

Sustainable Resource Management Applications: The research can be extended to support sustainable resource management in various domains, including water resource management, forestry, and agriculture. Predictive models can play a significant role in optimizing resource allocation, reducing waste, and promoting environmentally responsible practices.

These future research directions hold the potential to further advance the field of environmental and climate research. By harnessing the power of predictive modelling and meta-learning, researchers can address pressing challenges related to climate change, ecosystem health, sustainable resource management, and regional variations in hydrological dynamics.

Appendix A

Stakeholders

This appendix chapter provide and explain the main identified stake holders.

The Office of Public Works (OPW) The OPW is a service organisation. Their services cover two main areas. First, Estate Portfolio Management including Heritage Services. sSecond, Flood Risk Management (Corporate Services support all areas), The OPW, as the leading agency for flood risk management in Ireland, minimises the impacts of flooding through sustainable planning. The mission of the OPW is to provide innovative, effective and sustainable shared services to the public and their clients using their experience and expertise

Environmental Protection Agency (EPA) The Environmental Protection Agency is responsible for regulating the quality of drinking water and enforcing the Drinking Water Regulations for public water supplies.

Local Authority Waters Programme (LAWPRO) The Local Authority Waters Programme (LAWPRO) is tasked with identifying the problems influencing water quality in each of Ireland's counties.

Transport Infrastructure Ireland (TII) The TII is a state agency in Ireland dealing with road and public transport infrastructure. Its mission is to deliver transport infrastructure and services, which contribute to the quality of life for the people of Ireland and support the country's economic growth. Road projects may have impacts on the water environment. These include impacts on the quality of water bodies and on the existing hydrology of the catchments through which roads pass. Among TII interest Road Drainage and the Water Environment Roads are designed to drain freely to prevent the build-up of standing water on the carriageway whilst avoiding exposure to or causing flooding.

Inland Fisheries Ireland (IFI) IFI is responsible for managing and protecting Ireland's inland fisheries and sea angling resources, for example, setting conservation limits and issuing fishing licences. It is also involved in a broad range of fisheries research and monitoring activities and sectors.

Irish Water (IW) It is a water utility company in Ireland. Irish Water was established to deliver the Government's reform programme for the water services sector. *"The task assigned to Irish Water by the Government is to build a new national water utility to provide safe, affordable and environmentally compliant water services to all customers."* Irish Water has been established to take on the following challenges: Drinking Water Quality, Drinking Water, Wastewater Quality and Wastewater Capacity.

Waterways Ireland (WI) It is responsible for managing, maintaining, developing, and restoring Inland navigable waterways primarily for recreational purposes. Waterways Ireland's work includes: Water Management (Water Quality, pollution and water levels), Infrastructure (navigation, canals ...), Marine Notices to public and Boating (Boat Registration, permits, Mooring).

Appendix B

OPW Hydrometric Stations

Opw provides 432 hydrometric stations as presented in the Figure B.1 below.

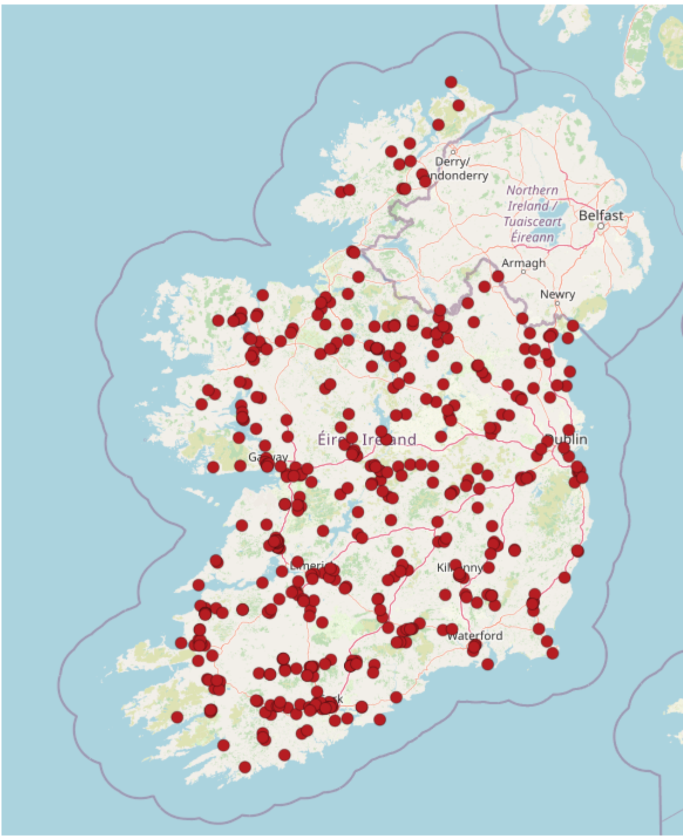

Figure B.1: OPW hydrometric stations

The table below provides the selected hydrometric station in *Experiment 1: Eight Machine Learning Models on 70 Hydrometric Stations* in Chapter 4.

Station name	Station number	RBD	River name	Catchment	Area	Latitude	Longitude
Aclint	6026	N	LAGAN (GLYDE)	GLYDE	144.00 km²	53.92476528	-6.640019444
Brewery Park	6015	N	RAMPARTS	MAIGUE	14.76 km²	53.993583	-6.416611
Cappoge Bridge	3058	N	BLACKWATER [MONAGHAN]	BLACKWATER [MONAGHAN]	65.00 km²	54.26680861	-7.021296944
Charleville	6013	N	DEE	—	307.00 km²	53.85584278	-6.413995833
Clarebane	6012	N	FANE	FANE	163.00 km²	54.09285639	-6.666055833
Glaslough	3055	N	MOUNTAIN WATER	BLACKWATER [MONAGHAN]	72.00 km²	54.32328083	-6.894343889
Ladyswell	6036	N	Ramparts	Castletown river	18.00 km²	53.99376162	-6.405584048
Mansfields town	6021	N	GLYDE	GLYDE	321.00 km²	53.89660361	-6.444490556
Moyles Mill	6011	N	FANE	FANE	230.00 km²	54.01157444	-6.596077222
Port Oriel	6060	N	IRISH SEA	—		53.79899	-6.221712778
Aghawoney	39009	NW	LEANNAN	LEANNAN	207.00 km²	55.04378556	-7.720692778
Anlore	36015	NW	FINN	FINN	98.27 km²	54.177075	-7.177330833
Ashfield	36018	NW	DROMORE	ERNE	233.00 km²	54.07216528	-7.121145556
Ballybofey	1043	NW	FINN [Donegal]	FINN	319.00 km²	54.79976861	-7.790749444
Ballyloskey	40008	NW	BALLYWILLY BROOK	Culdaff-Clomany Donagh-Coastal	2.69 km²	55.247114	-7.262751
Bellahillan	36011	NW	ERNE	ERNE	318.00 km²	53.96262361	-7.457176111
Butlers Bridge	36010	NW	ANNALEE	ERNE	774.00 km²	54.04187944	-7.377087778
Foalies Bridge	36171	NW	L. ERNE UPPER	ERNE	1520.00 km²	54.1387075	-7.436896111
Gartan Bridge	39008	NW	LEANNAN	LEANNAN	77.39 km²	55.00025333	-7.894476111
Keenagh Deel Bridge	34114	NW	DEEL [Crossmolina]	MOY	65.00 km²	54.08130375	-9.514569457
Abington	25003	S	MULKEAR	Shannon	397.00 km²	52.63186778	-8.421220833
Athlacca	24005	S	MORNINGSTAR	MAIGUE	131.95 km²	52.45876611	-8.65116583
Annacotty	25001	S	MULKEAR	SHANNON	646.00 km²	52.66927583	-8.529146944
Ballinamore	26001	S	SHIVEN	SHANNON	230.00 km²	53.48978639	-8.366001667
Banagher	25017	S	SHANNON	SHANNON	7989.00 km²	53.19378667	-7.993646667
Ballinasloe Town	26354	S	SUCK	Shannon	1428.26 km²	53.3295	-8.21493
Castleroberts	24008	S	MAIGUE	SHANNON	805.00 km²	52.54344083	-8.767416389
Clarecastle Barrage	27065	S	FERGUS ESTY	SHANNON	625.58 km²	52.8172475	-8.962748889
Clarianna	25029	S	NENAGH	SHANNON	292.67 km²	52.89162278	-8.207709722
Gourdeen	25027	S	OLLATRIM	SHANNON	118.00 km²	52.86841167	-8.168575
Aasleagh Bridge	32060	W	ERRIFF	ERRIFF	166.32 km²	53.61774917	-9.671156667

Station name	Station number	RBD	River name	Catchment	Area	Latitude	Longitude
Ballina	34061	W	MOY ESTY	MOY	1923.00 km²	54.11718417	-9.146781
Ballylahan	34004	W	MOY	MOY	935.50 km²	53.93794028	-9.102895
Clarinbridge	29004	W	CLARINBRIDGE	Galway Bay South East	123.00 km²	53.22934167	-8.874782778
Dangan	30098	W	CORRIB	CORRIB	3091.00 km²	53.29616944	-9.0765925
Kilcolgan	29011	W	DUNKELLIN	Kilcogan	373.00 km²	53.21410944	-8.871328889
Pontoon	34081	W	L. CULLIN	MOY	819.40 km²	53.97766528	-9.208075833
Riverside Close	23050	W	BRICK	FEALE	6.39 km²	52.344863	-9.688059
Terryland	30117	W	Terryland Sandy River	CORRIB	5.99 km²	53.285939	-9.041714
Wolfe Tone Bridge	30061	W	CORRIB ESTY.	CORRIB	3111.00 km²	53.26998	-9.05567
Anglesea Road	9369	E	Dodder	LIFFEY	106.22 km²	53.327878	-6.230943
Bluebell	9110	E	Naas Canal Supply Stream	LIFFEY	0.99 km²	53.20790295	-6.67846311
Cherry Wood	10048	E	LOUGHLINS TOWN	—	20.04 km²	53.246972	-6.137667
Fyanstown	7006	E	MOYNALTY	BOYNE	188.00 km²	53.72582306	-6.802883056
Giles Quay	6062	E	SEA	—	0.00 km²	53.98444001	-6.23990263
Killyon	7002	E	DEEL [Raharney]	BOYNE	285.00 km²	53.48777083	-6.970770833
Leixlip	9001	E	RYEWATER	LIFFEY	215.00 km²	53.36866472	-6.490438611
Mornington	7062	E	BOYNE ESTY	—	0.00 km²	53.71942667	-6.254402778
Navan Weir	7009	E	BOYNE	BOYNE	1610.00 km²	53.64355944	-6.6720575
Trim	7005	E	BOYNE	BOYNE	1282.00 km²	53.55640528	-6.791843889
Adelphi Quay	16160	SE	JOHN'S RIVER	SUIR	3508.00 km²	52.25966639	-7.102433056
Brownsbarn	15006	SE	NORE	NORE	2388.00 km²	52.50076972	-7.091699722
Cahir Park	16009	SE	SUIR	SUIR	1602.00 km²	52.35767333	-7.922936944
Dundalk Port	6061	SE	—	—	422.00 km²	54.00769064	-6.38555706
Joyces Lane	16147	SE	SUIR	—	2140.59 km²	52.35155588	-7.70670506
Killardry	16007	SE	AHERLOW	SUIR	273.35 km²	52.41727611	-7.975713611
Levitstown	14019	SE	BARROW	BARROW	1660.00 km²	52.93537778	-6.949797222
McMahons Bridge	15004	SE	NORE	NORE	491.00 km²	52.86702444	-7.379327222
Rathangan	14011	SE	SLATE	BARROW	163.00 km²	53.22061583	-6.992655278
Tullowbeg	12006	SE	SLANEY	SLANEY	251.00 km²	52.79606417	-6.738396944
Athea D/S	23051	SW	GALEY	FEALE	36.30 km²	52.461378	-9.286944
Athea U/S	23052	SW	GALEY	FEALE	36.30 km²	52.460679	-9.286718

Station name	Station number	RBD	River name	Catchment	Area	Latitude	Longitude
Ballydahin	18119	SW	BLACKWATER [MUNSTER]	BLACKWATER [MUNSTER]	1190.19 km²	52.13139014	-8.65421144
Ballyduff	18002	SW	BLACKWATER [MUNSTER]	BLACKWATER [MUNSTER]	2338.00 km²	52.14434472	-8.051951667
Coolmuckey Br	19112	SW	Bride (Cork)	LEE	70.21 km²	51.86084	-8.783723
Flesk Bridge	22006	SW	FLESK(LAUNE)	LAUNE	325.00 km²	52.04803417	-9.497946944
Glandalane	18053	SW	BLACKWATER [MUNSTER]	BLACKWATER [MUNSTER]	2275.00 km²	52.14981139	-8.220838333
Morris's Bridge	19104	SW	Laney	LEE	82.11 km²	51.92933	-8.93616
Riverville	22003	SW	MAINE	LAUNE	272.00 km²	52.19762111	-9.570731944
Waterworks Weir	19102	SW	LEE	LEE	1185.00 km²	51.893989	-8.510053

Appendix C

Data Warehouse Schema

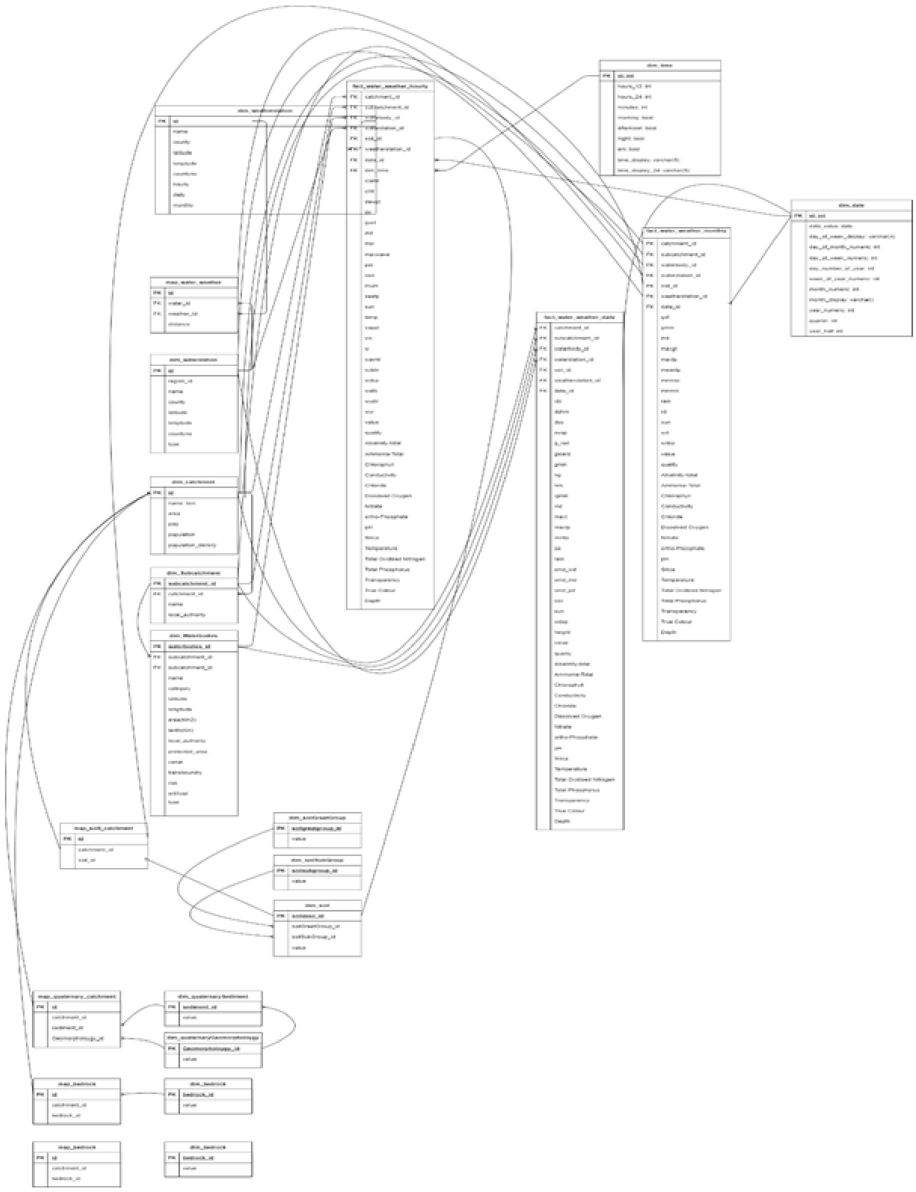

Figure C.1: Whole DW schema

Appendix D

Data Mapping Rules

Data Mapping Rules for Effective Data Integration

Data integration involves combining and harmonizing data from diverse sources. To achieve this, it's essential to establish clear data mapping rules. These rules ensure that data from different systems can be effectively combined and used coherently. Here are some essential data mapping rules for effective data integration:

1. **Consistent Data Naming:** Ensure that data elements (e.g., columns or fields) have consistent and meaningful names across all data sources. This consistency simplifies the mapping process.

2. **Unique Identifiers:** Identify unique identifiers within each dataset that can be used to match and join records across different sources. This might include primary keys or unique combinations of attributes.

3. **Data Type Consistency:** Make sure that data types for corresponding attributes match. For example, a numerical attribute in one source should be of the same data type (e.g., integer or float) in another source.

4. **Data Transformation Rules:** Define clear rules for transforming data from one format to another. This includes specifying how categorical values are mapped, units of measurement conversions, and date format transformations.

5. **Data Validation Rules:** Establish validation rules to check the integrity of data. These rules can identify data that doesn't conform to expected patterns and may indicate errors or inconsistencies.

6. **Hierarchical Data Mapping:** If dealing with hierarchical or nested data structures, define how parent-child relationships are mapped between systems. This is common in scenarios like XML or JSON data.

7. **Data Aggregation Rules:** Specify how data should be aggregated or summarized when combining records from multiple sources. Aggregation rules may include averaging, summing, or taking the maximum or minimum values.

8. **Data Cleaning Rules:** Define rules for handling missing data, outliers, or inconsistent data. Decide whether to impute missing values, remove outliers, or flag potentially problematic records.

9. **Data Privacy and Security Rules:** Ensure that data mapping rules respect privacy and security requirements, particularly when integrating sensitive or personally identifiable information (PII).

10. **Data Versioning and Change Tracking:** Establish mechanisms to track changes in data over time. This can include versioning data sources and recording change timestamps.

11. **Data Matching and Deduplication Rules:** Define how records from different sources are matched and deduplicated to avoid data redundancy. This is crucial when integrating data from multiple sources with overlapping information.

12. **Data Transformation Logs:** Keep logs of data transformations and mappings. These logs can be useful for auditing, debugging, and ensuring transparency.

13. **Error Handling Rules:** Specify how errors during data integration, such as data type mismatches or failed transformations, should be handled. Define fallback mechanisms or alerting procedures.

14. **Data Lineage Tracking:** Implement data lineage tracking to trace the origin and history of data elements, enabling users to understand how integrated data was derived.

15. **Mapping Documentation:** Thoroughly document data mapping rules, providing data consumers with insights into how data has been integrated and transformed.

16. **Scalability Rules:** Consider how the data mapping process can scale as new data sources are added or as data volume increases.

17. **Testing and Validation Procedures:** Establish procedures for testing and validating data mappings to ensure the integrated data meets quality and accuracy standards.

18. **Regular Review and Maintenance:** Schedule periodic reviews of data mapping rules to accommodate changing data sources and requirements.

19. **User Feedback Mechanism:** Create a mechanism for users to provide feedback on the integrated data, helping to identify issues and improve the mapping process.

These rules help ensure that data integration is conducted efficiently and produces high-quality integrated datasets that can be used for analysis and decision-making. The specifics of data mapping rules may vary depending on the nature of the data and the integration requirements.

Appendix E

Results

Table E.1: Experiment 1: ML prediction results for Neagh Bann RBD: Results using MAE Validation and Test

Station name	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM	Best Model
Validation									
Aclint	0.0665	0.0667	0.0716	0.0136	0.0209	0.0103	0.0039	0.0039	GRU
Brewery Park	0.0393	0.0463	0.0452	0.0235	0.0123	0.0141	0.0088	0.0064	LSTM
Cappoge Bridge	0.3659	0.3637	0.3475	0.0546	0.0555	0.0496	0.0378	0.0376	LSTM
Charleville	0.0942	0.094	0.0923	0.012	0.0554	0.0119	0.01	0.01	GRU
Clarebane	0.0436	0.2625	0.0453	0.1012	0.03	0.018	0.011	0.011	GRU
Glaslough	0.2907	0.3112	0.2837	0.0686	0.0629	0.0265	0.0337	0.0265	RNN
Ladyswell	0.0488	0.0581	0.0541	0.0125	0.0173	0.0133	0.0078	0.0078	GRU
Mansfieldstown	0.0798	0.0798	0.0826	0.014	0.0146	0.0107	0.0074	0.007	LSTM
Moyles Mill	0.054	0.054	0.0599	0.0164	0.0246	0.0154	0.013	0.013	GRU
Port Oriel	1.1293	0.8547	0.8539	0.0315	0.0345	0.0446	0.0448	0.0315	MultiDense
Aghawoney	0.1233	0.1224	0.1233	0.0125	0.0131	0.0088	0.0063	0.0063	GRU
Anlore	0.1907	0.1949	0.2038	0.0151	0.0184	0.0067	0.007	0.0061	LSTM
Ashfield	0.0515	0.0515	0.0517	0.03134	0.0209	0.0076	0.0058	0.0055	LSTM
Ballybofey	0.3734	0.385	0.3708	0.0529	0.0452	0.0421	0.0298	0.0298	GRU
Ballyloskey	0.1287	0.2775	0.1428	0.0345	0.0487	0.0465	0.0281	0.0281	GRU
Bellahillan	0.0415	0.0417	0.0393	0.0067	0.0091	0.0033	0.0111	0.0033	RNN
Butlers Bridge	0.1025	0.1023	0.1055	0.0143	0.0246	0.0118	0.0069	0.006	LSTM
Foalies Bridge	0.0716	0.0708	0.0698	0.0166	0.0298	0.0133	0.0102	0.0102	GRU
Gartan Bridge	0.1512	0.1534	0.1467	0.0155	0.0154	0.0136	0.0114	0.0111	LSTM
Keenagh Deel Bridge	0.3979	0.4308	0.434	0.0328	0.0426	0.0236	0.0217	0.0182	LSTM
Abington	0.2105	0.02074	0.02233	0.0226	0.0254	0.0171	0.0196	0.0158	LSTM

Continued on next page

Table E.1 Results using MAE (*continued from previous page*)

Station name	**Baseline**	**Linear**	**Dense**	**MultiDense**	**CNN**	**RNN**	**GRU**	**LSTM**	**Model**
Athlacca	0.1704	0.1702	0.1777	0.0185	0.0169	0.0137	0.0122	0.0115	LSTM
Annacotty	0.217	0.2247	0.226	0.0215	0.0247	0.015	0.0155	0.015	RNN
Ballinamore	0.2021	0.2031	0.2133	0.0438	0.0197	0.0153	0.0078	0.0051	LSTM
Banagher	0.0385	0.0383	0.0412	0.0123	0.0181	0.006	0.0041	0.0033	LSTM
Ballinasloe Town	0.0428	0.0438	0.0462	0.0079	0.0647	0.0093	0.0117	0.0073	LSTM
Castleroberts	0.1343	0.1347	0.1224	0.0206	0.0202	0.0112	0.011	0.0095	LSTM
Clarecastle Barrage	1.0161	0.7925	0.7918	0.0461	0.0554	0.0713	0.0669	0.0461	MultiDense
Clarianna	0.1537	0.1492	0.1608	0.0261	0.0392	0.017	0.0163	0.0163	GRU
Gourdeen	0.1423	0.139	0.1439	0.0403	0.0352	0.016	0.0147	0.0147	GRU
Aasleagh Bridge	0.3859	0.4034	0.4082	0.0637	0.0638	0.0511	0.0489	0.0479	LSTM
Ballina	0.4493	0.5083	0.5045	0.0611	0.0906	0.0898	0.0776	0.0611	MultiDense
Ballylahan	0.24	0.2424	0.2351	0.0115	0.0144	0.0096	0.0073	0.0064	LSTM
Clarinbridge	0.0899	0.0896	0.0888	0.0134	0.0233	0.0121	0.0113	0.0092	LSTM
Dangan	0.0475	0.0477	0.049	0.0205	0.0189	0.01	0.0088	0.0083	LSTM
Kilcolgan	0.1789	0.2022	0.2041	0.0481	0.0504	0.0594	0.0594	0.0481	MultiDense
Pontoon	0.0475	0.0472	0.0454	0.0133	0.0236	0.0115	0.0083	0.0083	GRU
Riverside Close	0.3643	0.5785	0.336	0.0955	0.1386	0.0583	0.0556	0.0556	GRU
Terryland	0.5826	0.6662	0.6594	0.1395	0.1742	0.1566	0.113	0.113	GRU
Wolfe Tone Bridge	0.7232	0.6677	0.635	0.0878	0.0811	0.1169	0.1125	0.0811	CNN
Anglesea Road	0.3673	0.341	0.3321	0.0791	0.1248	0.0637	0.0585	0.0536	LSTM
Bluebell	0.1502	0.1502	0.1526	0.0485	0.0663	0.0464	0.0441	0.0439	LSTM
Cherry Wood	0.1514	0.3185	0.1525	0.0429	0.0837	0.0437	0.0356	0.0356	GRU
Fyanstown	0.151	0.1509	0.1572	0.0153	0.0217	0.0112	0.0086	0.0086	GRU
Giles Quay	1.0408	1.0103	0.9583	0.1314	0.1902	0.1562	0.1512	0.1314	Multi_Dense
Killyon	0.0735	0.0704	0.0732	0.0162	0.0157	0.0103	0.014	0.0064	LSTM
Leixlip	0.1379	0.1341	0.1268	0.0156	0.0171	0.0179	0.012	0.011	LSTM
Mornington	0.4241	0.4298	0.4015	0.0992	0.0906	0.1037	0.1004	0.0906	CNN
Navan Weir	0.0863	0.0869	0.0843	0.0164	0.0237	0.0166	0.012	0.0097	LSTM
Trim	0.0832	0.082	0.0807	0.0136	0.0209	0.0128	0.0035	0.0035	GRU
Adelphi Quay	1.1466	0.8738	0.8726	0.0426	0.0454	0.0552	0.0552	0.0426	MultiDense
Brownsbarn	0.1215	0.1252	0.137	0.0208	0.0163	0.0232	0.0092	0.0086	LSTM
Cahir Park	0.1031	0.1041	0.1024	0.0307	0.0137	0.0103	0.0097	0.0074	LSTM
Dundalk Port	1.0689	1.017	0.9601	0.1128	0.1156	0.1641	0.2132	0.1128	Multi_Dense
Joyces Lane	0.1334	0.1325	0.1417	0.0168	0.0236	0.0195	0.0137	0.0137	GRU
Killardry	0.2254	0.2414	0.2412	0.0189	0.0189	0.0098	0.0094	0.0094	GRU
Levitstown	0.1222	0.1241	0.1231	0.0203	0.0133	0.0095	0.0094	0.0085	LSTM
McMahons Bridge	0.0997	0.1016	0.1029	0.0071	0.0082	0.011	0.0039	0.0039	GRU
Rathangan	0.1287	0.125	0.1268	0.0296	0.0108	0.0114	0.0104	0.0101	LSTM

Continued on next page

Table E.1 Results using MAE (*continued from previous page*)

Station name	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM	Model
Tullowbeg	0.3431	0.3587	0.3197	0.0765	0.0468	0.0309	0.0321	0.0272	LSTM
Athea D/S	0.502	0.6446	0.4722	0.1011	0.1964	0.0508	0.0458	0.0458	GRU
Athea U/S	0.4552	0.4212	0.4306	co0.0629	0.0587	0.0486	0.0494	0.0486	RNN
Ballydahin	0.2313	0.2489	0.255	0.0502	0.0441	0.0212	0.0176	0.0166	LSTM
Ballyduff	0.1788	0.1882	0.1811	0.0192	0.011	0.0092	0.0079	0.007	LSTM
Coolmuckey Br	0.2211	0.2064	0.2003	0.0502	0.0947	0.0392	0.0252	0.0252	GRU
Flesk Bridge	0.3698	0.3856	0.3926	0.0487	0.0272	0.226	0.0188	0.0188	GRU
Glandalane	0.1768	0.1797	0.181	0.0482	0.0214	0.0159	0.0131	0.0131	GRU
Morris's Bridge	0.1588	0.2261	0.1928	0.0218	0.033	0.0149	0.0293	0.0149	RNN
Riverville	0.2909	0.2908	0.2798	0.0345	0.0291	0.017	0.0148	0.0136	LSTM
Waterworks Weir	0.1427	0.217	0.2318	0.0487	0.0546	0.0404	0.0399	0.0368	LSTM
Test									
Aclint	0.0455	0.0463	0.048	0.0098	0.028	0.0085	0.0026	0.0077	GRU
Brewery Park	0.0858	0.0886	0.0878	0.02	0.0116	0.0134	0.0085	0.0077	LSTM
Cappoge Bridge	0.362	0.3932	0.3899	0.0583	0.0593	0.055	0.0476	0.0462	LSTM
Charleville	0.0627	0.0673	0.0639	0.009	0.0431	0.0111	0.0082	0.0125	GRU
Clarebane	0.0282	0.1907	0.0298	0.0763	0.0191	0.013	0.0074	0.0134	GRU
Glaslough	0.2644	0.3107	0.2584	0.0763	0.063	0.026	0.0324	0.0401	RNN
Ladyswell	0.056	0.0646	0.0607	0.0108	0.0178	0.0122	0.0076	0.0141	GRU
Mansfieldstown	0.0623	0.0624	0.0674	0.0119	0.0118	0.0109	0.0073	0.0072	LSTM
Moyles Mill	0.035	0.0357	0.0389	0.0113	0.023	0.0116	0.0091	0.0117	GRU
Port Oriel	1.1144	0.8433	0.843	0.0341	0.0378	0.0439	0.0447	0.0453	Multi_Dense
Aghawoney	0.0929	0.0924	0.0917	0.0122	0.0133	0.008	0.0052	0.0054	GRU
Anlore	0.1612	0.1707	0.1724	0.0165	0.0217	0.0076	0.0085	0.0072	LSTM
Ashfield	0.0401	0.0405	0.0411	0.0148	0.0233	0.0067	0.0066	0.0054	LSTM
Ballybofey	0.3602	0.3817	0.3638	0.0537	0.046	0.0431	0.0308	0.0343	GRU
Ballyloskey	0.3324	0.3182	0.3293	0.0803	0.0717	0.0665	0.0561	0.0582	GRU
Bellahillan	0.0357	0.0366	0.0347	0.0092	0.011	0.0028	0.0072	0.0041	RNN
Butlers Bridge	0.088	0.0882	0.0911	0.0201	0.021	0.0111	0.0068	0.0061	LSTM
Foalies Bridge	0.061	0.0609	0.0599	0.0125	0.0251	0.0113	0.0082	0.0085	GRU
Gartan Bridge	0.129	0.1324	0.1247	0.0164	0.0156	0.0128	0.0109	0.104	GRU
Keenagh Deel Bridge	0.366	0.3958	0.3957	0.0339	0.0436	0.024	0.023	0.0193	LSTM
Abington	0.1265	0.1387	0.148	0.0125	0.0175	0.0097	0.0136	0.0089	LSTM
Athlacca	0.083	0.0982	0.0923	0.01	0.0088	0.0084	0.0061	0.0064	GRU
Annacotty	0.1088	0.1298	0.1248	0.0175	0.0155	0.0097	0.0101	0.01	RNN
Ballinamore	0.1306	0.1399	0.1404	0.0324	0.0112	0.0152	0.0066	0.0041	LSTM
Banagher	0.0256	0.0256	0.0283	0.0068	0.0073	0.0038	0.003	0.0025	LSTM
Ballinasloe Town	0.0561	0.0566	0.0577	0.0097	0.0203	0.0089	0.0082	0.0087	GRU

Continued on next page

Table E.1 Results using MAE (*continued from previous page*)

Station name	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM	Model
Castleroberts	0.0719	0.0819	0.0736	0.0205	0.0223	0.0073	0.0075	0.0061	LSTM
Clarecastle Barrage	1.1257	0.8776	0.8785	0.0484	0.0593	0.078	0.0737	0.0721	Multi_Dense
Clarianna	0.0547	0.0733	0.0616	0.0159	0.0339	0.0127	0.0134	0.0117	LSTM
Gourdeen	0.0628	0.0799	0.0815	0.0175	0.0191	0.0126	0.0111	0.018	GRU
Aasleagh Bridge	0.3489	0.3728	0.3644	0.0691	0.0633	0.0523	0.0514	0.0505	LSTM
Ballina	0.5205	0.5127	0.5055	0.0585	0.0826	0.0962	0.0522	0.0898	GRU
Ballylahan	0.2101	0.2208	0.2088	0.0104	0.0129	0.0092	0.0065	0.0031	LSTM
Clarinbridge	0.0694	0.0698	0.20691	0.0108	0.0297	0.0127	0.0133	0.0094	LSTM
Dangan	0.0363	0.0366	0.0377	0.0127	0.0134	0.0081	0.007	0.0067	LSTM
Kilcolgan	0.1994	0.2215	0.2205	0.0618	0.0477	0.0715	0.0702	0.0667	CNN
Pontoon	0.0363	0.0364	0.0357	0.0093	0.0123	0.0091	0.0072	0.0073	GRU
Riverside Close	0.31	0.5008	0.2858	0.0853	0.1066	0.0468	0.0448	0.0488	GRU
Terryland	0.8598	0.8265	0.8264	0.171	0.1955	0.2095	0.1492	0.1652	GRU
Wolfe Tone Bridge	0.8965	0.8318	0.8108	0.0849	0.0778	0.1157	0.1109	0.1135	CNN
Anglesea Road	1.1624	1.0183	0.9891	0.2054	0.2905	0.1917	0.1389	0.1298	LSTM
Bluebell	0.1529	0.1548	0.1575	0.0498	0.0642	0.0504	0.0483	0.0477	LSTM
Cherry Wood	0.2495	0.3754	0.2178	0.0555	0.0958	0.05	0.0449	0.0576	GRU
Fyanstown	0.1105	0.1278	0.1151	0.0147	0.018	0.0094	0.0058	0.0089	GRU
Giles Quay	1.0846	1.0608	1.1149	0.2652	0.3113	0.2475	0.2519	0.282	RNN
Killyon	0.0474	0.0494	0.048	0.0082	0.0081	0.0104	0.008	0.0063	LSTM
Leixlip	0.3153	0.3113	0.3071	0.1006	0.0927	0.0813	0.0805	0.0792	LSTM
Mornington	0.2728	0.3047	0.2712	0.0789	0.0656	0.0763	0.0872	0.09	CNN
Navan Weir	0.0634	0.0738	0.0625	0.0321	0.0166	0.0215	0.0151	0.0094	LSTM
Trim	0.0473	0.0519	0.0491	0.0055	0.0213	0.0157	0.0034	0.0039	GRU
Adelphi Quay	1.1395	0.867	0.8665	0.0396	0.0418	0.0519	0.052	0.0518	Multi_Dense
Brownsbarn	0.0584	0.0717	0.064	0.0117	0.0086	0.0179	0.0052	0.0049	LSTM
Cahir Park	0.0716	0.0808	0.0751	0.0144	0.0159	0.0092	0.0061	0.0054	LSTM
Dundalk Port	0.8418	0.9117	0.8412	0.1602	0.1404	0.1996	0.276	0.2598	CNN
Joyces Lane	0.0899	0.0996	0.1107	0.0098	0.0134	0.0153	0.0136	0.0091	LSTM
Killardry	0.1452	0.1856	0.1658	0.0125	0.0162	0.0074	0.007	0.0081	GRU
Levitstown	0.0454	0.0527	0.0459	0.0248	0.0118	0.0068	0.006	0.0048	LSTM
McMahons Bridge	0.0852	0.0894	0.0919	0.0062	0.0119	0.0096	0.0037	0.0039	GRU
Rathangan	0.0708	0.0701	0.0732	0.0146	0.0067	0.0074	0.0076	0.0076	CNN
Tullowbeg	0.2495	0.3173	0.2383	0.0546	0.035	0.0294	0.0288	0.0232	LSTM
Athea D/S	0.3243	0.4759	0.308	0.0799	0.1525	0.0364	0.0327	0.0378	GRU
Athea U/S	0.2998	0.2801	0.2861	0.0454	0.0411	0.0408	0.0431	0.0433	RNN
Ballydahin	0.1429	0.1776	0.1715	0.0535	0.0279	0.0157	0.0119	0.0095	LSTM
Ballyduff	0.0998	0.1253	0.1018	0.0186	0.0072	0.0074	0.0056	0.005	LSTM

Continued on next page

Table E.1 Results using MAE (*continued from previous page*)

Station name	Baseline	Linear	Dense	MultiDense	CNN	RNN	GRU	LSTM	Model
Coolmuckey Br	0.1096	0.1081	0.1147	0.0401	0.0632	0.027	0.019	0.0214	GRU
Flesk Bridge	0.2212	0.2635	0.2581	0.0409	0.0201	0.0161	0.0125	0.0128	GRU
Glandalane	0.1135	0.1275	0.1183	0.0237	0.0143	0.0125	0.008	0.0083	GRU
Morris's Bridge	0.2191	0.2066	0.2376	0.0344	0.0412	0.0207	0.043	0.0212	RNN
Riverville	0.196	0.2146	0.1977	0.0276	0.0206	0.0148	0.0122	0.0109	LSTM
Waterworks Weir	0.3886	0.422	0.4425	0.0734	0.0608	0.0707	0.0665	0.0638	CNN

Table E.2: Experiment 2: ML prediction results for Neagh Bann RBD: Results using MAE Validation and Test

Station	Model	Sample Size	Execution Time	Validation Loss	Validation MSE	Validation MAE	Test Loss	Test MSE	Test MAE	Model Complexity	Learning Rate	Number of Layers	Number of Dense Layers	Number of Features	Skewness	Kurtosis
Aclint	Baseline	43305	20.75081277	5.67e-05	5.67e-05	0.003824759	8.56e-05	8.56e-05	0.005250855	0	0.001	0	0	1	0.99304277	0.099293655
Aclint	Linear	43305	28.44120145	0.000481884	0.000481884	0.018081751	0.000819241	0.000819241	0.021876695	2	0.001	1	1	1	0.99304277	0.099293655
Aclint	Dense	43305	36.33077049	7.25e-05	7.25e-05	0.005983175	0.000107227	0.000107227	0.007656832	97	0.001	2	2	1	0.99304277	0.099293655
Aclint	Conv	43305	31.88564372	6.64e-05	6.64e-05	0.005119384	9.38e-05	9.38e-05	0.006224512	1153	0.001	3	2	1	0.99304277	0.099293655
Aclint	RNN	43305	96.05474448	4.74e-05	4.74e-05	0.0038626	7.46e-05	7.46e-05	0.004712064	4289	0.001	2	1	1	0.99304277	0.099293655
Aclint	GRU	43305	190.7939789	2.57e-05	2.57e-05	0.002670779	3.57e-05	3.57e-05	0.003135516	12929	0.001	2	1	1	0.99304277	0.099293655
Aclint	LSTM	43305	157.2062078	4.3e-05	4.3e-05	0.003829513	5.69e-05	5.69e-05	0.004185662	16961	0.001	2	1	1	0.99304277	0.099293655
Aclint	LSTM Attention	43305	175.6318953	0.000385352	0.000385352	0.014306233	0.000563326	0.000563326	0.017681567	16961	0.001	3	1	1	0.99304277	0.099293655
Aclint	Autoencoder	43305	35.73760653	0.48196739	0.48196739	0.599996805	0.465708017	0.465707988	0.530306339	97	0.001	2	0	1	0.99304277	0.099293655
Aclint	Transformer	43305	317.260154	0.001197908	0.001197908	0.027872432	0.002252203	0.002252203	0.034853574	141825	0.001	2	1	1	0.99304277	0.099293655
Aclint	Variational Autoencoder	43305	44.05313993	0.481011987	0.481011987	0.594296515	0.464190543	0.464190334	0.518407285	1283	0.001	4	2	1	0.99304277	0.099293655
Burley	Baseline	43535	32.6529994	0.000339604	0.000339604	0.005943309	0.001229142	0.001229142	0.012585501	0	0.001	0	0	1	1.253711554	1.085049752
Burley	Linear	43535	36.40879488	0.002140671	0.002140671	0.038675021	0.00352739	0.00352739	0.041135874	2	0.001	1	1	1	1.253711554	1.085049752
Burley	Dense	43535	41.76133347	0.000341983	0.000341983	0.006157488	0.001230229	0.00123023	0.01295893	97	0.001	2	2	1	1.253711554	1.085049752
Burley	Conv	43535	22.93091559	0.000354437	0.000354437	0.005233133	0.001282477	0.001282477	0.010875169	1153	0.001	3	2	1	1.253711554	1.085049752
Burley	RNN	43535	72.11123419	0.000191218	0.000191218	0.006976678	0.000558594	0.000558594	0.009383216	4289	0.001	2	1	1	1.253711554	1.085049752
Burley	GRU	43535	157.6550214	0.000161397	0.000161397	0.009737604	0.000361502	0.000361502	0.011043481	12929	0.001	2	1	1	1.253711554	1.085049752
Burley	LSTM	43535	156.1097853	0.000180618	0.000180618	0.009922829	0.000340997	0.000340997	0.010266489	16961	0.001	2	1	1	1.253711554	1.085049752
Burley	LSTM Attention	43535	191.1764503	0.000159438	0.000159438	0.006478794	0.000412899	0.000412899	0.008916059	16961	0.001	3	1	1	1.253711554	1.085049752
Burley	Autoencoder	43535	36.75624537	0.674104452	0.674104095	0.732620776	0.55982393	0.55982399	0.59515661	97	0.001	2	0	1	1.253711554	1.085049752
Burley	Transformer	43535	232.7040086	0.001716456	0.001716456	0.039879858	0.001608527	0.001608527	0.036589049	141825	0.001	2	1	1	1.253711554	1.085049752
Burley	Variational Autoencoder	43535	46.72229743	0.673325181	0.673325002	0.728271186	0.558377028	0.558377087	0.584118962	1283	0.001	4	2	1	1.253711554	1.085049752
Mansfieldstown	Baseline	43539	33.16723752	0.000136089	0.000136089	0.007652212	0.000148214	0.000148215	0.004932125	0	0.001	0	0	1	0.877206669	-0.041430935
Mansfieldstown	Linear	43539	24.58413339	0.000136248	0.000136248	0.007678926	0.000148598	0.000148598	0.004974347	2	0.001	1	1	1	0.877206669	-0.041430935
Mansfieldstown	Dense	43539	41.54216337	0.000141712	0.000141712	0.007933453	0.000162461	0.000162461	0.00567618	97	0.001	2	2	1	0.877206669	-0.041430935
Mansfieldstown	Conv	43539	36.1080296	0.000142833	0.000142833	0.007975983	0.000157311	0.000157311	0.005856662	1153	0.001	3	2	1	0.877206669	-0.041430935
Mansfieldstown	RNN	43539	94.08911061	0.000144177	0.000144177	0.008033888	0.000155226	0.000155226	0.005825376	4289	0.001	2	1	1	0.877206669	-0.041430935
Mansfieldstown	GRU	43539	177.1435242	0.000125692	0.000125692	0.007055377	0.00015178	0.00015178	0.004627496	12929	0.001	2	1	1	0.877206669	-0.041430935
Mansfieldstown	LSTM	43539	168.032867	0.000150744	0.000150744	0.007770944	0.00018155	0.00018155	0.005118376	16961	0.001	2	1	1	0.877206669	-0.041430935
Mansfieldstown	LSTM Attention	43539	153.8194001	0.000525022	0.000525022	0.015321888	0.000480399	0.000480399	0.011820149	16961	0.001	3	1	1	0.877206669	-0.041430935
Mansfieldstown	Autoencoder	43539	47.01303506	0.622719467	0.622719467	0.71318239	0.611083269	0.611083329	0.661874294	97	0.001	2	0	1	0.877206669	-0.041430935
Mansfieldstown	Transformer	43539	366.5220332	0.0006926	0.0006926	0.019152004	0.001601163	0.001601163	0.027950013	141825	0.001	2	1	1	0.877206669	-0.041430935
Mansfieldstown	Variational Autoencoder	43539	47.57119393	0.622224331	0.622224629	0.710105181	0.611867547	0.611867726	0.66011709	1283	0.001	4	2	1	0.877206669	-0.041430935
Brewery Park	Baseline	29704	23.71990824	0.000535329	0.000535329	0.007721613	0.000792539	0.000792539	0.011605559	0	0.001	0	0	1	0.886812599	-0.114431344
Brewery Park	Linear	29704	24.64914775	0.000535649	0.000535649	0.007951909	0.000791911	0.000791911	0.011771226	2	0.001	1	1	1	0.886812599	-0.114431344
Brewery Park	Dense	29704	25.67712665	0.00053815	0.00053815	0.008265068	0.000795977	0.000795977	0.012167377	97	0.001	2	2	1	0.886812599	-0.114431344
Brewery Park	Conv	29704	24.17311978	0.000547121	0.000547121	0.00847131	0.000799785	0.000799785	0.013211026	1153	0.001	3	2	1	0.886812599	-0.114431344
Brewery Park	RNN	29704	47.74200177	0.00059428	0.00059428	0.012042418	0.000898602	0.000898602	0.013968144	4289	0.001	2	1	1	0.886812599	-0.114431344
Brewery Park	GRU	29704	94.27042389	0.000456839	0.000456839	0.007538928	0.000507226	0.000507226	0.010255939	12929	0.001	2	1	1	0.886812599	-0.114431344
Brewery Park	LSTM	29704	101.8199942	0.00049508	0.00049508	0.007775528	0.000575712	0.000575712	0.010127707	16961	0.001	2	1	1	0.886812599	-0.114431344
Brewery Park	LSTM Attention	29704	101.4415431	0.00089807	0.00089807	0.013616472	0.002140802	0.002140802	0.021583838	16961	0.001	3	1	1	0.886812599	-0.114431344
Brewery Park	Autoencoder	29704	31.24470353	0.413227916	0.413227856	0.552525997	0.619298398	0.619298518	0.53672421	97	0.001	2	0	1	0.886812599	-0.114431344
Brewery Park	Transformer	29704	292.7304616	0.001179589	0.001179589	0.027222352	0.002051272	0.002051272	0.037418462	141825	0.001	2	1	1	0.886812599	-0.114431344
Brewery Park	Variational Autoencoder	29704	42.07607698	0.429136425	0.429136485	0.561150908	0.726171911	0.726171911	0.576985717	1283	0.001	4	2	1	0.886812599	-0.114431344
Ladyswell	Baseline	34695	29.98807502	0.00029265	0.00029265	0.006305945	0.000544692	0.000544692	0.007939289	0	0.001	0	0	1	0.966726542	0.451453799
Ladyswell	Linear	34695	25.81446815	0.000292706	0.000292706	0.006302151	0.000545984	0.000545984	0.007957851	2	0.001	1	1	1	0.966726542	0.451453799
Ladyswell	Dense	34695	32.45640993	0.000298417	0.000298416	0.006913451	0.000551978	0.000551978	0.008366831	97	0.001	2	2	1	0.966726542	0.451453799
Ladyswell	Conv	34695	32.63690281	0.00029899	0.00029899	0.006492844	0.000561829	0.000561829	0.00840381	1153	0.001	3	2	1	0.966726542	0.451453799
Ladyswell	RNN	34695	53.75104165	0.000321055	0.000321055	0.008347066	0.000568015	0.000568014	0.009681939	4289	0.001	2	1	1	0.966726542	0.451453799
Ladyswell	GRU	34695	102.8241737	0.000303173	0.000303173	0.006661152	0.000546445	0.000546445	0.008030664	12929	0.001	2	1	1	0.966726542	0.451453799

Table E.2 ML prediction results for Neagh Bann RBD (*continued from previous page*)

Station	Model	Sample Size	Execution Time	Validation Loss	Validation MSE	Validation MAE	Test Loss	Test MSE	Test MAE	Model Complexity	Learning Rate	Number of Layers	Number of Dense Layers	Number of Features	Skewness	Kurtosis
Ladyswell	LSTM	34695	104.2000492	0.000281522	0.000281522	0.006641378	0.000501076	0.000501076	0.007810653	16961	0.001	2	1	1	0.966726542	0.451453799
Ladyswell	LSTM Attention	34695	147.7337027	0.000552984	0.000552983	0.013262113	0.000930588	0.000930588	0.014456107	16961	0.001	3	1	1	0.966726542	0.451453799
Ladyswell	Autoencoder	34695	30.99494863	0.519986391	0.519986451	0.625047743	0.302615255	0.302615196	0.410788208	97	0.001	2	0	1	0.966726542	0.451453799
Ladyswell	Transformer	34695	367.0310214	0.002342865	0.002342865	0.042944413	0.001693647	0.001693648	0.03303365	141825	0.001	2	1	1	0.966726542	0.451453799
Ladyswell	Variational Autoencoder	34695	43.8871727	0.518456101	0.518456042	0.615827203	0.300388008	0.300387979	0.394594371	1283	0.001	4	2	1	0.966726542	0.451453799
Port Oriel	Baseline	42027	33.33167386	0.201844081	0.201844186	0.387722999	0.181482613	0.181482553	0.375898123	0	0.001	0	0	1	-0.070104305	-1.167331052
Port Oriel	Linear	42027	35.46929479	0.190363511	0.19036369	0.380873203	0.170684338	0.170684382	0.36522916	2	0.001	1	1	1	-0.070104305	-1.167331052
Port Oriel	Dense	42027	38.61936021	0.189692006	0.189691946	0.379525393	0.170259759	0.170259729	0.364849299	97	0.001	2	2	1	-0.070104305	-1.167331052
Port Oriel	Conv	42027	37.71259856	0.189714506	0.189714506	0.379537553	0.170297429	0.170297489	0.36485374	1153	0.001	3	2	1	-0.070104305	-1.167331052
Port Oriel	RNN	42027	89.22507644	0.013431472	0.01343147	0.048657797	0.010190305	0.010190304	0.047067694	4289	0.001	2	1	1	-0.070104305	-1.167331052
Port Oriel	GRU	42027	137.9181316	0.013488624	0.013488625	0.04849913	0.010047764	0.010047762	0.046114303	12929	0.001	2	1	1	-0.070104305	-1.167331052
Port Oriel	LSTM	42027	154.1841776	0.013564277	0.013564276	0.046969332	0.010026383	0.010026379	0.044245392	16961	0.001	2	1	1	-0.070104305	-1.167331052
Port Oriel	LSTM Attention	42027	174.8574541	0.00599651	0.005996511	0.03594083	0.002206662	0.002206662	0.028886255	16961	0.001	3	1	1	-0.070104305	-1.167331052
Port Oriel	Autoencoder	42027	42.16922808	0.485760391	0.48576051	0.568965852	0.428613514	0.428613693	0.546203375	97	0.001	2	0	1	-0.070104305	-1.167331052
Port Oriel	Transformer	42027	393.0877821	0.002016096	0.002016096	0.022409286	0.001185233	0.001185233	0.021574201	141825	0.001	2	1	1	-0.070104305	-1.167331052
Port Oriel	Variational Autoencoder	42027	35.57532001	0.484951019	0.484951198	0.569397509	0.427671909	0.427671969	0.545612931	1283	0.001	4	2	1	-0.070104305	-1.167331052
Carlingford	Baseline	12742	6.265403271	0.257861495	0.257861495	0.452357888	0.247589141	0.247589082	0.446868539	0	0.001	0	0	1	-0.05573451	-1.150803599
Carlingford	Linear	12742	16.75966287	0.241920188	0.241920173	0.436284244	0.232862562	0.232862577	0.430553764	2	0.001	1	1	1	-0.05573451	-1.150803599
Carlingford	Dense	12742	12.63864565	0.241719171	0.241719112	0.436051458	0.232962817	0.232962832	0.430866867	97	0.001	2	2	1	-0.05573451	-1.150803599
Carlingford	Conv	12742	11.38276935	0.241645753	0.241645753	0.43579638	0.232686237	0.232686222	0.430499345	1153	0.001	3	2	1	-0.05573451	-1.150803599
Carlingford	RNN	12742	26.83644676	0.020642838	0.020642828	0.077205546	0.022034124	0.022034122	0.085745767	4289	0.001	2	1	1	-0.05573451	-1.150803599
Carlingford	GRU	12742	53.71510983	0.014088963	0.014088962	0.060534488	0.014621997	0.014621997	0.065104559	12929	0.001	2	1	1	-0.05573451	-1.150803599
Carlingford	LSTM	12742	46.92060184	0.01456914	0.014569135	0.062155254	0.015051666	0.015051663	0.067294054	16961	0.001	2	1	1	-0.05573451	-1.150803599
Carlingford	LSTM Attention	12742	48.20368266	0.004463012	0.004463012	0.04419817	0.006611352	0.006611353	0.052634839	16961	0.001	3	1	1	-0.05573451	-1.150803599
Carlingford	Autoencoder	12742	14.02024293	0.606231809	0.606231809	0.651987731	0.529120624	0.529120564	0.622214496	97	0.001	2	0	1	-0.05573451	-1.150803599
Carlingford	Transformer	12742	152.1122644	0.004930564	0.004930565	0.056622274	0.004994173	0.004994174	0.057077918	141825	0.001	2	1	1	-0.05573451	-1.150803599
Carlingford	Variational Autoencoder	12742	19.4215467	0.600597918	0.600597918	0.649502993	0.524113595	0.524113595	0.620079994	1283	0.001	4	2	1	-0.05573451	-1.150803599
Clarebane	Baseline	43056	21.78466392	8.29e-05	8.29e-05	0.00309814	4.86e-05	4.86e-05	0.003305455	0	0.001	0	0	1	1.058022868	29.3213842
Clarebane	Linear	43056	46.60821271	0.004963579	0.00496358	0.058773391	0.007473165	0.007473165	0.073304616	2	0.001	1	1	1	1.058022868	29.3213842
Clarebane	Dense	43056	36.92939758	0.000543619	0.000543619	0.015274812	0.000538173	0.000538173	0.017133802	97	0.001	2	2	1	1.058022868	29.3213842
Clarebane	Conv	43056	48.34004378	0.000176891	0.000176891	0.007869808	8.95e-05	8.95e-05	0.006911696	1153	0.001	3	2	1	1.058022868	29.3213842
Clarebane	RNN	43056	70.20056558	0.001564424	0.001564423	0.031662706	0.001889774	0.001889774	0.033835221	4289	0.001	2	1	1	1.058022868	29.3213842
Clarebane	GRU	43056	119.5320303	0.000510621	0.000510621	0.014254931	0.000698378	0.000698378	0.017772449	12929	0.001	2	1	1	1.058022868	29.3213842
Clarebane	LSTM	43056	178.5083411	0.000459681	0.000459681	0.012635123	0.000496562	0.000496562	0.014837924	16961	0.001	2	1	1	1.058022868	29.3213842
Clarebane	LSTM Attention	43056	84.1429646	0.00174249	0.00174249	0.032703143	0.001973307	0.001973307	0.035130404	16961	0.001	3	1	1	1.058022868	29.3213842
Clarebane	Autoencoder	43056	42.4030757	0.135049894	0.135049954	0.298473954	0.149043292	0.149043247	0.320153356	97	0.001	2	0	1	1.058022868	29.3213842
Clarebane	Transformer	43056	385.1858428	0.020875555	0.020875547	0.137075469	0.022443362	0.02244336	0.140797302	141825	0.001	2	1	1	1.058022868	29.3213842
Clarebane	Variational Autoencoder	43056	44.64393306	0.133395299	0.13339527	0.287773937	0.147015661	0.147015646	0.309861302	1283	0.001	4	2	1	1.058022868	29.3213842
Glaslough	Baseline	43587	29.57352066	0.008692763	0.00869276	0.028891882	0.025929488	0.025929485	0.043238092	0	0.001	0	0	1	3.188205448	14.53446235
Glaslough	Linear	43587	38.58968091	0.011314793	0.011314787	0.046871245	0.028917683	0.028917689	0.061369814	2	0.001	1	1	1	3.188205448	14.53446235
Glaslough	Dense	43587	34.47345662	0.008750587	0.008750588	0.028557425	0.025827929	0.025827926	0.043736015	97	0.001	2	2	1	3.188205448	14.53446235
Glaslough	Conv	43587	35.17149425	0.009017178	0.009017176	0.029073583	0.026087768	0.026087767	0.044603679	1153	0.001	3	2	1	3.188205448	14.53446235
Glaslough	RNN	43587	98.53522635	0.004800214	0.004800214	0.022851095	0.013076267	0.013076266	0.030725515	4289	0.001	2	1	1	3.188205448	14.53446235
Glaslough	GRU	43587	157.6609581	0.003717867	0.003717867	0.017945386	0.011420581	0.011420583	0.025574176	12929	0.001	2	1	1	3.188205448	14.53446235
Glaslough	LSTM	43587	172.9839382	0.003917307	0.003917309	0.02083662	0.011319174	0.011319175	0.027403736	16961	0.001	2	1	1	3.188205448	14.53446235
Glaslough	LSTM Attention	43587	146.3242633	0.000981499	0.000981499	0.014354339	0.002746611	0.002746612	0.017998224	16961	0.001	3	1	1	3.188205448	14.53446235
Glaslough	Autoencoder	43587	44.7206881	0.661308229	0.66130805	0.53752923	0.9584077	0.9584077	0.655721366	97	0.001	2	0	1	3.188205448	14.53446235
Glaslough	Transformer	43587	398.1736891	0.003301739	0.003301739	0.042772569	0.005876924	0.005876923	0.051206287	141825	0.001	2	1	1	3.188205448	14.53446235
Glaslough	Variational Autoencoder	43587	48.60704851	0.660084307	0.660084307	0.53114742	0.957161248	0.957161486	0.64930743	1283	0.001	4	2	1	3.188205448	14.53446235

Table E.2 ML prediction results for Neagh Bann RBD (*continued from previous page*)

Station	Model	Sample Size	Execution Time	Validation Loss	Validation MSE	Validation MAE	Test Loss	Test MSE	Test MAE	Model Complexity	Learning Rate	Number of Layers	Number of Dense Layers	Number of Features	Skewness	Kurtosis
Charleville	Baseline	43478	32.81112838	0.00030297	0.000302969	0.00809729	0.000540026	0.000540026	0.012119969	0	0.001	0	0	1	0.758117161	0.179997231
Charleville	Linear	43478	31.39916253	0.000303027	0.000303027	0.008108167	0.000541581	0.000541581	0.012127589	2	0.001	1	1	1	0.758117161	0.179997231
Charleville	Dense	43478	39.72204757	0.000304288	0.000304288	0.008245775	0.000543228	0.000543228	0.012181607	97	0.001	2	2	1	0.758117161	0.179997231
Charleville	Conv	43478	25.91568017	0.000316855	0.000316855	0.008525968	0.000570876	0.000570876	0.012403313	1153	0.001	3	2	1	0.758117161	0.179997231
Charleville	RNN	43478	85.59224939	0.000375349	0.000375349	0.010957185	0.00053976	0.000539759	0.014224254	4289	0.001	2	1	1	0.758117161	0.179997231
Charleville	GRU	43478	190.2062664	0.000291492	0.000291492	0.009736422	0.000451415	0.000451416	0.013100673	12929	0.001	2	1	1	0.758117161	0.179997231
Charleville	LSTM	43478	174.0981503	0.00027474	0.00027474	0.008531641	0.00041757	0.00041757	0.011542085	16961	0.001	2	1	1	0.758117161	0.179997231
Charleville	LSTM Attention	43478	191.8985741	0.000200566	0.000200566	0.007262905	0.00036181	0.00036181	0.009890033	16961	0.001	3	1	1	0.758117161	0.179997231
Charleville	Autoencoder	43478	53.46545243	0.700843394	0.700843155	0.68449831	0.661876738	0.661876619	0.634236991	97	0.001	2	0	1	0.758117161	0.179997231
Charleville	Transformer	43478	312.2117908	0.002759612	0.002759613	0.045547303	0.003607731	0.003607731	0.051549044	141825	0.001	2	1	1	0.758117161	0.179997231
Charleville	Variational Autoencoder	43478	44.59109354	0.699507833	0.699507773	0.675406575	0.6603508	0.660350621	0.623724639	1283	0.001	4	2	1	0.758117161	0.179997231
Dundalk Port	Baseline	40993	20.38635063	0.226469979	0.226470008	0.33771652	0.261600286	0.261600286	0.371932656	0	0.001	0	0	1	0.564556838	-1.068826998
Dundalk Port	Linear	40993	41.38325214	0.209302619	0.209302574	0.348151028	0.243017986	0.243017972	0.384632707	2	0.001	1	1	1	0.564556838	-1.068826998
Dundalk Port	Dense	40993	35.8876524	0.208666965	0.208666906	0.35741809	0.242545798	0.242545739	0.396284878	97	0.001	2	2	1	0.564556838	-1.068826998
Dundalk Port	Conv	40993	26.84384203	0.209893808	0.209893808	0.358898789	0.244568095	0.24456811	0.407453984	1153	0.001	3	2	1	0.564556838	-1.068826998
Dundalk Port	RNN	40993	55.29249048	0.055218995	0.055219017	0.159706548	0.037323441	0.037323441	0.130263925	4289	0.001	2	1	1	0.564556838	-1.068826998
Dundalk Port	GRU	40993	170.9850457	0.056748759	0.05674877	0.166618749	0.037464593	0.037464596	0.132148057	12929	0.001	2	1	1	0.564556838	-1.068826998
Dundalk Port	LSTM	40993	73.00692177	0.058582112	0.058582101	0.172089815	0.04121197	0.041211966	0.142196879	16961	0.001	2	1	1	0.564556838	-1.068826998
Dundalk Port	LSTM Attention	40993	175.6343107	0.026565118	0.02656514	0.105844259	0.0129521	0.012952098	0.083412424	16961	0.001	3	1	1	0.564556838	-1.068826998
Dundalk Port	Autoencoder	40993	39.96802664	0.346713245	0.346713185	0.525276542	0.416440994	0.416441083	0.588047624	97	0.001	2	0	1	0.564556838	-1.068826998
Dundalk Port	Transformer	40993	369.2215626	0.005847584	0.005847582	0.052568417	0.003999951	0.003999951	0.046903502	141825	0.001	2	1	1	0.564556838	-1.068826998
Dundalk Port	Variational Autoencoder	40993	49.52127814	0.344134241	0.344134063	0.523672879	0.41413945	0.41413936	0.586849988	1283	0.001	4	2	1	0.564556838	-1.068826998
Cappoge Bridge	Baseline	42492	35.83580923	0.01158676	0.011586758	0.040690478	0.031263843	0.031263847	0.055613771	0	0.001	0	0	1	2.364957909	11.27380204
Cappoge Bridge	Linear	42492	43.15126848	0.011569772	0.011569773	0.040690016	0.031095231	0.031095222	0.056517135	2	0.001	1	1	1	2.364957909	11.27380204
Cappoge Bridge	Dense	42492	29.58593273	0.011941154	0.01194115	0.047827397	0.031245634	0.03124563	0.062783331	97	0.001	2	2	1	2.364957909	11.27380204
Cappoge Bridge	Conv	42492	30.32320523	0.012373968	0.012373961	0.048562579	0.031440046	0.031440035	0.063941136	1153	0.001	3	2	1	2.364957909	11.27380204
Cappoge Bridge	RNN	42492	90.38880396	0.008530121	0.008530119	0.036177181	0.018189602	0.018189605	0.043098927	4289	0.001	2	1	1	2.364957909	11.27380204
Cappoge Bridge	GRU	42492	177.3822665	0.008094205	0.008094206	0.033904418	0.016859716	0.016859716	0.037030779	12929	0.001	2	1	1	2.364957909	11.27380204
Cappoge Bridge	LSTM	42492	161.7959967	0.008953923	0.008953917	0.039167073	0.017770439	0.017770439	0.042156123	16961	0.001	2	1	1	2.364957909	11.27380204
Cappoge Bridge	LSTM Attention	42492	166.3688288	0.002439416	0.002439416	0.020571822	0.004996877	0.004996876	0.024450215	16961	0.001	3	1	1	2.364957909	11.27380204
Cappoge Bridge	Autoencoder	42492	47.90084505	0.748915195	0.748915255	0.533188999	0.890533328	0.89053297	0.60853374	97	0.001	2	0	1	2.364957909	11.27380204
Cappoge Bridge	Transformer	42492	383.5963466	0.001962517	0.001962517	0.022965768	0.004257088	0.004257088	0.024875665	141825	0.001	2	1	1	2.364957909	11.27380204
Cappoge Bridge	Variational Autoencoder	42492	48.71194577	0.747071087	0.747071564	0.523630381	0.888841569	0.888841271	0.599836588	1283	0.001	4	2	1	2.364957909	11.27380204
Tallanstown	Baseline	43390	28.64182401	0.000142433	0.000142433	0.006918443	0.00058362	0.00058362	0.009424239	0	0.001	0	0	1	0.534508517	-0.430835051
Tallanstown	Linear	43390	28.83851147	0.000142606	0.000142606	0.007101301	0.000583402	0.000583402	0.009586784	2	0.001	1	1	1	0.534508517	-0.430835051
Tallanstown	Dense	43390	37.47582531	0.000142801	0.000142801	0.007015191	0.000665551	0.000665551	0.011159034	97	0.001	2	2	1	0.534508517	-0.430835051
Tallanstown	Conv	43390	41.07093549	0.000149592	0.000149592	0.007736923	0.001114933	0.001114933	0.014338262	1153	0.001	3	2	1	0.534508517	-0.430835051
Tallanstown	RNN	43390	92.00792265	0.000182823	0.000182823	0.008785757	0.000642022	0.000642022	0.011758314	4289	0.001	2	1	1	0.534508517	-0.430835051
Tallanstown	GRU	43390	190.9363039	0.000150367	0.000150367	0.007611895	0.000623519	0.000623519	0.010808625	12929	0.001	2	1	1	0.534508517	-0.430835051
Tallanstown	LSTM	43390	173.5929058	0.000141061	0.000141061	0.007321983	0.000632373	0.000632373	0.009942244	16961	0.001	2	1	1	0.534508517	-0.430835051
Tallanstown	LSTM Attention	43390	145.6702821	0.000291096	0.000291096	0.009987688	0.001039532	0.001039532	0.014274808	16961	0.001	3	1	1	0.534508517	-0.430835051
Tallanstown	Autoencoder	43390	42.71107721	0.64925915	0.64925921	0.661846399	0.833024681	0.8330248	0.722671688	97	0.001	2	0	1	0.534508517	-0.430835051
Tallanstown	Transformer	43390	411.0905769	0.000206296	0.000206296	0.011476566	0.000678323	0.000678323	0.014748177	141825	0.001	2	1	1	0.534508517	-0.430835051
Tallanstown	Variational Autoencoder	43390	47.98490381	0.647952795	0.647952676	0.652596593	0.831801713	0.831801832	0.71516794	1283	0.001	4	2	1	0.534508517	-0.430835051
Moyles Mill	Baseline	43585	35.34412932	0.000381091	0.000381091	0.009328621	0.000568795	0.000568795	0.013003197	0	0.001	0	0	1	0.651522724	-0.420442671
Moyles Mill	Linear	43585	29.38905764	0.000381163	0.000381162	0.009441935	0.000568598	0.000568597	0.013102532	2	0.001	1	1	1	0.651522724	-0.420442671
Moyles Mill	Dense	43585	49.09904075	0.000382723	0.000382723	0.009557544	0.000571226	0.000571225	0.013194066	97	0.001	2	2	1	0.651522724	-0.420442671
Moyles Mill	Conv	43585	26.10391951	0.000382451	0.000382451	0.009577347	0.000569873	0.000569873	0.013175994	1153	0.001	3	2	1	0.651522724	-0.420442671
Moyles Mill	RNN	43585	96.66727304	0.000395205	0.000395206	0.010279368	0.00056025	0.00056025	0.013708481	4289	0.001	2	1	1	0.651522724	-0.420442671
Moyles Mill	GRU	43585	158.6116581	0.000365342	0.000365342	0.009641827	0.000545456	0.000545456	0.013126716	12929	0.001	2	1	1	0.651522724	-0.420442671
Moyles Mill	LSTM	43585	193.4575806	0.000372983	0.000372983	0.009660706	0.000549075	0.000549075	0.01310628	16961	0.001	2	1	1	0.651522724	-0.420442671

Table E.2 ML prediction results for Neagh Bann RBD (*continued from previous page)*

Station	Model	Sample Size	Execution Time	Validation Loss	Validation MSE	Validation MAE	Test Loss	Test MSE	Test MAE	Model Complexity	Learning Rate	Number of Layers	Number of Dense Layers	Number of Features	Skewness	Kurtosis
Moyles Mill	LSTM Attention	43585	180.7173495	0.0005309	0.0005309	0.011202248	0.000841682	0.000841682	0.014763918	16961	0.001	3	1	1	0.651522724	-0.420442671
Moyles Mill	Autoencoder	43585	36.84814906	0.468096703	0.468096435	0.565416634	0.355982274	0.355982035	0.496932566	97	0.001	2	0	1	0.651522724	-0.420442671
Moyles Mill	Transformer	43585	409.7187788	0.001421257	0.001421258	0.030618394	0.001280538	0.001280537	0.02910566	141825	0.001	2	1	1	0.651522724	-0.420442671
Moyles Mill	Variational Autoencoder	43585	48.62589383	0.475770235	0.475770056	0.566487789	0.372253209	0.372253209	0.504831254	1283	0.001	4	2	1	0.651522724	-0.420442671

Appendix F

Detailed Classification Results for the Meta-Learners

Table F.1: Test Results : predicted vs true_label

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Random Forest	RNN	RNN
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM	LSTM
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Baseline	Baseline
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM	LSTM
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Baseline	Baseline
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM	LSTM
Original	Random Forest	RNN	RNN
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM	LSTM
Original	Random Forest	GRU	GRU
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM	LSTM
Original	Random Forest	GRU	GRU
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Random Forest	Baseline	Baseline
Original	Random Forest	Transformer	Transformer
Original	Random Forest	GRU	GRU
Original	Random Forest	Transformer	Transformer
Original	Random Forest	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	RNN	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	LSTM	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	LSTM	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM	LSTM_Attention
Original	Logistic Regression	RNN	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	Baseline	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	LSTM	LSTM_Attention
Original	Logistic Regression	GRU	Baseline
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	GRU
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM	LSTM_Attention
Original	Logistic Regression	GRU	GRU
Original	Logistic Regression	LSTM_Attention	GRU
Original	Logistic Regression	Baseline	GRU
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	GRU	LSTM_Attention
Original	Logistic Regression	Transformer	LSTM_Attention
Original	Logistic Regression	LSTM_Attention	LSTM_Attention
Original	Naive Bayes	RNN	RNN
Original	Naive Bayes	LSTM_Attention	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Naive Bayes	Baseline	LSTM_Attention
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Baseline	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	GRU	LSTM_Attention
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM	LSTM
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	LSTM_Attention	LSTM_Attention
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	LSTM_Attention
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	GRU	LSTM_Attention
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Baseline	LSTM_Attention
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	LSTM	LSTM
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	LSTM_Attention
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	GRU	LSTM_Attention
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM	LSTM
Original	Naive Bayes	RNN	RNN
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	LSTM_Attention	Transformer
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	GRU	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	GRU	LSTM
Original	Naive Bayes	LSTM	LSTM
Original	Naive Bayes	GRU	GRU
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	Baseline	Baseline
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	Naive Bayes	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Naive Bayes	LSTM	LSTM
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	LSTM_Attention	Baseline
Original	Naive Bayes	Baseline	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	GRU	Transformer
Original	Naive Bayes	Transformer	Transformer
Original	Naive Bayes	LSTM_Attention	LSTM
Original	K-Nearest Neighbors	RNN	RNN
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM	LSTM
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Baseline
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM	LSTM
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM	LSTM
Original	K-Nearest Neighbors	RNN	RNN
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	GRU	LSTM_Attention
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM	LSTM
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM	LSTM
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	K-Nearest Neighbors	Baseline	Baseline
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	GRU	GRU
Original	K-Nearest Neighbors	Transformer	Transformer
Original	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Original	Support Vector Machine	RNN	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	LSTM_Attention	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	LSTM_Attention
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM	GRU
Original	Support Vector Machine	RNN	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Transformer	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Support Vector Machine	Baseline	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	GRU	GRU
Original	Support Vector Machine	Transformer	GRU
Original	Support Vector Machine	LSTM_Attention	GRU
Original	Gradient Boosting	RNN	RNN
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM	LSTM
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM	LSTM
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM	LSTM
Original	Gradient Boosting	RNN	RNN
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM	LSTM
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM	LSTM
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	Gradient Boosting	Baseline	Baseline
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	GRU	GRU
Original	Gradient Boosting	Transformer	Transformer
Original	Gradient Boosting	LSTM_Attention	LSTM_Attention
Original	MLP	RNN	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Baseline	GRU
Original	MLP	Transformer	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Baseline
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	LSTM_Attention
Original	MLP	GRU	GRU
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Transformer	GRU
Original	MLP	GRU	Transformer
Original	MLP	GRU	GRU
Original	MLP	GRU	GRU
Original	MLP	Transformer	RNN
Original	MLP	GRU	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	GRU
Original	MLP	GRU	RNN
Original	MLP	Transformer	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	GRU	Transformer
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	GRU
Original	MLP	Baseline	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	Transformer	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	RNN
Original	MLP	Transformer	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Transformer	LSTM_Attention
Original	MLP	Baseline	RNN
Original	MLP	Transformer	LSTM_Attention
Original	MLP	GRU	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	Transformer	RNN
Original	MLP	Baseline	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	Baseline	GRU
Original	MLP	LSTM_Attention	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	LSTM	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Baseline	GRU
Original	MLP	Transformer	GRU
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	GRU	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	MLP	Transformer	RNN
Original	MLP	GRU	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	Baseline
Original	MLP	GRU	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Baseline	Baseline
Original	MLP	Baseline	GRU
Original	MLP	Transformer	Transformer
Original	MLP	GRU	RNN
Original	MLP	Baseline	GRU
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Baseline	LSTM_Attention
Original	MLP	LSTM_Attention	RNN
Original	MLP	LSTM	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Transformer	Transformer
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	GRU
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Baseline	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	MLP	LSTM_Attention	GRU
Original	MLP	GRU	GRU
Original	MLP	Transformer	Transformer
Original	MLP	LSTM	LSTM_Attention
Original	MLP	RNN	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	GRU	RNN
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	GRU	RNN
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	Baseline
Original	MLP	GRU	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	GRU	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Transformer	RNN
Original	MLP	GRU	RNN
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM	LSTM_Attention
Original	MLP	GRU	GRU
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Transformer	Transformer
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Transformer	LSTM_Attention
Original	MLP	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	MLP	GRU	LSTM_Attention
Original	MLP	LSTM	LSTM_Attention
Original	MLP	GRU	Transformer
Original	MLP	LSTM_Attention	LSTM_Attention
Original	MLP	Baseline	LSTM_Attention
Original	MLP	Transformer	GRU
Original	MLP	GRU	Transformer
Original	MLP	Transformer	RNN
Original	MLP	LSTM_Attention	Transformer
Original	XGBoost	RNN	RNN
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM	LSTM
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Baseline	Baseline
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM	LSTM
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Baseline	Baseline
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM	LSTM
Original	XGBoost	RNN	RNN
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM	LSTM
Original	XGBoost	GRU	GRU
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM	LSTM
Original	XGBoost	GRU	GRU
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	XGBoost	Baseline	Baseline
Original	XGBoost	Transformer	Transformer
Original	XGBoost	GRU	GRU
Original	XGBoost	Transformer	Transformer
Original	XGBoost	LSTM_Attention	LSTM_Attention
Original	LightGBM	RNN	RNN
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM	LSTM
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Baseline	Baseline
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM	LSTM
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Baseline	Baseline
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM	LSTM
Original	LightGBM	RNN	RNN
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM	LSTM
Original	LightGBM	GRU	GRU
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM	LSTM
Original	LightGBM	GRU	GRU
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	LightGBM	Baseline	Baseline
Original	LightGBM	Transformer	Transformer
Original	LightGBM	GRU	GRU
Original	LightGBM	Transformer	Transformer
Original	LightGBM	LSTM_Attention	LSTM_Attention
Original	CatBoost	RNN	RNN
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM	LSTM
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Baseline	Baseline
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM	LSTM
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Baseline	Baseline
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM	LSTM
Original	CatBoost	RNN	RNN
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM	LSTM
Original	CatBoost	GRU	GRU
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM	LSTM
Original	CatBoost	GRU	GRU
Original	CatBoost	LSTM_Attention	LSTM_Attention
Original	CatBoost	Baseline	Baseline
Original	CatBoost	Transformer	Transformer
Original	CatBoost	GRU	GRU
Original	CatBoost	Transformer	Transformer
Original	CatBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Isolation Forest	RNN	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Outlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Outlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM	Inlier
Original	Isolation Forest	RNN	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Transformer	Outlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	Baseline	Outlier
Original	Isolation Forest	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	Isolation Forest	Baseline	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	GRU	Inlier
Original	Isolation Forest	Transformer	Inlier
Original	Isolation Forest	LSTM_Attention	Inlier
Original	OneClassSVM	RNN	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	LSTM	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM	Inlier
Original	OneClassSVM	RNN	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	GRU	Inlier
Original	OneClassSVM	LSTM	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	LSTM_Attention	Inlier
Original	OneClassSVM	Baseline	Inlier
Original	OneClassSVM	Transformer	Inlier
Original	OneClassSVM	GRU	Outlier
Original	OneClassSVM	Transformer	Outlier
Original	OneClassSVM	LSTM_Attention	Outlier
Original	Dummy Classifier	RNN	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Baseline	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM	GRU
Original	Dummy Classifier	RNN	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	Dummy Classifier	Baseline	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	GRU	GRU
Original	Dummy Classifier	Transformer	GRU
Original	Dummy Classifier	LSTM_Attention	GRU
Original	LDA	RNN	RNN
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	Baseline
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	GRU	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	GRU
Original	LDA	GRU	Baseline
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	GRU
Original	LDA	GRU	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	GRU	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	LSTM_Attention
Original	LDA	Baseline	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Baseline	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	Transformer
Original	LDA	GRU	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	LSTM	LSTM
Original	LDA	GRU	GRU
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	GRU
Original	LDA	Transformer	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Transformer	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	GRU
Original	LDA	Baseline	GRU
Original	LDA	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	Transformer	LSTM_Attention
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Baseline	Baseline
Original	LDA	Baseline	GRU
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Baseline	Baseline
Original	LDA	Baseline	Baseline
Original	LDA	Baseline	Baseline
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	LSTM	LSTM
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	GRU
Original	LDA	Transformer	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	Baseline
Original	LDA	Baseline	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	LSTM	GRU
Original	LDA	RNN	RNN
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	GRU
Original	LDA	GRU	GRU
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	LSTM_Attention	GRU
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	LSTM_Attention
Original	LDA	GRU	GRU
Original	LDA	GRU	GRU
Original	LDA	LSTM	LSTM
Original	LDA	GRU	GRU
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	LSTM_Attention	GRU
Original	LDA	GRU	GRU
Original	LDA	LSTM	LSTM
Original	LDA	GRU	GRU
Original	LDA	LSTM_Attention	LSTM_Attention
Original	LDA	Baseline	Baseline
Original	LDA	Transformer	Transformer
Original	LDA	GRU	GRU
Original	LDA	Transformer	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Original	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	RNN	RNN
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM	LSTM
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM	LSTM
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM	LSTM
Random Oversampling	Random Forest	RNN	RNN
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM	LSTM
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM	LSTM
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Random Forest	Baseline	Baseline
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	GRU	GRU
Random Oversampling	Random Forest	Transformer	Transformer
Random Oversampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Oversampling	Logistic Regression	RNN	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Baseline	Baseline
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	Transformer	Baseline
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	Transformer
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	Transformer
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Baseline	Baseline
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	Baseline
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM	LSTM
Random Oversampling	Logistic Regression	RNN	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM	LSTM
Random Oversampling	Logistic Regression	GRU	Baseline
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	Baseline	GRU
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	LSTM_Attention	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM	LSTM
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Logistic Regression	Baseline	LSTM
Random Oversampling	Logistic Regression	Transformer	GRU
Random Oversampling	Logistic Regression	GRU	LSTM
Random Oversampling	Logistic Regression	Transformer	LSTM
Random Oversampling	Logistic Regression	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	RNN	RNN
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	Baseline	LSTM_Attention
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	GRU	Baseline
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Baseline	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	GRU	LSTM_Attention
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Baseline
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM	LSTM
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM_Attention
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM_Attention
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	GRU	LSTM_Attention
Random Oversampling	Naive Bayes	Baseline	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Baseline	LSTM_Attention
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM_Attention
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	LSTM
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM_Attention
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM_Attention
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM	LSTM
Random Oversampling	Naive Bayes	RNN	RNN
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	GRU	Baseline
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	GRU	LSTM
Random Oversampling	Naive Bayes	LSTM	LSTM
Random Oversampling	Naive Bayes	GRU	GRU
Random Oversampling	Naive Bayes	Baseline	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	Baseline	LSTM
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	GRU	Baseline
Random Oversampling	Naive Bayes	LSTM	LSTM
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	LSTM
Random Oversampling	Naive Bayes	Baseline	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	GRU	Transformer
Random Oversampling	Naive Bayes	Transformer	Transformer
Random Oversampling	Naive Bayes	LSTM_Attention	Baseline
Random Oversampling	K-Nearest Neighbors	RNN	RNN
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM	LSTM
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	LSTM
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM	LSTM
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM	LSTM
Random Oversampling	K-Nearest Neighbors	RNN	RNN
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM	LSTM
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM	LSTM
Random Oversampling	K-Nearest Neighbors	GRU	GRU
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	K-Nearest Neighbors	Baseline	Baseline
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	K-Nearest Neighbors	Transformer	Transformer
Random Oversampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Oversampling	Support Vector Machine	RNN	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	Baseline	Baseline
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM	LSTM
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Baseline	Baseline
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	LSTM	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Baseline	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM_Attention
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM	RNN
Random Oversampling	Support Vector Machine	RNN	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Support Vector Machine	GRU	LSTM
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	Baseline	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	LSTM
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM	RNN
Random Oversampling	Support Vector Machine	GRU	Baseline
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	GRU	RNN
Random Oversampling	Support Vector Machine	LSTM	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	RNN
Random Oversampling	Support Vector Machine	Baseline	RNN
Random Oversampling	Support Vector Machine	Transformer	RNN
Random Oversampling	Support Vector Machine	GRU	Transformer
Random Oversampling	Support Vector Machine	Transformer	Transformer
Random Oversampling	Support Vector Machine	LSTM_Attention	Transformer
Random Oversampling	Gradient Boosting	RNN	RNN
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM	LSTM
Random Oversampling	Gradient Boosting	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM	LSTM
Random Oversampling	Gradient Boosting	RNN	RNN
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM	LSTM
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM	LSTM
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	Gradient Boosting	Baseline	Baseline
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	GRU	GRU
Random Oversampling	Gradient Boosting	Transformer	Transformer
Random Oversampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Oversampling	MLP	RNN	LSTM
Random Oversampling	MLP	LSTM_Attention	LSTM
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	GRU	LSTM
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	Transformer	LSTM_Attention
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	LSTM_Attention	LSTM
Random Oversampling	MLP	GRU	LSTM
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	Baseline
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Baseline	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	LSTM_Attention	Baseline
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	MLP	Transformer	Baseline
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	LSTM_Attention	LSTM_Attention
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	LSTM
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	LSTM_Attention
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	Baseline
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	Transformer	Baseline
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Baseline	Baseline
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	Baseline
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	LSTM	LSTM
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	Baseline
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	LSTM_Attention	Baseline
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM	LSTM
Random Oversampling	MLP	RNN	LSTM
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	Transformer
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	LSTM_Attention	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	Baseline
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	LSTM
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	GRU	LSTM
Random Oversampling	MLP	LSTM	LSTM
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	Baseline	GRU
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	LSTM	LSTM
Random Oversampling	MLP	GRU	GRU
Random Oversampling	MLP	LSTM_Attention	GRU
Random Oversampling	MLP	Baseline	Transformer
Random Oversampling	MLP	Transformer	GRU
Random Oversampling	MLP	GRU	Transformer
Random Oversampling	MLP	Transformer	Transformer
Random Oversampling	MLP	LSTM_Attention	LSTM
Random Oversampling	XGBoost	RNN	RNN
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM	LSTM
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM	LSTM
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM	LSTM
Random Oversampling	XGBoost	RNN	RNN
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM	LSTM
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM	LSTM
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	XGBoost	Baseline	Baseline
Random Oversampling	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	XGBoost	GRU	GRU
Random Oversampling	XGBoost	Transformer	Transformer
Random Oversampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	RNN	RNN
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM	LSTM
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM	LSTM
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM	LSTM
Random Oversampling	LightGBM	RNN	RNN
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM	LSTM
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM	LSTM
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	LightGBM	Baseline	Baseline
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	GRU	GRU
Random Oversampling	LightGBM	Transformer	Transformer
Random Oversampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	RNN	RNN
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	CatBoost	LSTM	LSTM
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM	LSTM
Random Oversampling	CatBoost	RNN	RNN
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM	LSTM
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM	LSTM
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	CatBoost	Baseline	Baseline
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	GRU	GRU
Random Oversampling	CatBoost	Transformer	Transformer
Random Oversampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Oversampling	Isolation Forest	RNN	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Outlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Outlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Outlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Outlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Outlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Baseline	Outlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM	Inlier
Random Oversampling	Isolation Forest	RNN	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Baseline	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	Baseline	Outlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM	Inlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	Isolation Forest	Baseline	Outlier
Random Oversampling	Isolation Forest	Transformer	Outlier
Random Oversampling	Isolation Forest	GRU	Inlier
Random Oversampling	Isolation Forest	Transformer	Inlier
Random Oversampling	Isolation Forest	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	RNN	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	LSTM	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM	Inlier
Random Oversampling	OneClassSVM	RNN	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	Baseline	Inlier
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	GRU	Inlier
Random Oversampling	OneClassSVM	LSTM	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Inlier
Random Oversampling	OneClassSVM	Baseline	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	OneClassSVM	Transformer	Inlier
Random Oversampling	OneClassSVM	GRU	Outlier
Random Oversampling	OneClassSVM	Transformer	Outlier
Random Oversampling	OneClassSVM	LSTM_Attention	Outlier
Random Oversampling	Dummy Classifier	RNN	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM	Baseline
Random Oversampling	Dummy Classifier	RNN	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	Dummy Classifier	Baseline	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	GRU	Baseline
Random Oversampling	Dummy Classifier	Transformer	Baseline
Random Oversampling	Dummy Classifier	LSTM_Attention	Baseline
Random Oversampling	LDA	RNN	RNN
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	Baseline
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	RNN
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	GRU	Baseline
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	GRU	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	GRU	LSTM_Attention
Random Oversampling	LDA	GRU	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	LSTM_Attention
Random Oversampling	LDA	Baseline	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	Transformer
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	Transformer
Random Oversampling	LDA	GRU	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	LDA	LSTM	LSTM
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	GRU
Random Oversampling	LDA	GRU	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Transformer	GRU
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Transformer	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	LSTM	LSTM
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	GRU
Random Oversampling	LDA	Transformer	GRU
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	RNN
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Baseline	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM	LSTM
Random Oversampling	LDA	RNN	RNN
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	LSTM_Attention
Random Oversampling	LDA	GRU	Transformer
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	LSTM_Attention
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	LSTM	LSTM
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	RNN
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	LSTM_Attention	GRU
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	LSTM	LSTM
Random Oversampling	LDA	GRU	GRU
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Oversampling	LDA	Baseline	Baseline
Random Oversampling	LDA	Transformer	Transformer
Random Oversampling	LDA	GRU	LSTM
Random Oversampling	LDA	Transformer	LSTM_Attention
Random Oversampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	RNN	RNN
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Random Forest	GRU	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM	LSTM
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	Transformer	LSTM
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM	LSTM
Random Undersampling	Random Forest	RNN	RNN
Random Undersampling	Random Forest	LSTM_Attention	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	Transformer
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	Transformer
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	LSTM
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	Baseline
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM	LSTM
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Random Forest	Baseline	Baseline
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	GRU	GRU
Random Undersampling	Random Forest	Transformer	Transformer
Random Undersampling	Random Forest	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	RNN	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	Baseline	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	LSTM_Attention	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Baseline	GRU
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Baseline	Baseline
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Baseline	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	Baseline	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	GRU
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	LSTM	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Baseline	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Baseline	Baseline
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Baseline	Transformer
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	GRU	GRU
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	Baseline	Transformer
Random Undersampling	Logistic Regression	Baseline	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Baseline	LSTM_Attention
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	LSTM	LSTM
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	GRU
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	Baseline	Baseline
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	Baseline	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	Transformer
Random Undersampling	Logistic Regression	GRU	GRU
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	LSTM	LSTM

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Logistic Regression	RNN	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	Transformer	Baseline
Random Undersampling	Logistic Regression	GRU	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	Baseline	Transformer
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	GRU
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	Baseline	GRU
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM	LSTM
Random Undersampling	Logistic Regression	GRU	Baseline
Random Undersampling	Logistic Regression	Baseline	LSTM
Random Undersampling	Logistic Regression	Transformer	LSTM
Random Undersampling	Logistic Regression	Baseline	GRU
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	LSTM
Random Undersampling	Logistic Regression	LSTM	LSTM
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Logistic Regression	Baseline	GRU
Random Undersampling	Logistic Regression	Transformer	LSTM_Attention
Random Undersampling	Logistic Regression	GRU	Transformer
Random Undersampling	Logistic Regression	Transformer	Transformer
Random Undersampling	Logistic Regression	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	RNN	RNN
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	Baseline	LSTM_Attention
Random Undersampling	Naive Bayes	Transformer	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Transformer	LSTM
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Baseline	Baseline
Random Undersampling	Naive Bayes	Transformer	Baseline
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	GRU	LSTM_Attention
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	LSTM
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM	LSTM
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	LSTM_Attention	GRU
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	LSTM
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	LSTM
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM_Attention
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	GRU	LSTM_Attention
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Baseline	LSTM_Attention
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	LSTM	LSTM
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	GRU
Random Undersampling	Naive Bayes	Transformer	LSTM
Random Undersampling	Naive Bayes	Baseline	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Baseline
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM_Attention
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM_Attention
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM	LSTM
Random Undersampling	Naive Bayes	RNN	RNN
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	GRU	LSTM
Random Undersampling	Naive Bayes	LSTM	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Naive Bayes	GRU	GRU
Random Undersampling	Naive Bayes	Baseline	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	Baseline	Baseline
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	Naive Bayes	GRU	Baseline
Random Undersampling	Naive Bayes	LSTM	LSTM
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	Baseline
Random Undersampling	Naive Bayes	Baseline	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	GRU	Transformer
Random Undersampling	Naive Bayes	Transformer	Transformer
Random Undersampling	Naive Bayes	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	RNN	RNN
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	LSTM
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	GRU	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Transformer
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	LSTM
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM
Random Undersampling	K-Nearest Neighbors	Baseline	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Baseline	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM	LSTM
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM	LSTM
Random Undersampling	K-Nearest Neighbors	RNN	RNN
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	Transformer	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Baseline	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	LSTM
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	GRU	LSTM
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	Transformer
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	LSTM	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	GRU
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	GRU	LSTM_Attention
Random Undersampling	K-Nearest Neighbors	LSTM	LSTM
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	GRU
Random Undersampling	K-Nearest Neighbors	Baseline	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Baseline
Random Undersampling	K-Nearest Neighbors	GRU	Baseline
Random Undersampling	K-Nearest Neighbors	Transformer	Transformer
Random Undersampling	K-Nearest Neighbors	LSTM_Attention	Baseline
Random Undersampling	Support Vector Machine	RNN	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	Baseline	GRU
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	LSTM
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	RNN

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM	LSTM
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	LSTM
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	LSTM
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Baseline	GRU
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	LSTM	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	LSTM
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	LSTM
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM_Attention
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM	RNN
Random Undersampling	Support Vector Machine	RNN	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Support Vector Machine	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	LSTM
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM	RNN
Random Undersampling	Support Vector Machine	GRU	GRU
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	GRU	RNN
Random Undersampling	Support Vector Machine	LSTM	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	RNN
Random Undersampling	Support Vector Machine	Baseline	RNN
Random Undersampling	Support Vector Machine	Transformer	RNN
Random Undersampling	Support Vector Machine	GRU	Transformer
Random Undersampling	Support Vector Machine	Transformer	Transformer
Random Undersampling	Support Vector Machine	LSTM_Attention	Transformer
Random Undersampling	Gradient Boosting	RNN	RNN
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	Baseline
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	Baseline
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Gradient Boosting	GRU	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	GRU
Random Undersampling	Gradient Boosting	Baseline	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	RNN
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	LSTM
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Baseline	GRU
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM	LSTM
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	GRU
Random Undersampling	Gradient Boosting	Transformer	GRU
Random Undersampling	Gradient Boosting	GRU	LSTM
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Gradient Boosting	LSTM	LSTM
Random Undersampling	Gradient Boosting	RNN	RNN
Random Undersampling	Gradient Boosting	LSTM_Attention	RNN
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	RNN
Random Undersampling	Gradient Boosting	GRU	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	Transformer
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	LSTM	LSTM
Random Undersampling	Gradient Boosting	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	Gradient Boosting	Baseline	Baseline
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	GRU	GRU
Random Undersampling	Gradient Boosting	Transformer	Transformer
Random Undersampling	Gradient Boosting	LSTM_Attention	LSTM_Attention
Random Undersampling	MLP	RNN	GRU
Random Undersampling	MLP	LSTM_Attention	Baseline
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	LSTM_Attention	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	LSTM
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	Transformer
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	Transformer
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	Transformer
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	Transformer
Random Undersampling	MLP	Baseline	LSTM_Attention
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	LSTM_Attention	LSTM
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	Baseline	Baseline
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	LSTM_Attention	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	Transformer	Transformer
Random Undersampling	MLP	LSTM	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	LSTM_Attention	LSTM
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Baseline	LSTM_Attention
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	Baseline	LSTM
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	LSTM_Attention	LSTM_Attention
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	Baseline
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Baseline	LSTM
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	LSTM
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	MLP	Baseline	LSTM_Attention
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Baseline	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	LSTM
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	LSTM	GRU
Random Undersampling	MLP	RNN	GRU
Random Undersampling	MLP	LSTM_Attention	LSTM_Attention
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM_Attention	LSTM_Attention
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	Baseline	LSTM
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Transformer	Baseline
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	MLP	LSTM	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	Transformer
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	GRU	GRU
Random Undersampling	MLP	LSTM	GRU
Random Undersampling	MLP	GRU	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	GRU
Random Undersampling	MLP	Baseline	GRU
Random Undersampling	MLP	Transformer	GRU
Random Undersampling	MLP	GRU	Baseline
Random Undersampling	MLP	Transformer	LSTM_Attention
Random Undersampling	MLP	LSTM_Attention	Baseline
Random Undersampling	XGBoost	RNN	RNN
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM	LSTM
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM	LSTM
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	LSTM_Attention
Random Undersampling	XGBoost	Transformer	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM	LSTM
Random Undersampling	XGBoost	RNN	RNN
Random Undersampling	XGBoost	LSTM_Attention	RNN
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	Transformer
Random Undersampling	XGBoost	GRU	LSTM
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	Transformer
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	Baseline
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM	LSTM
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM	LSTM
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	XGBoost	Baseline	Baseline
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	GRU	GRU
Random Undersampling	XGBoost	Transformer	Transformer
Random Undersampling	XGBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	RNN	RNN
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	Baseline
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	LSTM
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	Baseline
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM	LSTM
Random Undersampling	LightGBM	GRU	LSTM
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	Transformer
Random Undersampling	LightGBM	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Baseline	GRU
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM	LSTM
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM	LSTM
Random Undersampling	LightGBM	RNN	RNN
Random Undersampling	LightGBM	LSTM_Attention	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LightGBM	GRU	LSTM
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	Transformer
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	GRU
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	LSTM
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	Baseline
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM	LSTM
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM	LSTM
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	LightGBM	Baseline	Baseline
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	GRU	GRU
Random Undersampling	LightGBM	Transformer	Transformer
Random Undersampling	LightGBM	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	RNN	RNN
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	Baseline
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	Baseline
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM	LSTM
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	LSTM
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	LSTM
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	LSTM
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	LSTM
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Baseline	GRU
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM	LSTM
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	LSTM
Random Undersampling	CatBoost	Transformer	LSTM
Random Undersampling	CatBoost	GRU	LSTM
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	LSTM
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM	LSTM
Random Undersampling	CatBoost	RNN	RNN
Random Undersampling	CatBoost	LSTM_Attention	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	Transformer
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM	LSTM
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	CatBoost	Baseline	Baseline
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	GRU	GRU
Random Undersampling	CatBoost	Transformer	Transformer
Random Undersampling	CatBoost	LSTM_Attention	LSTM_Attention
Random Undersampling	Isolation Forest	RNN	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Outlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Outlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM	Inlier
Random Undersampling	Isolation Forest	RNN	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Baseline	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	Baseline	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	LSTM	Inlier
Random Undersampling	Isolation Forest	GRU	Outlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	Isolation Forest	Baseline	Outlier
Random Undersampling	Isolation Forest	Transformer	Outlier
Random Undersampling	Isolation Forest	GRU	Inlier
Random Undersampling	Isolation Forest	Transformer	Inlier
Random Undersampling	Isolation Forest	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	RNN	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	LSTM	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Baseline	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM	Inlier
Random Undersampling	OneClassSVM	RNN	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Outlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	GRU	Inlier
Random Undersampling	OneClassSVM	LSTM	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Inlier
Random Undersampling	OneClassSVM	Baseline	Inlier
Random Undersampling	OneClassSVM	Transformer	Inlier
Random Undersampling	OneClassSVM	GRU	Outlier
Random Undersampling	OneClassSVM	Transformer	Outlier
Random Undersampling	OneClassSVM	LSTM_Attention	Outlier
Random Undersampling	Dummy Classifier	RNN	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM	Baseline
Random Undersampling	Dummy Classifier	RNN	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	Dummy Classifier	Baseline	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	GRU	Baseline
Random Undersampling	Dummy Classifier	Transformer	Baseline
Random Undersampling	Dummy Classifier	LSTM_Attention	Baseline
Random Undersampling	LDA	RNN	RNN
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	LSTM_Attention
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Transformer	RNN
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	LSTM_Attention
Random Undersampling	LDA	Transformer	RNN
Random Undersampling	LDA	LSTM	LSTM
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	LDA	GRU	LSTM_Attention
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	LSTM	LSTM
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Transformer	Baseline
Random Undersampling	LDA	Transformer	Baseline
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	Transformer	RNN
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM	LSTM
Random Undersampling	LDA	RNN	RNN
Random Undersampling	LDA	LSTM_Attention	RNN
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	LSTM_Attention	Baseline
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	GRU	Transformer
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	GRU	Baseline
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	LSTM	LSTM
Random Undersampling	LDA	GRU	GRU
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	RNN
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	GRU
Random Undersampling	LDA	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
Random Undersampling	LDA	LSTM	LSTM
Random Undersampling	LDA	GRU	LSTM_Attention
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
Random Undersampling	LDA	Baseline	Baseline
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	GRU	LSTM
Random Undersampling	LDA	Transformer	Transformer
Random Undersampling	LDA	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	RNN	RNN
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM	LSTM
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM	LSTM
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM	LSTM
SMOTE	Random Forest	RNN	RNN
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM	LSTM
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM	LSTM
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Random Forest	Baseline	Baseline
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	GRU	GRU
SMOTE	Random Forest	Transformer	Transformer
SMOTE	Random Forest	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	RNN	LSTM
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	Baseline
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	LSTM_Attention	LSTM
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	Baseline
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	LSTM_Attention	LSTM
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	LSTM_Attention	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM	LSTM
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	Transformer	Baseline
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	LSTM_Attention	LSTM
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	Transformer	LSTM_Attention
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Baseline	LSTM
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	Baseline	LSTM
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	LSTM	LSTM
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	Baseline
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM	LSTM
SMOTE	Logistic Regression	RNN	LSTM
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	LSTM_Attention
SMOTE	Logistic Regression	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Logistic Regression	Baseline	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	LSTM_Attention	GRU
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	GRU	LSTM
SMOTE	Logistic Regression	LSTM	LSTM
SMOTE	Logistic Regression	GRU	Baseline
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	GRU
SMOTE	Logistic Regression	LSTM_Attention	LSTM_Attention
SMOTE	Logistic Regression	GRU	GRU
SMOTE	Logistic Regression	LSTM	LSTM
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	LSTM_Attention	LSTM
SMOTE	Logistic Regression	Baseline	GRU
SMOTE	Logistic Regression	Transformer	Baseline
SMOTE	Logistic Regression	GRU	Transformer
SMOTE	Logistic Regression	Transformer	Transformer
SMOTE	Logistic Regression	LSTM_Attention	Transformer
SMOTE	Naive Bayes	RNN	RNN
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Baseline	LSTM_Attention
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Baseline
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Baseline	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	GRU	LSTM_Attention
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Baseline
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM	LSTM
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM_Attention
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	LSTM_Attention
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Baseline
SMOTE	Naive Bayes	GRU	LSTM_Attention
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Baseline	LSTM_Attention
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Baseline
SMOTE	Naive Bayes	LSTM	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	LSTM_Attention
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	LSTM
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	GRU	LSTM_Attention
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM_Attention
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM	LSTM
SMOTE	Naive Bayes	RNN	RNN
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	LSTM_Attention	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	Baseline
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM	LSTM
SMOTE	Naive Bayes	GRU	GRU
SMOTE	Naive Bayes	Baseline	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	Baseline	LSTM
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	GRU	LSTM
SMOTE	Naive Bayes	LSTM	LSTM
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	Naive Bayes	Baseline	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	GRU	Transformer
SMOTE	Naive Bayes	Transformer	Transformer
SMOTE	Naive Bayes	LSTM_Attention	LSTM
SMOTE	K-Nearest Neighbors	RNN	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM	LSTM
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Baseline
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM	LSTM
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM	LSTM
SMOTE	K-Nearest Neighbors	RNN	RNN
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	GRU	LSTM_Attention
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM	LSTM
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM	LSTM
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	K-Nearest Neighbors	Baseline	Baseline
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	GRU	GRU
SMOTE	K-Nearest Neighbors	Transformer	Transformer
SMOTE	K-Nearest Neighbors	LSTM_Attention	LSTM_Attention
SMOTE	Support Vector Machine	RNN	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	Baseline	Baseline
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	LSTM_Attention	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM	LSTM
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	RNN

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Baseline	Baseline
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	LSTM	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	LSTM_Attention
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM	RNN
SMOTE	Support Vector Machine	RNN	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Support Vector Machine	GRU	LSTM
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	LSTM
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	Baseline	RNN

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM	RNN
SMOTE	Support Vector Machine	GRU	Baseline
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	GRU	RNN
SMOTE	Support Vector Machine	LSTM	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	RNN
SMOTE	Support Vector Machine	Baseline	RNN
SMOTE	Support Vector Machine	Transformer	RNN
SMOTE	Support Vector Machine	GRU	Transformer
SMOTE	Support Vector Machine	Transformer	Transformer
SMOTE	Support Vector Machine	LSTM_Attention	Transformer
SMOTE	Gradient Boosting	RNN	RNN
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM	LSTM
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM	LSTM
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM	LSTM
SMOTE	Gradient Boosting	RNN	RNN
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM	LSTM
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM	LSTM
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	Gradient Boosting	Baseline	Baseline
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	GRU	GRU
SMOTE	Gradient Boosting	Transformer	Transformer
SMOTE	Gradient Boosting	LSTM_Attention	LSTM_Attention
SMOTE	MLP	RNN	RNN
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	GRU	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Transformer	LSTM
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	LSTM_Attention	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	LSTM	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM_Attention	RNN
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	LSTM
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM_Attention	RNN
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Transformer	LSTM
SMOTE	MLP	LSTM_Attention	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	LSTM
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Baseline	LSTM
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	LSTM_Attention	LSTM_Attention
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	LSTM	Baseline
SMOTE	MLP	RNN	GRU
SMOTE	MLP	LSTM_Attention	RNN
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM_Attention	RNN
SMOTE	MLP	GRU	GRU
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	Baseline	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Transformer	Baseline
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	LSTM
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	LSTM	Baseline
SMOTE	MLP	GRU	GRU
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	Transformer
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	GRU	GRU
SMOTE	MLP	LSTM	Baseline
SMOTE	MLP	GRU	Baseline
SMOTE	MLP	LSTM_Attention	GRU
SMOTE	MLP	Baseline	GRU
SMOTE	MLP	Transformer	GRU
SMOTE	MLP	GRU	Transformer
SMOTE	MLP	Transformer	LSTM
SMOTE	MLP	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	XGBoost	RNN	RNN
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM	LSTM
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM	LSTM
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM	LSTM
SMOTE	XGBoost	RNN	RNN
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM	LSTM
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM	LSTM
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	XGBoost	Baseline	Baseline
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	GRU	GRU
SMOTE	XGBoost	Transformer	Transformer
SMOTE	XGBoost	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	RNN	RNN
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM	LSTM
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM	LSTM
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM	LSTM
SMOTE	LightGBM	RNN	RNN
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM	LSTM
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM	LSTM
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	LightGBM	Baseline	Baseline
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	GRU	GRU
SMOTE	LightGBM	Transformer	Transformer
SMOTE	LightGBM	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	RNN	RNN
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM	LSTM
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM	LSTM
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM	LSTM
SMOTE	CatBoost	RNN	RNN
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM	LSTM
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM	LSTM
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	CatBoost	Baseline	Baseline
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	GRU	GRU
SMOTE	CatBoost	Transformer	Transformer
SMOTE	CatBoost	LSTM_Attention	LSTM_Attention
SMOTE	Isolation Forest	RNN	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	GRU	Outlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	Baseline	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Baseline	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Baseline	Outlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM	Inlier
SMOTE	Isolation Forest	RNN	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Outlier
SMOTE	Isolation Forest	Transformer	Outlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Baseline	Inlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	Baseline	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	LSTM	Inlier
SMOTE	Isolation Forest	GRU	Outlier
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	Isolation Forest	Baseline	Outlier
SMOTE	Isolation Forest	Transformer	Inlier
SMOTE	Isolation Forest	GRU	Inlier
SMOTE	Isolation Forest	Transformer	Inlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Isolation Forest	LSTM_Attention	Inlier
SMOTE	OneClassSVM	RNN	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Baseline	Outlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Transformer	Outlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	LSTM	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM	Inlier
SMOTE	OneClassSVM	RNN	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	Baseline	Inlier

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	GRU	Inlier
SMOTE	OneClassSVM	LSTM	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Inlier
SMOTE	OneClassSVM	Baseline	Inlier
SMOTE	OneClassSVM	Transformer	Inlier
SMOTE	OneClassSVM	GRU	Outlier
SMOTE	OneClassSVM	Transformer	Outlier
SMOTE	OneClassSVM	LSTM_Attention	Outlier
SMOTE	Dummy Classifier	RNN	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM	Baseline
SMOTE	Dummy Classifier	RNN	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	Dummy Classifier	Baseline	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	GRU	Baseline
SMOTE	Dummy Classifier	Transformer	Baseline
SMOTE	Dummy Classifier	LSTM_Attention	Baseline
SMOTE	LDA	RNN	RNN
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	Baseline
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	GRU	Baseline
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	GRU	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	GRU	LSTM_Attention
SMOTE	LDA	GRU	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	LSTM_Attention
SMOTE	LDA	Baseline	GRU
SMOTE	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LDA	GRU	Transformer
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	Transformer
SMOTE	LDA	GRU	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM	LSTM
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	GRU
SMOTE	LDA	GRU	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Transformer	GRU
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Transformer	LSTM_Attention
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Baseline	Baseline

Table F.1 continued from previous page

Resampling Technique	**Classifier**	**True_Labels**	**Predicted_Labels**
SMOTE	LDA	Baseline	GRU
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	LSTM	LSTM
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	GRU
SMOTE	LDA	Transformer	GRU
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	RNN
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM	LSTM
SMOTE	LDA	RNN	RNN
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	LSTM_Attention
SMOTE	LDA	GRU	Transformer
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer

Table F.1 continued from previous page

Resampling Technique	Classifier	True_Labels	Predicted_Labels
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	LSTM_Attention
SMOTE	LDA	GRU	GRU
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	LSTM	LSTM
SMOTE	LDA	GRU	GRU
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	RNN
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	LSTM_Attention	GRU
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM	LSTM
SMOTE	LDA	GRU	GRU
SMOTE	LDA	LSTM_Attention	LSTM_Attention
SMOTE	LDA	Baseline	Baseline
SMOTE	LDA	Transformer	Transformer
SMOTE	LDA	GRU	LSTM
SMOTE	LDA	Transformer	LSTM_Attention
SMOTE	LDA	LSTM_Attention	LSTM_Attention

Appendix G

Detailed Results for water level prediction second experiment

Appendix H

Meta Features

Features for training

- Model
- Sample Size
- Execution Time
- Validation Loss
- Validation MSE
- Validation MAE
- Test Loss
- Test MSE
- Test MAE
- Model Complexity
- Learning Rate
- Number of Layers
- Number of Dense Layers
- Number of Features
- Mean
- Median
- Variance

- Standard Deviation
- Range
- Skewness_y
- Kurtosis_y
- Seasonal Strength
- ACF1
- PACF1
- ADF_PValue
- KPSS_PValue
- Alpha
- Beta
- UR_PP_TestStat
- UR_PP_PValue
- UR_KPSS_TestStat
- UR_KPSS_PValue
- LMResiduals_ACF1
- BED500K_ITM_UnitName
- BED500K_ITM_natmapcode
- BED500K_ITM_unit_label
- BED500K_ITM_AgeBracket
- BED500K_ITM_Area_km2
- BED500K_ITM_Formation
- IE_GSI_Q_SEDIMENTS_50K_ITM_LITHOLOGY
- IE_GSI_Q_SEDIMENTS_50K_ITM_QUAT_SED
- IE_GSI_Q_SEDIMENTS_50K_ITM_SYMBOLOGY
- IE_GSI_Q_SEDIMENTS_50K_ITM_ORIG_FID
- IE_GSI_Q_SEDIMENTS_50K_ITM_SHAPE_AREA

- IE_GSI_Q_SEDIMENTS_50K_ITM_SHAPE_LEN
- SOIL_SISNationalSoils_Associatio
- SOIL_SISNationalSoils_Associat_1
- SOIL_SISNationalSoils_Associat_2
- SOIL_SISNationalSoils_Texture_Su
- SOIL_SISNationalSoils_ha
- SOIL_SISNationalSoils_DRAINAGE
- SOIL_SISNationalSoils_TEXTURE
- SOIL_SISNationalSoils_DEPTH
- SOIL_SISNationalSoils_SHAPE_Leng

Features for testing

- Mean
- Median
- Variance
- Standard Deviation
- Range
- Skewness_y
- Kurtosis_y
- Seasonal Strength
- ACF1
- PACF1
- ADF_PValue
- KPSS_PValue
- Alpha
- Beta
- UR_PP_TestStat
- UR_PP_PValue

- UR_KPSS_TestStat
- UR_KPSS_PValue
- LMResiduals_ACF1
- BED500K_ITM_UnitName
- BED500K_ITM_natmapcode
- BED500K_ITM_unit_label
- BED500K_ITM_AgeBracket
- BED500K_ITM_Area_km2
- BED500K_ITM_Formation
- IE_GSI_Q_SEDIMENTS_50K_ITM_LITHOLOGY
- IE_GSI_Q_SEDIMENTS_50K_ITM_QUAT_SED
- IE_GSI_Q_SEDIMENTS_50K_ITM_SYMBOLOGY
- IE_GSI_Q_SEDIMENTS_50K_ITM_ORIG_FID
- IE_GSI_Q_SEDIMENTS_50K_ITM_SHAPE_AREA
- IE_GSI_Q_SEDIMENTS_50K_ITM_SHAPE_LEN
- SOIL_SISNationalSoils_Associatio
- SOIL_SISNationalSoils_Associat_1
- SOIL_SISNationalSoils_Associat_2
- SOIL_SISNationalSoils_Texture_Su
- SOIL_SISNationalSoils_ha
- SOIL_SISNationalSoils_DRAINAGE
- SOIL_SISNationalSoils_TEXTURE
- SOIL_SISNationalSoils_DEPTH
- SOIL_SISNationalSoils_SHAPE_Leng

Bibliography

[1] Mahdi Abbasi et al. “A hybrid of Random Forest and Deep Auto-Encoder with support vector regression methods for accuracy improvement and uncertainty reduction of long-term streamflow prediction”. In: *Journal of Hydrology* 597 (2021), p. 125717.

[2] John Aber et al. “Forest processes and global environmental change: predicting the effects of individual and multiple stressors: we review the effects of several rapidly changing environmental drivers on ecosystem function, discuss interactions among them, and summarize predicted changes in productivity, carbon storage, and water balance”. In: *BioScience* 51.9 (2001), pp. 735–751.

[3] Tarek Abdel-Latif Aboul-Atta and Mourad Medhat Elzeiny. “Towards a Comprehensive Approach to Assess and Measure Regional Integration.” In: *Journal of African Union Studies* 12.3 (2023).

[4] Taher Omran Ahmed and Maryvonne Miquel. “Multidimensional structures dedicated to continuous spatiotemporal phenomena”. In: *Lecture Notes in Computer Science* 3567 (2005), pp. 29–40. ISSN: 03029743. DOI: 10.1007/11511854_3.

[5] Elisabeth Albertini. “Does environmental management improve financial performance? A meta-analytical review”. In: *Organization & Environment* 26.4 (2013), pp. 431–457.

[6] Eman Alduweib, Muhammad Abu Arqoub, and Waseem Alromema, eds. *Automated Machine Learning for Information Management and Information Systems: An Overview.* 2024. DOI: 10.1109/iccr61006.2024.10532828.

[7] Zainab Al-Ali Hussain Al-Ali and Roland N. Horne. "Meta Learning Using Deep N-BEATS Model for Production Forecasting with Limited History". In: *SPE Reservoir Evaluation & Engineering* (2023). DOI: 10.2118/214214-ms.

[8] Diana M Allen et al. "Data integration and standardization in cross-border hydrogeological studies: a novel approach to hydrostratigraphic model development". In: *Environmental geology* 53 (2008), pp. 1441–1453.

[9] B Arinze. "Selecting appropriate forecasting models using rule induction". In: *Omega* 22.6 (1994), pp. 647–658. ISSN: 0305-0483. DOI: https://doi.org/10.1016/0305-0483(94)90054-X. URL: https://www.sciencedirect.com/science/article/pii/030504839490054X.

[10] J Scott Armstrong. "Principles of forecasting: A handbook for researchers and practitioners". In: *Springer Science & Business Media* (2001).

[11] A. Azarnivand and B. Gharabaghi. "Machine learning for soil erosion prediction: a review". In: *Earth-Science Reviews* 201 (2020), p. 103067.

[12] Fouad Bahrpeyma, Mark Roantree, and Andrew Mccarren. "Multistep-ahead Prediction: A Comparison of Analytical and Algorithmic Approaches". In: *20th International Conference, DaWaK 2018, Regensburg, Germany, September 3–6, 2018, Proceedings* (2018), pp. 345–354. DOI: 10.1007/978-3-319-98539-8_26.

[13] Ramón C. Barquín and Herb. Edelstein, eds. *Planning and Designing the Data Warehouse*. Prentice Hall, July 1996, p. 311. ISBN: 978-0-13-255746-7.

[14] Gabor Bartha and Sandor Kocsis. "Standardization of geographic data: The european inspire directive". In: *European Journal of Geography* 2.2 (2011).

[15] I. A. Basheer and M. Hajmeer. "Artificial Neural Networks: Fundamentals, Computing, Design, and Application". In: *Journal of Microbiological Methods* 43.1 (2000), pp. 3–31.

[16] Carlo Batini et al. "Methodologies for data quality assessment and improvement". In: *ACM Computing Surveys (CSUR)* 41.3 (2009), p. 16.

[17] Punam Bedi, Vinita Jindal, and Anjali Gautam. "Beginning with big data simplified". In: *2014 International Conference on Data Mining and Intelligent Computing, ICDMIC 2014* (Nov. 2014). DOI: 10.1109/ICDMIC.2014.6954229.

[18] Yoshua Bengio and et al. "Learning long-term dependencies with gradient descent is difficult". In: *IEEE Transactions on Neural Networks* 5.2 (1992), pp. 157–166.

[19] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. "Learning long-term dependencies with gradient descent is difficult". In: *IEEE transactions on neural networks* 5.2 (1994), pp. 157–166.

[20] Günter Blöschl, Marc F.P. Bierkens, Antonio Chambel, et al. "Twenty-three unsolved problems in hydrology (UPH)–a community perspective". In: *Hydrological Sciences Journal* 64.10 (2019), pp. 1141–1158. DOI: 10.1080/02626667.2019.1620507.

[21] Günter Blöschl, Julia Hall, Alberto Viglione, et al. "Changing climate both increases and decreases European river floods". In: *Nature* 573.7772 (2019), pp. 108–111. DOI: 10.1038/s41586-019-1495-6.

[22] A. Botta, D. De Donno, and M. G. Perrucci. *The Internet of Things: 20th Tyrrhenian Workshop on Digital Communications*. Springer, 2016.

[23] Gary J. Bowden and Anthony W. Bowman. *Modeling and Analysis of Stochastic Systems*. Chapman and Hall/CRC, 2002.

[24] George E. P. Box and Gwilym M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, 1970.

[25] George E. P. Box et al. *Time Series Analysis: Forecasting and Control, 5th Edition*. Wiley, 2015, p. 712. ISBN: 978-1-118-67502-1.

[26] Pavel Brazdil et al. "Meta-Learning: Applications to Data Mining". In: *Scientific Research* 1.3 (2008), pp. 155–160.

[27] D Brillinger et al. *Robust and Nonlinear Time Series Analysis, vol. 26*. 1984.

[28] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting*. Springer, 2016.

[29] Maria Antonia Brovelli et al. "An overview of current free and open source desktop GIS software". In: *ISPRS International Journal of Geo-Information* 5.5 (2016), p. 66.

[30] Weijian Cao et al. "A combined model of dissolved oxygen prediction in the pond based on multiple-factor analysis and multi-scale feature extraction". In: *Aquacultural Engineering* 84 (Feb. 2019), pp. 50–59. ISSN: 0144-8609. DOI: 10.1016/J.AQUAENG.2018.12.003.

[31] Gema Casal, Clara Cordeiro, and Tim McCarthy. "Using Satellite-Based Data to Facilitate Consistent Monitoring of the Marine Environment around Ireland". In: *Remote Sensing* 14.7 (2022). ISSN: 2072-4292. DOI: 10.3390/rs14071749. URL: https://www.mdpi.com/2072-4292/14/7/1749.

[32] Tianfeng Chai and Roland R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature". In: *Geoscientific Model Development* 7.3 (2014), pp. 1247–1250.

[33] F. J. Chang, Y. H. Chang, and W. C. Lin. "Mining the Customer Credit Using Classification and Regression Tree and Multivariate Adaptive Regression Splines". In: *Expert Systems with Applications* 36.3 (2009), pp. 6365–6372.

[34] Fi-John Chang and Li-Chiu Chang. "Adaptive Neuro-Fuzzy Inference System for Prediction of Water Levels in Rivers". In: *Journal of Hydroinformatics* 14.2 (2012), pp. 292–305.

[35] E. Charou et al. *Integration of Intelligent Techniques for Environmental Data Processing*. Springer, 2003. DOI: 10.1007/978-94-010-0231-8_23.

[36] Chris Chatfield. *The Analysis of Time Series: An Introduction*. CRC Press, 2016.

[37] Surajit Chaudhuri and Umeshwar Dayal. "An overview of data warehousing and OLAP technology". In: *ACM SIGMOD Record* 26 (1 Mar. 1997), pp. 65–74. ISSN: 01635808. DOI: 10.1145/248603.248616. URL: https://dl.acm.org/doi/abs/10.1145/248603.248616.

[38] Nitesh V. Chawla et al. "SMOTE: Synthetic Minority Over-sampling Technique". In: *arXiv: Artificial Intelligence* (June 2011). DOI: 10.1613/JAIR.953.

[39] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: KDD '16 (2016), pp. 785–794. DOI: 10.1145/2939672.2939785. URL: https://doi.org/10.1145/2939672.2939785.

[40] Yimeng Chen et al. "Explore and Exploit the Diverse Knowledge in Model Zoo for Domain Generalization". In: *arXiv.org* (2023).

[41] Kyunghyun Cho et al. "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation". In: *arXiv preprint arXiv:1406.1078* (2014).

[42] Minwoo Cho et al. "Water level prediction model applying a long short-term memory (lstm)–gated recurrent unit (gru) method for flood prediction". In: *Water* 14.14 (2022), p. 2221.

[43] L. Z. Chua, V. M. Babovic, and H. Madsen. "A comparative study of artificial neural networks, decision trees and support vector machines for rainfall–runoff modeling". In: *Advances in Water Resources* 32.6 (2009), pp. 725–738.

[44] Ting Fung Chui and Yiu Chung Law. "Adaptive Neuro-Fuzzy Inference System for Predicting Water Quality Time Series". In: *Engineering Applications of Artificial Intelligence* 20.2 (2007), pp. 193–207.

[45] George Colliat. "OLAP, relational, and multidimensional database systems". In: *ACM Sigmod Record* 25.3 (1996), pp. 64–69.

[46] Fred Collopy and J Scott Armstrong. "Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations". In: *Management science* 38.9 (1992), pp. 1394–1414.

[47] C. B. Cooper et al. "Citizen science as a tool for conservation in residential ecosystems". In: *Ecology and Society* 17.4 (2012), p. 11.

[48] Simon Cox. "OGC Implementation Specification 07-022r1: Observations and Measurements- Part 1 - Observation schema". In: (Jan. 2007).

[49] Can Cui et al. "A recommendation system for meta-modeling: A meta-learning based approach". In: *Expert Systems with Applications* 46 (2016), pp. 33–44.

[50] Christopher W. Dawson, R. J. Abrahart, and Linda M. See. "HydroTest: A Web-Based Toolbox of Statistical Tests for Hydrological Data". In: *Environmental Modelling & Software* 21.12 (2006), pp. 1702–1710.

[51] Anne Denton and Arighna Roy. "Cluster-overlap algorithm for assessing preprocessing choices in environmental sustainability". In: (2017). DOI: 10.1109/BIGDATA.2017.8258447.

[52] Rodolphe Devillers et al. "Thirty years of research on spatial data quality: achievements, failures, and opportunities". In: *Transactions in GIS* 17.4 (2013), pp. 467–484. DOI: 10.1111/tgis.12038.

[53] Kaize Ding et al. "Few-shot network anomaly detection via cross-network meta-learning". In: (2021), pp. 2448–2456.

[54] Q. Y. Duan, S. Sorooshian, and V. K. Gupta. "Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models". In: *Water Resources Research* 28.4 (2012), pp. 1015–1031.

[55] Linh NK Duong et al. "A review of robotics and autonomous systems in the food industry: From the supply chains perspective". In: *Trends in Food Science & Technology* 106 (2020), pp. 355–364.

[56] Ömer Faruk Durdu. "A hybrid neural network and ARIMA model for water quality time series prediction". In: *Engineering Applications of Artificial Intelligence* 23 (4 2010), pp. 586–594. ISSN: 09521976. DOI: 10.1016/j.engappai.2009.09.015.

[57] D. R. Easterling et al. "Climate extremes: Observations, modeling, and impacts". In: *Science* 289.5487 (1997), pp. 2068–2074.

[58] Roghieh Eskandari et al. "Meta-analysis of unmanned aerial vehicle (UAV) imagery for agro-environmental monitoring using machine learning and statistical models". In: *Remote Sensing* 12.21 (2020), p. 3511.

[59] Philippe Esling and Carlos Agon. "Time-series data mining". In: *ACM Computing Surveys (CSUR)* 45.1 (2012), pp. 1–34.

[60] European Space Agency. *Copernicus*. [Website] https://www.esa.int/Applications/Observing_the_Earth/Copernicus. 2020.

[61] B. D. Fulcher and N. S. Jones. "Highly Comparative Feature-Based Time-Series Classification". In: *IEEE Transactions on Knowledge and Data Engineering* 26.12 (2014), pp. 3026–3037.

[62] Salvador Garcıéa, Julián Luengo, and Francisco Herrera. *Data preprocessing in data mining*. Vol. 72. Springer, 2015.

[63] Andrew Gelman et al. *Bayesian Data Analysis*. CRC Press, 2014.

[64] Amin Gharehbaghi et al. "Groundwater level prediction with meteorologically sensitive Gated Recurrent Unit (GRU) neural networks". In: *Journal of Hydrology* 612 (2022), p. 128262.

[65] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT press, 2016.

[66] Dhruvi Gosai and Minal Patel. "Exploring AutoML Libraries: Comparative Evaluation of AutoWEKA, TPOT, H2O, and Auto-Sklearn for Automated Model Development". In: *Journal of Electrical Systems* (2024). DOI: 10.52783/jes.1377.

[67] Rao S. Govindaraju. "Artificial Neural Networks in Hydrology. I: Preliminary Concepts". In: *Journal of Hydrologic Engineering* 5.2 (2000), pp. 115–123.

[68] Alex Graves. "Generating Sequences With Recurrent Neural Networks". In: *arXiv preprint arXiv:1308.0850* (2013).

[69] Jim Gray et al. "Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals". In: *Data Mining and Knowledge Discovery* 1 (1 Jan. 1997), pp. 29–53. ISSN: 13845810. DOI: 10.1023/A:1009726021843. URL: https://dl.acm.org/doi/abs/10.1023/A%3A1009726021843.

[70] James D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

[71] Jiawei Han, Jian Pei, and Hanghang Tong. *Data mining: concepts and techniques*. Morgan kaufmann, 2022.

[72] T. M. Harris and S. J. Ventura. *Fundamentals of remote sensing for GIS users*. ESRI Press, 2017.

[73] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[74] M.A. Hearst et al. "Support vector machines". In: *IEEE Intelligent Systems and their Applications* 13.4 (1998), pp. 18–28. DOI: 10.1109/5254.708428.

[75] S Hegarty et al. "Citizen science and water quality monitoring: evidence from Dublin and Beyond". In: *Filho, W, et al.(eds.), Clean Water and Sanitation* (2020), pp. 1–13.

[76] M. J. Hill, S. L. Johnson, and G. F. McCracken. "Citizen science reveals a continental-scale ecological upheaval". In: *Ecological Monographs* 90.1 (2020), e01368.

[77] Keith W. Hipel and A. Ian McLeod. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, 1994.

[78] Tin Kam Ho. "Random decision forests". In: 1 (1995), pp. 278–282.

[79] Y. C. Ho. "Time series prediction with a hybrid radial basis function neural network model". In: *European Journal of Operational Research* 134.3 (2002), pp. 535–548.

[80] Joachim M. Hoch et al. "Evaluating the impact of model complexity on flood wave propagation and inundation extent with a hydrologic–hydrodynamic model coupling framework". In: *Natural Hazards and Earth System Sciences* 19 (2019), pp. 1723–1735. DOI: 10.5194/nhess-19-1723-2019.

[81] Sepp Hochreiter and Jürgen Schmidhuber. "The vanishing gradient problem during learning recurrent neural nets and problem solutions". In: *Neural Computation* 9.8 (1998), pp. 1735–1780.

[82] Paul R. Houser et al. "Integration of soil moisture remote sensing and hydrologic modeling using data assimilation". In: *Water Resources Research* 34 (12 Dec. 1998), pp. 3405–3420. ISSN: 1944-7973. DOI: 10.1029/1998WR900001.

[83] Tao Hu et al. "A Spatiotemporal Data Fusion Model for Integrating Multiple Remotely Sensed Observations to Monitor Land Cover and Land Use Change". In: *Remote Sensing of Environment* 231 (2019), p. 111222. DOI: 10.1016/j.rse.2019.111222.

[84] William A. Hyman et al. "Summary of research on data collection systems for maintenance management (abridgment)". In: *Transportation Research Record* (1990). URL: https://typeset.io/papers/summary-of-research-on-data-collection-systems-for-2v4zbcds2i.

[85] R. J. Hyndman. "It's Time to Move from What to Why". In: *International Journal of Forecasting* 17.1 (2001), pp. 567–570.

[86] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2021.

[87] Rob J Hyndman and Yeasmin Khandakar. “Automatic time series forecasting: the forecast package for R”. In: *Journal of statistical software* 26.3 (2008), pp. 1–22.

[88] Rob J Hyndman et al. “State space models”. In: *Journal of forecasting* 21.5 (2002), pp. 385–404.

[89] Rob J. Hyndman and George Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, 2018.

[90] Rob J. Hyndman and Anne B. Koehler. “Another look at measures of forecast accuracy”. In: *International Journal of Forecasting* 22.4 (2006), pp. 679–688.

[91] William H. Inmon. *Building the Data Warehouse*. 3rd ed. J. Wiley, 2002, p. 412. ISBN: 978-0471081302.

[92] William H. Inmon. *Building the Data Warehouse*. John Wiley & Sons, 2005.

[93] Günter Jäger, Florian Battke, and Kay Nieselt. “TIALA—time series alignment analysis”. In: (2011), pp. 55–61.

[94] Zhu Jiafeng et al. “Resampling algorithm based on sample similarity and variation coefficient”. In: *Proceedings Article* (Oct. 2022). DOI: 10.1109/ICCASIT55263.2022.9986565.

[95] T. Johnson, P. Westerhoff, and J. Crittenden. “Rethinking water management: innovative approaches to contemporary issues”. In: *American Society of Civil Engineers* (2016).

[96] J. A. Jones et al. “The role of models in ecosystem management”. In: *BioScience* 67.6 (2017), pp. 495–507.

[97] Sven Erik Jørgensen. *Handbook of Environmental Data and Ecological Parameters: Environmental Sciences and Applications*. Vol. 6. Elsevier, 2013.

[98] Heikki Junninen et al. “Methods for imputation of missing values in air quality data sets”. In: *Atmospheric environment* 38.18 (2004), pp. 2895–2907.

[99] Alexandros Kalousis and Melanie Hilario. "Model selection via meta-learning: a comparative study". In: (2000), pp. 406–413.

[100] Alexandros Kalousis and Theoharis Theoharis. "NOEMON: Design, implementation and performance results of an intelligent assistant for classifier selection". In: *Intelligent Data Analysis* 3.5 (1999), pp. 319–337. ISSN: 1088-467X. DOI: https://doi.org/10.1016/S1088-467X(99)00026-8. URL: https://www.sciencedirect.com/science/article/pii/S1088467X99000268.

[101] Maged N Kamel Boulos et al. "Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples". In: *International journal of health geographics* 10 (2011), pp. 1–29.

[102] Yanfei Kang, Rob J. Hyndman, and Kate Smith-Miles. "Visualising forecasting algorithm performance using time series instance spaces". In: *International Journal of Forecasting* 33 (2 2017), pp. 345–358. DOI: 10.1016/j.ijforecast.2016.09.017.

[103] Dmitri Kavetski, Fabrizio Fenicia, and Hubert H. Savenije. "Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory". In: *Water Resources Research* 42.1 (2006).

[104] C. Melisa Kaya, Gokmen Tayfur, and Oguz Gungor. "Predicting flood plain inundation for natural channels having no upstream gauged stations". In: *Journal of Water and Climate Change* 10 (2 2019), pp. 360–372. ISSN: 24089354. DOI: 10.2166/wcc.2017.307.

[105] M. K. Khalil and I. A. Elkhider. "Applying learning theories and instructional design models for effective instruction". In: *Advances in Physiology Education* 40 (2 2016), pp. 147–156. DOI: 10.1152/advan.00138.2015.

[106] Ralph Kimball. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley, 2008.

[107] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[108] O. Kisi and J. Shiri. "Comparison of artificial neural network and wavelet transform models for river flow forecasting". In: *Journal of Hydrology* 349.1-2 (2008), pp. 132–146.

[109] Genshiro Kitagawa. "Non-gaussian state—space modeling of nonstationary time series". In: *Journal of the American statistical association* 82.400 (1987), pp. 1032–1041.

[110] Rob Kitchin. "Big Data, new epistemologies and paradigm shifts". In: *Big data & society* 1.1 (2014), p. 2053951714528481.

[111] Anja Klein and Wolfgang Lehner. "Representing data quality in sensor data streaming environments". In: *Journal of Data and Information Quality (JDIQ)* 1.2 (2009), pp. 1–28.

[112] Ralf Klinkenberg. "Meta-Learning, Model Selection, and Example Selection in Machine Learning Domains with Concept Drift." In: *Meta-learning, Model Selection, and Example Selection in Machine Learning Domains with Concept Drift* (Jan. 2005), pp. 164–171.

[113] Mirko Kück, Sven F. Crone, and Michael Freitag. "Meta-learning with neural networks and landmarking for forecasting model selection an empirical evaluation of different feature sets applied to industry data". In: (2016), pp. 1499–1506. DOI: 10.1109/IJCNN.2016.7727376.

[114] Moritz Kück, Sven F. Crone, and Michael Freitag. "Meta-learning for time series forecasting and forecasting combination". In: *International Journal of Production Research* 54.23 (2016), pp. 7000–7023.

[115] Shuichi Kure and Tadashi Yamada. "Theoretical Derivation of the Conceptual Rainfall-Runoff Models". In: *Journal of Japan Society of Hydrology and Water Resources* 22 (5 2009), pp. 386–400. ISSN: 0915-1389. DOI: 10.3178/JJSHWR.22.386.

[116] Michael H. Kutner, Christopher J. Nachtsheim, and John Neter. *Applied Linear Regression Models*. McGraw-Hill/Irwin, 2004.

[117] Holger Lange and Sebastian Sippel. "Machine learning applications in hydrology". In: *Forest-water interactions* (2020), pp. 233–257.

[118] M. Lawrence. "Why Another Study?" In: *International Journal of Forecasting* 17.1 (2001), pp. 574–575.

[119] Chien-Chang Lee, James Yeongjun Park, and Wan-Ting Hsu. "Bridging expertise with machine learning and automated machine learning in clinical medicine". In: *Annals Academy of Medicine Singapore* (2024). DOI: 10.47102/10.47102/annals-acadmedsg.202481.

[120] Christian Lemke and Bogdan Gabrys. "Meta-learning for time series forecasting and forecast combination". In: (2010), pp. 1079–1084.

[121] Christiane Lemke and Bogdan Gabrys. "Meta-learning for time series forecasting and forecast combination". In: *Neurocomputing* 73.10-12 (2010), pp. 2006–2016.

[122] Christiane Lemke and Bogdan Gabrys. "Meta-Learning for Time Series Model Selection: A Review". In: *Neurocomputing* 400 (2020), pp. 125–138.

[123] Jiawen Li and Tao Zhou. "Evolutionary Multi Agent Deep Meta Reinforcement Learning Method for Swarm Intelligence Energy Management of Isolated Multi Area Microgrid with Internet of Things". In: *IEEE Internet of Things Journal* (2023).

[124] Menggang Li et al. "Multi-source data fusion for economic data analysis". In: *Neural Computing and Applications* 33 (2021), pp. 4729–4739.

[125] Ming Li and Xiaogang Liu. "A Survey of Multi-view Time Series Data Clustering". In: *Artificial Intelligence Review* 30.3-4 (2008), pp. 305–321.

[126] Xin Li et al. "Response of vegetation activity dynamic to climatic change and ecological restoration programs in Inner Mongolia from 2000 to 2012". In: *Ecological Engineering* 120 (2018), pp. 352–362.

[127] Z. Li et al. "Remote sensing of vegetation by fusing multi-source and multi-temporal satellite data: A review". In: *Remote Sensing* 11.9 (2019), p. 1036.

[128] Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*. Vol. 793. John Wiley & Sons, 2019.

[129] Jianguo Liu et al. "Complexity of coupled human and natural systems". In: *Science* 347.6225 (2015), p. 1259865.

[130] X. Liu and J. Xia. "Time series modeling and prediction of river flow data using soft computing". In: *Journal of Hydrology* 344.3-4 (2007), pp. 111–124.

[131] Pasquale De Luca, Ardelio Galletti, and Livia Marcellino. "A GPU-Based Algorithm for Environmental Data Filtering". In: *Lecture Notes in Computer Science* (2022). DOI: 10.1007/978-3-031-08760-8_4.

[132] Wo Shun Luk and Chao Li. "A Partial Pre-aggregation Scheme for HOLAP Engines". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 3181 (2004), pp. 129–137. ISSN: 16113349. DOI: 10.1007/978-3-540-30076-2_13. URL: https://link.springer.com/chapter/10.1007/978-3-540-30076-2_13.

[133] Helmut Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, 2005.

[134] H. R. Maier, G. C. Dandy, and W. K. Foong. "Review of Urban Water Demand Forecasting: Methods and Models". In: *Water Resources Research* 39.9 (2003), p. 1181.

[135] Spyros Makridakis and Michèle Hibon. "The M3-Competition: results, conclusions and implications". In: *International Journal of Forecasting* 16.4 (2000). The M3- Competition, pp. 451–476. ISSN: 0169-2070. DOI: https://doi.org/10.1016/S0169-2070(00)00057-1. URL: https://www.sciencedirect.com/science/article/pii/S0169207000000571.

[136] Pablo Michel Marıén-Ortega et al. “ELTA: new approach in designing business intelligence solutions in era of big data”. In: *Procedia technology* 16 (2014), pp. 667–674.

[137] J. J. McDonnell et al. “Debates-The future of hydrology: An evolving science for a changing world”. In: *Water Resources Research* 50.6 (2014), pp. 5342–5350.

[138] G. A. Meehl et al. “The WCRP CMIP3 multimodel dataset: A new era in climate change research”. In: *Bulletin of the American Meteorological Society* 88.9 (2007), pp. 1383–1394.

[139] Farid Ghareh Mohammadi, Hamid R Arabnia, and M Hadi Amini. “On parameter tuning in meta-learning for computer vision”. In: (2019), pp. 300–305.

[140] Soheila Mehr Molaei and Mohammad Reza Keyvanpour. “An analytical review for event prediction system on time series”. In: (2015), pp. 1–6. DOI: `10.1109/PRIA.2015.7161635`.

[141] Barend Mons et al. “The value of data”. In: *Nature genetics* 43.4 (2011), pp. 281–283.

[142] Alberto Montanari et al. “”Panta Rhei-Everything Flows”: Change in hydrology and society-The IAHS Scientific Decade 2013–2022”. In: *Hydrological Sciences Journal* 58.6 (2013), pp. 1256–1275.

[143] Pablo Montero-Manso et al. “The M5 competition: A proposal for ensemble time series forecasting”. In: *International Journal of Forecasting* 36 (1 2020), pp. 54–74. DOI: `10.1016/j.ijforecast.2019.05.004`.

[144] Douglas C. Montgomery, L. A. Johnson, and John S. Gardiner. *Forecasting and Time Series Analysis*. 2nd. McGraw-Hill, New York, 1990.

[145] Daniel N. Moriasi et al. “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations”. In: *Transactions of the ASABE* 50.3 (2007), pp. 885–900.

[146] Gen Nagatani et al. “Application of Distributed Rainfall Runoff Model to Slope Failure Simulation”. In: *Journal of Japan Society of Civil Engineers, Ser. F5 (Professional Practices in Civil Engineering)* 68 (1 2012), pp. 16–26. ISSN: 2185-6613. DOI: 10.2208/JSCEJPPCE.68.16.

[147] Stefano Nativi et al. “Integration of Earth Observation and in situ data: The GEOSS experience”. In: *ISPRS International Journal of Geo-Information* 4.4 (2015), pp. 2610–2622.

[148] Madhusudanan Navinchandran et al. “Studies to predict maintenance time duration and important factors from maintenance workorder data”. In: (2019).

[149] A. Nguyen and S. Wang. “A review on recurrent neural networks for time series forecasting”. In: *IEEE Access* 7 (2019), pp. 65797–65817.

[150] Fahima Noor et al. “Water level forecasting using spatiotemporal attention-based long short-term memory network”. In: *Water* 14.4 (2022), p. 612.

[151] V. Nourani et al. “A review of hybrid models for daily rainfall–runoff modeling”. In: *Journal of Hydrology* 519 (2015), pp. 1353–1367.

[152] OPW. *Hydro-Data: the Hydrometic Website of the Office of Public Works.* URL: https://waterlevel.ie/hydro-data/.

[153] F.Y. Osisanwo et al. “Supervised Machine Learning Algorithms: Classification and Comparison”. In: *International Journal of Computer Trends and Technology (IJCTT)* 48.3 (June 2017), pp. 128–138. DOI: 10.14445/22312803/IJCTT-V48P126.

[154] R. C. Paiva et al. “A review on the use of hydrologic models in the context of public policies for hydrographic basins”. In: *Revista Brasileira de Recursos Hídricos* 16.2 (2011), pp. 57–68.

[155] Mingyang Pan et al. “Water level prediction model based on GRU and CNN”. In: *Ieee Access* 8 (2020), pp. 60090–60100.

[156] John Paparrizos and Luis Gravano. “k-shape: Efficient and accurate clustering of time series”. In: (2015), pp. 1855–1870.

[157] C. Pfister et al. “A historical perspective on the evolution of the hydrological sciences”. In: *WIREs Water* 4.2 (2017), e1185.

[158] David MW Powers. “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation”. In: *arXiv preprint arXiv:2010.16061* (2020).

[159] Aji Teguh Prihatno, Ida Bagus Krishna Yoga Utama, and Yeong Min Jang. “OneM2M-enabled prediction of high particulate matter data based on Multi-Dense Layer BiLSTM model”. In: *Applied Sciences* 12.4 (2022), p. 2260.

[160] Liudmila Prokhorenkova et al. “CatBoost: unbiased boosting with categorical features”. In: 31 (2018). Ed. by S. Bengio et al.

[161] Ricardo B Prudencio and Teresa B Ludermir. “Meta-learning with decision trees for selection of time series models”. In: (2004), pp. 844–848.

[162] Ricardo B. C. Prudêncio and Teresa B. Ludermir. “Selection of time series forecasting models using meta-learning”. In: 1 (2004), pp. 469–476.

[163] Ricardo BC Prudêncio and Teresa B Ludermir. “Meta-learning approaches to selecting time series models”. In: *Neurocomputing* 61 (2004), pp. 121–137.

[164] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2022. URL: `https://www.R-project.org/`.

[165] Jeffrey S Racine. “Consistent cross-validatory model-selection for dependent data: hv-block cross-validation”. In: *Econometrica* 68.4 (2000), pp. 885–909.

[166] D. J. Reid. *A Comparison of Forecasting Techniques on Economic Time Series*. Birmingham, UK: Operational Research Society and the Society for Long Range Planning, 1972.

[167] Lei Ren et al. “Deep learning for time-series prediction in IIoT: progress, challenges, and prospects”. In: *IEEE transactions on neural networks and learning systems* (2023).

[168] Tomás Robles et al. "An IoT based reference architecture for smart water management processes". In: *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications* 6 (1 2015), pp. 4–23. ISSN: 20935382. DOI: 10.22667/JOWUA.2015.03.31.004.

[169] Richard Roe and Anne Lee. "Handling imbalanced classes in machine learning". In: *Data Science Journal* 8.1 (2019), pp. 67–89.

[170] Yu. I. Rusinovich et al. "Classification of anatomic patterns of peripheral artery disease with automated machine learning (AutoML)". In: *Vascular* (2024). DOI: 10.1177/17085381241236571.

[171] Adam Santoro and et al. "Meta-learning with memory-augmented neural networks". In: 48 (2016), pp. 1842–1850.

[172] Moisés Rocha dos Santos, Leandro Resende Mundim, and André C. P. L. F. de Carvalho. "Evaluation of Error Metrics for Meta-learning Label Definition in the Forecasting Task". In: *Hybrid Artificial Intelligence Systems* (2020). DOI: 10.1007/978-3-030-61705-9_33.

[173] Nicholas I. Sapankevych and Ravi Sankar. "Time Series Prediction Using Support Vector Machines: A Survey". In: *IEEE Computational Intelligence Magazine* 4.2 (2009), pp. 24–38. DOI: 10.1109/MCI.2009.932254.

[174] Shaker H. Ali El-Sappagh, Abdeltawab M. Ahmed Hendawi, and Ali Hamed El Bastawissy. "A proposed model for data warehouse ETL processes". In: *Journal of King Saud University - Computer and Information Sciences* 23.2 (2011), pp. 91–104. ISSN: 1319-1578. DOI: https://doi.org/10.1016/j.jksuci.2011.05.005. URL: https://www.sciencedirect.com/science/article/pii/S131915781100019X.

[175] Jürgen Schmidhuber. "Evolutionary principles in self-referential learning, or on learning how to learn: The meta-meta-... hook". PhD thesis. Technische Universität München, 1987.

[176] Shayle R Searle. *Linear models*. Vol. 65. John Wiley & Sons, 1997.

[177] C. E. Seltzer, S. A. Markolf, and C. Nagy. "Social media as a tool for monitoring and predicting serious social phenomena". In: *PLoS ONE* 14.1 (2019), e0205593.

[178] Nihar Shah and Edmund K Burke. "Forecasting with neural networks: an application using bankruptcy data". In: *Omega* 25.6 (1997), pp. 691–697.

[179] S. Shah. "Metalearning algorithms for real-time strategy game AI". In: 2 (1997), pp. 746–751.

[180] Chaopeng Shen. "A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists". In: *Water Resources Research* 54 (11 Nov. 2018), pp. 8558–8593. ISSN: 1944-7973. DOI: 10.1029/2018WR022643. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2018WR022643%20https://onlinelibrary.wiley.com/doi/abs/10.1029/2018WR022643%20https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2018WR022643.

[181] J. Shiri and A. Akhondi. "Machine learning applications in hydrology: Introduction to the special issue". In: *Journal of Hydrology* 573 (2019), pp. 130–135.

[182] Pavel Shumkovskii et al. "MetaSieve: Performance vs. Complexity Sieve for Time Series Forecasting". In: *IEEE International Conference on Data Mining Workshops (ICDMW)* (2022). DOI: 10.1109/ICDMW58026.2022.00037.

[183] Dimitri P. Solomatine and Krishna N. Dulal. "Model Trees as an Alternative to Neural Networks in Hydrological Modelling". In: *Hydrological Sciences Journal* 53.2 (2008), pp. 247–262.

[184] Robert G. Steel, James H. Torrie, and David A. Dickey. *Principles and Procedures of Statistics: A Biometrical Approach*. McGraw-Hill, 2004.

[185] Stefan Steiniger and Andrew J. Hunter. "Free and open source GIS software for building a spatial data infrastructure". In: *ISPRS International Journal of Geo-Information* 2.2 (2013), pp. 337–359.

[186] Stefan Steiniger and Andrew J.S. Hunter. "The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption". In: *Computers, Environment and Urban Systems* 39 (2013), pp. 136–150.

[187] Stefan Steiniger and Robert Weibel. "An overview on current free and open source desktop GIS developments". In: *International Journal of Geographical Information Science* 23.10 (2009), pp. 1345–1370.

[188] Graeme L. Stephens and Christian D. Kummerow. "The Remote Sensing of Clouds and Precipitation From Space: A Review". In: *Journal of the Atmospheric Sciences* 64 (11 2007), pp. 3742–3765. ISSN: 00224928. DOI: 10.1175/2006JAS2375.1.

[189] "Streaming Data Preprocessing via Online Tensor Recovery for Large Environmental Sensor Networks". In: *ACM Transactions on Knowledge Discovery From Data* (2022). DOI: 10.1145/3532189.

[190] Deqing Sun et al. "A fully-connected layered model of foreground and background flow". In: (2013), pp. 2451–2458.

[191] Ying Sun et al. "Using Machine Learning Algorithms for Estimating Runoff and Sediment Yield in a Semi-Arid Watershed of the Loess Plateau, China". In: *Catena* 135 (2015), pp. 27–40.

[192] Ilya Sutskever, James Martens, and Geoffrey E Hinton. "Generating Text with Recurrent Neural Networks". In: *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*. 2011, pp. 1017–1024.

[193] Souhaib Ben Taieb and Rob J Hyndman. "Bias-Variance Decomposition for Multi-Step-Ahead Time Series Forecasting". In: *Journal of Time Series Analysis* 37.4 (2016), pp. 433–455.

[194] Thiyanga S. Talagala, Feng Li, and Yanfei Kang. "FFORMPP: Feature-based forecast model performance prediction". In: *arXiv: Applications* (2019).

[195] Aqil Tariq and Shujing Qin. "Spatio-temporal variation in surface water in Punjab, Pakistan from 1985 to 2020 using machine-learning methods with time-series remote sensing data and driving factors". In: *Agricultural Water Management* 280 (2023), p. 108228. ISSN: 0378-3774. DOI: https://doi.org/10.1016/j.agwat.2023.108228.

[196] C. J. Tennant et al. "Regional-scale hydrological modelling: representing effects of hydrological connectivity and lateral groundwater flows". In: *Hydrological Processes* 31.20 (2017), pp. 3559–3572.

[197] Sebastian Thrun and Lorien Pratt. *Learning to learn*. Springer Science & Business Media, 2012.

[198] Yingjie Tian, Xiaoxi Zhao, and Wei Huang. "Meta-learning approaches for learning-to-learn in deep learning: A survey". In: *Neurocomputing* 494 (2022), pp. 203–223.

[199] United States Environmental Protection Agency. *Environmental Data Sources*. [Website] https://www.epa.gov/enviro/environmental-data-sources. 2021.

[200] Joaquin Vanschoren and et al. "Meta-learning: A survey". In: *arXiv preprint arXiv:1810.03548* (2018).

[201] Panos Vassiliadis. "A survey of extract–transform–load technology". In: *International Journal of Data Warehousing and Mining (IJDWM)* 5.3 (2009), pp. 1–27.

[202] Anna Vettoruzzo et al. *Advances and Challenges in Meta-Learning: A Technical Review*. 2023. arXiv: 2307.04722 [cs.LG].

[203] Daniel Viviroli et al. "Increasing dependence of lowland populations on mountain water resources". In: *Nature Sustainability* 3.12 (2020), pp. 1058–1065.

[204] Vladimir Vujović. "Development of a custom Data Acquisition System based on Internet of Things". In: (2015).

[205] Hongyu Wang, Kate Smith-Miles, and Rob J Hyndman. "A meta-learning framework for time series forecasting". In: (2009), pp. 629–637.

[206] Jui Teng Wang. "Oversampling-Based Combining Under ISI Channels". In: *IEEE Wireless Communications Letters* (Mar. 2023). DOI: 10.1109/LWC.2023.3234209.

[207] Na Wang et al. "Learning Generalizable Models via Disentangling Spurious and Enhancing Potential Correlations". In: *arXiv.org* (2024).

[208] Xiaoming Wang et al. "Design and Implementation of Spatial Data Integration System Based on Hybrid Storage". In: *Advances in Computer Science and Ubiquitous Computing* 1.1 (2019), pp. 1–7. DOI: 10.22606/acsu.2019.11001.

[209] Yanrong Wang et al. "Clinical outcome of 55 asymptomatic cases at the time of hospital admission infected with SARS-Coronavirus-2 in Shenzhen, China". In: (2020).

[210] Xu Wei et al. "Environment data processing method and device". 2020.

[211] Achmad Widodo and Imam Budi. "Meta-learning for time series forecasting using reduced feature sets". In: *Expert Systems with Applications* 40.5 (2013), pp. 1686–1693.

[212] Daniel S. Wilks. *Statistical Methods in the Atmospheric Sciences*. 3rd. Academic Press, 2011.

[213] J. Xia et al. "Joint control of terrestrial gross primary productivity by plant phenology and physiology". In: *Proc. Natl. Acad. Sci.* 112.9 (2015), pp. 2788–2793. DOI: 10.1073/pnas.1413090112.

[214] Jiaxing Xu et al. "Transformer Based Water Level Prediction in Poyang Lake, China". In: *Water* 15.3 (2023), p. 576.

[215] Min Xu et al. "A Spatiotemporal Data Integration Framework for Water Quality Monitoring in Lakes". In: *Journal of Hydroinformatics* 22.3 (2020), pp. 601–617. DOI: 10.2166/hydro.2020.061.

[216] Tianfang Xu and Feng Liang. “Machine learning for hydrologic sciences: An introductory overview”. In: *Wiley Interdisciplinary Reviews: Water* 8.5 (2021), e1533.

[217] Jiexi Yan et al. “Learning with Diversity: Self-Expanded Equalization for Better Generalized Deep Metric Learning”. In: *Proceedings Article* (2023).

[218] Chengsheng Yang et al. “A Review of Spatio-Temporal Data Models”. In: *ISPRS International Journal of Geo-Information* 7.10 (2018), p. 390. DOI: 10.3390/ijgi7100390.

[219] Z. M. Yaseen et al. “Modelling daily pan evaporation using extreme learning machine, M5 model tree and deep neural network techniques”. In: *Journal of Hydrology* 556 (2018), pp. 865–880.

[220] Wenpeng Yin. “Meta-learning for few-shot natural language processing: A survey”. In: *arXiv preprint arXiv:2007.09604* (2020).

[221] Jaesik Yoon et al. “Bayesian model-agnostic meta-learning”. In: *Advances in neural information processing systems* 31 (2018).

[222] Hao Yu, Yingxiao Du, and Jianxin Wu. “Reviving Undersampling for Long-Tailed Learning”. In: *arXiv.org* (Jan. 2024). DOI: 10.48550/arxiv.2401.16811.

[223] Duo Zhang, Geir Lindholm, and Harsha Ratnaweera. “Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring”. In: *Journal of Hydrology* 556 (Jan. 2018), pp. 409–418. ISSN: 00221694. DOI: 10.1016/J.JHYDROL.2017.11.018.

[224] Y-K Zhang and KE Schilling. “Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: a field observation and analysis”. In: *Journal of Hydrology* 319.1-4 (2006), pp. 328–338.

[225] Hui Zhao et al. “The Design and Implementation of an Environmental Data Integration System Based on OGC Web Services”. In: *Journal of Ambient*

Intelligence and Humanized Computing 9.2 (2018), pp. 515–526. DOI: 10.1007/s12652-017-0501-1.

[226] Xiang Zhou, Yichen Jiang, and Mohit Bansal. "Data Factors for Better Compositional Generalization". In: *arXiv.org* (2023).