Prediction of Blood Glucose using Contextual LifeLog Data

Tenzin Palbar, Manoj Kesavulu, Cathal Gurrin, and Renaat Verbruggen

Dublin City University, Dublin, Ireland
manoj.kesavulu@dcu.ie

Abstract. In this paper, we describe a novel approach to the prediction of human blood glucose levels by analysing rich biometric human contextual data from a pioneering lifelog dataset. Numerous prediction models (RF, SVM, XGBoost and Elastic-Net) along with different combinations of input attributes are compared. An efficient ensemble method of stacking of multiple combination of prediction models was also implemented as our contribution. It was found that XGBoost outperformed three other models and that a stacking ensemble method further improved the performance.

Keywords: Blood Glucose · Lifelogging · Human Context.

1 Introduction

Blood Glucose (BG) level also known as Blood Sugar level, is the concentration of glucose present in the blood. BG value is an important health indicator and when it is in high or low range for a long duration can lead to serious health conditions. There are different human contextual and environmental factors that influence BG level, such as diet, physical activity, emotional state, and other environmental factors. Analysing these parameters and including them in a prediction model can provide valuable insights into the life activities and health context of an individual. Lifelogging [10] and the Quantified Self [15] practice has led to the gathering of vast varieties of contextual human data in different forms, such as PoV images, personal biometrics, physical activity logs, and various other sources [10]. Using such data for personalized prediction models supporting health behaviour recommendations are not yet widely explored in the healthcare and wellness domains. Prior work explores prediction model using limited pre-defined attributes or continuous data from blood glucose monitors, whereas in this paper we use a rich multimodal contextual lifeLog dataset from two LifeLoggers which includes data from various sources and types such as physical activity logs, food consumption images, locations, heart rate and blood glucose monitors. The contribution of this paper is in the evaluation of prediction models such as RF, SVM, XGBoost and Elastic Net Regression along with a novel stacking ensemble approach in the first study of BG prediction using lifelog data.

2 Background and Related Research

The quantified-self movement supports the individual to monitor and enhance their wellness using data [18] and typically such data is gathered with with a health and wellness goal. A lifelog [10] on the other hand, aims to gather a rich multimodal archive of the totality of an individual's life experience, and consequently, data from various sources are combined to form the lifelog, which promises to bring health benefits and lifestyle benefits. One such source of quantified self or lifelog data is BG level. In past work, a variety of approaches has been employed to predict BG values from various wearable sensor sources [21]. Examining only historical BG data, Marcus et al. [13], performed Kernel Ridge Regression (KRR) on Continuous Glucose Monitor (CGM) data and Martinsson et al. [14], presented a BG prediction model based on a Recurrent Neural Network (RNN) with a prediction horizon of 60 minute. Alfian et al. [2], utilized time-domain features as additional attributes for the proposed Artificial Neural Networks (ANN) based prediction model to improve accuracy with CGM as the single input.

Recently, we have seen the that various other sensor sources, such as food intake, physical activity, and stress levels have been demonstrated effective in studies [12, 25]. Takeuchi et al. [19], used time-series blood-sugar level data to analyse the relationship to lifestyle events (food ingestion, alcohol intake, and exercise) using manual logging and wearable sensors. El Idrissi et al. [11], presented a deep learning neural network (NN) model for BG prediction using a Long-Short-Term Memory (LSTM) layer with two fully connected dense layers showing promising results. Zecchin et al. [24], proposed a jump neural network prediction algorithm with a 30-minute horizon that exploits historical CGM data and consumed carbohydrate information. The prediction model showed accurate and comparable results to those obtained by the previously proposed feed-forward neural network (FNN) and first-order polynomial model approach [26]. Zarkogianni et al. [23], performed a comparative assessment of four glucose prediction models based on FNN, SOM, a neurofuzzy network with wavelets as activation functions (WFNN), and a linear regression model (LRM), using BG and physical activity data, with SOM demonstrating best performance. Munoz-Organero et al.[16], used simulated and real datasets consisting of CGM, meals and insulin boluses (slow and fast acting insulin). LSTM based RNN was applied to learn the carbohydrate digestion and insulin absorption processes from each parameter. Finally Georga et al. [6], used multiple input variables such as glucose concentration, the energy expenditure, the time of the day, meal, and insulin intake. Summarising past work, NNs and Auto Regression (AR) are the most popular techniques.

In this paper, we will evaluate the effect on BG prediction of employing a wide set of multimodal features from a lifelog dataset, which we consider to be the ultimate form of user context at present. We will evaluate the significance of the input variables and perform comparison of prediction models such as RF, SVR, XGBoost, Elestic-Net Regression and our own contribution in terms of a stacked ensemble model. To the best of our knowledge, this is first BG prediction study using LifeLog data.

3 Dataset

In this work we use the NTCIR-14 dataset, which is a rich multimodal lifelog dataset that was created for the NTCIR-14 Lifelog Retrieval challenge in 2019. The NTCIR-14 Lifelog dataset [9] consists of multimodal lifelog data over 42 days from two active lifeloggers. The data consists of multimedia data with wearable camera PoV images (two per minute, about 1,500 images per day). It includes a time-stamped record of music listened to and an archive of all conventional digital photos of food consumed (for examples, see Fig. 1). Additionally all-day time-aligned biometric data (heart rate, calorie burn, steps taken and distance moved) was provided using FitBit fitness trackers, continuous blood glucose monitoring (CGM) with readings of every 15 minutes was captured using a wearable sensor. Additionally daily activities of the lifeloggers were captured in terms of the semantic locations visited and time-stamped diet log of all food and drink consumed by the lifeloggers throughout the day were manually annotated. This dataset has been widely used for tasks such as activity detection, semantic event retrieval and analysis to gain insights into the lifelogger daily life routine. Some outliers and missing values were observed in the dataset, which are described in the next section. Both univariate and multivariate imputation approaches were used to estimate features with missing values. Ordinary Least Squares (OLS) was used to analyse the significance of each attribute to the missing data containing attribute.

Fig. 1. Examples of Wearable Camera Images of Food from the NTCIR-14 Dataset [9].

4 Data Processing

As the dataset consists of a rich number of attributes that could have high influence on BG variability, we will be using a variety data sources to bring new insights on the factors influencing BG level and possible improvements of prediction models.

4.1 Data Understanding and Cleaning

As the dataset consists of real-world data gathered from a LifeLogger wearable devices and manual records, it contains some missing values. Multiple techniques were compared and used to handle missing data in the dataset. Univariate and Multivariate imputation techniques were applied in this work. Multivariate imputation strategy imputes missing values by modelling each feature with missing values as a function of other features. A univariate approach, which utilizes the attribute itself, was used for location name feature imputation. The following imputation methods were used to handle the observed missing values.

Heart rate The heart rate attribute in LifeLogger 1 dataset had 26% missing values. We split the non-missing heart_rate values into training (70%) and test set (30%), and used the test portion to compare multiple imputation techniques (Bayesian Ridge, Linear interpolation, Decision Tree, Extra Trees and k-Nearest Neighbors).

Table 1. Heart Rate Imputation Methods.

Imputation Method	MSE	RMSE	R-Squared
Bayesian Ridge	176.09	13.27	0.22
k-Nearest Neighbors	41.69	6.46	0.81
Decision Tree Regressor	34.01	5.83	0.85
ExtraTrees Regressor	**32.44**	**5.69**	**0.86**

We evaluated the performance of four imputation methods using non-missing features (heart rate, steps, calories, distance and carbohydrate intake) and found that Extra Trees Regressors performed best, which aligns with similar efforts [17]. Table 1 shows the four approaches and we used an Extra Trees method to impute heart rate missing values.

As the historic glucose attribute was recorded on average every 15 minutes, the remaining attributes with continuous variables were averaged according to historic glucose except for carbohydrate intake. The sum of carbohydrate intake in grams (g) every 15 minutes were considered.

4.2 Carbohydrate Estimation

Images with time stamp for everyday meal intake were used to extract information about carbohydrate intake in grams (g). To calculate the carbohydrate values of each meal in the dataset, we consulted the Food and Nutrient Database for Dietary Studies (FNDDS) 2017-2018 issued by U.S. Department of Agriculture, Agricultural Research Service [1]. The database consists of the nutrient amount per 100g of edible portion for 64 nutrients, for more than 7,000 main food descriptions. The food portion estimation was performed using portions and weights estimate guidelines provided in the FNDDS database. Fig. 2, outlines the food annotation and Table 2 presents the food description and other variables used from the FNDDS database for carbohydrate estimation. In future work, acquiring a food intake dataset with accurate portion and food class information can provide better estimation of carbohydrate values.

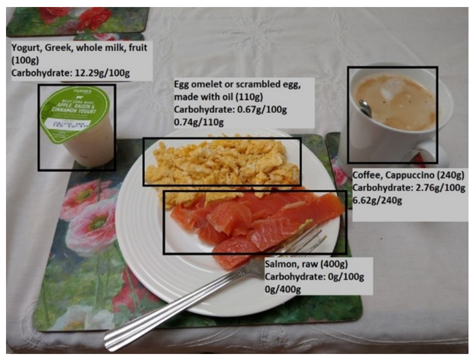

Fig. 2. Diet Log Image food annotation for Carbohydrate Estimation [9].

Table 2. Carbohydrate Intake Estimation Variable.

Variable	**Description**	**Example Values**
main_food_description	Identified from diet log image	Coffee, Potato, Chocolate
amt_g	Food amount in grams	240, 400
amt_consumed	Amount Consumed (%)	80, 100
carbs_100g	Carbohydrate per 100grams	2.76, 12.29
total_carbs	Total carbohydrate (grams)	12.29, 6.62

4.3 Additional Feature Extraction

For physical activity metadata, activities were segmented the data into three categories, 'physical activity' which represented physical activities and transport activities, 'sedentary', which was derived from distance and step features, and a sleep category which was derived from motion and (lack-of) camera features. Additionally, each minute of each day was used to derive three categorical features, ‘day of the week’, ‘time of the day’ and ‘day of the month’ to capture the potential daily and weekly periodicity of BG level. Binarization of all categorical features were performed with the one hot encoding technique. The features such as heart rate, calories burned, steps, distance, carbohydrate intake, location, physical activity, time of the day, day of the week, day of the month were used as predictor variables for the models.

5 Prediction Methods

As mentioned earlier, we employ four basic (pre-existing) models from the related literature to represent and we also propose our own new ensemble technique (Stacked BG Detector).

5.1 State-of-the-Art Prediction Methods

Random Forest There may be potentially irrelevant and redundant features in the proposed dataset either from raw data (distance and steps count) or derived features (day of the week and time of the day). Moreover, the relevant features set may differ from individual to individual, so feature selection with pre-defined rules would not advisable. Hence Random Forest (RF), which has been proved to be robust with redundant and irrelevant features [4] was employed. RF is an ensemble learning method that creates a forest with multiple decision trees. It averages the prediction of the individual tree, providing a better result compared to single decision trees, which are prone to over-fitting to the training set.

Support Vector Regression Support Vector Regression (SVR) uses the same principle as Support Vector Machine (SVM), which is a margin-based classifier that uses a hyperplane to separate feature classes. SVR helps in deciding a decision boundary at some distance away from the hyperplane where the data points within the decision boundary has least error rate are considered. SVR acknowledges the presence of non-linearity in the data and provides a proficient prediction model. We choose SVR kernel-based method as one of our prediction models as they have been shown to perform well over the full range of glucose values in previous studies [6, 7].

XGBoost XGBoost (XGB) is short for eXtreme Gradient Boosting. It implements the gradient boosting decision tree algorithm, an ensemble technique where new models are added to correct the errors made by the prior models [5]. Gradient Boosting uses a gradient descent algorithm to minimize the loss when adding new models. Alfian et. al. [3] developed a model based on XGB to predict BG. The model was shown to outperform Multi-Layer Perceptron (MLP), KNN, DT, SVR, RF, and AdaBoost models.

Elastic-Net Elastic-Net Regression uses regularization to reduce the magnitude of coefficients of features in a regression model to avoid prediction dependency on specific features with extreme coefficient values, and is especially useful at dealing with the presence of correlation between features. Zanon et al. [22] used regularization-based techniques such as Lasso, Ridge, and Elastic-Net regression, which proved effective at predicting BG levels. Thus, we used Elastic-Net in this work with hyper parameter tuning using 10-fold cross validation on the training set. Cost functions of the regularization techniques are calculated as follows.

Elestic-Net Regression:

$$min\left(\left||Y - X\theta|\right|_2^2 + \lambda_1||\theta||_1 + \lambda_2||\theta||_2^2\right) \tag{1}$$

Ridge Regression:

$$min\left(\left||Y - X\theta|\right|_2^2 + \lambda||\theta||_2^2\right) \tag{2}$$

Lasso Regression:

$$min\left(\left||Y - X\theta|\right|_2^2 + \lambda||\theta||_1\right) \tag{3}$$

Stacked BG Detector Stacking or Stacked Generalization is an ensemble machine learning algorithm which involves using a meta learner which is a machine learning model to learn how best to combine the output of the base learners [20]. A stacking implementation is not widely used in previous BG prediction studies, thus this paper introduces stacking ensemble approach as the primary contribution of this work.

Preprocessing pipelines were designed for each prediction model that included one-hot encoding for categorical data and standardization for numeric data where applicable. Hyperparameter tuning was performed on each model with girdSearch 10-fold cross validation with repetition on the training dataset. In this study, we compared multiple combinations of base models and observed the model performance. Out of all combinations, two stacking models comprised of a combination of basic models are outlined in this paper. We decided to use the above mentioned four models as base models for the stacking model. The first stacking model with simple Linear Regression (LR) model as meta-model because most of the work is already done by the base learners and simple linear model usually works well as meta-learners [20]. For the second stacking model, we used Ridge Regression, a variant of LR as a meta-model because Ridge Regression can restrict the influence of predictor variables (output of base-model) over the output variable by compressing their coefficients.

5.2 Evaluation

The dataset was partitioned such that 80% of the data was used for training, and the rest of the data (20%) was used for testing. The training set was used to perform hyperparameter tuning on each predictive model using 10-fold cross-validation. The developed models were then utilized to predict the test data. Two performance metrics were used in this study to evaluate each model. Root mean square error (RMSE) and mean square error (MSE) were used as evaluation metrics. RMSE is better in terms of reflecting performance when dealing with large error values and when dealing with domains where slight increase in fall in error can have a huge impact. MSE score was used for hyperparameter tuning evaluation on the training dataset. It is widely accepted in the diabetes research community. [3, 22].

6 Results Analysis and Discussion

This section presents the evaluation results for both the basic models and stacking systems. For each prediction model, grid search was performed to tune the hyperparameters.The hyperparameters rendering the model with the lowest MSE across all 10-fold cross-validated training dataset was then saved, and the corresponding model was used later to perform prediction on the testing dataset. In the below discussion we present the average performance of each model as well as considering the performance for each of the two lifeloggers separately, since BG prediction can differ across individuals.

Base Methods The results of the RMSE and MAE of the basic predictive models for the prediction on both validation and test sets are displayed in Table 3 In addition to model performance on the individual LifeLogger dataset, the

Table 3. Evaluation Table of the Basic Prediction Models.

Life Logger	**Basic Model**	**Evaluation**			
		MSE(mmol/L)		**MSE(mmol/L)**	
		Validation	**Test**	**Validation**	**Test**
L1	RF	0.56	0.56	0.75	0.75
	SVR	0.77	0.83	0.88	0.91
	XGBoost	**0.51**	**0.5**	**0.72**	**0.71**
	Elastic-Net	1.08	1.07	1.04	1.03
L2	RF	**0.55**	**0.39**	**0.74**	**0.63**
	SVR	0.73	0.67	0.85	0.82
	XGBoost	0.61	0.47	0.78	0.68
	Elastic-Net	0.99	0.94	0.99	0.97
Average	RF	**0.56**	**0.48**	**0.75**	**0.69**
	SVR	0.75	0.75	0.86	0.86
	XGBoost	0.56	0.49	0.75	1
	Elastic-Net	1.04	1	1.02	0.69

average of RMSE and MSE of both L-1 and L-2 are outlined. XGBoost showed the best performance among the basic models as it combines the advantages from both RF and gradient boosting. XGBoost was also observed to outperform other models in literature [22]. RF model showed better result with L-2 dataset due to lesser number of features and samples available compared to L-1.

Novel Stacking Model In Table 4, we can observe that the evaluation result of the stacking ensemble method shows a performance improvement compared to individual base models. As the stacked regressor combines the strength of different base regressors, there was a decrease in both MSE, the average squared difference between the predicted and the actual value and RMSE, the squared root of MSE. Lower MSE and RMSE score indicates better performance of our model with unseen test data. We have used k-fold cross validation to handle possible overfitting. Fig. 3, illustrates the MSE and RMSE values for all four base model and stacked model 2 on both LifeLogger datasets.

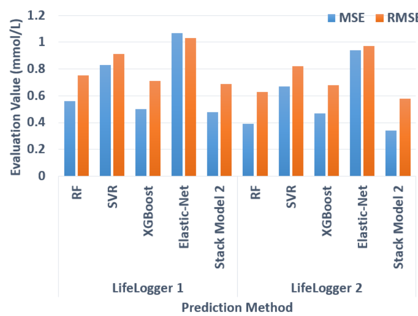

Fig. 3. Evaluation Result.

The overall performance was further improved by considering the predictions from all base models as features to train a meta-learner L-1 (MSE: 0.48, RMSE: 0.69) and L-2 (MSE: 0.34, RMSE: 0.58). Thus, the combined learner was able to make more accurate predictions on the test set compared to individual basic models. Figure 3 shows the results of all models for each lifelogger.

Observations on Performance We observed that three readily available data sources from off-the-shelf fitness trackers (heart rate, calories burned and step count) to have highest contribution to model performance and therefore are the most useful features. Carbohydrate intake also (unsurprisingly) has a very high influence on BG variation. Since we had performed food and portion classification and used the FNDDS database [1] for carbohydrate estimation, there is a significant opportunity to fully-automate this process using wearable cameras.

7 Conclusion and Future Work

In this work we compared the effectiveness of various basic BG predictor models to a new ensemble model when operating over the rich contextual lifelog data from two lifeloggers. The LifeLog dataset includes many forms of multimodal data captured continuously during 42 days. To the best of our knowledge, this is

Table 4. Model Evaluation Table for Stacked system.

Life Logger	Stacking System	Evaluation			
		MSE(mmol/L)		MSE(mmol/L)	
		Validation	Test	Validation	Test
L1	Stack Model 1	**0.50**	**0.48**	**0.71**	**0.69**
	Stack Model 2	**0.50**	**0.48**	**0.71**	**0.69**
L2	Stack Model 1	0.49	0.34	0.71	0.58
	Stack Model 2	**0.49**	**0.34**	**0.70**	**0.58**
Average	Stack Model 1	0.49	0.41	0.71	0.64
	Stack Model 2	0.49	0.41	0.70	0.64

the first BG prediction study on a rich contextual LifeLog dataset. Unlike previous studies that used predefined attributes or only CGM, we have performed statistical analysis on a multimodal attributes available in the LifeLog dataset and enriched the dataset by extracting features from LifeLog images. Multiple prediction models were implemented on the LifeLog extracted attributes and their performance were compared using RMSE and MSE evaluation metrics. Initially, four regression models were trained and validated on the training dataset using 10-fold cross validation, resulting in selection of hyperparameters for each model. The four models were then used to make final predictions on the test set. Predictions from the basic models were fed as features to a regression to build two stacked ensemble approaches, which improved BG prediction and forms the main contribution of this paper.

This work has some limitations, such as a small dataset, but it is the only one available which has such rich contextual data. Data from multiple participants could provide us better insight on intra-individual variability of influence of attributes on individual BG. Another issue is the lack of a well-defined approach to estimate carbohydrate intake, which can reduce the performance of the models, and lead to errors, as was also found in [21]. In future work, inclusion of richer contextual features and a larger dataset can give us more scope to perform analysis. In addition, stress is known to have a high correlation to BG variation [8], so it can be included as another feature. Finally, a better approach for carbohydrate estimation adoption can improve predictive performance [21] and recent advances in wearable cameras (from Facebook) can facilitate more data for better carbohydrate estimation.

References

1. U.S. Department of Agriculture, A.R.S.: Food and nutrient database for dietary studies (fndds). In: FoodData Central. Food Surveys Research Group, Beltsville Human Nutrition Research Center (2017), http://www.ars.usda.gov/nea/bhnrc/fsrg
2. Alfian, G., Syafrudin, M., Anshari, M., Benes, F., Atmaji, F.T.D., Fahrurrozi, I., Hidayatullah, A.F., Rhee, J.: Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features. Biocybernetics and Biomedical Engineering **40**(4), 1586–

1599 (2020). https://doi.org/https://doi.org/10.1016/j.bbe.2020.10.004, https://www.sciencedirect.com/science/article/pii/S0208521620301248

3. Alfian, G., Syafrudin, M., Rhee, J., Anshari, M., Mustakim, M., Fahrurrozi, I.: Blood glucose prediction model for type 1 diabetes based on extreme gradient boosting. IOP Conference Series: Materials Science and Engineering **803**, 012012 (may 2020). https://doi.org/10.1088/1757-899x/803/1/012012, https://doi.org/10.1088/1757-899x/803/1/012012
4. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001). https://doi.org/10.1023/A:1010933404324, http://dx.doi.org/10.1023/A%3A1010933404324
5. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. p. 785–794. KDD '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2939672.2939785, https://doi.org/10.1145/2939672.2939785
6. Georga, E., Protopappas, V., Polyzos, D., Fotiadis, D.: Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models. Medical & biological engineering & computing **53** (03 2015). https://doi.org/10.1007/s11517-015-1263-1
7. Georga, E.I., Protopappas, V.C., Ardigò, D., Polyzos, D., Fotiadis, D.I.: A glucose model based on support vector regression for the prediction of hypoglycemic events under free-living conditions. Diabetes Technology & Therapeutics **15**(8), 634–643 (2013). https://doi.org/10.1089/dia.2012.0285, https://doi.org/10.1089/dia.2012.0285, pMID: 23848178
8. Goetsch, V.L., Wiebe, D.J., Veltum, L.G., van Dorsten, B.: Stress and blood glucose in type ii diabetes mellitus. Behaviour Research and Therapy **28**(6), 531–537 (1990). https://doi.org/https://doi.org/10.1016/0005-7967(90)90140-E, https://www.sciencedirect.com/science/article/pii/000579679090140E
9. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Ninh, V.T., Le, T.K., Albatal, R., Dang-Nguyen, D.T., Healy, G.: Advances in lifelog data organisation and retrieval at the ntcir-14 lifelog-3 task. In: Kato, M.P., Liu, Y., Kando, N., Clarke, C.L.A. (eds.) NII Testbeds and Community for Information Access Research. pp. 16–28. Springer International Publishing, Cham (2019)
10. Gurrin, C., Smeaton, A.F., Doherty, A.R.: Lifelogging: Personal big data. Foundations and Trends® in Information Retrieval **8**(1), 1–125 (2014). https://doi.org/10.1561/1500000033, http://dx.doi.org/10.1561/1500000033
11. Idriss, T.E., Idri, A., Abnane, I., Bakkoury, Z.: Predicting blood glucose using an lstm neural network. In: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS). pp. 35–41 (2019). https://doi.org/10.15439/2019F159
12. Manohar, C., Levine, J.A., Nandy, D.K., Saad, A., Dalla Man, C., McCrady-Spitzer, S.K., Basu, R., Cobelli, C., Carter, R.E., Basu, A., Kudva, Y.C.: The effect of walking on postprandial glycemic excursion in patients with type 1 diabetes and healthy people. Diabetes Care **35**(12), 2493–2499 (2012). https://doi.org/10.2337/dc11-2381, https://care.diabetesjournals.org/content/35/12/2493
13. Marcus, Y., Eldor, R., Yaron, M., Shaklai, S., Ish-Shalom, M., Shefer, G., Stern, N., Golan, N., Dvir, A.Z., Pele, O., Gonen, M.: Improving blood glucose level predictability using machine learning. Diabetes/Metabolism Research and Reviews **36**(8), e3348 (2020). https://doi.org/https://doi.org/10.1002/dmrr.3348, https://onlinelibrary.wiley.com/doi/abs/10.1002/dmrr.3348

14. Martinsson, J., Schliep, A., Eliasson, B., Mogren, O.: Blood glucose prediction with variance estimation using recurrent neural networks. Journal of Healthcare Informatics Research **4** (03 2020). https://doi.org/10.1007/s41666-019-00059-y
15. Meyer, J., Simske, S., Siek, K.A., Gurrin, C.G., Hermens, H.: Beyond quantified self: Data for wellbeing. In: CHI '14 Extended Abstracts on Human Factors in Computing Systems. p. 95–98. CHI EA '14, Association for Computing Machinery, New York, NY, USA (2014)
16. Munoz-Organero, M.: Deep physiological model for blood glucose prediction in t1dm patients. Sensors **20**(14) (2020). https://doi.org/10.3390/s20143896, https://www.mdpi.com/1424-8220/20/14/3896
17. Suresh, M., Taib, R., Zhao, Y., Jin, W.: Sharpening the blade: Missing data imputation using supervised machine learning. In: Liu, J., Bailey, J. (eds.) AI 2019: Advances in Artificial Intelligence. pp. 215–227 (2019)
18. Swan, M.: The quantified self: Fundamental disruption in big data science and biological discovery. Big data **1 2**, 85–99 (2013)
19. Takeuchi, H., Kodama, N., Tsurumi, K.: Time-series data analysis of blood-sugar level of a diabetic in relationship to lifestyle events. In: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 5195–5198 (2009). https://doi.org/10.1109/IEMBS.2009.5334582
20. Witten, I.H., Frank, E.: Data mining: Practical machine learning tools and techniques with java implementations. SIGMOD Rec. **31**(1), 76–77 (Mar 2002). https://doi.org/10.1145/507338.507355, https://doi.org/10.1145/507338.507355
21. Woldaregay, A.Z., Årsand, E., Walderhaug, S., Albers, D., Mamykina, L., Botsis, T., Hartvigsen, G.: Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. Artificial Intelligence in Medicine (2019)
22. Zanon, M., Sparacino, G., Facchinetti, A., Talary, M.S., Caduff, A., Cobelli, C.: Regularised Model Identification Improves Accuracy of Multisensor Systems for Noninvasive Continuous Glucose Monitoring in Diabetes Management. Journal of Applied Mathematics **2013**(SI05), 1 – 10 (2013). https://doi.org/10.1155/2013/793869, https://doi.org/10.1155/2013/793869
23. Zarkogianni, K., Mitsis, K., Litsa, E., Arredondo, M., Fico, G., Fioravanti, A., Nikita, K.: Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring. Medical & biological engineering & computing **53** (06 2015). https://doi.org/10.1007/s11517-015-1320-9
24. Zecchin, C., Facchinetti, A., Sparacino, G., Cobelli, C.: Jump neural network for online short-time prediction of blood glucose from continuous monitoring sensors and meal information. Computer Methods and Programs in Biomedicine **113**(1), 144–152 (2014). https://doi.org/https://doi.org/10.1016/j.cmpb.2013.09.016, https://www.sciencedirect.com/science/article/pii/S0169260713003234
25. Zecchin, C., Facchinetti, A., Sparacino, G., Dalla Man, C., Manohar, C., Levine, J.A., Basu, A., Kudva, Y.C., Cobelli, C.: Physical activity measured by physical activity monitoring system correlates with glucose trends reconstructed from continuous glucose monitoring. Diabetes Technology & Therapeutics **15**(10), 836–844 (2013). https://doi.org/10.1089/dia.2013.0105, https://doi.org/10.1089/dia.2013.0105, pMID: 23944973
26. Zecchin, C., Facchinetti, A., Sparacino, G., De Nicolao, G., Cobelli, C.: Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration. IEEE Transactions on Biomedical Engineering **59**(6), 1550–1560 (2012). https://doi.org/10.1109/TBME.2012.2188893