

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

Customer Purchase Behavior Prediction in E-commerce: A Conceptual Framework and Research Agenda *

Douglas Cirqueira¹[0000-0002-1283-0453], Markus Hofer², Dietmar Nedbal³[0000-0002-7596-0917],
Markus Helfert⁴[0000-0001-6546-6408], Marija Bezbradica¹[0000-0001-9366-5113]

¹ Dublin City University, Dublin, Ireland

`douglas.darochacirqueira2@mail.dcu.ie`

² Raiffeisenlandesbank Oberösterreich, Linz, Austria

³ University of Applied Sciences Upper Austria, Steyr, Austria

⁴ Maynooth University, Maynooth, Ireland

Abstract. Digital retailers are experiencing an increasing number of transactions coming from their consumers online, a consequence of the convenience in buying goods via E-commerce platforms. Such interactions compose complex behavioral patterns which can be analyzed through predictive analytics to enable businesses to understand consumer needs. In this abundance of big data and possible tools to analyze them, a systematic review of the literature is missing. Therefore, this paper presents a systematic literature review of recent research dealing with customer purchase prediction in the E-commerce context. The main contributions are a novel analytical framework and a research agenda in the field. The framework reveals three main tasks in this review, namely, the prediction of customer intents, buying sessions, and purchase decisions. Those are followed by their employed predictive methodologies and are analyzed from three perspectives. Finally, the research agenda provides major existing issues for further research in the field of purchase behavior prediction online.

Keywords: Consumer Behavior, Purchase Prediction, Behavior Analytics, Machine Learning, Data Mining, E-commerce, Digital Retail.

1 Introduction

Daily online activities generate plenty of opportunities for businesses to understand their consumer behavior in E-commerce platforms [1]. Indeed, consumers around the globe purchased \$2.86 trillion on the web in 2018, which represented an 18% growth¹ in online sales compared to the \$2.43 trillion sold in 2017. According to predictions

* This research was supported by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 765395; the industry partner Raiffeisenlandesbank Oberösterreich AG; and supported, in part, by Science Foundation Ireland grant 13/RC/2094.

¹ Digital Commerce 360, Global E-commerce Sales 2019. <https://www.digitalcommerce360.com/article/global-ecommerce-sales/>

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

of the purchasing behavior of customers, companies aim to anticipate their needs and provide personalized services [2, 3].

However, consumer behavior itself is well known as a complex pattern among the data mining community [4]. Aiming to predict the likelihood of such patterns, researchers were applying multiple probabilistic and machine learning (ML) statistical models to historical online customer's data, resulting in somewhat reliable probabilities to predict the next customer's steps [5, 6]. That has also increased the complexity of analyzing this literature, given the multiple approaches and datasets available. Previous reviews and surveys related to this topic have usually focused on the specific literature of recommendation systems [7, 8, 9, 10]. On the other hand, our focus is on reducing complexity for understanding the step before recommendations, which is the prediction of customer's next purchases, and in visualizing research opportunities in the field.

Therefore, these paper contributions are a novel conceptual framework for analysis and a research agenda. The framework systematically maps this literature regarding datasets adopted, predictive methods, and tasks with their applications. Specifically, the framework reveals three main tasks, namely, prediction of buying sessions, purchase decisions, and customer intents. Next, it provides eight applications enabled by each task. Finally, it illustrates three perspectives on predictive methodologies, and a research agenda with future work opportunities in the field.

The rest of this paper is organized as follows: Section 2 describes the research methodology of the literature review; Section 3 presents results and the main contributions, followed by final remarks in Section 4.

2 Research Methodology

To provide the framework and research agenda proposed, we performed a literature review following systematic guidelines from Watson (2002) [11] and Kitchenham, et al. (2009) [12]. Inspired by [13], two research questions and a search query were developed to collect comprehensive literature within the research scope of purchase prediction in E-commerce. Then, the search query was applied in the following scientific databases, well known for containing literature in the field of behavior analytics: Scopus, Web of Science, Science Direct, EBSCO Host (Business Source Complete and Academic Search Complete), Emerald, IEEE Xplore, Association of Information Systems (AIS) library and ACM Digital Library.

— **Search Query:** "(consumer or customer) AND (purchas* OR buy* OR sale* OR shop* OR behavi*) AND (predict* OR forecast*)"

The searches were performed in the abstract field, except for the Web of Knowledge (abstract title and keywords were used) and AIS libraries (full text was used), due to the characteristics of their search engines. The search period has covered papers from 2014 to 2019, only in the English language, which has provided a total of 9824 exported proposals. The next step removed duplicates and had an inclusion filter

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

only to retrieve papers focused on the problem of consumer purchase behavior prediction. That has provided a total of 429 papers.

Next, the exclusion criteria were applied to remove papers not focused on the E-commerce context. At this stage, the total of papers kept was 35. Based on those proposals, backward and forward searches were conducted via Google Scholar, adding 18 and 10 studies, respectively. The final number of papers for extraction and mapping steps was 63. All those results are available at a Github repository (<https://github.com/dougcirqueira/literature-review-purchase-prediction>).

3 Results

Tables 1 and 2 provide non-exhaustive lists of the proposals selected for this literature review. Table 1 brings single task proposals (prediction of one outcome), while Table 2 provides multi-task proposals (prediction of multiple outcomes).

Table 1. Selected proposals in single task settings (A: Aggregation; R: Rule; P: Personalized Function; L: Learning; CDM: Classical Data Mining; PC: Probabilistic Classifier; DLC: Deep Learning Classifier; CF: Collaborative Filtering)

| Ref | Task | Focused Applications | Feature Construction | | | | Predictive Method | Contribution and Targeted Research Gap | |
|------|----------------------|----------------------|-----------------------|-------|---|---|---|---|---|
| | | | A | R | P | L | | | |
| [14] | Buying Session (PBS) | B, C, D, F | x | | | | CDM | Real-time predictions in single visits | |
| [15] | | | x | | | | | Feature engineering for clickstream | |
| [16] | | | x | x | | | | Association rules for fast predictions | |
| [17] | | | x | | x | | | Feature engineering for popular products | |
| [18] | | | x | x | x | | | Feature engineering from customer search | |
| [19] | | | x | | | | | Benchmark over multiple online shops | |
| [20] | | | x | x | | | | Feature engineering for multiple products | |
| [21] | | | x | x | | | | Feature engineering with graph metrics | |
| [22] | | | x | x | | | | K-Nearest Neighbor for fast predictions | |
| [23] | | | x | | x | | | Feature engineering with motifs in single sessions | |
| [24] | | | x | x | | | | PC | Prediction over multiple online visits |
| [25] | | | x | | | x | | DLC | Feature learning for automatic feature construction |
| [26] | | | x | x | | x | | | Real-time predictions in single visits |
| [27] | | | x | x | | | | | Handle of data unbalancing |
| [28] | Product (PPD) | A,B,C,E,F,G | x | | | | CDM | Feature engineering with recency and frequency of page views | |
| [29] | | | x | x | | | CF | Combination of features from clickstream and transactions for collaborative filtering | |
| [30] | | | x | | | | | Feature engineering for clickstream | |
| [31] | | | x | | | | | Combination of features from clickstream and transactions for collaborative filtering | |
| [32] | | | x | | | | | Feature engineering with product heterogeneity for collaborative filtering | |
| [33] | | | x | x | | x | DLC | Real-time predictions with ensemble and deep learning | |
| [34] | | | | | | x | | Recommendation of bundles of products considering quality and diversity criteria | |
| [35] | | | Purchase Intent (PCI) | B,D,H | x | | x | | CDM |
| [36] | x | | | | | | Predict the intensity of user intent | | |
| [37] | x | x | | | | | Predict intent before an online visit | | |
| [38] | x | | | | x | x | Predict intent before/during the online visit | | |

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

Table 2. Selected proposals in multi-task settings (A: Aggregation; R: Rule; P: Personalized Function; L: Learning; CDM: Classical Data Mining; PC: Probabilistic Classifier; DLC: Deep Learning Classifier; CF: Collaborative Filtering)

| Ref | Task | Focused Applications | Feature Construction | | | | Predictive Method | Contribution and Targeted Research Gap |
|------|-------------|----------------------|----------------------|---|---|---|-------------------|---|
| | | | A | R | P | L | | |
| [39] | (PBS & PPD) | A,B | x | | | | CDM | Ensemble learning for buying session and product prediction |
| [40] | | | x | | | x | DLC | Feature learning for buying session and product prediction |
| [41] | | | x | | x | x | | |
| [42] | (PPD) | A,E | x | | | | CDM | Feature engineering for product and customer interdependency |
| [43] | | | x | x | | | | Feature engineering and ensemble learning for product and time prediction |
| [44] | (PPD) | E | x | | x | | CDM | Prediction of repeated buying patterns over multiple sessions |
| [45] | | E | | | | x | DLC | Prediction of repeated buying patterns over multiple sessions |
| [46] | (PPD) | B,E | x | | | | PC | Feature engineering from purchase emails to predict next time and amount |
| [47] | (PBS & PPD) | A,B,E | x | | | | PC | Feature engineering for predicting buying session and next amount |

A Conceptual Framework of Analysis for Customer Purchase Prediction in E-commerce. A conceptual framework of analysis aims to optimize the understanding of a complex topic by breaking it down into smaller and comprehensive components [48]. We adopted a systematic literature review approach to developing the conceptual framework of analysis proposed and illustrated in Figure 1.

The framework has six components. Component 1 defines the dataset types adopted in this literature. Component 2 classifies in dimensions the input data present in those datasets. Component 3 shows the methodologies adopted for constructing features out of the input data, illustrating how consumer behavior is modeled to predictive analytics. Component 4 introduces the predictive methods summarized into four categories. Component 5 shows which tasks enable what applications from component 6, as identified in Subsection 3.1. Details on each component will be given under the research questions developed in the literature review.

The two research questions developed to conduct the systematic literature review were the guidance for scoping our findings. The results will be presented, reflecting those questions in Subsections 3.1 and 3.2.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

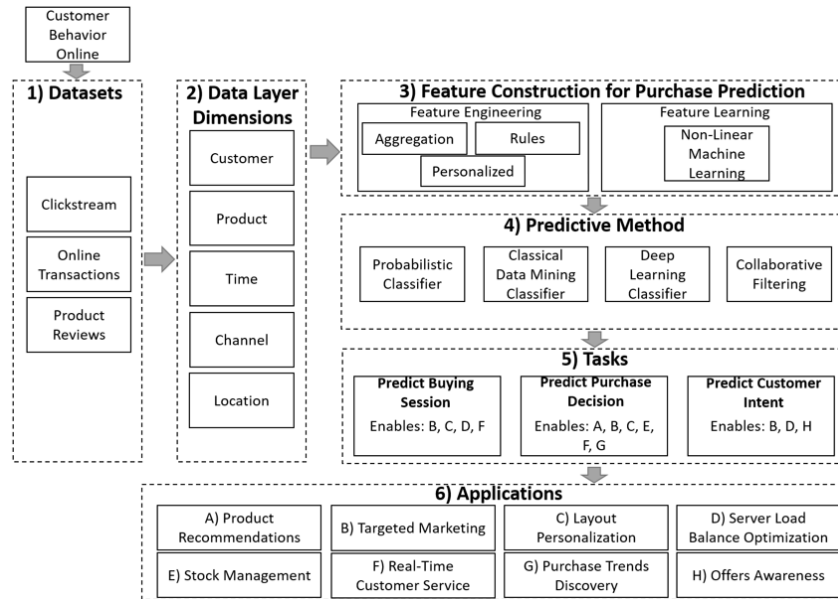


Fig. 1. A conceptual framework of analysis for the literature in behavior and predictive analytics for customer purchase prediction online. (Legends for applications enabled by tasks: A = Product Recommendations; B = Targeted Marketing; C = Layout Personalization; D = Server Load Balance Optimization; E = Stock Management; F = Real-time Customer Service; G = Purchase Trends Discovery; H = Offers Awareness)

3.1 RQ 1. What tasks and applications have been addressed in the problem of consumer purchase behavior prediction in E-commerce?

This research question addresses components 5 and 6 of the proposed framework. It reveals the literature targeting three main tasks within the online purchase prediction problem. Every task has a different prediction outcome, described as follows:

- **Predict Customer Intent (PCI):** Predict the intention of customer visits online. Examples of intention types reported in the literature are purchase oriented or general [35], browsing, searching, purchasing, and bouncing [37]. This task is essential for identifying similar groups of customers, and for applications in which customer segmentation is needed.
- **Predict Buying Session (PBS):** Predict if a current user online session will end up with a purchase or not. This task is interesting for applications that need to capture the general likelihood of the user conversion during his visit online, without details regarding preferences for specific products.
- **Predict Purchase Decisions (PPD):** Predict customers purchase behavior concerning their buying decisions. For instance, to foresee what product or category a customer will buy; to predict the time or period likely to witness a purchase; to predict

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

the next amount customers are likely to spend in their purchases. PPD is the most complex task, as the aim is to predict fine-grained decisions. That is the ideal task for recommending specific products or services to customers.

Those three identified tasks enable a variety of business intelligence applications for online retailers, such as: **A)** Product Recommendations [29]; **B)** Targeted Marketing [16, 42]; **C)** Layout Personalization of E-commerce Landing Pages [17]; **D)** Load balance Optimization to Prioritize Quality of Service for Likely Buyers [14]; **E)** Stock Management Optimization of Products [28, 32]; **F)** Real-time Customer Service [49]; **G)** Purchase Trends Discovery [15]; **H)** Offers Awareness Based on the Detected Intention of Consumers [35].

3.2 RQ 2. What methodologies have been adopted to predict consumer purchase behavior online?

This research question addresses the components from 1 to 4 of the framework proposed. It provides three perspectives in the predictive methodologies adopted in this literature.

Online Customer Behavior Datasets and their Features. Customer behavior in E-commerce is captured through datasets of past online sessions and shopping logs, which are described in Table 3:

Table 3. Dataset types identified in the literature

| Dataset | Description | Data Layer Dimensions Involved |
|--------------|--|--|
| Clickstream | Sequences of clicks performed by consumers during their online visits. | Customer |
| Transactions | Purchases executed by customers within the E-commerce | Customer, Product, Time, Channel, Location |
| Reviews | Text and rating reviews given by customers to specific products | Products |

The input data is further classified in the data layer, inspired by [2], in dimensions, which have specific input data features. Every dimension and its features support in explaining and predicting customer behavior from different perspectives, which bring some benefits for predictive tasks on that data, as illustrated in Table 4.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

Table 4. Classification of E-commerce data in dimensions and its benefits

| E-commerce data Dimensions | Input Data Features $feat_{in}$ | Description |
|----------------------------|--|---|
| Customer | <ul style="list-style-type: none"> • Demographics • Clicks • Session Variables | Reveals the profile of every consumer, and enables their segmentation. Benefit: tackling the cold start problem. |
| Product | <ul style="list-style-type: none"> • Value • Description • Status of Availability | Relates to the raw characteristics of products online. Benefit: supports the detection of preferences according to product characteristics. |
| Time | <ul style="list-style-type: none"> • Timestamp • Season | Timestamps of consumer transactions. Benefit: supports the prediction of when events can happen based on previous timestamps and seasonal patterns. |
| Channel | <ul style="list-style-type: none"> • Customer Device • Visit Source | Characteristics of touchpoints between consumers and an E-commerce platform. Benefit: assessment of influences on customer purchase likelihood in different channels. |
| Location | <ul style="list-style-type: none"> • Neighborhood • City | Information on location of consumers. Benefits: helps in identifying patterns according to the spatial placement of consumers [43, 46]. |

Feature Construction for Purchase Prediction. In this Subsection, we use a formal notation to explain the feature construction process. The input data features $feat_{in}$ described previously serve as the basis for feature construction, from which is derived new descriptive features $feat_{eng_out}$ to capture historical patterns, which can indicate the probability of purchase. Two methodologies are adopted to create descriptive features. The first is Feature Engineering, where domain expertise is used to think of a function or rule f_{eng} to apply on input data features $feat_{in}$ present in a dataset, which are related to a current customer transaction Ti . This process can be shaped by conditions $cond_n$ to capture relationships between multiple input data features. The Feature Engineering process can be described in equation 1.

$$feat_{eng_out} = f_{eng}(D, Ti, feat_{in}, cond_1, cond_n) \quad (1)$$

The second methodology for feature construction is Feature Learning, in which a function f_{learn} to create new features is an unsupervised ML model, which automatically derives new explanatory features. For instance, researchers extract Latent Representations, or hidden layer weights $feat_{learn_out}$ learned during training time of a Recurrent Neural Network or Autoencoder model, carrying hidden correlations and relationships between variables. This learning process is conditioned by the target outcome $targ_{out}$ and a cost function $cost_f$, which represent the desired outcome of the learned representation, and how the weights of the hidden layer will be learned. The desired outcome is, for instance, a binary label for predicting buying sessions, or a multi-category label for predicting purchase decisions regarding products. The Feature Learning process is described in equation 2.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

$$feat_{learn_out} = f_{learn}(D, Ti, feat_{in}, targ_{out}, cost_f) \quad (2)$$

Table 5 illustrates examples of those methodologies in action.

Table 5. Methodologies for Customer behavior Feature Construction

| Methodology | Function | Description | Examples of New Features $feat_{eng_out}$ / $feat_{learn_out}$ |
|---------------------|------------------------------------|---|---|
| Feature Engineering | Aggregation | Examples of Functions: Count, Sum, Timing, Average, Variance, Ratio $f_{eng} = Average \quad \frac{\sum_{i=1}^n Amount_i}{n}$ $feat_{in} = clicks$ $cond_1 = product\ category$ | Example of Feature: Average.Number.Clicks.On.Category [27] |
| | Rules | Example of Rule: Is Purchase in a Shopping Holiday? $f_{eng} = binary\ rule$ $feat_{in} = time$ $cond_1 = is\ time.\ weekday\ a\ holiday?$ | Examples of Feature: 1 or 0 indicating if the rule condition is satisfied or not [26] |
| | Personalized Functions | Example of Personalized Function: Entropy to detect how diverse are customer purchases $f_{eng} = Entropy \quad -\sum_{i \in C} p_i \log_2 p_i$ $feat_{in} = page\ views$ $cond_1 = product\ category$ | Examples of Feature: Diversity.Browsed.Categories [18] |
| Feature Learning | Non-Linear Machine Learning Models | Examples of Functions: Autoencoders, Recurrent Neural Networks, $f_{eng} = Autoencoder$ $feat_{in} = (feat_{in_1}, \dots, feat_{in_m})$ $targ_{out} = (feat_{in_1}, \dots, feat_{in_m})$ $cost_f = Mean\ Squared\ Error$ | Examples of Features: Compressed representation of input data by the weights of a hidden Autoencoder layer $(weight_1, \dots, weight_{m/x})$ |

Predictive Methods. Researchers have been working with ML and probabilistic methods to predict the complex customer purchase behavior online [5]. Based on the conceptual framework, we summarize the predictive models adopted into four categories, with their advantages and disadvantages. It is provided examples of particular methods within each category, specifically for purchase prediction in E-commerce. We illustrate in Table 4 how those models compare concerning their characteristics and suitability for tasks identified in Subsection 3.1.

The characteristics analyzed are a) *Suitability for Real-Time*: concerning usual time required for training, if any, and for providing predictions in production settings; b) *Interpretability*: concerning the capacity of providing explanations for why a predicted outcome is given by the model; c) *Sequential Modeling*: it illustrates if a predictive method is able to model the customer activities sequentially. That is important when researchers want to explicitly analyze the influence of past purchases in current

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

customer actions; d) *Feature Construction Function*: reveals what methodology and function are usually adopted for feature construction when applying the predictive method analyzed.

Table 6. Predictive Methodologies

| Predictive Method | Example | Characteristics of Predictive Method | | | | Suitability for Purchase Prediction Tasks | | |
|-----------------------------------|---|--------------------------------------|------------------|---------------------|---|---|-----|-----|
| | | Suitability for Real-Time | Interpretability | Sequential Modeling | Feature Construction Function | PCI | PBS | PPD |
| Probabilistic Classifier | Bayesian Classifier, Hidden Markov Model | High | High | High | Aggregation & Rules | | X | X |
| | Unsupervised Clustering: K-Means | Low | High | Low | Aggregation & Rules & Personalized | X | | |
| Classical Data Mining Classifiers | Association Rules: Apriori Algorithm | High | High | Low | | | X | |
| | Instance-Based: K-Nearest Neighbor | High | High | Low | | | X | |
| | Linear ML: Logistic Regression, Decision Tree | High | High | Low | | | X | X |
| | Ensemble Learning: XGBoost, Adaboost, Majority Voting | Low | Low | Low | | | X | X |
| Deep Learning Classifier | Non-Linear ML: RNNs, LSTMs, GRUs | Low | Low | High | Aggregation & Non-Linear Feature Learning | | X | X |
| Collaborative Filtering | Matrix Factorization and Factorization Machines | High | Low | Low | Aggregation & Rules | | | X |

Details regarding each predictive methodology are provided as follows.

- **Probabilistic Classifier:** A model that uses probability theory to model the uncertainty in the data. **Advantage:** Usually, it requires a few numbers of engineered features, which makes them a feasible choice for real-time settings, as well as the natural capacity of sequentially modeling short-term patterns in events. **Disadvantage:** It is difficult to capture the effects of long-term patterns in customer behavior. However, this capacity can be achieved in the cost of increasing model complexity and processing time.
 - *Bayesian Classifier:* Estimates conditional probability distributions based on the influence of given features to output a specific prediction. In [42], authors predict purchase decisions by analyzing the influence of sequential purchases, number, and duration of visits to compute probabilities for the customer choice of a specific product or time of purchase.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

- *Hidden Markov Model*: A generalization of a probabilistic mixture model, where the probability of an event, such as a purchase, depends on the occurrence of hidden variables through a sequential Markov process modeling a previous customer action [24].
- **Classical Data Mining Classifiers**: Those models work by learning similarities between feature vectors of buying sessions, intents, and purchase decisions. **Advantage**: Most of the approaches in this category perform well even with small or medium dataset sizes, which makes some of them suitable for real-time settings. **Disadvantage**: Authors adopting this methodology usually need to perform extensive Feature Engineering to achieve good prediction results, also for detecting sequential patterns.
 - *Unsupervised Clustering*: Unlabeled sessions and purchase transactions are input to a model which will discover patterns in similar instances and group them for providing predictions. For example, [37, 38] adopt the K-means algorithm to segment customers based on variables regarding their clickstream behavior.
 - *Association Rules*: Enables the discovery of associations between features, which can reveal rules with high confidence to indicate probabilities of sessions ending up with a purchase [16].
 - *Instance-Based*: Model which classify new data instances based on similar cases and their features. In [22], authors employ K-Nearest Neighbor to predict buying sessions according to previous examples of sessions, with similar features, which ended up with a purchase.
 - *Linear ML*: Machine learning models which assume a linear decision boundary between buying and non-buying sessions, or feature vectors representing purchase decisions of customers. However, the kernel trick can be adopted to detect non-linear relationships between features [50], or Feature Engineering to create combinations between multiple features [27].
 - *Ensemble Learning*: stacking of various weak predictive models together to build up a robust model for providing predictions [20].
- **Deep Learning Classifiers**: ML models which can naturally learn complex and non-linear decision boundaries and relationships in the dataset. **Advantage**: These models can be powerful in modeling long-term influences of past customer events on current decisions [25], and do not require extensive Feature Engineering, as they have Feature Learning built-in. **Disadvantage**: This method usually requires massive amounts of data, which makes it hard for usage with new customers and a few purchases [40, 41]. The interpretability of predictions is also an issue.
- **Collaborative Filtering**: Classical model applied in recommendation systems. This approach models customers and products in a utility matrix based on their clicks, views, reviews, or purchases, which is then factorized to provide latent factors representing the likelihood of customers choosing similar products [29, 30, 44]. That is the favorite model adopted by researchers focusing on predicting purchase decisions, but it is also utilized in predicting buying sessions [14]. **Advantage**: One of the most flexible approaches for multiple types of features in different E-commerce platforms. It also scales well with more customers and products being added in a dataset. **Disadvantage**: The utility matrix is usually sparse, as

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

most of the customers have not viewed many of the products available in an E-commerce platform. Therefore, it is challenging to predict purchases for new customers, and it is important to think of Feature Engineering for creating features that can overcome such issues.

3.3 State-of-the-art Performance

To have a fair comparison between the identified predictive methodologies, for every specific task, we grouped the existing proposals by the predictive methodology adopted. We evaluated only the F1 score and Area Under Receiver Operating Characteristic Curve (AUC) reported by those. Our choice for those metrics considers the fact that datasets in this literature are usually unbalanced, with few occurrences of purchases, and it is well known that F1 and AUC scores are the ideal metrics in unbalanced scenarios [51]. Table 4 illustrates the average results obtained from predictive methodologies for suitable tasks where they can be applied. It is not reported performance for predicting customer intent as the authors did not adopt the mentioned metrics.

Table 7. State of the Art Results for Predicting Buying Sessions and Purchase Decisions

| Predict Buying Session | | | |
|---------------------------------------|--------|---------------------------------------|--------|
| F1 | | AUC | |
| Classical Data Mining Classifier [20] | 97.20% | Classical Data Mining Classifier [19] | 84% |
| Deep Learning Classifier [26] | 87.94% | Deep Learning Classifier [25] | 83.90% |
| Classical Data Mining Classifier [15] | 82.91% | Classical Data Mining Classifier [14] | 75% |
| Predict Purchase Decision | | | |
| F1 | | AUC | |
| Collaborative Filtering [30] | 53% | Collaborative Filtering [31] | 87.94% |
| Classical Data Mining Classifier [28] | 43.62% | Deep Learning Classifier [41] | 86% |
| Collaborative Filtering [32] | 42% | Classical Data Mining Classifier [39] | 85% |

Classical Data Mining Classifiers are the current state-of-the-art for Predicting Buying Sessions, specifically Ensemble learners [20] and Support Vector Machines [19]. Those are followed by Deep Learning classifiers. It is interesting to observe the drop in performance when going to the task of Predicting Purchase Decision, which proves it is the most complex task due to the fine-grained predictions aimed at it. Concerning performance, the classical Collaborative Filtering approach is the most robust, comprised of a Latent Factor Model [30] and Matrix Factorization [31]. Those are followed by Classical Data Mining and Deep Learning classifiers.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

3.4 Research Agenda

We derive a research agenda based on the targeted research gaps and findings of this review, containing the following directions:

- **Sequential Learning:** Few proposals have explored sequential ML models in this literature. Examples are recurrent neural networks, which are only adopted in three studies [25, 33, 40]. Such models are indicated to learn the evolving consumer behavior over time, and sequential patterns such as “She is buying a phone case after purchasing a smartphone”.
- **Interpretability:** It is noticed the majority of authors reporting higher performance as those applying Classical Data Mining and Deep Learning classifiers, which also have a black-box nature. Indeed, interpretability seems not to be the focus of this recent literature.
- **Customer Data and General Data Protection Regulation (GDPR):** Given the rise of privacy policies with GDPR in Europe, it is needed more research on the trade-off between the amount of data required and protection of customers’ privacy, regarding the performance of purchase prediction tasks.
- **Dataset for benchmarking:** There is no clear consensus on datasets for state-of-the-art comparison in this literature, as many studies have used private data. However, we observed a significant adoption of the Recsys 2015 challenge data [17, 25, 31, 39, 40, 42], which suggests this dataset as a candidate in this regard.
- **Evaluation in Multiple E-commerce Platforms:** Most researchers evaluate their proposed predictive methods in a single dataset, or focus on specific E-commerce settings. Therefore it is hard to argue their methodologies are general for multiple E-commerce platforms, such as general-purpose and specialized marketplaces.
- **Feature Engineering and Feature Learning:** It was noticed that the well-performing proposals adopting Classical ML models had been heavily investing in Feature Engineering. However, more investigation in the field of Feature Learning is recommended in this area, or the combination of those two methodologies in purchase prediction online.
- **Creation Process of Personalized Feature Engineering Functions:** Some researchers explore the creation of personalized functions in Feature Engineering, such as the popularity of a product [17], the diversity of customer behavior [18, 35] and graph metrics [21]. It could be relevant to map this creation process, and help other researchers in establishing such novel features for customer behavior online.
- **A Framework for Purchase Prediction Tasks in E-commerce:** Existing proposals focus on one of the three tasks identified, but there is a lack of a view into how those tasks can work together. Therefore, further research could be taken to provide a framework which aligns the identified tasks in this review.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

4 Final Remarks

This study presents a systematic literature review of recent proposals in consumer purchase prediction in E-commerce. A novel conceptual framework provides lenses in the state-of-the-art of this field. It is noticed that, despite the broad literature, there is still a need for an in-depth investigation of specific directions. Therefore, a research agenda is provided, illustrating potential future work demands.

A next step would be to adopt a benchmark dataset, and evaluate predictive methodologies in multi-task settings, such as to forecast the next product, purchase time, or amount a customer will likely buy. Therefore, it is relevant to investigate the construction of a framework for purchase prediction, which considers the combination of the three tasks identified in this review.

References

1. Agnihotri, R., Dingus, R., Hu, M.Y., Krush, M.T.: Social media: Influencing customer satisfaction in b2b sales. *Industrial Marketing Management* 53 (2016) 172–180
2. Bradlow, E.T., Gangwar, M., Kopalle, P., Voleti, S.: The role of big data and predictive analytics in retailing. *Journal of Retailing* 93(1) (2017) 79–95
3. Le, D.T., Fang, Y., Lauw, H.W.: Modeling sequential preferences with dynamic user and context factors. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer (2016) 145–161
4. Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.
5. Shmueli, G., et al.: To explain or to predict? *Statistical science* 25(3) (2010) 289–310
6. Martens, D., Provost, F., Clark, J., de Fortuny, E.J.: Mining massive fine-grained behavior data to improve predictive analytics. *MIS quarterly* 40(4) (2016)
7. Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Introduction to recommender systems handbook." *Recommender systems handbook*. Springer, Boston, MA, 2011. 1-35.
8. Bobadilla, Jesús, et al. "Recommender systems survey." *Knowledge-based systems* 46 (2013): 109-132.
9. Lu, Jie, et al. "Recommender system application developments: a survey." *Decision Support Systems* 74 (2015): 12-32.
10. Isinkaye, F. O., Y. O. Folajimi, and B. A. Ojokoh. "Recommendation systems: Principles, methods and evaluation." *Egyptian Informatics Journal* 16.3 (2015): 261-273.
11. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly* (2002) xiii–xxiii
12. Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology* 51 (1) (2009) 7–15
13. Akter, S., Wamba, S.F.: Big data analytics in e-commerce: a systematic review and agenda for future research. *Electronic Markets* 26(2) (2016) 173–194
14. Zeng, M., Cao, H., Chen, M., Li, Y.: User behaviour modeling, recommendations, and purchase prediction during shopping festivals. *Electronic Markets* (2018) 1–12

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

15. Jia, R., Li, R., Yu, M., Wang, S.: E-commerce purchase prediction approach by user behavior data. In: 2017 International Conference on Computer, Information and Telecommunication Systems (CITS), IEEE (2017) 1–5
16. Suchacka, G., Chodak, G.: Using association rules to assess purchase probability in online stores. *Information Systems and e-Business Management* 15(3) (2017) 751–780
17. Chen, C., Xiao, J., Hou, C., Yuan, X.: Improving purchase behavior prediction with most popular items. *IEICE TRANSACTIONS on Information and Systems* 100(2) (2017) 367–370
18. Niu, X., Li, C., Yu, X.: Predictive analytics of e-commerce search behavior for conversion. In: Twenty-third Americas Conference on Information Systems. (2017)
19. Lee, M., Ha, T., Han, J., Rha, J.Y., Kwon, T.T.: Online footsteps to purchase: Exploring consumer behaviors on online shopping sites. In: Proceedings of the ACM Web Science Conference, ACM (2015) 15
20. Boroujerdi, E.G., Mehri, S., Garmaroudi, S.S., Pezeshki, M., Mehrabadi, F.R., Malakouti, S., Khadivi, S.: A study on prediction of user's tendency toward purchases in websites based on behavior models. In: 2014 6th Conference on Information and Knowledge Technology (IKT), IEEE (2014) 61–66
21. Baumann, A., Haupt, J., Gebert, F., Lessmann, S.: Changing perspectives: Using graph metrics to predict purchase probabilities. *Expert Systems with Applications* 94 (2018) 137–148
22. Suchacka, G., Skolimowska-Kulig, M., Potempa, A.: A k-nearest neighbors method for classifying user sessions in e-commerce scenario. *Journal of Telecommunications and Information Technology* (2015)
23. Lin, W., Milic-Frayling, N., Zhou, K. and Ch'ng, E., 2019, October. Predicting Outcomes of Active Sessions Using Multi-action Motifs. In *IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 9-17).
24. Park, C.H., Park, Y.H.: Investigating purchase conversion by uncovering online visit patterns. *Marketing Science* 35(6) (2016) 894–914
25. Sheil, H., Rana, O. and Reilly, R., 2018. Predicting purchasing intent: Automatic feature learning using recurrent neural networks. arXiv preprint arXiv:1807.08207.
26. Sakar, C.O., Polat, S.O., Katircioglu, M. and Castro, Y., 2019. Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications*, 31(10), pp.6893-6908.
27. Li, Q., Gu, M., Zhou, K., Sun, X.: Multi-classes feature engineering with sliding window for purchase prediction in mobile commerce. In: 2015 IEEE International Conference on Data Mining Workshop (ICDMW), IEEE (2015) 1048–1054
28. Iwanaga, J., Nishimura, N., Sukegawa, N. and Takano, Y., 2016. Estimating product-choice probabilities from recency and frequency of page views. *Knowledge-Based Systems*, 99, pp.157-167.
29. He, T., Yin, H., Chen, Z., Zhou, X., Luo, B.: Predicting users purchasing behaviors using their browsing history. In: Australasian Database Conference, Springer (2015) 129–141
30. Jia, R., Li, R.: Modeling user purchase preference based on implicit feedback. In: CSCWD, IEEE (2018) 832–836
31. Park, C., Kim, D., Yang, M.C., Lee, J.T. and Yu, H., 2017. Your Click Knows It: Predicting User Purchase through Improved User-Item Pairwise Relationship. arXiv preprint arXiv:1706.06716.
32. Nishimura, N., Sukegawa, N., Takano, Y. and Iwanaga, J., 2018. A latent-class model for estimating product-choice probabilities from clickstream data. *Information Sciences*, 429, pp.406-420.

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

33. Singhal, R., Shroff, G., Kumar, M., Choudhury, S.R., Kadarkar, S., Virk, R., Verma, S., Tewari, V.: Fast online 'next best offers' using deep learning. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. CoDS-COMAD '19, New York, NY, USA, ACM (2019) 217–223
34. Bai, Jinze, et al. "Personalized Bundle List Recommendation." The World Wide Web Conference. ACM, 2019.
35. Zheng, B., Liu, B.: A scalable purchase intention prediction system using extreme gradient boosting machines with browsing content entropy. In: 2018 IEEE International Conference on Consumer Electronics (ICCE), IEEE (2018) 1–4
36. Minjing, P., Xinglin, L., Ximing, L., Mingliang, Z., Xianyong, Z., Xiangming, D., Mingfen, W.: Recognizing intentions of e-commerce consumers based on ant colony optimization simulation. *Journal of Intelligent & Fuzzy Systems* 33(5) (2017) 2687–2697
37. Schellong, D., Kemper, J., Brettel, M.: Generating consumer insights from big data clickstream information and the link with transaction-related shopping behavior. In: Proceedings of the 25th European Conference on Information Systems (ECIS). (2017)
38. Schellong, D., Kemper, J., Brettel, M.: Clickstream data as a source to uncover consumer shopping types in a large-scale online setting. In: ECIS. (2016) Research Paper 1
39. Romov, P., Sokolov, E.: Recsys challenge 2015: Ensemble learning with categorical features. In: Proceedings of the 2015 International ACM Recommender Systems Challenge, ACM (2015) 1
40. Wu, Z., Tan, B.H., Duan, R., Liu, Y., Mong Goh, R.S.: Neural modeling of buying behaviour for e-commerce from clicking patterns. In: Proceedings of the 2015 International ACM Recommender Systems Challenge, ACM (2015) 12
41. Vieira, A.: Predicting online user behaviour using deep learning algorithms. arXiv preprint arXiv:1511.06247 (2015)
42. Yeo, J., Kim, S., Koh, E., Hwang, S.w., Lipka, N.: Predicting online purchase conversion for retargeting. In: Proceedings of the Tenth ACM international conference on web search and data mining, ACM (2017) 591–600
43. Li, D., Zhao, G., Wang, Z., Ma, W., Liu, Y.: A method of purchase prediction based on user behavior log. In: 2015 IEEE International Conference on Data Mining Workshop (ICDMW), IEEE (2015) 1031–1039
44. Liu, G., Nguyen, T.T., Zhao, G., Zha, W., Yang, J., Cao, J., Wu, M., Zhao, P., Chen, W.: Repeat buyer prediction for e-commerce. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM (2016) 155–164
45. Guo, L., Hua, L., Jia, R., Zhao, B., Wang, X. and Cui, B., 2019, July. Buying or Browsing?: Predicting Real-time Purchasing Intent using Attention-based Deep Network with Multiple Behavior. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1984-1992).
46. Kooti, F., Lerman, K., Aiello, L.M., Grbovic, M., Djuric, N., Radosavljevic, V.: Portrait of an online shopper: Understanding and predicting consumer behavior. In: Proceedings of the ninth ACM international conference on web search and data mining, ACM (2016) 205–214
47. Panagiotelis, A., Smith, M.S., Danaher, P.J.: From amazon to apple: Modeling online retail sales, purchase incidence, and visit behavior. *Journal of Business & Economic Statistics* 32(1) (2014) 14–29
48. Green, Helen Elise. "Use of theoretical and conceptual frameworks in qualitative research." *Nurse researcher* 21.6 (2014).

This is an open access post-print version; the final authenticated version is available online at https://link.springer.com/chapter/10.1007%2F978-3-030-48861-1_8 by Springer Nature Switzerland AG.

49. Tang, L., Wang, A., Xu, Z., Li, J.: Online-purchasing behavior forecasting with a firefly algorithm-based svm model considering shopping cart use. *Eurasia Journal of Mathematics, Science and Technology Education* 13(12) (2017) 7967–7983
50. Schölkopf, B., 2001. The kernel trick for distances. In *Advances in neural information processing systems* (pp. 301-307).
51. Jeni, L.A., Cohn, J.F. and De La Torre, F., 2013, September. Facing imbalanced data--recommendations for the use of performance metrics. In *2013 Humaine association conference on affective computing and intelligent interaction* (pp. 245-251). IEEE.