# Machine Learning Solutions in Retail eCommerce to Increase Marketing Efficiency

Maciej Pondel, Jolanta Pondel

# Machine Learning Solutions in Retail eCommerce to Increase Marketing Efficiency

Maciej Pondel[1] and Jolanta Pondel[2]

[1]Wrocław University of Economics,
Komandorska 118/120, Wrocław, Poland

Email: maciej.pondel@ue.wroc.pl
ORCID: 0000-0002-1978-6571

[2]WSB University in Wrocław,
Fabryczna St. 29-31, 53-609 Wrocław, Poland

Email: jolanta.pondel@wsb.wroclaw.pl
ORCID: 0000-0002-9562-7783

***Abstract.*** Retail companies operate in a highly competitive market. Marketing specialists should focus their activities on keeping customers loyalty on a high level, which is critical to achieving success in the retail business. The efficiency of marketing campaigns impacts the performance of the whole business directly. Electronic channels of customer contact help in gathering massive data sets describing customers and their behaviour. Companies can use the data to gain knowledge about their customers and valuable insights about patterns of customer behaviour. Such knowledge enables the application of actions improving the loyalty of customers and satisfaction of customers. This paper elaborates on customer loyalty and shows how to use data in marketing and sales processes in retail. Authors present a method of advanced data analysis based on machine learning techniques to generate tailored offers for clients applicable in an eCommerce company.

**Keywords:** Machine Learning • marketing • retail • eCommerce • customer churn.

## 1 Introduction

Strong market competition and dynamic changes in the markets force enterprises to take actions directed at improving their relationships with customers. Those customers are increasingly choosing online stores as a place to buy [Doligalski 2018]. The success of retail organisations depends on how adequately they can analyse the data about their clients' behaviour and draw valuable conclusions.

It is essential to acquire a new customer, but even more critical to keep the customer with good relation. The most critical priority impacting the future of the company and its financial results is to keep the customer satisfaction on a high level, improve the relationship with the customers and thus build long-term and profitable relationships. Such an approach requires many efforts to be taken, including financial and time engagements. It is also necessary to continually monitor the initiated activities, inspect the results, and take improving actions to ensure the conducted activities are proper and bring measurable benefits to the enterprise.

The need to establish a lasting relationship with the customer becomes the foundation for market activities for enterprises, particularly in the retail market.

We can define customer loyalty as "... re-purchasing only one brand, without considering the purchase of other brands" [Newman, Werbel 1973] or as "... a persistent and consistent readiness to re-purchase or re-promote your preferred product or brand in the future, resulting in repeat purchases of the same brand despite market influences and marketing efforts that have the potential to change behaviour" [Oliver 1999]. Two aspects of loyalty can be distinguished based on the definition:

- a behavioural aspect related to specific behaviours,
- the emotional element, expressed by the customer's attachment and "... attitude towards specific objects related to the supplier, leading to the expression of loyalty behaviour" [Urban, Siemieniako 2008].

There is also an essential phenomenon of emotional attachment, also called "the client's emotional loyalty". Customers with apparent loyalty are more susceptible to changing the supplier under the influence of favourable economic factors, and reshopping is not related to the positive attitude of the customer and how they are devoted to the company. An essential feature of loyalty is that the purchase is planned, not accidental. Loyalty is the result of a deliberate choice, not a case or alternative offer [Bloomer, Kasper 1995].

Customer loyalty is closely related to their satisfaction. A satisfied customer is likely to make purchases again. Satisfaction is a prerequisite, but not always enough to achieve true loyalty [Sudolska 2011]. Research shows that among customers who decide to change suppliers, about 65-85% are satisfied with the received product [Stum, Thirty 1991]. The relationship between satisfaction and loyalty depends on the degree of competition in a given market. To maintain the customer and provide them with favourable conditions, the company must measure customer loyalty and customer satisfaction.

The full availability of communication and sales channels provided by the Internet and social media has radically changed the traditional buying model. Customers expect a higher and higher quality of service and convenience provided by the eCommerce channel. It is essential in such an approach to allow customers to interact in a convenient way and at the time they prefer. Additional occasions, such as providing tools and information, allow building even closer relations.

Rawson states that many companies excel in individual interactions with customers, but they fail to pay adequate attention to the customer's complete experience on the way to purchase and after. Solution for the problem is that companies need to combine top-down, judgment-driven evaluations and bottom-up, data-driven analysis to identify key journeys, and then engage the entire organization in redesigning the cus-

tomer experience. This requires shifting from siloed to cross-functional approaches and changing from a touchpoint to a journey orientation [Rawson and others 2013].
An important factor determining customer's satisfaction and loyalty is the personalization of approach and customized offers directed to clients. To be able to generate such, the company must conduct in-depth and multidimensional customer analysis.
The main goal of this paper is to elaborate customers loyalty and satisfaction and known methods involving the acquisition of knowledge about customers to apply a customized approach to a client. This paper's aim is also to propose the innovative method of B2C eCommerce originated data analysis to generate recommendations of marketing actions.
This paper is structured as follows. In the next section, the related articles are analyzed and basing on other research findings, authors provide the foundation of the proposed data analysis method. The research gap is also identified. The following section describes research method and next the results of research are presented, and the results are discussed.

## 2 Related Works

Techniques of collecting information from clients can be based, for example, on the surveys, interviews, reports, or direct conversations with customers. You can use for this [Jones, Sasser 1995], [Ciok 2006]:

- customer satisfaction indicators,
- questionnaire surveys,
- feedback from the customer (comments, complaints and questions from customers),
- market research,
- staff who have direct contact with the client,
- strategic activities (hiring employees in the same market as the target group, inviting clients to work related to the improvement of products and services,
- apparent shopping (mysterious customer),
- analysis of customer loss (contact with customers who stopped buying products to identify the problem).

However, the information obtained from these sources may be biased, as the customers may not want to express their real opinion (fear, lack of time, different perception of product or service). Authors of this paper would like to elaborate on different sources of data storing recorded real actions of all customers to analyze the whole range of customers and acquire actual shopping patterns.
The problem of customer-related data analysis is widely discussed in the literature. Most of the applications regard to customer churn [Burez, Van den Poel 2009], [De Caigny and others 2018]. Customer churn [Amin and others 2019] is commonly related to service industries, and rarely considered as a problem in eCommerce-oriented retail sector.
Customers who stop using products/services of the company are considered as churn customers. We consider customers' churn as a lost opportunity for profit. One of the

main reasons to undergo the customer's preservation process is that keeping a customer is far less expensive than finding a new customer. That is, the sale cost for new customers is five times of that of old customers [Fathian and others 2016]. Bhattacharya claims that the cost of obtaining a new customer is usually five to even six times higher than the costs of retaining an existing one [Bhattacharya, 1998].

As a result, efforts done by marketing specialists to sustain market share switched from focusing on acquiring new customers to retaining existing - reducing customer churn. For this reason, customer churn, also known as customer turnover, customer attrition, or customer deflection, is a significant concern for several industries. That is why customer churn is especially relevant in the e-commerce context, where consumers can easily compare products or services and change the merchant with minimal effort. Therefore, by spotting churned customers and understanding drivers, companies may have a better chance to minimize customers' churn.

During the last decade, customer churn prediction has received a growing consideration in order to survive in an increasingly competitive and global marketplace (Gordini 2010). Concretely, in customer churn prediction, a scoring model allows the estimation of a future churn probability for every customer based on the historical knowledge of the customer. In practice, these scores can be used to select targeted customers for a retention campaign [De Caigny and others 2018].

We can find several approaches for telecom industry [Amin and others 2019], [Hung and others 2006], [Lu and others 2012], [Tsai, Lu 2009] financial industry (banking and insurance) [Zhao, Dang 2008], [Prasad, Madhavi 2012], subscription services like pay-TV [Burez, Poel 2007] and cellular network services [Sharma and others 2013]. In service industries like those mentioned, the definition of customer churn is clear. There is usually a contract between customer and service provider. When such a contract expires or is terminated, we can clearly say that customer abandoned company (is a churned customer). Few publications elaborating on churn analysis in retail business can be found. In [Yu and others 2011] we can find a recommendation to predict eCommerce customer churn based on SVM algorithm. A practical drawback of such approach is a premise that we can apply binary classification (churn and non-churn) while eCommerce reality is much more complicated. In retail and eCommerce, we focus more on customer recency than churn. It is difficult to identify the exact moment when clients discontinue their relationship with companies [Miguéis and others 2012]. If a customer has not placed an order for 100 days, we cannot clearly state that he or she is a churned customer. We can only predict the probability or the customer's return. In [Miguéis and others 2012] we can find an interesting approach based on sequence mining. Authors grouped the purchases in periods of three months and by classifying as churners those customers who, from a certain period, did not buy anything else or those who in all subsequent periods spent less than 40% of the amount spent in the reference period. Companies should strive for models that can accurately identify potential churners, and this becomes even more important in the digital economy context. Over the last decade, this issue was mentioned and researched by many practitioners and academics.

In the present literature, we can observe two main trends concerning customer churn. According to [Gordini 2017] the first branch includes traditional classification methods such as decision tree and logistic regression [Burez, Van den Poel, 2007, Gordini, Veglio, 2013, 2014; Verbeke et al., 2011]. Next approach is based on the artificial intelligence methods such as neural networks (Gordini, Veglio, 2013; Sharma et al., 2011).

In [Sulistiani, Tjahyanto 2017] we can find an experiment concerning the prediction of customer loyalty based on a data gathered from customer survey where they declare their satisfaction, but as mentioned earlier, this paper concentrates on researching real customer's behavior rather than their declarations.

There are several factors impacting customer loyalty in eCommerce. One of them is the customization of offers. In [Clauss and others 2018] the factor is listed among such as reliability, price value, information quality, and others. Personalization of offers is understood as a positive answer to the following questions:

- This platform makes purchase recommendations that match my needs.
- The advertisements and promotions that this platform sends to me are tailored to my situation.
- I believe that this platform is customized to my needs.

In [Pappas and others 2017] authors identify personalization as an essential factor in the area of marketing. Online shopping personalization is a strategy that may aid in persuading customers to select a product or service and lead to a purchase. The conclusion is that a key factor in increasing customer loyalty is a proper personalisation, but authors also claim that traditional techniques in personalized online shopping (e.g., recommendations based on previous purchases, tailored messages based on browsing history) are not enough to lead customers to an online purchase, when customers are on a shopping mission. Importance of identification of customer's price sensitivity and promotion sensitivity in building positive customer experience and loyalty is also highlighted.

The positive recommendation based algorithms are used in a vast amount of websites, such as the movie recommendation algorithms on Netflix, the music recommendations on Spotify, video recommendations on Youtube and the product recommendations on Amazon [Boström, Filipsson 2017]. Online product recommendation mechanism is becoming increasingly available on websites to assist consumers in reducing information overload, provide advice in finding suitable products, and facilitate online consumer decision-making. Recommendation mechanisms have been found to help consumers efficiently filter available alternatives, increase the quality of their consideration set and increase their product choice satisfaction [Lee, Kwon 2008].

In this paper, authors concentrate on methods preventing customer churn by purchase patterns detection combined with customer segmentation to define best suited marketing messages and tailored offers adjusted to actual customer's needs and expectations to build positive experience increasing loyalty

## 3 Research method

The most common methods of data analysis used in marketing are:

- Frequent Pattern Mining (Association rules and sequence rules).

- Collaborative Filtering.
- Clustering.
- Classification and regression.

The concept of frequent itemset was first introduced for mining transaction databases [Han and others 2007]. Frequent patterns are itemsets, subsequences, or substructures that appear in a data set with frequency no less than a user-specified threshold. Frequent pattern mining was first proposed by Agrawal [Agrawal and others 1993] for market basket analysis in the form of association rule mining. It analyses customer buying habits by finding associations between the different items that customers place in their "shopping baskets". There are many algorithms searching for association rules such as Apriori, Charm, FP-Growth and others. They differ in computational complexity, and thus in resource demand and execution time. Their operation is deterministic, so the obtained results will be the same. The source data for association rules is a set of transactions / orders of a customer in eCommerce store. The analysis is also referred to as "Market basket analysis" because it searches for patterns in shopping baskets. The result of an algorithm is a list of frequent sets (products appearing frequently). Having a list of sets we are able to construct rules answering a question which products are frequently purchased together. Example of association rules generated using FP-Growth algorithm is presented in Figure 1.

| First product | Next product | Confidence | Frequency | Lift |
|---|---|---|---|---|
| Calvin Klein / Gloves and scarves / Unisex | Calvin Klein / Winter hats / Unisex | 32,13% | 392 | 17,08 |
| Calvin Klein / Winter hats / Unisex | Calvin Klein / Gloves and scarves / Unisex | 17,21% | 392 | 17,08 |
| Calvin Klein / Shirts / Women | DKNY / Shirts / Women | 20,55% | 598 | 10,10 |
| DKNY / Shirts / Women | Calvin Klein / Shirts / Women | 24,28% | 598 | 10,10 |
| DKNY / Socks / Unisex | Lauren / Socks / Unisex | 9,68% | 224 | 9,72 |
| Lauren / Socks / Unisex | DKNY / Socks / Unisex | 18,59% | 224 | 9,72 |
| DKNY / Socks / Unisex | DKNY / Socks / Women | 6,48% | 150 | 9,54 |
| DKNY / Socks / Women | DKNY / Socks / Unisex | 18,25% | 150 | 9,54 |

*Figure 1. Example of association rules*

Each rule is evaluated by the following measures:

Confidence – this is a percentage value that shows the probability of consequent product purchase in a situation where an antecedent product has been already placed in a basket.

Frequency – number of transactions including both products. It helps to asses if a rule is common or perhaps is only an exception. If confidence is high and frequency is low, it means that two very seldom bought products were purchased together only a few times. In such case, the overall value of a rule is considered as low.

Lift – this is the ratio determining independence of antecedent and consequent product. In some cases, confidence and frequency are not enough to evaluate a rule. If e.g. a plastic bag is added to nearly all orders, the rules with a plastic bag as a consequent product will have a very high confidence as well as high frequency, but the business value of those rules will be very low. To eliminate those rules from further considering lift can help. If the rule had a lift of 1, it would imply that the probability of occurrence of the antecedent and that of the consequent are independent of each other. When two events are independent of each other, no rule can be drawn involving those two events. If the lift is > 1, that lets us know the degree to which those two occurrences are dependent on one another and makes those rules potentially useful for predicting the consequent in future data sets. If the lift is < 1, that lets us know the items are a substitute to each other. This means that the presence of one item has a negative effect on the presence of another item and vice versa.

There are more measures to value the association rule like support, coverage, strength or leverage. In our experiment, we find it enough to analyze those 3 measures explained and presented in Figure 1.

Sequence rules algorithm is based on a similar foundation. A sequence database consists of ordered elements or events, recorded with or without a concrete notion of time. There are many applications involving sequence data, such as customer shopping sequences, Web clickstreams, and biological sequences.

Agrawal and Srikant first introduced the sequential pattern mining in [Srikant, Agrawal 1996]. Given a set of sequences, where each sequence consists of a list of elements and each element consists of a set of items, and given a user-specified min-support threshold, sequential pattern mining is to find all the frequent subsequences, i.e., the subsequences whose occurrence frequency in the set of sequences is no less than minsupport.

The result of sequence mining is a list of rules valued with similar measures like in case of association rules. The most important difference is that sequence mining help with the prediction of future purchases of customers basing on their previously bought products. Antecedent and consequent products may be the same in a situation where customers regularly buys the same products that end or wear out and then need to be re-purchsed. In case of association rules such situation never happens. One product cannot be purchased together with itself. If customer buys a several items of the same product it is treated by algorithm as one product.

Example of sequence rules generated by PrefixSpan algorithm is presented in Figure 2.

| Antecedent | Consequent | Frequency | Confidence | Support |
|---|---|---|---|---|
| Hilfiger / Shirts / Men | Calvin Klein / Shirts / Men | 174 | 14,09% | 0,064% |
| Calvin Klein / Polo shirts / Men | Calvin Klein / Shirts / Men | 208 | 13,00% | 0,076% |
| Calvin Klein / Shirts / Men | Calvin Klein / Shirts / Men | 1237 | 11,78% | 0,453% |
| Calvin Klein / Polo shirts / Men | Calvin Klein / Sweatshirts / Men | 168 | 10,50% | 0,062% |
| Versace / Shirts / Men | Calvin Klein / Shirts / Men | 145 | 10,36% | 0,053% |
| DKNY / Shirts / Men | Calvin Klein / Shirts / Men | 748 | 10,35% | 0,274% |
| Calvin Klein / Autumn jackets / Men | Calvin Klein / Shirts / Men | 262 | 9,72% | 0,096% |
| Calvin Klein / Shorts / Men | Calvin Klein / Shirts / Men | 347 | 9,53% | 0,127% |
| Calvin Klein / Polo shirts / Men | DKNY / Shirts / Men | 147 | 9,19% | 0,054% |
| DKNY / Shirts / Men | DKNY / Shirts / Men | 646 | 8,94% | 0,236% |
| Calvin Klein / Autumn jackets / Men | Calvin Klein / Sweatshirts / Men | 239 | 8,87% | 0,087% |
| DKNY / Polo shirts / Men | Calvin Klein / Shirts / Men | 144 | 8,79% | 0,053% |
| Calvin Klein / Sweatshirts / Men | Calvin Klein / Sweatshirts / Men | 923 | 8,73% | 0,338% |

*Figure 2.Sequence rules generated by PrefixSpan algorithm*

Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people [Schafer and others 2007].

Collaborative Filtering is the most widely used technique for Recommender Systems. The biggest advantage of Collaborative Filtering over content-based systems is that explicit content description is not required. Instead Collaborative Filtering only relies on opinions expressed by users on items. Instead of calculating the similarity between an item description and a user profile as a content-based recommender would do, a Collaborative Filtering system searches for similar users (neighbours) and then uses ratings from this set of users to predict items that will be liked by the current user [Massa, Avesani 2004].

Unfortunately, in regular eCommerce in contrast to content delivery services (music or movie providers) users are not used to express opinions about the product they have purchased that is why collaborative filtering has limited usage in the topic of this paper.

The goal of clustering is to find clusters of similar observations in a set of customers, products, transactions, and customer contacts with store web pages. In our experiment, we would like to determine segments of customers to whom we can send an offer or whom we can target when promoting specific products. We require clusters to have specific statistical characteristics (such as minimum variance) and usefulness in marketing decision making (e.g. determining loyal customer groups).

There are several algorithms supporting clustering. In our experiments we concentrate on the following:

k-means based on the Euclidean distance between observations,

Bisecting k-means acting on a similar basis to k-means, however, starting with all the observations in one cluster and then dividing the cluster into 2 sub-clusters, using the k-means algorithm,

Gaussian Mixture Model (GMM), which is a probabilistic model based on the assumption that a particular feature has a finite number of normal distributions,

DBSCAN identifying clusters by measuring density as the number of observations in the designated area. If the density is greater than the density of observations belonging to other clusters, then the defined area is identified as a cluster.

To define the goal of customer clustering, we used RFM method widely used in the marketing of retail companies. The acronym RFM comes from the words “recency” (a period from the last purchase), “frequency”, and “monetary value”. In this type of analysis, customers are divided into groups, based on information on time which has passed from last purchases, how often they make purchases, and how much money they spent.

The following observations explain why RFM is interesting for retail companies:

- Customers who have recently made purchases are more likely to make a new purchase soon.
- Customers who frequently make purchases are more likely to do more shopping.
- Customers who spend a lot of money are more likely to spend more money.

Each of these observations corresponds to one of the dimensions of RFM.

An example of RFM clustering is presented in Figure 3. One customer is one dot and color presents the number of a cluster to which customer was assigned.

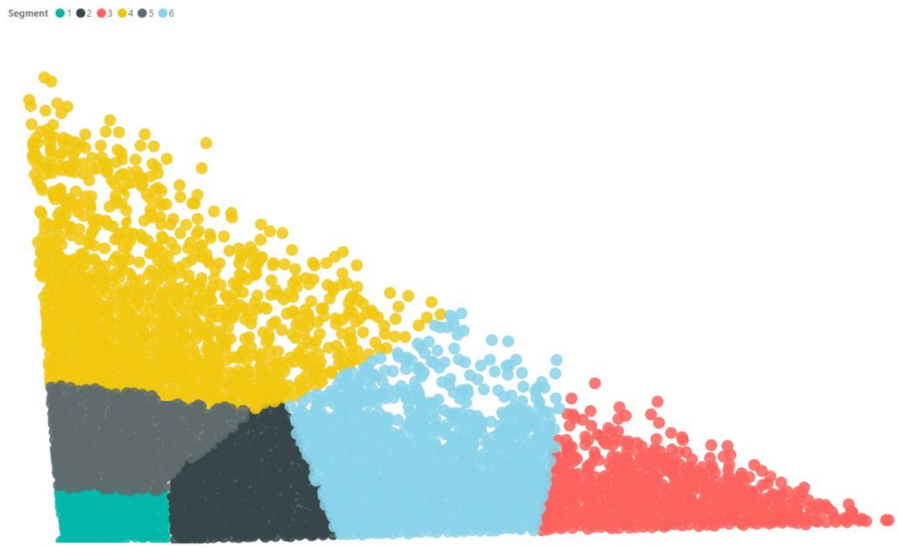


*Figure 3. RFM clustering using k-means algorithm.*

Another set of characteristics describing customer behaviour was used in the next experiment. Personal interviews with eCommerce managers inspired another proposed segmentation. Mangers autonomously observed two major in terms of profit generation but also contrary segments of customers. One of the segments brings together fashion-driven customers (they are interested in mainly new and fashionable items). The second one is “bargain hunters” – discount-driven customers who are eager to purchase products present on the market for a longer time, but they expect significantly high discounts. This segmentation is referred to as “fashion vs. discount”. In such segmentation we take into account the following dimensions:

- Average discount of a customer.
- The average number of days from product launch to transaction.
- Average order value.
- Number of orders.

The distribution of average discount used by a customer in their transactions is presented in Figure 4. Another dimension showing if customers purchase new or relatively old products is shown in Figure 5. Those graphs present that there are customers (cluster no 5) where the average discount is 2 per cent only, which means that in fact, they do not expect any discount. At the same time, they buy relatively new products so we can state that cluster 5 is fashion-driven. The opposite clusters are 1 and 4. Those clients expect about 50% discount, which is exceptionally high, and they buy old products so they can be referred to as “bargain hunters”. Every identified segment of customers expect separate pricing policy to achieve customer satisfaction when it comes to response for the offer.

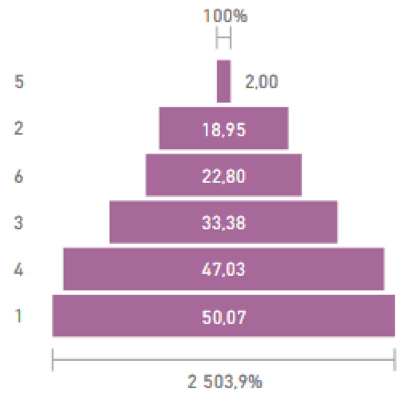


*Figure 4.Average discount distribution among clusters of customers.*

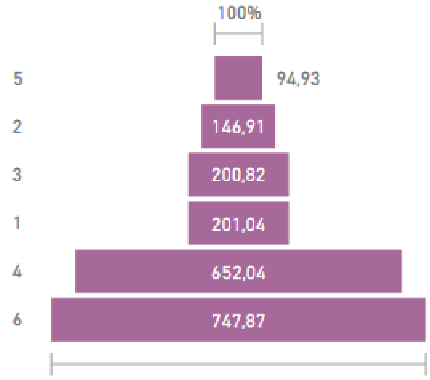


*Figure 5. Average number of days from product launch to transaction distribution between clusters.*

To provide the customer with an attractive offer, we have to classify the user in terms of the advertised product and the proposed purchase conditions (price, discount, additional value). Many marketers are moving away from traditional aggregate-level mass marketing programs and are searching for accurate methods of identifying their most promising customers to focus specifically on those individuals [Kaefer and others 2005]. In contrast to the mass marketing approach, a direct marketing approach evaluates the profitability of current and potential customers and uses this information to focus marketing resources on their best prospects. Various studies have found that a consumer's past purchases can act as a good predictor of future preferences and choice outcomes. Classification is a technique where we categorize data into a given number of classes (e.g. purchasing /non-purchasing category). The main goal of a classification problem is to identify the category/class to which a new data should be assigned, basing on the whole available characteristic of customers (demographic profile, previously purchased products, selected conditions). The most challenging issue in preparation of the classification model is to choose correct features to train a model.

# 4 Results of the research

Our research aims to find a method enabling providing each customer of internet store most accurate offer in terms of:

- product category,
- price level,
- discount level,
- moment in time to buy.

Such a proposal generation considers the history of customer purchases. We analyze what products he or she has purchased, but also the conditions of the purchase (price and discount). Among the whole population of customers, we are searching for purchase patterns. Those patterns are identified separately for various subsets of the population, eg. we identify them for both genders separately, but those subsets are also based on other demographic characteristics that we can identify. Having eCommerce data in most cases, we know the customer's gender as well as location (delivery place), but we do not know the customer's age.
To achieve such accurate proposal the combination of previously mentioned methods is required.

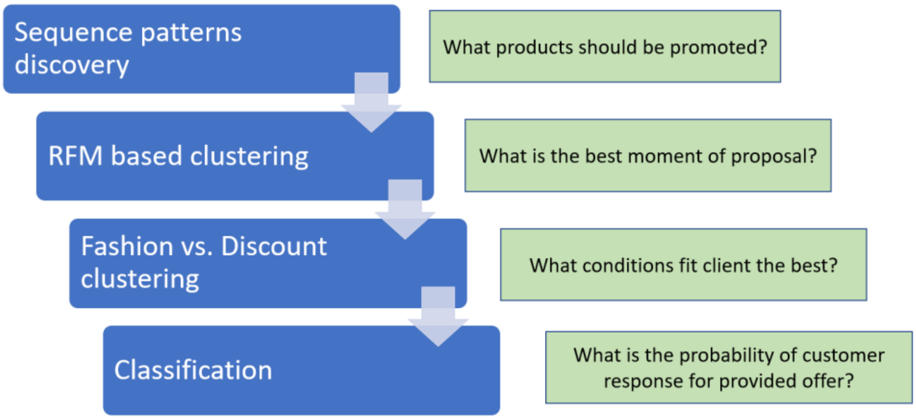


*Figure 6. Procedure to predict of the best offer for a customer.*

The procedure of the tailored offer generation for each individual customer is proposed in Figure 6. Basing on the first 3 steps we generate various options for proproposals. Lasts step (the classification model) is used to evaluate generated offers and selected only this / those with highest probability. The step is to avoid the situation where customer is overwhelmed with delivered offers. We would like to provide them with accurate number of messages not to discourage them.
In the last stage we build classification models where the target variable takes values: will buy / will not buy. The classification model is always generated in the context of a product category that was pointed out by sequence rules. In that case,

we know that there is an evident reference between the historical purchases and a current one.

Such an approach generates several classification models depending on the number of sequence rules that have been discovered. We train every model using the data of customers who bought product categories indicated as an antecedent of a sequence rule (their demographic profile, history of orders) and analyze if they have purchased the consequent product. Such an approach predicts the recommended product more accurate than only sequence rules because it includes the context of each individual customer.

The moment in time for offer delivery and proposed conditions in terms of discount result from the clustering applied to a customer.

Classification algorithm selected for this task is XGBoost, which is nowadays considered as very efficient and has been widely recognized in many machine learning and data mining challenges [Chen, Guestrin 2016]. This algorithm uses a gradient tree boosting technique.

Having trained and tested algorithm, we can apply it on every single customer data to estimate the probability of a selected product category to be purchased by a given client. As we can see in Figure 7, such an approach for rules with 8% confidence was validated in case of a specific client to give us more than 40% probability of a customer's next purchase.

| Customer | Antecedent | Rule confidence | Customer purchase probability |
|---|---|---|---|
| 4b8e9135013449fbc572622f3a1fa5ac@unity.pl | DKNY / Sweatshirts / Women | 5,041% | 48,900% |
| 12fdedad5670aa94cc236df62592bd8d@unity.pl | DKNY / Sweatshirts / Women | 5,041% | 48,700% |
| 23b07a69b044c4381d329a716f0a3152@unity.pl | DKNY / Trousers / Women | 6,714% | 48,400% |
| 20bf61d7ffd975b7d69289fd9c922852@unity.pl | Lauren / Socks / Unisex | 8,285% | 47,500% |
| d51a6cb734ee215e2e66d39e3ca7d7c0@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,900% |
| 05996c40624b776c6fed80771e6f6312@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,800% |
| 9a5453c0fb7043be3899e79c924af0e8@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,800% |
| 266421292d70aa3936079f75cb14956e@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,300% |
| cc30a9f4a7dfcd480a4faccf848c4538@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,300% |
| e6212112d8ecb6f1796db4b752b59794@unity.pl | Lauren / Socks / Unisex | 8,285% | 44,300% |
| 09332ea8b8442cd3f9e18f0750be2285@unity.pl | Calvin Klein / Lifestyle shoes / Women | 4,877% | 43,300% |

*Figure 7. Results of classification algorithm applied of a list of customers*

Having that knowledge, we can prepare a tailored offer for every single customer.

The proposed approach seems to give valuable insights regarding marketing activities. Authors will validate the real business value of the proposed method in the next phases of this research study.

# 5 Discussion

eCommerce is a very efficient sale channel for both retailers and their customers. Due to its convenience customers can easily switch between providers. Well adjusted offers and personalization can be considered as a critical factor in the process of building long-lasting relationships that limit customer churn phenomenon. The proposed method of combining several machine learning techniques in one comprehensive process for tailored offers generation brings promising results. Purchase probability rated for 40% is very high. When we compare it to the average conversion rate is between 1% and 4% [Digitalmarketinginstitute, 2020], the achieved number is impressive. We have to remember that predicted probability cannot be treated as expected campaign conversion rate. Several factors lower the final conversion (customer may miss the offer; they can purchase from other vendor and others). Such a method was tested in experimental campaigns resulted in 4 times higher Click-through rate and 3 times higher conversion when comparing to traditional campaign prepared by marketing analysts. Those results prove that the proposed approach can be efficient. Further research will be concentrated on tuning and adjusting the method to achieve higher efficiency of campaigns. Another direction of work will regard to store and include Big Data use cases including web events like clicks, page views, searches and others.

# 6 Conclusions

Data analysis based on machine learning can be applied in marketing and customer-related data. The topic is widely discussed in the literature however research regards mostly to customer churn problem and in terms of customer segmentation, satisfaction, and loyalty base on surveys and more psychological grounds. Authors proposed the method of data analysis that is a combination of various techniques and algorithms of machine learning that is aimed to give a better result than typical methods to raise customer loyalty in the retail business.

# References


1. Agrawal, R., Imieliński, T., Swami, A. (1993): Mining association rules between sets of items in large databases. In Acm sigmod record (Vol. 22, No. 2). ACM.
2. Amin, A., Shah, B., Khattak, A.M., Moreira, F.J.L., Ali G., Rocha Á., Anwar S. (2019).: Cross-company customer churn prediction in telecommunication: a comparison of data transformation methods. International Journal of Information Management, 46.
3. Bhattacharya, C. B. (1998). When customers are members: Customer retention in paid membership contexts. Journal of the Academy of Marketing Science, 26(1), 31–44.
4. Bloemer J.M.M., Kasper H.D.P. (1995): The complex relationship between consumer satisfaction and brand loyalty, Journal of Economic Psychology vol.16.
5. Boström P., Filipsson M. (2017): Comparison of User Based and Item Based Collaborative Filtering Recommendation Services.

6. Burez J., Van den Poel D. (2007). CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. Expert Systems with Applications, 32(2).
7. Burez J., Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. Expert Systems with Applications, 36(3).
8. Chen T., Guestrin C. (2016): Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. ACM.
9. Ciok E. (2006): Trust, but control - basic information about Mystery Shopping (pol. Ufaj, ale kontroluj -czyli podstawowe informacje na temat Mystery Shopping), Zarządzanie jakością, nr 1.
10. Clauss T., Harengel P., Hock M. (2018): The perception of value of platform-based business models in the sharing economy: determining the drivers of user loyalty. Review of Managerial Science.
11. De Caigny A., Coussement K., De Bock K.W. (2018): A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269 (2).
12. Doligalski T. (2018): E-commerce 2018 - convenience of customers, sweat and tears of sellers (pol. E-commerce 2018 – wygoda klientów, pot i łzy sprzedawców), Gazeta SGH.
13. https://digitalmarketinginstitute.com/blog/what-is-a-good-conversion-rate-for-ecommerce, access date, 2020-08-01
14. Fathian M., Hoseinpoor Y., Minaei-Bidgoli B. (2016): Offering a hybrid approach of data mining to predict the customer churn based on bagging and boosting methods. Kybernetes, 45(5).
15. Gordini, N.,: Market Driven Management (2017). A critical literature overview. Symphonya Emerging Issues in Management, 2, 97-109. http://dx.doi.org/10.4468/2010.2.08gordini
16. Gordini, N., Veglio, V. (2013). Using neural networks for customer churn predictionmodeling: preliminary findings from the italian electricity industry (pp. 1–13)Proceedings de X° Convegno Annuale della Società Italiana Marketing: "Smart Life. Dall'Innovazione Tecnologica al Mercato", 3–4 Ottobre 2013obre 2013. Milano, Italy: Università degli Studi di Milano-Bicocca.
17. Han J., Cheng H., Xin D., Yan X. (2007): Frequent pattern mining: current status and future directions. Data mining and knowledge discovery, 15(1).
18. Hung S.Y., Yen D.C., Wang, H. Y. (2006): Applying data mining to telecom churn management. Expert Systems with Applications, 31(3).
19. Jones T.O., Sasser W.E. (1995): Why Satisfied Customers Defect, Harward Business Review, nr 6.
20. Kaefer F., Heilman C.M., Ramenofsky S.D. (2005): A neural network application to consumer classification to improve the timing of direct marketing activities. Computers & Operations Research, 32(10).
21. Lee K.C., Kwon S. (2008): Online shopping recommendation mechanism and its influence on consumer decisions and behaviors: A causal map approach. Expert Systems with Applications, 35(4).
22. Lu N., Lin H., Lu J., Zhang G. (2012): A customer churn prediction model in telecom industry using boosting. IEEE Transactions on Industrial Informatics, 10 (2).
23. Massa P., Avesani P. (2004): Trust-aware collaborative filtering for recommender systems. In OTM Confederated International Conferences "On the Move to Meaningful Internet Systems". Springer, Berlin, Heidelberg.

24. Miguéis, V.L., Van den Poel, D., Camanho, A.S., Cunha, J.F. (2012): Modeling partial customer churn: On the value of first product-category purchase sequences. Expert systems with applications, 39(12).
25. Newman J.W., Werbel R.A. (1973): Analysis of brand loyalty for major household appliances, Journal of Marketing Research, Vol. 10.
26. Oliver R.L. (1999): Whence consumer loyalty, Journal of Marketing, vol.63.
27. Pappas I.O., Kourouthanassis P.E., Giannakos M.N., Lekakos G. (2017): The interplay of online shopping motivations and experiential factors on personalized e-commerce: A complexity theory approach. Telematics and Informatics, 34(5).
28. Prasad U.D., Madhavi S. (2012): Prediction of churn behavior of bank customers using data mining tools. Business Intelligence Journal, 5(1).
29. Rawson A., Duncan E., Jones C. (2013): The truth about customer experience. Harvard Business Review, 91(9).
30. Schafer J.B., Frankowski D., Herlocker J., Sen S. (2007): Collaborative filtering recommender systems. In The adaptive web. Springer, Berlin, Heidelberg.
31. Sharma A., Panigrahi D., Kumar P. (2013): A neural network based approach for predicting customer churn in cellular network services. arXiv preprint.
32. Srikant R., Agrawal R. (1996): Mining sequential patterns: Generalizations and performance improvements. In International Conference on Extending Database Technology. Springer, Berlin, Heidelberg.
33. Stum D., Thiry A. (1991): Building Customer loyalty, Training and Development Journal.
34. Sudolska A. (2011): Managing customer experience as a key factor in the process of building their loyalty (pol. Zarządzanie doświadczeniem klientów jako kluczowy czynnik w procesie budowania ich lojalności), Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 660, Ekonomiczne problemy usług nr 72, Uniwersytet Szczeciński, Szczecin.
35. Sulistiani, H., Tjahyanto, A. (2017): Comparative Analysis of Feature Selection Method to Predict Customer Loyalty. IPTEK The Journal of Engineering, 3(1).
36. Tsai C.F., Lu Y.H. (2009): Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10).
37. Urban W., Siemieniako D. (2008): Customer loyalty. Models, motivation and measurement (pol. Lojalność klientów. Modele, motywacja i pomiar), PWN, Warszawa.
38. Verbeke, W., Martens, D., Mues, C., & Baesens, B. (2011). Building comprehensible customer churn prediction models with advanced rule induction techniques. Expert Systems with Applications, 38(3), 2354–2364. http://dx.doi.org/10.1016/j.eswa.2010.08.023
39. Yu X., Guo S., Guo J., Huang X. (2011): An extended support vector machine forecasting framework for customer churn in e-commerce. Expert Systems with Applications, 38(3).
40. Zhao J., Dang X. H. (2008): Bank customer churn prediction based on support vector machine: Taking a commercial bank's VIP customer churn as the example. In 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing, IEEE.