ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 12 (2022) No. 4, pp. 57-68
DOI:10.33168/JSMS.2022.0404

A Proposed Churn Window for Non-Contractual Purchases

Sunther Ganeson, Lew Sook Ling, Siti Fatimah Abdul Razak

Faculty Information Science and Technology, Multimedia University, Malaysia

sllew@mmu.edu.my

Abstract. The identification of retainable online non-contractual customers is pertinent for the operations and growth of non-contractual online businesses, since there are no obligations for customers to be loyal to a particular online business. This research paper, therefore, aimed at proposing a churn window for non-contractual purchases. Firstly, a prediction model based on the average customer historical buying pattern was presented. Secondly, based on customers' purchase history trends, a new churn window was proposed. Generally, customers have their own unique churn window based on their purchasing behaviour. This is different from the traditional definition of churn window, whereby a defined churn window is applied across customers regardless of their individual purchasing history. Results revealed that the proposed churn window model has better accuracy compared with the traditionally defined window, which is generically applied to all customers. This leads to the conclusion that the proposed prediction model and the churn window model can be useful to support marketing strategies and activities of non-contractual online businesses.

Keywords: non-contractual, online customer, churn prediction, churn window.

1. Introduction

The Covid-19-wrought pandemic has heralded a new era in the retail industry worldwide, with lockdowns and movement control orders creating formidable barriers to both producers and users. Consequently, many entities from both sides of the divide have resorted to going online to supply and procure good and services, thereby leading to an increase of global retail trade via e-commerce from 14% in 2019, to 17% in 2020 (Unctad, 2021). It is estimated that ecommerce will take up 22% of global sales by the year 2023. Every customer is unique and by understanding online customer's behavior can assist to create an effective digital strategy for an online business (Estay, 2022).

If businesses hope to survive in today's catch phrase aptly termed as the "new normal", there is an urgent need to embrace and adopt new and innovative strategies to reach out to their customers among the wide masses amidst competition that has risen to levels previously unheard of. One of these includes the identification of retainable customers, that is seen as essential for the functioning and growth of any online business (Ahn et al., 2020). This research, therefore, is aimed at proposing a prediction model and a churn window for non-contractual purchases. This new window that is proposed will be used as a gauge to determine whether a customer has churned or not churned. Data from an online superstore from the US consisting of non-contractual customers' dataset will be used for testing the proposed prediction model and churn window. The range of customers consists of both corporate and individual consumers.

The underlying tenet in this study is that each customer is treated as unique, whereby each customer's particular buying behaviour is used as a determining factor to ascertain his or her churn status. The main aim is to produce a usable prediction model with an appropriate churn window definition that can be used for all the non-contractual customers.

2. Literature Review

Previous research and studies related to non-contractual purchases and churn window concepts are discussed in this section.

2.1. Non-contractual purchase

The concept of churn in an online business is a tricky one, since in a regular online purchase, there are normally no contracts involved (Banday & Khan, 2021; Merchie & Ernst, 2022). When a customer decides to purchase an item online, he or she can proceed with minimal registration, without necessitating the need for a contract. Customers involved in these transactions are usually termed "non-contractual customers". The absence of a contract makes it difficult to ascertain when or whether a customer has churned. To date, there is little research on customer churn prediction for non-contractual relationships (Xia & He, 2018) . Based on (Xia & He, 2018) an online customer can be permanently or intermittently lost. The two definitions based on their research are as follows.

2.1.1. Intermittent loss

Intermittent loss means that customers did not buy the enterprise's goods within the time threshold. However, it does not mean that the enterprise has completely lost the customer, as he or she may continue to purchase the product or service offered beyond the time threshold. What matters is the decrease in trading frequency (Tran & Thai, 2022).

2.1.2. Permanent loss

Permanent loss refers to situations where no future purchases may take place from particular customers (Tran & Thai, 2022). In such a scenario, although customers' accounts may lie dormant for a substantial period, the seller will not write off the customer's account, simply because they have no way of knowing whether these customers are lost permanently or will possibly use their registered accounts to log in at another time. Permanent loss also means that the customer has exhausted their purchasing behaviour with that particular enterprise, possibly due to many reasons, such as a change in the customer's spending habits, the absence of the need for the product or service from the enterprise, or the customer's natural death.

This paper has adopted customer inactivity for a period of time to determine the status of the customer. This period of time is defined as the churn window. If there are no activities during the churn window, the customer's status is defined as churned; the presence of activities puts their status as non-churned.

2.2. Traditional churn window

A common way to determine whether a customer has churned is by using a time window to track the customer's activity. To start the process, a timer starts whenever a customer makes a purchase. Based on a set period timeline, the customer is then monitored for any activity. If there is no activity, the customer is considered as churned (Caigny et al., 2018) as illustrated in Figure 1.

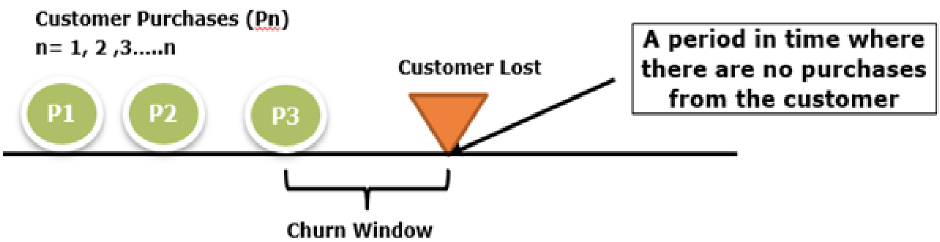

Fig. 1: Traditional churn window.

The churn window is not specifically defined and is generally the same for all customers (TowardsDataScience, 2017). If the window period is set as a three-month period which is called Trimester [T] (Migueis et al., 2012), customers will be labelled as churners if they do not purchase any item during the pre-defined window period or spend less than forty percent compared with a previous T which is taken as a reference. This narrative is pictured in Figure 2, where there are two types of hypothetical customers based on their distribution of purchases according to T. In Figure 2, the customer who spent 35 Euros in Trimester 4, which is 38% less compared with the purchase in Trimester 3, is be labelled as a churner. As for the non-churner, there is no subsequent trimesters where the sale is less than forty percent. For example, the non-churner spent 40 Euros in Trimester 4, which is only 44% percent compared with Trimester 3. Thus, this customer is labelled as a non-churner. Both the inactivity and reduction of sale within a time frame are some typical indications that have been used to determine churn.

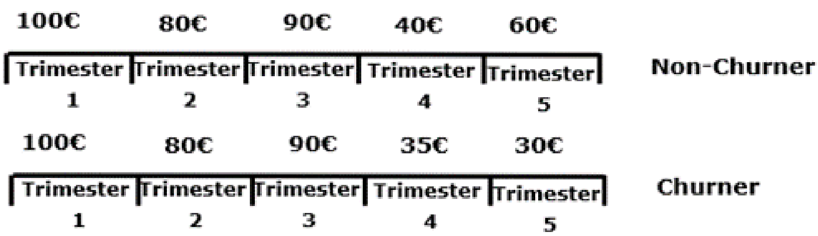

Fig. 2: Churn window based on expenditure by trimester

This paper therefore adopts the time frame concept for the churn window definition. Nevertheless, the time frame was analyzed according to months, since the data used in this research contained invoices that are monthly in nature. The new churn window proposed requires all prior purchases in determining the length of the churn window.

2.3. Sliding window concept

The sliding window concept uses two-user-configurable parameters to define the window width. One parameter is called the churn definition period which refers to a period of customer inactivity (i.e., not making a purchase) in successive months (e.g., customer did not make a purchase in three successive months) (Mirković et al., 2018) as shown in Figure 3. The other parameter refers to monthly features, in which various features are examined (examples are amount spent in March versus amount spent in February).

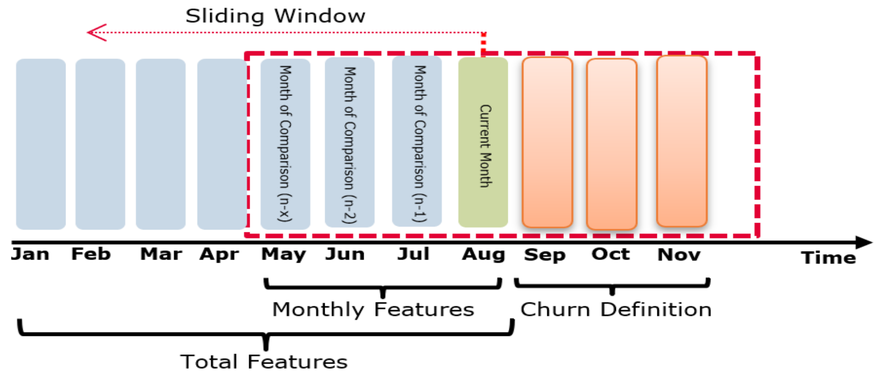

Fig. 3: Sliding window concept

This paper has adopted the general churn window definition which is based on a period of time. The next section will discuss the methodology that has been adopted.

3. Methodology

The adopted methodology is using the CRoss Industry Standard Process for Data Mining (CRISP-DM) process in the research lifecycle. This research has adopted the concept of "window period" as presented in Figure 3 to determine the churn window for non-contractual online customers (Mirković et al., 2018). Nevertheless, there are some adaptations made to the window concept. The main change is that the window is unique to each customer based on each customer's own purchase history. These are some of the improvements that fill the research gap in the area of customer churn prediction for non-contractual relationships (Xia & He, 2018).

3.1. Prediction model

In this study, the churn window for every customer was determined using the purchase data history for each customer, available from the dataset from the Superstore, comprising 793 customers and 4922 orders, in the years spanning 2015 to 2018.

The main criteria to determine the churn window is that there must be a minimum of two transactions prior to the cut-off date which will be called the check point (CP) for each customer analyzed. Based on the number of transactions, the amount of time taken from one purchase to the next purchase is added and then averaged out. For example, assuming customer A made 4 purchases (T1, T2, T3 and T4) at different intervals during the period in which data were obtained until the time of the CP, there would be three intervals called Period-One (P1), Period-Two (P2) and Period-Three (P3). The average tabulated for these 3 intervals would be known as Periodicity (P) and is unique for customer A. Next the calculated P is then added to the latest purchase date (T4). This new date is now called Next Projected Transaction (NPT). Subsequently, the following logic is applied. If NPT is after the CP, the customer is assumed as not churned and given the value "1". If NPT is before the CP, the customer is assumed as churned and given the value "0". This process is then applied to all customers individually.

3.2. n-th Month Churn Window (nMCW)

The first window concept which was adopted based on previous research papers, is called the n-th month churn window (nMCW) in this research paper. The window period is based on a multiple of n-th months where n can be a value ranging from 1 to 3 months (Malyar, MV Mykola, & Maryana, 2020). The nMCW starts from the CP. The validation is done from the CP until the end date of nMCW. Based on the difference of the average value of the nMCW and average predicted value, the ideal number of months (n value) has been determined in this study.

Figure 4 depicts the details of the churn window model using nMCW for customers predicted as churned. If the next transaction happens outside the validation period, then the prediction "0" is correct. Examples are customer A1 and customer A3. If a transaction happens within the validation period, the prediction is then noted as incorrect, as indicated by customer A2.

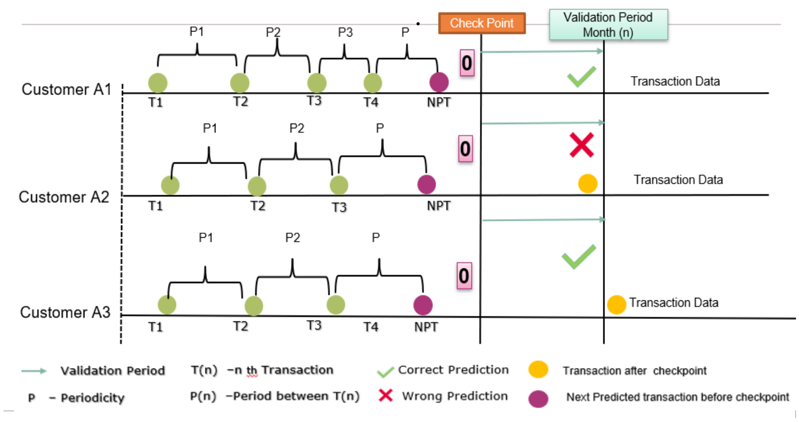

Fig. 4: *n*-Months churn window (predicted as churned)

Figure 5 describes customers who have been predicted as not churned. The NPT for these customers is after the CP. Customers B1, B2 and B3 have been predicted correctly. As for customer B4, the next transaction happens outside of the validation period. Thus, the prediction for customer H is depicted as incorrect. In summary if the next transaction happens after the CP and there is a transaction within the window period, the prediction is correct.

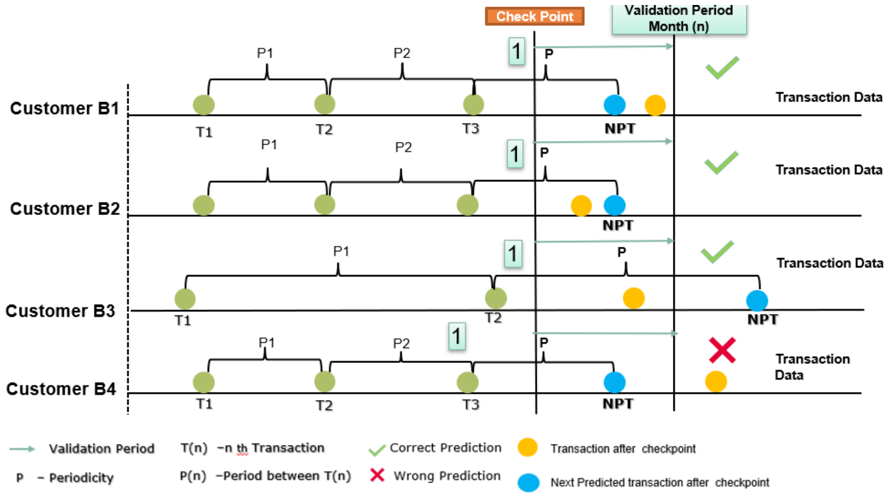

Fig. 5: n-Months churn window (predicted as not churned).

3.3. Variable Churn Window (VCW)

The proposed churn window is known as the Variable Churn Window (VCW). The definition is as follows: VCW = NP, where N can carry a value of 1, 2, 3 or 4. The N value has been determined according to this study. This churn window will be unique for each customer, where for each customer analyzed, the VCW is added to their last transaction. Figure 6 depicts customers who have been predicted as churned. Subsequently, a check is done within the VCW for any transactions. If the next transaction happens outside the validation period, then the prediction "0" is correct, as seen in the case of customer C2. If a transaction happens within the validation period like in the case of customer C3, the prediction is then noted as incorrect. As for customer C1, the VCW is before the CP, thus the prediction is denoted as correct. Figure 7 shows examples of customers where the prediction value is "1" which denotes that the customer has not churned. If the next transaction happens within the validation period, then the prediction "1" is correct. This is shown in example of customer D1 and customer D2 in Figure 7. If a transaction happens outside validation period, the prediction is then noted as incorrect. Customers who have only a single transaction before the CP are out of the prediction scope.

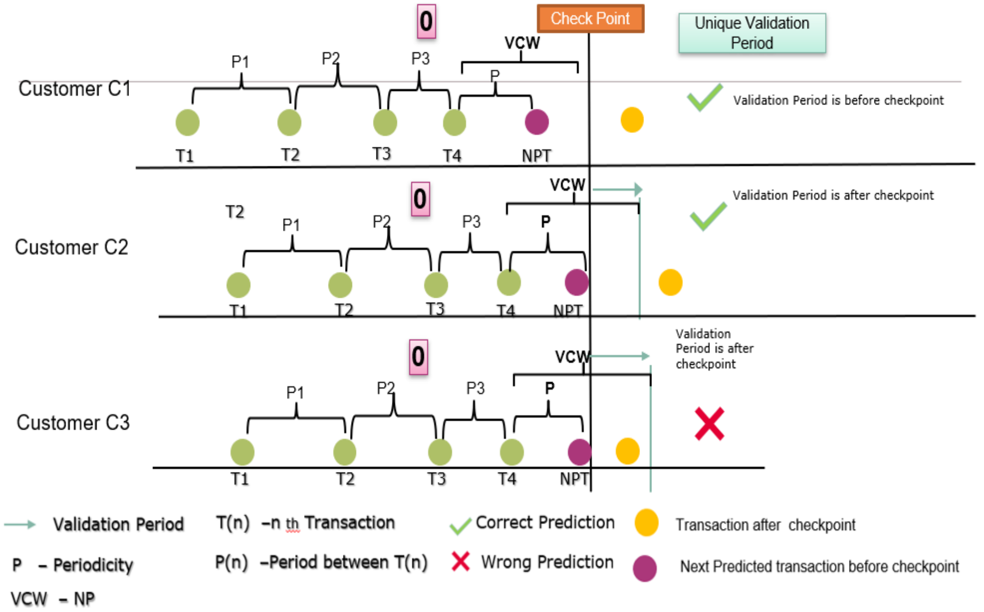

Fig. 6: Unique validation period for predicted churned.

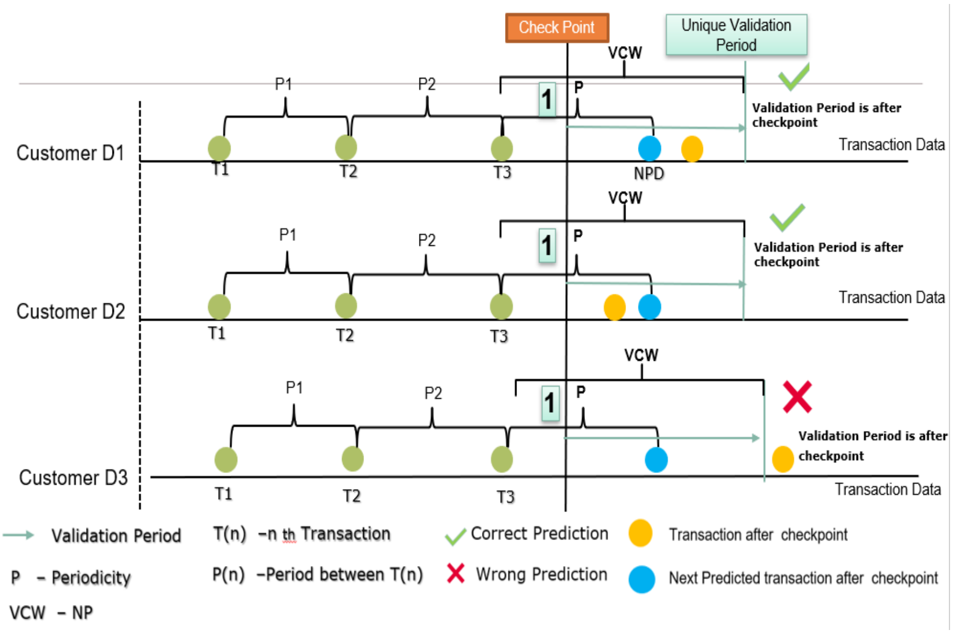

Fig. 7: Unique validation period for customer predicted as not churned.

4. Discussions

The churn window for the verification of data is based on data collected from the online sale of a Superstore from 2015 to 2018. The data were used to verify the accuracy of the predicted churn value.

If the predicted value is churned, there will be verification during the churn window, indicating whether the customer has really churned. This was done for all customers who were identified as churned and not churned. A chart was then plotted based on the predicted value and the actual value using the nMCW and VCW(NP) from each month. The chart was plotted for every month from January 2015 to December 2018 as shown in Figure 8.

As a next step, the average value of each plotted line was computed, in this case, the average value for non-churned prediction, 1MCW, 2MCW, 3MCW and VCW (1P,2P,3P,4P). A comparison was then done between the predicted non-churned average value and the different churn window average value. This was tested for different window periods, after which, a recommendation was given for the window values for n and N. The difference between the average values for the different windows and predicted value was the determining factor in choosing the proposed window type for this research. Two kinds of churn windows were validated - the churn windows nMCW and the proposed VCW. Figure 8 shows the results of comparison between what was predicted, and the verification that was done with real data. The graph shows the results for customers that will not churn (denoted as 1) compared with real results from the different churn windows. Based on nMCW churn window concept, 3MCW has the closest value to the predicted non-churned value compared with 1MCW and 2MCW. Thus, the ideal number of n in nMCW is 3. As for the VCW churn window concept, 3P has the closest value to the predicted value. This is in comparison to 1P, 2P and 4 P. Thus N=3 is ideal for the VCW churn window.

Overall, when VCW and nMCW results are compared, the differences between the average value of what is predicted and the average of VCW= 3P has the smallest value. The smallest value means the predicted value is the closest to the churn window that has been examined among all the windows concept. The percentage difference between what is predicted and what is verified is summarised in Table 1, based on the different churn windows used in this study. The difference between average prediction value (57.7%) and different churn windows average values are shown in the third column of Table 1.

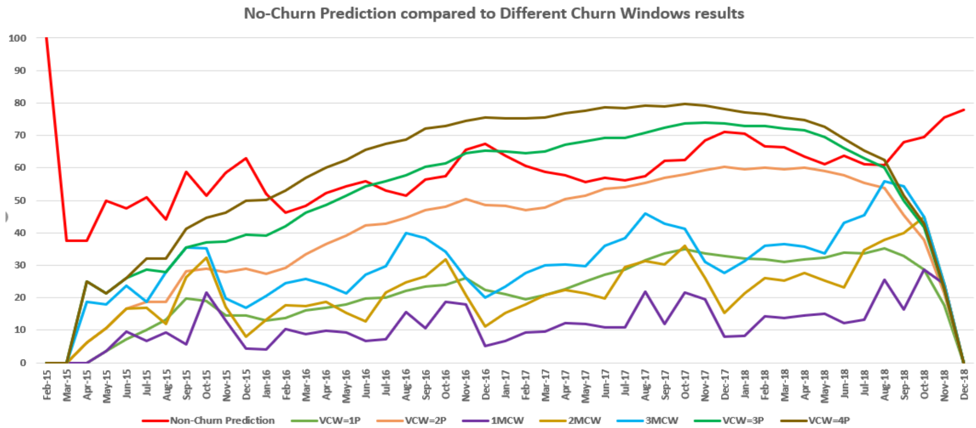

Fig. 8: Comparison of prediction of customers who will not churn against nMCW and VCW.

In contrast with previous studies, this paper has proposed a novel generic churn window called VCW, which is unique for every customer based on their own unique historical spending data. The window for verification starts from the last transaction plus the VCW period. Previous studies used the same standard window, which was applied across all customers and starts from the CP as shown in this paper using nMCW. The new churn window proposed focuses on the fact that each customer is to be treated as unique with his or her own unique buying pattern. It does not assume that each customer will have the same churn window period (examples are 1 month or 2 months). Rather the P for each customer is derived based on each customer's individual buying pattern. In this paper, it is found that VCW=3P is preferred based on the data analyzed. 3P is closest to the predicted value compared to 3MCW. This paper has also introduced a prediction model concept which is then verified within the churn window that has been discussed above.

Table 1: Value difference between what is predicted and different churn windows

Churn window	Average Value (%)	Difference
VCW=1P	22.7	35.0
VCW=2P	43.5	14.3
*VCW=3P	56.3	1.5 *
VCW=4P	63.5	-5.8
1MCW	11.5	46.3
2MCW	21.6	36.1
3MCW	30.9	26.8

VCW: Variable Churn Window = N x Periodicity
nMCW: n-th Months Churn Window
*VCW=3 average value is the closest to the Predicted non-churn value

5. Conclusion

Preventing Customer churn is a pertinent (Jaiswal et al., 2018) factor in the success of retailers involved in online purchases. While churn cannot be avoided, reducing it brings benefit and profitability (Zhao et al., 2021). This research contributes to further expansion of the prediction model and churn window by using existing data to determine churn window which is relative to customers' buying behaviour. The other domain that will benefit from this research consists of practitioners, specifically those in the marketing area (Owinat, 2021). The marketing department will be able to use the prediction data for customers who are deemed as churned. The marketers can then take appropriate actions or steps to entice these customers back to the online platform.

As a next step and future research, the new attribute called P can be further explored by machine learners especially in in the area supervised machine learning.

Authors' Contribution

Sunther, G., Lew, S. L., & Razak, S F. A comprehended the idea and contributed to the research article. All authors contributed to the writing, editing, and consent of the final manuscript.

References

Ahn, J., Hwang, D., Kim, D., Choi, H., & Kang, S. (2020). A survey on churn analysis in various business domains. *IEEE Access*, 8, 220816-220839.

Banday, S. & Khan, S. (2021). evaluation study of churn prediction models for business intelligence. In Big Data Analytics. *Auerbach Publications*, 201-213.

De Caigny , A., Coussement, K., & De Bock, K. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, 269(2), 760-772.

Estay, B. (2022). https://www.bigcommerce.com/blog/online-shopping-statistics/#5-essential-online-shopping-statistics.

Jaiswal, A., Niraj, R., Park, C., & Agarwal, M. (2018). The effect of relationship and transactional characteristics on customer retention in emerging online markets. *Journal of Business Research*, 92, 25-35.

Malyar, M., MV Mykola, R., & Maryana, S. (2020). Churn prediction estimation based on machine learning methods. 1-4.

Merchie, F. & Ernst, D. (2022). Churn prediction in online gambling. *arXiv preprint arXiv:*2201.02463.

Migueis, V., Poel, D. V., & Camanho, A. (2012). Modeling partial customer churn: On the value of first product-category purchase sequences. *Expert Systems with Applications*, 11250-11256.

Mirković , M., Milisavljević, S., & Gračanin, D. (2018). A framework based on open-source technologies for automated churn prediction in non-contractual business settings. 2nd International Scientific Conference ITEMA.

Owinat, A. (2021). Relationship marketing. *Customer Churn Prediction.* Customer Churn Prediction.

TowardsDataScience. (2017). How to calculate churn rates across industries. Retrieved from https://towardsdatascience.com/how-to-calculate-churn-rates-across-industries-68c692b64605.

Tran, T. & Thai, K. (2022). Rider churn prediction model for ride-hailing service: A machine learning approach. *International Conference on Artificial Intelligence and Big Data in Digital Era*, 301-314. Springer, Cham.

Unctad. (2021, March 21). how-covid-19-triggered-digital-and-e-commerce-turning-point. Retrieved from Unctad.org: https://unctad.org/news/how-covid-19-triggered-digital-and-e-commerce-turning-point.

Xia, G. & He, Q. (2018). The research of online shopping customer churn prediction based on integrated learning. *International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018)*, 259-267. Atlantis Press.

Zhao, M., Zeng, Q., Chang, M., Tong, Q., & Su, J. (2021). A prediction model of customer churn considering customer value: An empirical research of telecom industry in China. *Discrete Dynamics in Nature and Society*.