Journal of Information Systems and Management Research
Bilişim Sistemleri ve Yönetim Araştırmaları Dergisi

http://dergipark.gov.tr/jismar

Araştırma Makalesi / Research Article

Sadakat Programında Müşteri Kayıp Tahmini: Bir Vaka Çalışması

Yeşim KOCA[*,a], Burak Emre SÖĞÜT[b], Sona MARDIKYAN[c]

[a,*] *Boğaziçi Üniversitesi, Yönetim Bilişim Sistemleri Bölümü, İSTANBUL, 34342, TÜRKİYE*
[b] *Boğaziçi Üniversitesi, Yönetim Bilişim Sistemleri Bölümü, İSTANBUL, 34342, TÜRKİYE*
[c] *Boğaziçi Üniversitesi, Yönetim Bilişim Sistemleri Bölümü, İSTANBUL, 34342, TÜRKİYE*

MAKALE BİLGİSİ

Alınma: 20.11.2019
Kabul: 25.12.2019

Anahtar Kelimeler:
Müşteri Kayıp Tahmini,
Sadakat Programı,
Lojistik Regresyon,
Yapay Sinir Ağları

Sorumlu Yazar
e-posta:
yesim.koca@boun.edu.tr

ÖZET

Artan rekabetle birlikte, yeni müşteri edinmek her geçen gün daha maliyetli bir hale geldi ve şirketler var olan müşterilerini elde tutmak için daha fazla çaba harcar oldu. Tahminleyici modellerin geliştirilmesi ve bilgisayar teknolojisindeki gelişmeler sayesinde büyük miktarda veriyi analiz edebilme kabiliyeti, şirketlere, hangi müşterilerinin müşterileri olarak kalmaya devam edeceğini ve hangilerinin terk etmeye meyilli olabileceğini güvenilir bir şekilde tahmin etme imkanı verdi. Bu çalışmada, Türkiye'de mobil bir sadakat uygulamasından elde edilen veriler; restoran, perakende ve e-ticaret olmak üzere faaliyet gösterilen üç farklı sektör üzerinden analiz edilecektir. Müşterilerin gelecek çeyrekteki aktiflik durumlarını tahmin etmek için her sektörde iki farklı tahminleyici model geliştirilmiştir. Geliştirilen modellerde Lojistik Regresyon ve Yapay Sinir Ağları yöntemleri kullanılmıştır. Geliştirilen tüm bu altı modelde, bütün sektörlerde genel olarak %90'ın üzerinde doğruluk oranına ulaşılmasının yanı sıra, Yapay Sinir Ağları doğruluk, hassasiyet ve özgüllük ölçütleri açısından Lojistik Regresyona kıyasla daha iyi bir performans sergiledi.

Churn Prediction in Customer Loyalty Program: A Case Study

ARTICLE INFO

Received: 20.11.2019
Accepted: 25.12.2019

Keywords:
Churn Prediction,
Loyalty Program,
Logistic Regression,
Artificial Neural Network

Corresponding Authors
e-mail:
yesim.koca@boun.edu.tr

ABSTRACT

With the ever-growing competition, globally and acquisitioning new customers becoming more expensive, more and more companies put effort on retaining their existing customers. Development of predictive models and the ability to analyze large chunks of data thanks to the advancements in computer technology gave companies the ability to reliably predict which of their customers are likely to keep being their customers and which ones are likely to churn. In this study, data taken from a mobile loyalty application in Turkey will be analyzed on three different sectors restaurant, retail and e-commerce they operate in. Two different predictive models for each sector are developed to estimate the activity status of customers in the last period. The methods that will be used in models are chosen as Logistic Regression and Artificial Neural Network. While all these six models developed resulted in over 90% overall accuracy rates across the board, Artificial Neural Networks performed much better in overall accuracy as well as in sensitivity and specificity measures.

1. INTRODUCTION *(GİRİŞ)*

The ability to store much bigger and complex data had been one of the factors that launched the Information Age. These advancements changed the way companies operate in a way that can be similar to Industrial Revolution. With data becoming one of the major driving forces for a company to be more and more successful, several different ways of utilizing this data has been proposed over the years, one of them being data analytics. Another change that Information Age has brought to the business world was the immense competition on a global scale compared to before. As a result, data analytics has become one of the most important tasks that a company has to perform for sake of their business. An important part of this data analytics appliances in the business world is the retention of existing customers and avoiding churn. Different articles written [1-3] found out that acquisition of a new customer costs up to 5 times compared to the retention of an existing customer. At this point, Ling and Yen, in [4], defined Customer Relationship Management (CRM) as an enterprise approach to better understand the customer needs so that acquisition and retention of customers can be improved. Berson, Smith, and Thearling [5] split the CRM framework in to operational and analytical CRM. Furthermore, they stated that appropriate data mining tools utilized by analytical CRM could help a company to make better strategic decisions regarding their business.
One of the endeavours of analytical CRM is determining which customers are still active and which ones have churned. The term “churn” has been defined differently in the literature. While Glady, Baesens, and Croux [6] gave the definition for a churned customer as “the customers whose transactions are lower than the appointed threshold”, Van den Poel and Larivière [7] defined it as “the customers who closed their accounts with the firm”. As a result, retention of customers and avoiding their churn has become an important issue for managers of CRM [8,9]. In addition to this, in their paper, they stated that prediction of churners is even more difficult in a non-contractual setting, meaning that since no contracts between company and the customers exist, determining the point at which customers stop being ones for the company is an issue of high complexity.

Prediction of future revenue streams as well as the active customer count a company will have within a certain period of time is fundamentally important for any company since these information will give them insight on their business on top of allowing them to formulate much more sound business strategies. Since with the computational power one can have with the current technologies, these predictions can be made much more effectively and quickly. On top of the ever-growing competition in the economy forces companies to make such predictions in order for them to see the customers are likely to churn and therefore develop marketing strategies to avoid such a situation. In the literature, various classification techniques are used for this purpose. Nie, Rowe, Zhang, Tian, and Shi, [10] applied and compared logistic regression and decision tree algorithms to build a churn prediction model using credit card data collected from a real Chinese bank. According to Dreisetl & Ohno-Machado [11], one of the key advantages of the logistics regression method is the ability to test the statistical significance of coefficients since this test can then be used to build model incrementally. Because of the nature of neural networks, same could not be said for it. [12] and [13] compared neural networks with decision trees and both found that there was not much difference in accuracy between these methods. [14] also compared neural networks with logistic regression and decision trees in simulated direct marketing data. Artificial neural networks performed better than the other two techniques when the sample size is small. However, in large sized samples, decision trees and logistic regression performed better results– with logistic regression being generally superior to decision trees. [15] combined demographic, billing and usage data to predict the churners in the internet service provider setting and evaluated the accuracy of three different data mining methods: logistic regression, decision tree and artificial neural networks. All the three models performed similarly with artificial neural network showing a little higher accuracy rate of 89.08%, logistic regression 89.01% and lastly decision tree 87.74%. [16] also used the same data mining techniques (decision tree, artificial neural network, logistic regression). These techniques are compared according to the misclassification rate and lift chart and it is concluded that logistic regression is the best model to use to predict the churn rate of customers.

The main focus of this research is to accurately predict the customers which will make a transaction within the next quarter or churn within that time period based on their past transaction history as well as their demographic information and other behavioral traits. In this context, several different models have been proposed to predict whether or not a customer is likely to defect from the company based on their past behavior. Some of the well-known techniques such as artificial neural networks, logistic regression are used to make this prediction.

2. METHODOLOGY AND DATA *(METHODOLOJİ VE VERİ)*

In this study, logistic regression, and artificial neural networks are used for predictive modelling. For decades, regression analysis has become one of the major components of data analytics regarding the description of the relationship between a response variable and its explanatory variables. Since linear regression encounters some problems with a dichotomous dependent variable, logistic regression was developed to better explain this relationship [17]. Artificial neural networks are, as the name implies, methods that are inspired by attempts to simulate the process in networks of nerve cells of the biological neural systems. Artificial neural networks use nonlinear mathematical equations in order to develop meaningful relationships between input and output variables through a learning process.

2.1. Data *(Veri)*

In this study, the data which is provided from a mobile loyalty application in Turkey, will be used to predict whether or not a customer is likely to be active in the following period. While the geographical scope of the company consists of only Turkey, the business sectors that in which it operates contains a wide variety from automobile industry to retail, holiday or restaurants. The company provides some discounts for its customers while they are making purchases from retail stores, using e-commerce platforms, spending in restaurants etc. The sample data which belongs to the period between 1 January 2018 and 31 March 2019 was delivered by two parts: the transaction based data and customer based data.

Due to legal reasons concerning KVKK, the company name the data has been taken cannot be given and the data itself did not include any variables that may indicate any personal information. The company, through partnerships it forms with other brands, offers its customers special offers in these brands from different sectors such retail, restaurants, cosmetics, tourism and many others as a loyalty-hospitality program in which the more actively the customer engages the better offers they receive. Through the years they have been operating, company have been able to obtain more than 2 million customers. The customer based data consisted of a randomized unique identifier, age as a continuous variable and gender as categorical. The variables included in the transaction based data can be summed up as, the transaction unique identifier, customer identifier that will be used to merge this table with customer based table, the amount paid and discount received by the customer, brand in which the transaction has taken place has also been made anonymous as an integer value, and the sector name of which the brand belongs in which will be a crucial point in this paper. The transformation of the raw data into variables that can be used in both models employed in this paper is explained below in the “Data Preprocessing” section.

2.1.1. Data preprocessing *(Verinin Hazırlanması)*

Data preprocessing is one of the most significant steps since ensuring a final predictive model with a high accuracy rate cannot be achieved without this step. In this study, data preprocessing steps were performed in SAS Enterprise Guide tool while predictive models were developed in IBM SPSS tool. Since the company itself also employs several predictive models periodically, they have created a lot of fields dedicated to serve those models. And yet, most of these fields which are not relevant to the model that is being developed in this study, should not be included. As a result of this, from the total count of 74 fields, from both transactional data and customer data, more than 50 of them were removed. Finally, transaction based data mainly consisted of a unique transaction identifier, customer, brand and store identifiers in which the transaction occurred, amount paid and the discount enjoyed by the user. In the meantime, customer based dataset consisted of a unique customer identifier, customer age and gender, and other behavioral history of the customers such as if they have been following entries on certain brands or sectors, they have ever visited such pages on the application, and so on.

After this step, the remaining data was checked to see whether or not there are any transaction or customer records that are duplicated since these will further damage our accuracy. Confirming that, no duplicate records were found, data was then searched for missing values. Due to the way data was first prepared for our use in this study, considering the nature of “left-join” mechanic of SQL language, if a record for a customer did not exist on the adjoining table, it would return “null” as a value. Considering that in the context at which the data is stored at the company, these null values would in fact represent a “0” value. To achieve data consistency, these null values were replaced by a 0.

After the data was cleaned, the next step includes data transformation steps. Since one part of the data at hand consisted of transactional data and the other customer data, joining them by an SQL query would result in a customer identifier duplicating over and over. As a result, this transactional data had to be transformed into customer level data so that the duplication issue would be resolved. Another reason was that the fields at transactional data could not be directly used in a model. It is determined to summarize data for 4 quarters of 2018 so that our model would take a customer’s quarterly transaction history into account as well as their other demographic information. Flagging transactions based on which quarter of the year they occurred was necessary. To solve both these issues at once, firstly a total of 12 columns were created at the transactional dataset. These 12 columns, 3 for each quarter, consist of transaction flag column, brand flag column and the amount spent column, were created so that in each transaction record in a given quarter, there would be a matching column for them. Therefore 4 quarters data of 2018 will be used to estimate whether the customer will be active or not in the first quarter of 2019. On top of this, customer's overall Recency, Frequency, Monetary (RFM) characteristics were generated by the SAS EG's built-in RFM module for 2018. These parameters would later be used in model as well as the ones created and would be useful for comparisons. After these columns were created, transactional data could then be transformed into customer level data by utilizing SQL functions such as SUM or COUNT DISTINCT. And lastly, it is decided to look at the three sectors (retail, e-commerce, and restaurants), which have the most transactions in the given period. Data have been filtered to only include customers that made transactions in these sectors.

3. FINDINGS *(BULGULAR)*

3.1. Logistic Regression Results *(Lojistik Regresyon Sonuçları)*

For three sectors retail, restaurant and e-commerce logistic regression method is used to predict the customer status in the first period of 2019. The dependent variable is the binary variable stating the customer is made a transaction or not in this period. The independent variables are; age, gender of customer, and for each 4 quarters of 2018; total count of transactions of customer, total amount spent by the customer, total count of different brands customer made transactions, whether or not customer visited a page about retail sector in 2018. Additionally, Recency, Frequency and Monetary scores (scale of 1 to 5, 5 being the highest score) of customers in each 4 periods are added as independent variables.

3.1.1. Retail *(Perakende)*

Data consist of 2745 transactions, which include purchases made in retail stores. The Omnibus test shows that our model is more effective than the null model, yielding a χ^2 of 1362.412 which is significant ($p<0.05$). Nagelkerke R^2 value is found as 0.865 which shows that most of the variance in dependent variable can be explained by our model here and therefore it provides a standardized measure for the practical significance of the model. Hosmer-Lemeshow Test is a goodness of fit test, which determines how well the data fits the model. In this case, the test produced a χ^2 of 9.264 which is insignificant ($p>0.05$), the null hypothesis is rejected and it can be declared that the data fits the model.

According to the Wald Chi-Squared Statistic, count of different brands in Q2, Q3, Q4, and recency score in Q4 are significant variables ($p<0.05$) in predicting the outcome. The most significant variable is the total count of different brands customer made in Q4 with an $e^{ß}$ value of 10.266. Therefore, it can be said that a customer with one more count of different brands in Q4, is 10.26 times more likely to make a transaction in Q1 of 2019.

During the testing period, the model has predicted 92.2% of customers who made a transaction and 98.8% of customers who did not make any transaction during the first quarter of 2019 accurately. The overall accuracy of the model is 98.3%.

3.1.2. Restaurant *(Restoran)*

Data consist of 15739 transactions made in restaurants. The Omnibus test shows that our model is more effective than the null model, yielding a χ^2 of 9991.3 which is significant ($p<0.05$). Nagelkerke R^2 value is found as 0.747 which shows the practical significance of the model is acceptable. According to Goodness of fit test, Hosmer-Lemeshow Test, the null hypothesis is rejected and it can be declared that the data fits the model.

According to the Wald Chi-Squared Statistic, age of customer, total amount spent in Q1 and Q3, total count of different brands in Q1, Q2, Q3, and Q4, total count of transaction in Q3 and Q4, recency score in Q4 and frequency score in Q2 are significant variables ($p<0.05$) in predicting the outcome. The most significant variable is the total count of different brands customer made in Q3 with an $e^{ß}$ value of 2.191. Therefore, it can be said that a customer with one more count of different brands in Q3, is 2.191 times more likely to make a transaction in Q1 of 2019.
During the testing period, the model has predicted 74.5% of customers who made a transaction and 96.1% of customers who did not make any transaction during fifth quarter accurately. The overall accuracy of the model is 91.8%.

3.1.3. E-Commerce *(E-Ticaret)*

Data consist of 1111 shopping transactions made by using online channels. The Omnibus test shows that our model is more effective than the null model, yielding a χ^2 of 391.775 which is significant ($p<0.05$). Nagelkerke R^2 value is found as 0.858 which shows that most of the variance in dependent variable can be explained by our model here and therefore it provides a standardized measure for the practical significance of the model. Hosmer-Lemeshow Test produced a χ^2 of 0.27 which is insignificant ($p>0.05$), the null hypothesis is rejected and it can be declared that the data fits the model.

According to the Wald Chi-Squared Statistic, age of customer, total count of different brands in Q3, and Q4, total count of transaction in Q4, whether or not customer visited a page about E-commerce sector are significant are significant variables ($p<0.05$) in predicting the outcome. The most significant variable is the total count of different brands customer made in Q3 with an $e^{ß}$ value of 52.043. Therefore, it can be said that a customer with one more count of different brands in Q3, is 52 times more likely to make a transaction in Q1 of 2019.
During the testing period, the model has predicted 90.2% of customers who made a transaction and 99.2% of customers who did not make any transaction during fifth quarter accurately. The overall accuracy of the model is 98.7%.

3.2. Artificial Neural Network Results *(Yapay Sinir Ağları Sonuçları)*

In this part of the study, multilayer perceptron artificial neural network models are applied to the dataset for the three sectors. In all models, the number of hidden layers is selected as “1” because it is the optimal number of layer in which it is gained the best accuracy rate, and the activation function is selected as “Hyperbolic Tangent”. Since artificial neural networks can deal with both categorical and continuous variables, all of the independent variables obtained from preprocessing step used in these three models as parameters. Number of units are 36 that is computed as "best" number of units in the hidden layer for each model by SPSS. Our dependent variable is still whether or not a customer will make a transaction in the 1st quarter of 2019 or not. Here in all models, almost 70% of the data is reserved for training while the rest is left for testing the model.

3.2.1. Retail *(Perakende)*

Classification table (Table 1) allows us to see both overall and partial accuracies of our model. This is quite important since if it was not for this matrix provided, one could easily assume the overall 98.8% accuracy to apply in any situation. But when one examines the matrix, our true positive, successfully guessing the customers that will make transaction, prediction rate, which is also called sensitivity is 91.8%. Our model’s true negative, successfully guessing the customers that will churn, rate, which is also called specificity is 99.35%. These statistics can be even more important than the overall rate depending on what it is wanted to find out. If the main purpose of this project was to see who will churn then specificity rate would be much more important when the model were evaluated.

Table 1. Classification table for retail
(Perakende için sınıflandırma tablosu)

		Predicted		
Sample	*Observed*	*0*	*1*	*Percent Correct*
Training	*0*	1716	14	99.2%
	1	4	180	97.8%
	Overall Percent	89.9%	10.1%	99.1%
Testing	*0*	765	5	99.4%
	1	5	56	91.8%
	Overall Percent	92.7%	7.3%	98.8%

ROC (Receiver Operating Characteristics) curve and AUC(Area Under The Curve) help understanding performance of models visually. The ROC curve (Fig. 1) is plotted with sensitivity which is equal to true positive rate against the 1-specificity which is equal to false positive rate where sensitivity is on y-axis and 1-specificity is on the x-axis. Additionally, a perfect model has AUC near to the 1 which means it has good measure of separability. In this case AUC is 0.997 for both active and inactive customers, it means there is 99.7% probability that model can distinguish between customers who make a transaction and who do not.

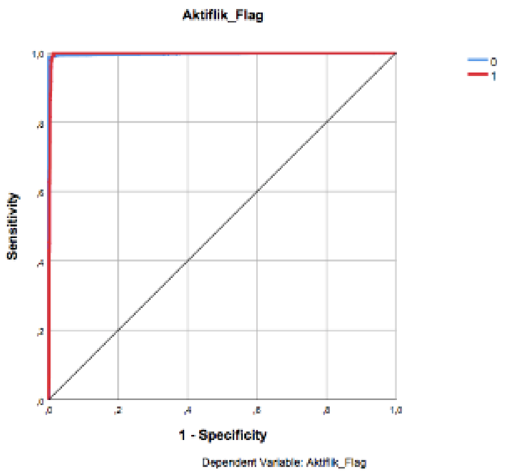

Figure 1. ROC curve for retail
(Perakende için ROC eğrisi)

3.2.2. Restaurant *(Restoran)*

The overall performance of the model can be evaluated by looking classification table (Table 2). During the testing period, the model has predicted 93.2% of active customers and 95.7% of inactive customers accurately. The overall accuracy of the model is 95.2%.

Table 2. Classification table for restaurant
(Restoran için sınıflandırma tablosu)

	Predicted			
Sample	*Observed*	*0*	*1*	*Percent Correct*
Training	*0*	8487	329	96.3%
	1	115	2056	94.7%
	Overall Percent	78.3%	21.7%	96.0%
Testing	*0*	3661	165	95.7%
	1	63	863	93.2%
	Overall Percent	78.4%	21.6%	95.2%

In the Figure 2, AUC is 0.988 for both active and inactive customers, it means there is 98.8% probability that model can distinguish between customers who make a transaction and who do not.

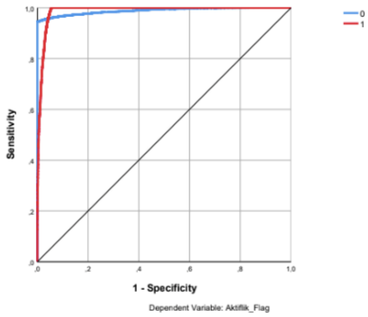

Figure 2. ROC curve for restaurant
(Restoran için ROC eğrisi)

3.2.3. E-Commerce *(E-Ticaret)*

The overall performance of the model can be evaluated by looking classification table (Table 3). During the testing period, the model has predicted 86.7% of active customers and 99.7% of inactive customers accurately. The overall accuracy of the model is 99.1%.

Table 3. Classification table for e-commerce
(E-ticaret için sınıflandırma tablosu)

	Predicted			
Sample	*Observed*	*0*	*1*	*Percent Correct*
Training	*0*	711	5	99.3%
	1	7	39	84.8%
	Overall Percent	94.2%	5.8%	98.4%
Testing	*0*	333	1	99.7%
	1	2	13	86.7%
	Overall Percent	96.0%	4.0%	99.1%

In this case AUC is 0.997 for both active and inactive customers, it means there is 99.7% probability that model can distinguish between customers who make a transaction and who do not.

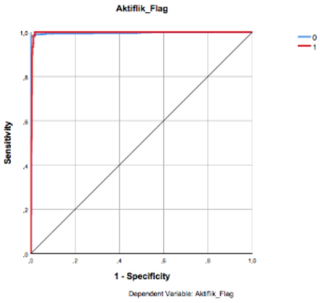

Figure 3. ROC curve for e-commerce
(E-ticaret için ROC eğrisi)

4. CONCLUSION *(SONUÇ)*

Due to the different dynamics in each sector, even the same variables are used in models for each of them, the impact ratio of each variable is quite diverse in each sector. For instance, in e-commerce sector, whether or not a customer visited any page about the sector is significant in determining the prediction results while it is not valid for other sectors. In restaurant setting, the recency and frequency variables that are calculated with RFM scoring technique are one of the significant factors effecting the results.
However, there are also predictor variables that are common in all three sectors. These are, surprisingly, total count of different brands customer made transactions in third and fourth quarters. Especially in e-commerce setting, total count of different brands customer made transactions in fourth quarter, with odds ratio of 52, is a huge determinant of whether customers make transaction or not. It can be stated that, customers who make purchases including a large number of brands are much more likely to be active in the coming period.

Table 4. Comparing overall performance of models
(Modellerin performans kıyaslaması)

Data Mining Technique	Business Sector	Sensiti-vity	Speci-ficity	Accu-racy
Logistic Regression	*Retail*	92.2%	98.8%	98.3%
	Restaurant	74.5%	96.1%	91.8%
	E-Commerce	90.2%	99.2%	98.7%
Artificial Neural Network	*Retail*	91.8%	99.4%	98.8%
	Restaurant	95.7%	93.2%	95.2%
	E-Commerce	86.7%	99.7%	99.1%

In order to compare the overall performance of logistic regression and artificial neural network in this study, it is crucial to calculate other measures of accuracy such as sensitivity and specificity as well. As can be seen in the table above, while predicting customer churn, artificial neural network models performed much better than logistic regression in each business sector in terms of sensitivity, specificity and accuracy.

REFERENCES *(KAYNAKLAR)*

[1] Dixon, M. "39 Experts predict the future." *America's Community Banker*, vol 8(7), pp. 20–31, 1999

[2] Floyd, T. "Creating a New Customer Experience". Bank Systems and Technology, vol 37(1), R8–R13, 2000

[3] Slater, Stanley & Narver, John. "Intelligence and Superior Customer Value." Journal of the Academy of Marketing Science, ol 28, pp. 120-127, 2012 Doi: 10.1177/0092070300281011.

[4] Ling, R., & Yen, D.C. "Customer relationship management: An analysis framework and implementation strategies". *Journal of Computer Information Systems*, vol 41(3), pp. 82–97, 2001

[5] Berson, Alex, Stephen Smith, and Kurt Thearling. *Building Data Mining Applications for CRM*. London: McGraw-Hill, 2000.

[6] Glady, Nicolas & Baesens, Bart & Croux, Christophe. "Modeling Churn Using Customer Lifetime Value". *European Journal of Operational Research*, vol 197, pp. 402-411, 2009 Doi: 10.1016/j.ejor.2008.06.027.

[7] Van den Poel, Dirk & Lariviere, Bart. "Customer Attrition Analysis for Financial Services Using Proportional Hazard Models". *European Journal of Operational Research*, vol 157, pp. 196-217, 2004 Doi: 10.1016/S0377-2217(03)00069-9.

[8] Buckinx W. and Van den Poel, D. "Customer base analysis: Partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting." *European Journal of Operational Research*, vol 164, pp. 252-268, July 2005 Doi: 10.1016/j.ejor.2003.12.010.

[9] Tamaddoni, Ali & Sepehri, Mohammad & Teimourpour, Babak & Choobdar, Sarvenaz. "Modeling customer churn in a non-contractual setting: The case of telecommunications service providers". *Journal of Strategic Marketing*, vol 18, pp. 587-598, 2010. Doi: 10.1080/0965254X.2010.529158.

[10] Nie, Guangli & Rowe, Wei & Zhang, Lingling & Tian, Yingjie & Shi, Yong. "Credit card churn forecasting by logistic regression and decision tree". *Expert Syst. Appl*., vol 38, pp. 15273-15285, 2011 Doi: 10.1016/j.eswa.2011.06.028.

[11] Dreiseitl, Stephan & Ohno-Machado, Lucila. Logistic regression and artificial neural network classification models: A methodology review. *Journal of biomedical informatics*, vol 35(5-6), pp. 352-359, 2002, Doi: 10.1016/S1532-0464(03)00034-0.

[12] L. Atlas *et al*., "A performance comparison of trained multilayer perceptrons and trained classification trees," *Conference Proceedings., IEEE International Conference on Systems, Man and Cybernetics*, Cambridge, MA, USA, 1989, pp. 915-920 vol.3.

[13] Brown, Donald E., Vincent Corruble and Clarence Louis Pittard. "A comparison of decision tree classifiers with backpropagation neural networks for multimodal classification problems." *Pattern Recognition*, vol 26, pp. 953-961, 1992 Doi: 10.1016/0031-3203(93)90060-A

[14] Linder, Roland & Geier, Jeannine & Kölliker, Mathias. "Artificial neural networks, classification trees and regression: Which method for which customer base?". *The Journal of Database Marketing & Customer Strategy Management*, vol 11, pp. 344-356, 2003 Doi: 10.1057/palgrave.dbm.3240233.

[15] Khan, A.A & Jamwal, Sanjay & Sepehri, Mohammad Mehdi. "Applying Data Mining to Customer Churn Prediction in an Internet Service Provider". *International Journal of Computer Applications*, vol 9(7), pp. 8-14, 2010 Doi: 10.5120/1400-1889.

[16] Hwang, Hyunseok & Jung, Taesoo & Suh, Euiho. "An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry". *Expert Systems with Applications*, vol 26, pp. 181-188, 2004 Doi: 10.1016/S0957-4174(03)00133-7.

[17] Hosmer, David W., Stanley Lemeshow, and Rodney X. Sturdivant. *Applied Logistic Regression*. 3rd ed. Hoboken, NJ: Wiley, 2013.