Predicting Customer Churn Based on Deep Learning, Neural Networks and Logistic Regression

Masome Motevali
Master's degree in Information Technology Engineering , Electronic Commerce, Mehralborz University, Tehran, Iran.
motevalyz@yahoo.com

Article Info
Page Number: 2180-2190
Publication Issue:
Vol. 72 No. 1 (2023)

Abstract: Today's competitive world has made keeping old customers one of the most important goals of economic enterprises. Customer attrition management in order to minimize losses from customer attrition and maximize profits from retaining valuable customers, as a powerful tool, analyze customer behavior using available data and identify and target customers prone to attrition. This category of customers plans and implements appropriate and effective strategies to maintain them. Therefore, the main goal of this study is to predict customer churn in Koroosh chain stores in Tehran by using the deep learning method and comparing it with the results of neural networks and logistic regression. The results of the study showed that the deep learning model has succeeded in better classification and prediction than the other introduced models, according to the key criteria of model evaluation.

Keywords: deep learning, customer churn, logistic regression, neural networks.

Article History
***Article Received**: 15 November 2022*
***Revised:** 24 December 2022*
***Accepted:** 18 January 2023*

Introduction

The customers are considered one of the most important asset for a business in numerous dynamic and competitive companies within a marketplace. In competitive market, companies in which the customers have numerous choice of service providers they can easily switch a service or even the provider. Such customers are referred to as churned customer. The causes of customer churn can be due to dissatisfaction, higher cost, low quality, lack of features, and privacy concerns (Lalwani, 2022). Many organizations e.g., financial service airline ticketing services, social network analysis, online gaming, banking sector and telecommunication sector, are ever more focusing on establishing and maintaining the long-term relationships with their existing customers Loyal customers can be considered long-term customers that are not only profitable for the company but also are great ambassadors in the market (Amin et al, 2019).

Churn is the event when a customer consciously decides to terminate its relationship with a service provider or vice versa. Customer churn prediction (CCP) is the study of identifying customers who are at risk of churn. There are three main types of churn - voluntary, involuntary and unavoidable (Sudharsan and Ganesh, 2022). Voluntary churn is when a customer decides to terminate the service relationship with provider. This may be due to switching to a competitor of the service provider. Involuntary churn is the case when the service provider terminates the relationship with customer. An example of this type of churn may be poor payment history of the customer. Unavoidable churn is when the relationship is terminated due to unavoidable reasons like when a customer is relocating to a new place or when a region faces natural disaster

(Fujo, Subramanian & Khder, 2022). It is possible to highlight "at risk" customers well in advance through in-depth analysis of differential behavioral patterns in the past and mapping these differential characteristics back to new customers. Clearly, churn prediction problem is a supervised classification problem (Ahmad et al. 2019). Customers are one of the most important business assets in many dynamic and competitive companies in the market. In a competitive market, companies whose customers have different choices of service providers can easily change a service or even a service provider (Amin et al, 2019). In an era of increasingly saturated markets that have intensified competition between companies, customer defection poses a real problem (Dingli et al, 2019).

Therefore, it has become clear to companies and managers that the historical customer information, which can be used to create models, in the existing customer base is one of the most important assets to combat customer churn (Zhang et al, 2022). The search and identification of customers who show a high inclination to abandon the company or customer churn prediction is of crucial importance (Momin et al, 2020, Lalwani et al, 2022) as part of a customer-oriented retention strategy that aims to reduce customer churn (Sudharsan and Ganesh, 2022). 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 (Shirazi & Mohammadi, 2019).

Customer churn has been tackled from two different angles in previous research. On the one hand, researchers focus on improving customer churn prediction models in which more complex models are being developed and proposed in order to boost the predictive performance (Rahman & Kumar, 2022). On the other hand, researchers want to understand what drives customer churn and defined important drivers of customer churn such as customer satisfaction (Shirazi & Mohammadi, 2019). They consider customer churn prediction as a managerial problem that is driven by the customer's individual choice. Therefore, action ability of customer churn prediction models is a key concern in which researchers can help managers to better understand the drivers of customer churn in order to make better informed decisions in combatting customer churn (Momin et al, 2020, Jain, 2020). Hereby many authors point out the managerial value for customer segmentation (De Caigny et al, 2020). By taking into account the main concerns of these two research angles, customer churn prediction models should have good predictive performance and lead to actionable insights.

In several studies, customer attrition and its influencing factors have been investigated and various statistical techniques such as data mining (pattern recognition) and machine learning (ML) have been used to predict attrition. In more developed models to investigate customer attrition, the information obtained from the previous purchases of customers are analyzed one by one for each customer and predictions are made (Fujo, Subramanian & Khder, 2022) In a general summary, we can summarize the studies that used algorithms and ML techniques to predict customer attrition. classified into three main categories. These three categories are: traditional machine learning techniques, group learning and deep learning (Chegunian, 2018). Among the existing studies in this regard, we can include studies with the support vector machine method (Li et al.2021; Jain et al, 2020; Sudharsan and Ganesh, 2022), logistic regression (Zhang et al, 2022). decision tree (Fujo, Subramanian & Patil et al, 2022), artificial neural networks (Cao et al, 2019), nearest neighbor (Xiahou & Harada, 2022), simple Bayes classifier algorithm (Rahman & Kumar, 2022), random forest algorithm (Martinez et al, 2020; De Caigny et al, 2020), pointed out.

In recent years, the use of hybrid models has been expanded to optimize forecasting models. A hybrid model usually uses two or more techniques. For example, classification and clustering models can be combined with each other (Xiahou & Harada, 2022). Also, the combination of decision tree and regression in building the turn prediction model (Momin et al, 2020), or the combination of the neural network model with multiple regression have been combined in order to increase the prediction accuracy (Cao et al, 2019). Researchers believe that the use of hybrid models can create better performance than simple models (Ahmad et al. 2019). Researchers in recent studies state that there are various variables that are essential in customer behavioral modeling that should also be considered. The set of variables of recency, frequency and monetary value (RFM) is one of the most common sets of variables used in customer behavioral modeling. (Shirazi & Mohammadi, 2019). In general, it is decided to measure the value of a customer based on its past purchase history (Rahman & Kumar, 2022). This approach was first introduced by Hughes (2000), because he describes important customers as customers who simultaneously They have low novelty, high frequency and high monetary value. Recency indicates the interval between the last transaction time and the current transaction, while frequency describes the repetition of the event in the past transaction and monetary ness describes the total amount of money spent on the company (Patil et al, 2022). Many studies in recent years successfully combined RFM analysis with data mining process (Xiahou & Harada, 2022; Martinez et al, 2020). Some studies first segment customers using RFM analysis and then apply data mining methods to create patterns for a set of important customers (Sudharsan and Ganesh, 2022; Ahmad et al. 2019). Today, according to the researchers, developing practical models for rainfall prediction is a vital task, which includes investigating and identifying the best predictor variables until choosing a powerful prediction method suitable for these variables. Since the retail industry collects a huge amount of data about customers, using past purchase data, it is possible to calculate the probability of customer churn to decide whether to leave or stay. Therefore, due to the necessity of processing a huge amount of data in chain stores and supermarkets, deep learning methods have an important place among the latest machine learning techniques for modeling customer deviance. The performance of a rainfall prediction model strongly depends on the variables selected from the data set (Li et al, 2021). Traditional approaches usually have two important obstacles and weaknesses:

a) feature extraction method among thousands of customer features, which is mostly tedious and time-consuming, and often needs to be checked and selected by an expert.

b) Traditional fallout prediction models are often applicable to a specific data set. Since deep learning methods can automatically find useful features, a deep learning algorithm can be applied to predict customer churn in the retail industry (Hou et al.2021). Especially in studies that have many and important features and input variables, the use of current machine learning methods such as deep learning for prediction bring better results and provide solutions as good as traditional models (Shirazi & Mohammadi, 2019).

Therefore, according to the appropriate performance and high predictive accuracy of machine learning algorithms and computational models that can be used for deep learning for marketing studies, in this study, predicting customer churn for Cyrus chain stores and comparing its performance with other methods of churn modeling, is the main goal of this research. For this purpose, in this study, deep learning methods and neural networks are used and an attempt is made to find a suitable answer to the question: the drop of customers in Korosh's chain stores, whose customers bought in conditions of uncertainty and many uncontrollable factors on the behavior Their purchase is effective, how was it and how are they different from each other?

Literature and research background

Deep learning is a subdomain in artificial neural networks (ANN) that consists of many hidden layers. Each layer of neurons in deep learning algorithms is influenced by the data features of the previous layer. Therefore, a feature hierarchy of increasing abstraction can be created, which is necessary in discovering patterns in a good dataset (Umayaparvathi and Iyakutti, 2019). The use of data related to past transactions has been widely used in previous studies and is considered an advantage (Li et al, 2021). In a study, Dingli et al. (2019) compared two deep learning algorithms, convolutional neural network and constrained Boltzmann machine, to predict churn in their retail sector. The data set and features were unknown, but it was stated that RFM was used. In their article, Rajablo and Hedayati (2020) predicted customer churn using a data mining algorithm. The results of this study enable managers and marketers of organizations to make policies based on the discovered patterns and have a better prediction of the current and future behaviors of their customers. In his thesis, Chegunian (2018) used aggregate methods and different algorithms to predict customer churn. The results of this research indicate that parameters such as the number of invoices delivered with a time delay and the number of invoices returned from sales are among the influencing factors in customer dissatisfaction.

Amin et al. (2019) in their study published in the Marketing Research Journal investigated the causes of customer churn and identified the reasons for customers turning to dissatisfaction, higher cost, low quality, lack of desired features and concerns about privacy. pointed out. In a study, Postigo-Boix & Melus-Moreno (2018) stated that customer churn occurs for various reasons. Competitive prices, service quality, discounts and promotions are among the most important reasons. Some other reasons that lead to customer churn are related to marketing campaigns, availability or quality of services received, complaints, etc. In 2019, studying various researches in the mobile phone industry, Lee et al stated that the causes of customer churn are related to service quality, demographic factors, customer satisfaction/dissatisfaction, and economic value. These studies are related to customer service such as quality, pricing, pricing options, billing information, usage data, switching costs, mobile service quality, poor voice quality, billing issues, minutes and calls, issues related to Subscriber demographics Customer demographic data, customer status, number of referrals, number of predicted features in the dataset, number of observations, number of mobile phone subscribers, etc., customer satisfaction/dissatisfaction (emotional value and economic value) and predicting the drop pay attention.

In their research in 2019, Ravi et al concluded that one of the other reasons for the loss of customers is the lack of attention and handling of customer complaints. These researchers divide the customer life cycle into three parts. At each stage of the customer's life cycle, sentiment analysis from electronic word-of-mouth advertising can be used, and it is stated that negative electronic word-of-mouth advertising of customers is considered more important than its positive type. In a study, Martinez et al. (2020) proposed customer forecasting with a new set of customer-related features that are updated every month from previous purchases and using advanced machine learning algorithms. They found that the gradient tree boosting method is superior to other methods with high accuracy and classification performance. In a study, Alboukaey et al, (2020) proposed a daily churn prediction model and suggested daily customer behaviors as a multivariate time to predict customer churn. In this study, a statistical model, an RFM model, a long-term short-term memory model (LSTM) and a convolutional neural network (CNN) model were applied to the mobile telecommunication data set and the results of the models were compared. The results of this study showed that the prediction of daily rainfall was better than the monthly predictions.

materials and methods

The current research is in terms of practical purpose, in terms of descriptive-survey data collection method and in terms of information collection tools, it is library and field. The current research population includes all the real customers of two branches of Koroosh chain stores in Tehran city, who were randomly selected as a statistical sample to collect their 14-month retail sales data, and the relevant officials were assured that their information would be completely confidential. will remain and will be used for research purposes. The collected data set includes features such as customer membership identification numbers, transaction dates, and costs of each product category. Database created with eight main product categories including spices, dairy products and breakfast food; canned food; drinks; health products; Margarine, butter and cooking oils, beans and rice were determined.

Table 1. Characteristics of collected data

The number of features	**Description**	**Customer characteristics**
12	Last day of shopping for a month	novelty
12	Number of times customers have shopped for a month	periodicity
12	The total monthly cost of a customer	Monetary value
91	Monthly fee based on each product category	The costs of each category of products

Source: research findings

The dataset contains 1550 customer transactions from two large branches. Table 1 presents the features (91 in total) for each month in the dataset. The number of clients and features is sufficient to run a deep learning neural network. The dataset contains 14 months of data, but only 12 months were used to train the model. The last two months of data were excluded, as this part of the data indicates whether a customer was a churner. The data were first cleaned and sorted by removing missing values and then normalized to ensure data integrity. The value of recent costs, frequency, monetary value and product after normalization were between 0-1. During a contractual decision, customers who have no transactions for two consecutive months or periods are considered lapsed customers. Although in some previous studies, such as Migos (2018), customers who spend less than a certain spending threshold for a consecutive three-month period have been considered as churn customers, but in this study, only customers who They did not buy at all, they are considered. After adjusting and arranging the data, a deep learning model and a neural network model will be implemented using the same data and the results will be compared. RapidMiner Studio software will be used to run prediction models.

Deep learning model

Deep learning algorithms are a subset of machine learning algorithms that aim to discover multiple levels of distributed representations of input data. Recently, many deep learning algorithms have been proposed to solve traditional artificial intelligence problems. In recent years, deep learning has been widely studied in the field of computer vision, and for this reason, a large number of related methods have been developed. In general, these methods are divided into 4 different categories based on the basic method from which they are derived, which are: convolutional neural networks, limited Boltzmann machines, automatic encoder, sparse coding.

Convolutional Neural Networks (CNN) are one of the most important deep learning methods in which several layers are trained in a powerful way. This method is very efficient and is one of the most common methods in various computer vision applications. These types of networks consist of neurons with learnable (adjustable) weights and biases. Each neuron receives a number of inputs and then calculates the product of the weights in the inputs and finally presents a result using a non-linear transformation (activation) function. The whole network still provides a derivable score function, which has the raw pixels of the input image on one side and the scores for each category on the other side.

These types of networks still have a cost function like (SVM, Softmax) in the last layer, and all the points mentioned about normal neural networks are true here.

In general, a CNN consists of three main layers, which are: convolution layer, pooling layer and fully connected layer. Different layers perform different tasks. In each convolutional neural network, there are two steps for training. feed forward stage and backpropagation stage. In the first stage, the input image is fed to the network, and this operation is nothing but a point multiplication between the input and the parameters of each neuron, and finally applying the convolution operation in each layer. Then the output of the network is calculated. Here, in order to adjust the network parameters or in other words the network training, the output result is used to calculate the network error. For this, the output of the network is compared with the correct answer using an error function (loss function) and thus the amount of error is calculated.

In the next step, based on the calculated error, the backpropagation step begins. In this step, the gradient of each parameter is calculated according to the chain rule, and all parameters are changed according to the effect they have on the error created in the network. After the parameters are updated, the next stage of feed-forward starts. After repeating a suitable number of these steps, the network training ends.
Deep learning is a subset of machine learning algorithms based on artificial neural networks that uses multiple layers to continuously extract higher-level features from the raw input. The learning of this algorithm can be supervised, semi-supervised or unsupervised. In deep learning, each level learns to represent its input data, a little more summarized and compounded. Most importantly, a deep learning process can learn at what level the optimal features should be placed on its own (Sudharsan and Ganesh, 2022). The word deep in deep learning refers to the number of layers through which data is transformed into output. Deep learning models are able to extract better features than shallow models and hence additional layers help in learning features (Rahman & Kumar, 2022). The deep neural network improves its skill and experience with each repetition of a task. The difference between deep learning and neural network is that deep learning has a wider scope than neural network and includes reinforcement learning algorithms.

Neural network modeling

Neural networks are a class of learning methods that have been developed in the fields of statistics and artificial intelligence and have regression and classification capabilities. In the neural network model, the explanatory variables and the response variable can be quantitative or qualitative, and the response variable has a non-linear and indirect relationship with the explanatory variables. The framework of neural networks consists of several layers called input layer, hidden layer(s) and output layer. Neural networks are divided into forward and backward neural networks. In forward networks, the nodes of each layer are connected only to the nodes of the next layer (Figure 1), but in backward networks, the nodes of each layer are connected to the nodes of the next layers or to themselves.
One of the most common architectural styles of neural networks, in which the communication pattern between layers is specified, is multilayer perceptron. In this structure, each neuron is a function of the linear combination of explanatory variables, and the response variable is also a function of the linear combination of neurons. that's mean:

$$Z_i = \sigma\left(a_{0,i} + \sum_{j=1}^{n} \beta_{j,i} X_j\right); i = 1,2, \ldots\ldots, m. \qquad y = f(\Upsilon + \sum_{i=1}^{m} \theta_i Z_i \qquad (1)$$

where theta is called the activation function and the graph of this function is non-linear and s-shaped, including hyperbolic tangent, arc tangent, and sine. Also, the f function is linear or non-linear and the face of non-linearity can have one of the mentioned forms like theta. In estimating the parameters of the model, first the initial values of these parameters are randomly selected and then during a learning process, these values are constantly updated until the amount of classification error reaches a relatively stable state (Lin et al, 2019). Finally, the neural network predicts the customer drop by calculating the conditional probability of the new observation X_0 being placed in class 0 or 1, which is called the posterior probability.

Logistic Regression

Another very useful statistical tool for classifying the response variable is logistic regression. If the response variable has two levels; It is called binary logistic regression and if it has more than two levels, depending on whether the order between the levels is important or not, it is called ordinal or nominal logistic regression. The response variable in the prediction of customer dropout is considered as dropout (1) or no dropout (0), so to observe $X = (x1, x2, \ldots, xn)$, the prediction model of the response variable class will be as follows:

$$P(X) = E(Y|X = x) = P(Y = 1|X = x) = \frac{e^{-M}}{1+e^{-M}} \quad (2)$$

$$Q(x) = P(Y = 0\left|X = x\right) = -P(x) = \frac{1}{1+e^{-M}}\Big| \quad (3)$$

where alpha and beta will be constant coefficients of explanatory variables obtained by minimizing the corresponding likelihood function.

Logistic regression modeling is a well-known and very attractive technique because:

1) A solution is available for posterior probabilities (unlike probit).

2) The basic assumption of logit (the logarithm of the ratio of group conditional densities is linear in parameters) is satisfied by many families of distributions (Amin et al, 2019).

3) It is easy to use and provides fast and strong results.

Model Evaluation Criteria

Various criteria are used to measure the accuracy and efficiency of data classification models. If it is assumed that observations with class 1 are considered as positive observations (P) and negative observations with class 0 are considered as negative observations (N). If FP is the number of negative observations that are predicted positively, and FN is the number of positive observations that are predicted negatively, and conversely TP is the number of positive observations that are predicted positively, the matrix related to the different prediction states of classification problems with two classes can be shown as described in the following table:

Table 2. Matrix of prediction states of classification problems with two classes

Negative view	**Positive view**	
FP	TP	Positive forecast
TN	FN	negative forecast

Some of the evaluation criteria of the classification models used in this study can be shown as follows:

(4) $Recall = TP\,rate = TP/TP + FN$

(5) $Precision = TP/\,TP + FP$

(6) $FN\ rate\ =\ FN/\ TP + FN$
(7) $Accuracy\ =\ TP + TN\ /\ TP + TN + FP + FN$

In Equation 4, a proportion of all positive observations that are correctly classified is calculated. In equation 5, a ratio of positive predictions that were really positive is calculated and in equation 6, a ratio of all positive observations that are wrongly classified is determined and calculated. Equation 7 also represents the proportion of observations that are correctly predicted.

Another criterion is the area under the curve (AUC). This criterion is also one of the most widely used criteria for model evaluation. It is generally used for binary classification problems. This criterion allows a classifier to distinguish between classes and is used as a summary of the receiver operating characteristic (ROC) curve. The higher the AUC, the better the performance of the model is assumed to be in distinguishing between positive and negative classes. The ROC curve can be used to select a threshold for a classifier that maximizes true positives and in turn minimizes false positives.

Research Findings

After the implementation of the models, the introduced four criteria (accuracy, correctness, recall and exact distinction between positive and negative classes) to evaluate the three models of this study are set and shown in Table 3 in a comparative manner. As can be seen, the deep learning model performed better in all three criteria related to relationships 7, 4 and the surface under the curve than the two logistic models and neural networks and has higher evaluation values, and only in the criterion related to relationship 5 (the ratio of prediction positives that were really positive) that the corresponding numerical value is smaller compared to the neural network model.

Therefore, these results show that deep learning is a very successful technique in predicting churn and purchases, which has better capabilities in the majority of model evaluation criteria, and therefore can be used as a reliable method to study customer churn in other branches or chain stores. The country is also used.

Table 3. Evaluation of the criteria of the studied models

Deep learning model	Neural network model	Logistic regression model	Evaluated criteria
0.916	0.899	0.835	Proportion of observations that are correctly predicted (Relation 7)
0.908	0.871	0.859	Area under the curve
0.895	0.905	0.878	A ratio of positive predictions that were really positive (Relation 5)
0.880	0.872	0.841	Proportion of total positive observations that are correctly classified (Relation 4)

Discussion and Conclusion:

So far, several studies in the field of relationship marketing and customer behavior have been conducted on predicting customer churn to determine which customers are likely to leave the company. In these studies, several forecasting models and techniques have been used, each of which has special abilities. Deep learning methods, as one of the recent developments in the field of artificial neural networks, are frequently used in image processing and image definition, which can also be used in customer churn studies. Therefore, in this study, a deep learning model was used to predict customer churn with the characteristics of customer consumption basket analysis, and the results were compared with certain criteria and compared to other competing models.

The data of this study has been collected from two branches of a retail supermarket chain in Tehran, which includes the transactions of its customers according to the eight main groups included in their consumption basket and includes almost a large percentage of household purchases. The results of implementing three different prediction models using deep learning, logistic regression and neural network techniques and comparing the results with appropriate comparative techniques showed that among the four main criteria, the deep learning model has a higher prediction in the estimation of three criteria. and estimates. This study, like other studies, also has limitations. The size of the data obtained from chain supermarkets is relatively less than other studies related to customer churn in the literature in this regard, and therefore it is suggested that in future studies, the increase in the number of branches or the number of customers and transactions should be considered so that after Analyzing the customers' consumption portfolio to provide better and more accurate forecasting performance. In addition, increasing the period of time and months under study will not be ineffective in achieving reliable and accurate results.

References

[1] Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 6(1), 1-24.

[2] Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. Journal of Business Research, 94, 290-301.

[3] Cao, S., Liu, W., Chen, Y., & Zhu, X. (2019). Deep learning based customer churn analysis. In 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP) (pp. 1-6). IEEE.

[4] Chegunian, N, (2018), predicting customer churn using data mining method (case study of Arad printing company), master's thesis in business management, Sheikh Bahai University, supervisor: Bahram Izadi.

[5] De Caigny, A., Coussement, K., De Bock, K. W., & Lessmann, S. (2020). Incorporating textual information in customer churn prediction models based on a convolutional neural network. International Journal of Forecasting, 36(4), 1563-1578.

[6] Dingli, A., Marmara, V., Fournier, N.S. (2019). Comparison of deep learning algorithms to predict customer churn within a local retail industry. International journal of machine learning and computing 7(5).

[7] Fujo, S. W., Subramanian, S., & Khder, M. A. (2022). Customer churn prediction in telecommunication industry using deep learning. Information Sciences Letters, 11(1), 24.

[8] Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit boost. Procedia Computer Science, 167, 101-112.

[9] Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. Computing, 1-24.

[10] Li, Y., Hou, B., Wu, Y., Zhao, D., Xie, A., & Zou, P. (2021). Giant fight: Customer churn prediction in traditional broadcast industry. Journal of Business Research, 131, 630-639.
[11] Martinez, A., Schmuck, C., Pereverzyev Jr, S., Pirker, C., Haltmeier, M. (2020). A machine learning framework for customer purchase prediction in the non-contractual
[12] Momin, S., Bohra, T., & Raut, P. (2020). Prediction of customer churn using machine learning. In EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing: BDCC 2018 (pp. 203-212). Springer International Publishing.
[13] Patil, K., Patil, S., Danve, R., & Patil, R. (2022). Machine Learning and Neural Network Models for Customer Churn Prediction in Banking and Telecom Sectors. In Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems: ICACECS 2021 (pp. 241-253). Singapore: Springer Nature Singapore.
[14] Postigo-Boix, M., & Melus-Moreno, J. (2018). A Social Model Based on Customers' Profiles for Analyzing the Churning Process in the Mobile Market of Data Plans. Physica A.
[15] Rahman, M., & Kumar, V. (2020, November). Machine learning based customer churn prediction in banking. In 2020 4th international conference on electronics, communication and aerospace technology (ICECA) (pp. 1196-1201). IEEE.
[16] Rajablo, Z and Hedayati, A, (2020), predicting customer churn using data mining algorithms, the second national conference on new technologies in electrical and computer engineering, Isfahan.
[17] Ravi, K., Ravi, V., & Krishna Prasad, P. (2019). Fuzzy Formal Concept Analaysis Based Opinion Mining for CRM in Financial Services. Applied Soft Computing.
[18] Shirazi, F., & Mohammadi, M. (2019). A big data analytics model for customer churn prediction in the retiree segment. International Journal of Information Management, 48, 238-253.
[19] Sudharsan, R., & Ganesh, E. N. (2022). A Swish RNN based customer churn prediction for the telecom industry with a novel feature selection strategy. Connection Science, 34(1), 1855-1876.
[20] Umayaparvathi, V., Iyakutti, K. (2019). Automated feature selection and churn prediction using deep learning models. International Research Journal of Engineering and Technology (IRJET) 4(3), 1846.
[21] Vo, N. N., Liu, S., Li, X., & Xu, G. (2021). Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems, 212, 106586.
[22] Xiahou, X., & Harada, Y. (2022). B2C E-commerce customer churn prediction based on K-means and SVM. Journal of Theoretical and Applied Electronic Commerce Research, 17(2), 458-475.
[23] Yu, R., An, X., Jin, B., Shi, J., Move, O. A., & Liu, Y. (2018). Particle classification optimization-based BP network for telecommunication customer churn prediction. Neural Computing and Applications, 29, 707-720.
[24] Zhang, T., Moro, S., & Ramos, R. F. (2022). A data-driven approach to improve customer churn prediction based on telecom customer segmentation. Future Internet, 14(3), 94.