EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Prediction of User Attrition in Telecommunication Using Neural Network

Mageshkumar N[1*], Vijayaraj A[2], Subba Reddy Chavva[3], Gururama Senthilvel[4]

[1]Department of Computer Science & Technology, Madanapalle Institute of Technology & Science, Madanapalle
[2]Department of Information Technology, R.M.K Engineering College, Chennai
[3]Department of Information Technology, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada
[4]Department of Computer Science and Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai

Abstract

INTRODUCTION: The telecommunications industry faces significant challenges due to customer attrition, which directly impacts revenue. To better understand and address this issue, Companies are looking into techniques to determine the internal issues that affect customer churn.
OBJECTIVES: This article offers an overview of customer attrition within the telecommunications sector.
METHODS: It introduces an advanced churn prediction model harnessing state-of-the-art technologies, including neural networks, machine learning, and other cutting-edge innovations, to achieve remarkably high accuracy rates. By analyzing diverse parameters and datasets collected from multiple telecom companies, valuable insights can be gained.
RESULTS: The model's performance on test data can be evaluated using metrics such as Accuracy Score, Area under Curve (AUC), Sensitivity, Specificity, and other performance indicators.
CONCLUSION: In order to effectively manage extensive datasets, organizations can leverage Big Data technology. This empowers them to forecast the probability of customer churn and put in place proactive strategies to retain their customer base.

Keywords: Component, Churn in Telecom, Feature selection, Data Analysis, Telecom Industry, Data Mining, Classification

Received on 04 December 2023, accepted on 21 February 2024, published on 29 February 2024

doi: 10.4108/eetsis.5242

*Corresponding author. Email: mageshkumarns19@gmail.com

1. Introduction

In both established and emerging countries, the telecommunications sector has grown to become one of the most important businesses. Due to advances in technology and a growth in the number of service providers, the level of competition has risen [1],and operators are scrambling to stay afloat in the current competitive market by employing various techniques. Here are three recommended strategies for boosting earnings, as outlined in [2]

- Attracting new customers
- Upselling to existing customers
- Extending customer engagement with the business

Among these approaches, the third one is notably the most cost-effective. This is because retaining existing customers is significantly more advantageous than acquiring new ones. [3].

For example, in Syria, it costs six times more to get a new client than to keep one who is most likely to leave the operator. Furthermore, this strategy is far easier than up selling the pre-existing customer [4]. To achieve the third strategy, businesses must reduce customer churn, which is defined as "the transfer of a customer from one service provider to another [5].

The telecommunications industry faces a significant challenge in the form of customer churn due to fierce competition. Identifying the customers with the highest likelihood of switching networks can assist a business in preventing potential revenue loss. [6]. Also, many researchers believe that implementing NN and ML algorithms for prediction could be very efficient in terms of accuracy from the previous data [7,8] and the three leading causes of customer churn is mentioned in Fig. 1.

Figure 1. Three major causes of Customer Churn

2. Related Works

In the telecom industry, researchers have employed various methodologies to predict customer attrition. AI algorithms, Data Mining, and Big Data technologies have been widely used for this purpose. While many studies have traditionally concentrated on the application of one or two algorithms to their datasets, a subset of researchers have ventured into the realm of employing multiple algorithms drawn from the domains of Machine Learning and Neural Networks [9]. Their objective is to enhance the scope of comparison and prediction accuracy. For instance, Ahmad et al. [10] developed a model for predicting customer churn that combines a variety of machine learning models, includingtree-based techniques like Decision Trees, Random Forest, Gradient Boost Machine Tree, and XGBoost, demonstrating a more varied modeling approach. To handle the large dataset of 70 Terabytes from telecom companies Syriatel and MTN in Syria, Using a Big Data platform, the authors, specifically the "Hadoop Distributed File System." These datasets contained structured, semi-structured data, and numerous features, which required feature engineering and data analysis. Leveraging Big Data technology allowed the researchers to efficiently store, retrieve, and process massive amounts of data. Furthermore, it facilitated the extraction of crucial elements, including Social Network Analysis (SNA), thereby enhancing the predictive model's precision. The authors selected the XGBoost prediction model after thorough evaluation due to its outperformance compared to other competing algorithms. In the past, initiatives to employ data warehouse systems in Syria's telecom sector to lower customer churn have had underwhelming success. The vastness of the database made it difficult for the system to efficiently store, retrieve, and process enormous amounts of data. Additionally, integrating diverse data sources into the Data Warehouse proved challenging, leading to long processing times, high computational demands, and substantial storage requirements when adding new features for Data Mining techniques.

To address these challenges, the concept of Big Data emerged as a solution for efficiently processing large datasets. Gavril et al. [11] conducted a study using call records from approximately 3333 prepaid customers, involving over 20 features. The features encompassed information about sent and received messages, along with the duration of both incoming and outgoing calls for each customer. Due to the dataset's extensive feature set, the author opted for Principal Component Analysis (PCA) as a dimensionality reduction method to preserve critical information. After dimensionality reduction, three distinct AI techniques were employed to analyze the dataset: neural networks, support vector machines, and Bayesian networks. The performance of each model was assessed using the AUC metric. With an amazing AUC score of 99.7%, the Bayes Networks model stood out for its exceptional performance. Additionally, the study enjoyed the advantage of having complete input data, further bolstering its robustness and accuracy. To overcome the Customer Churn problem, He Y et al. [12] used Neural Networks. The statistics came from a Chinese cellular provider with approximately 5.3 million subscribers. The model's accuracy score was 91 percent. On two distinct datasets, Idris [13] proposed Genetic Programming with the Ada Boost method to construct a customer churn model and achieved an accuracy score of 89 percent. Makhtar et al. [14] The author employed a Rough Set classification approach and proposed the application of Rough Set theory to build a customer churn model. In addition to offering several models, such as linear regression, decision trees, and voted perceptron neural networks, it was observed that Rough Set Classification excelled in both accuracy and performance.

This paper presents the development of a customer churn model utilizing neural networks and various machine learning techniques, including XG Boost and Random Forest Classifier, among others. Alboukaey et al. [15] The predictive capabilities of these models were examined by subjecting them to an evaluation using a dataset spanning 150 days, sourced from the MTN operator in our country. The outcomes highlighted a noteworthy trend. Pustokhina et al. [16] The benchmark customer churn prediction dataset was used to conduct an extensive simulation investigation. The trial outcomes amply demonstrated the improved performance of our suggested model in comparison to other approaches.Vo, Nhi NY, et al[17] The findings indicate that this model effectively predicts the likelihood of client churn and offers insightful information by combining personality traits and client segmentation into interpretable machine learning techniques., Irina V., et al[18] An array of simulations were run, and the ISMOTE-OWELM model shown remarkable predictive ability, reaching accuracy rates of 0.94, 0.92, and 0.909 on the dataset under consideration.

3. Existing System

In the existing system, datasets were gathered from a number of Telecom firms and open-source datasets. Each dataset was created by a different service provider, and each service provider had different requirements, resulting in a diverse set of attributes to develop effective strategies for customer retention. The attributes from different datasets include Churn: Identifies customers who terminated their service withthe provider in the previous month.

- Account Length (or Tenure): This term denotes the duration for which a customer has been a client of the company.
- Various service-related attributes, including Phone Service, Multiple Lines, Internet Service, Online Security, Online Backup, Device Protection, Tech Support, Streaming TV, and Streaming Movies.
- Aspects related to billing and payments, such as Paperless Billing, Payment Method, Monthly Charges, and Total Charges.
- Gender, age, dependents, and other demographic information elements are some instances.
- Features that record the quantity of calls, minutes, messages, and other activities initiated, received, and sent by a customer on a given day.
- Whether a customer has opted for an international plan is indicated by the entry under International Plan.

Additionally, each customer's total call duration throughout all time segments—morning, afternoon, evening, and night—is included in the dataset. Among the various features present in these datasets, the ones listed above have been identified as the most influential factors affecting the "Churn" output parameter.

3.1. Data Analysis

The focus of this study was data analysis rather than just developing models to forecast customer churn in the telecom sector. The data analysis provided valuable insights to organizations, enabling them to identify specific attributes in which they may be lacking and take appropriate actions to retain existing customers and attract new ones. The datasets underwent analysis using both categorical and numerical univariate and bivariate features. Each individual feature was subjected to univariate analysis, while each input feature was analyzed in relation to the output feature "Churn" through bivariate analysis.

3.2. Uni-Variate Analysis

Fig. 2 shows that nearly 49.6% of customers reside in area code 415; therefore, For clients in area code 415, it is preferable to offer more services and much better plans in order to prolong customer retention.

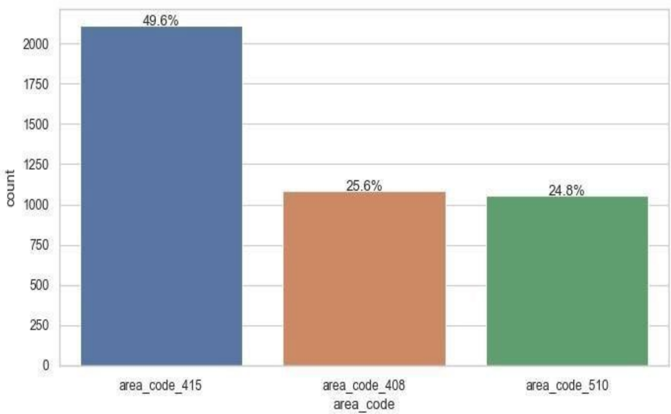

Figure 2. Population of customers at different Areas

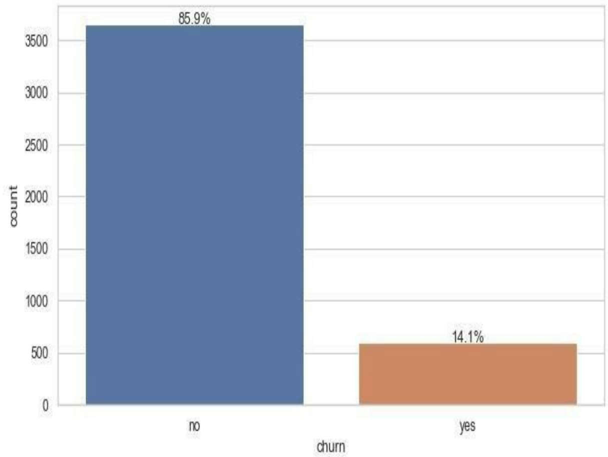

Figure 3. Distribution of Churn in the dataset

Additionally, as seen in Fig.3, there are considerably fewer consumers leaving the telecom company than there are customers staying. It's essential to employ the Upsampling strategy on the minority class, or data points with the churn value "yes," because to the imbalanced classes. The Synthetic Minority Oversampling Technique could therefore be used to assist the minority class in resolving their problem.

3.3. Bi-Variate Analysis

Clients who have been with the telecom firm for more than 50 months are less likely to churn, less than ten months of service, however, increases the likelihood that a customer would leave. As a result, in order to enhance customer retention, the telecom business must offer better deals and more data to new consumers.

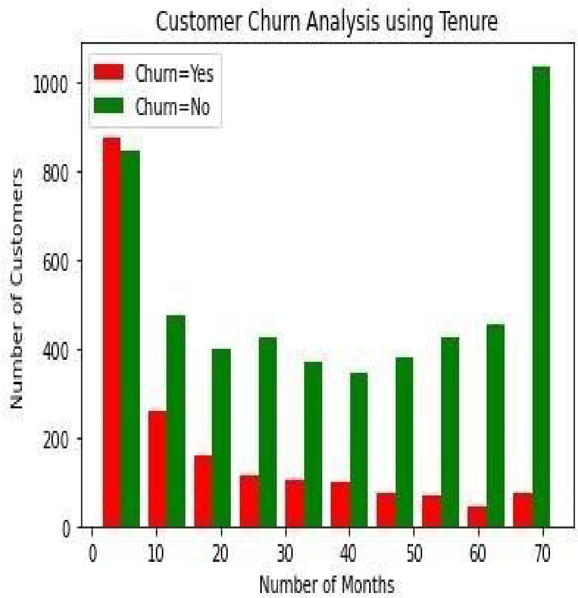

Figure 4. Customer Churn analysis (vs) Tenure

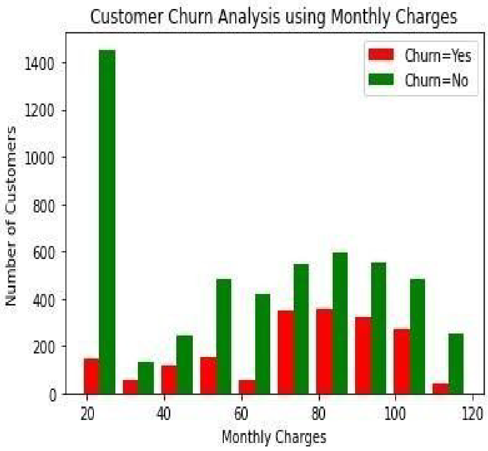

Figure 5. Customer Churn analysis (vs) Monthly charges

Customers who pay very little or high monthly fees do not leave the company, but as the graph above illustrates, customers who pay monthly fees between $70 and $100 have a 50% likelihood of doing so. The business could also concentrate more on clients who make monthly payments of about $70. Fig 6 and 7 demonstrate that a high churn rate occurs when "Total day calls" are between 85 and 115, and a high churn rate occurs when "Number of voice mail messages" is zero.

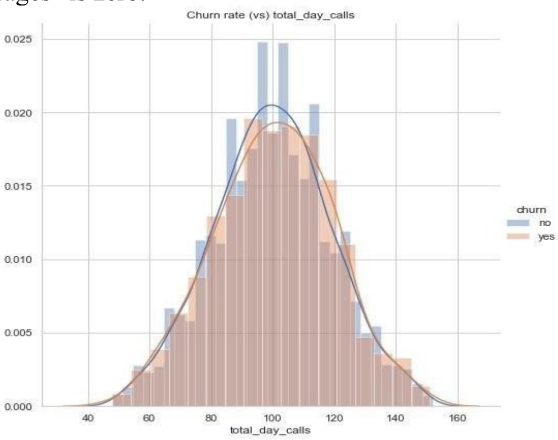

Figure 6. Churn rate (vs) Total number of calls in a day

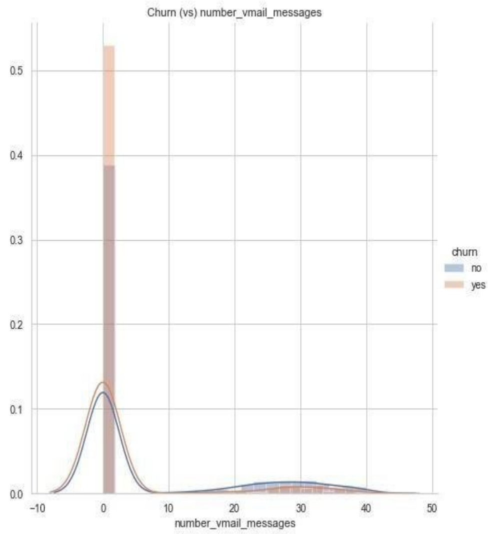

Figure 7. Churn rate (vs) Number of Voice mail messages

4. Proposed Method

This research introduced neural networks-based and machine learning-based approaches for predicting customer attrition in a company [Fig.9]. Several researchers have explored various AI algorithms for this purpose and concluded that AI technology generally outperforms other methods when applied to datasets. Furthermore, big data technology proves to be valuable in storing, analyzing, and retrieving data from large datasets efficiently.

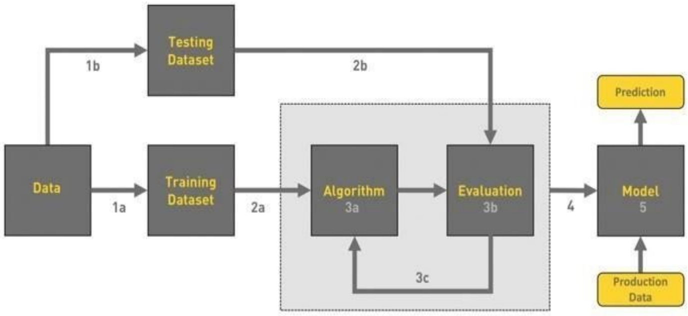

Figure 8. Workflow of a Supervised Learning Algorithm

To develop a better predictive model, various weight initialization strategies, optimisers, and hidden layers are used. The proposed strategies are basically composed of three number one processes, which are as follows: Using supervised classification algorithms on the train and test data and comparing the results to real-world data [Fig.8]. Implementing Supervised classification algorithms on the train and test data, comparing the results with actual data. Loading the dataset and performing pre-processing techniques for feature extraction. In this study, the neural network architecture comprised three to five hidden layers with different datasets, incorporating diverse activation functions and weight initialization techniques. For neurons with the activation function "Relu," the He Normal weight initialization method was applied, while for neurons with the activation function "Sigmoid," the Glorot Normal (Xavier Normal) weight initialization method was used. This selection was made based on the observation that He Normal performs well with Relu, whereas Glorot Normal is not as effective with Relu.

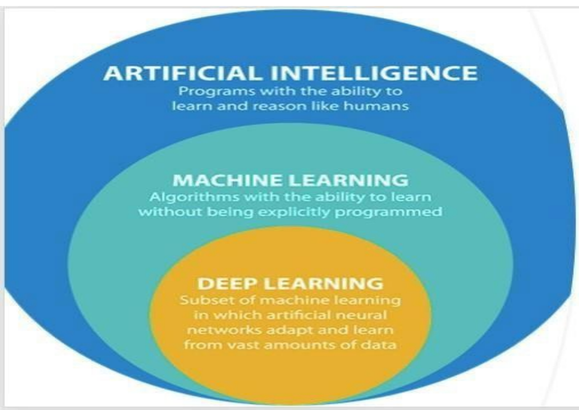

Figure 9. Neural Networks (vs) Machine Learning

In the initial attempt to build a neural network using random weights, the accuracy score was unsatisfactory, prompting us to seek improvements. To address this, the output layer was constructed with the Sigmoid activation function, as it scales the input values between 0 and 1. However, using the Sigmoid activation function in the hidden layers led to the Vanishing Gradient problem, as its derivative is limited to a range between 0 and 0.25. To overcome this issue, the Relu activation function was employed in the hidden layers instead. The structure of the neural network is depicted in Fig. 10.

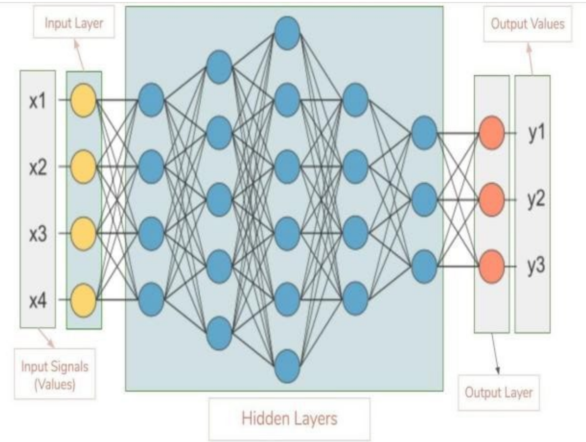

Figure 10. Structure of a Neural Network

5. Result and analysis

At the beginning of the experiment, the utilization of Machine Learning classification algorithms such as Random Forest classifier and KNN classifier on the Telco Customer Churn dataset did not meet the expected performance levels, even after conducting Hyperparameter tuning and Principal Component Analysis to account for feature variability. The study involved a comprehensive analysis of the obtained results, focusing on the evaluation of performance across various sizes of training data. The research also delved into addressing the challenges posed by an imbalanced dataset through three distinct scenarios. The initial primary focus was on choosing the best sliding window for data extraction, with a specific focus on statistical and Social Network Analysis (SNA) features. In the case of SNA features, a unique approach was adopted, with a particular emphasis on identifying the ideal sliding window duration for creating the social graph and extracting pertinent SNA characteristics.

The findings indicated that the optimal window spanned the final four months preceding the baseline. Another significant aspect considered in this study was the issue of dataset imbalance. To tackle this problem, three distinct experiments were conducted across all classification algorithms. These experiments encompassed techniques such as under sampling, oversampling, and exploring classification without employing balancing techniques. Lastly, an important observation emerged regarding the overall balance of customer accounts. It was noted that most churned customers exhibited lower account balances in comparison to their active counterparts, irrespective of the specific reasons for their churn. The reason behind this underperformance was the absence of a significant correlation between the input and output features. A thorough analysis of the Telco customer churn dataset using the Pearson correlation coefficient indicated that none of the features exhibited a strong association with the output parameter "Churn." Consequently, a deep neural network was constructed to better capture the input data, resulting in the following outcomes.

Table 1. Telco Customer Churn database Results

Algorithm	Results
	Accuracy Score
Random Forest Classifier	78%
KNN Classifier	74%
Neural Network	Training Accuracy Score-84% Testing Accuracy Score-76%

On the second dataset, Customer Churn Prediction 2020, AI systems surpassed machine learning techniques with a score of over 90% accuracy. I also looked into utilizing Artificial Neural Networks on the dataset to forecast client churn using various epochs and weight initialization procedures. The accuracy scores are as follows: -

Table 2. Customer Churn Prediction 2020 using NN

Number of Epochs	Results
	Accuracy Score
20	86%
40	92%(batch =size=100)
40	94%(batch =size=100)
40	94%(batch =size=100)
100	Training->95%,Testing->87% (batch size=1000)
100	Training->95% , Testing->87% (batch size=1000)

Utilizing different weight initialization methods to construct a Neural network proves to be more advantageous than randomly initializing the weights, as demonstrated in the table provided. The conclusive outcomes, including the ROC curve and accuracy scores, are detailed in Table 3 and visually represented in Fig. 11.

Table 3. Customer Churn Prediction 2020 Results

Algorithm	Results
	Accuracy Score
RandomForest classifier	91.5%
Neural Network	>87%(batch size=100) using weight initializationtechniques , no of epochs=100

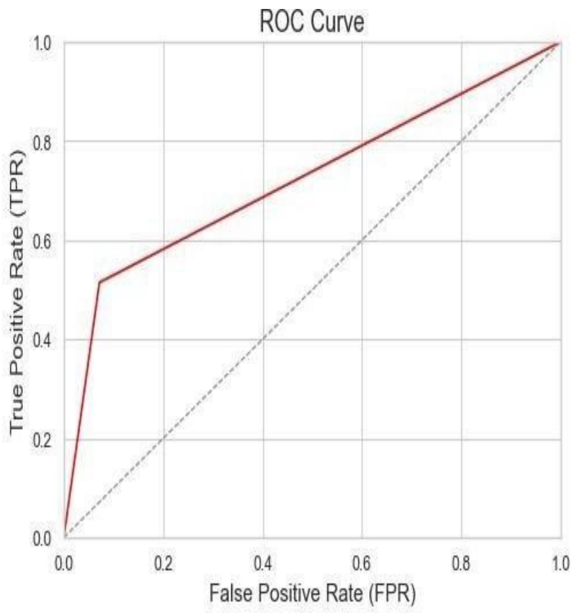

Figure.11 ROC Curve for Customer Churn Prediction 2020 Dataset

6. Conclusion

The telecommunications industry places significant reliance on customer feedback. In the contemporary business landscape, addressing customer dissatisfaction is not a matter to be overlooked. Given the paramount importance of customer retention within the telecommunications sector, organizations must prioritize efforts to comprehend the underlying reasons for customer churn, enabling them to implement necessary adjustments. Leveraging machine learning technologies in the telecom industry can substantially enhance the company's ability to predict and mitigate customer attrition. In this research, Artificial Intelligence was used to forecast a customer's churn from a telecom firm, and two different datasets were analyzed utilizing Feature Engineering and Data Analysis methodologies to extract relevant information. Overall, Deep Learning did an excellent job of predicting customer turnover, although Machine Learning did a little better. However, there are a lot of variables to consider when estimating the efficacy of different algorithms. Different AI methods could be used in the future to reduce the error value even more.

Acknowledgements.

I thank my coauthors Dr.Vijayaraj , Dr.Subba reddy Chavva , Dr.Gururama senthilvel for their expertise and assistance throughout all aspects of our study and for their help in writing the manuscript.

References

[1] Gerpott TJ, Rams W, Schindler A. Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. Telecommunications policy. 2001 May 1;25(4):249-69.

[2] Wei CP, Chiu IT. Turning telecommunications call details to churn prediction: a data mining approach. Expert systems with applications. 2002 Aug 1;23(2):103-12.

[3] Qureshi SA, Rehman AS, Qamar AM, Kamal A, Rehman A. Telecommunication subscribers' churn prediction model using machine learning. InEighth international conference on digital information management (ICDIM 2013) 2013 Sep 10 (pp. 131-136). IEEE..

[4] Ascarza E, Iyengar R, Schleicher M. The perils of proactive churn prevention using plan recommendations: Evidence from a field experiment. Journal of Marketing Research. 2016 Feb;53(1):46-60.

[5] Adwan O, Faris H, Jaradat K, Harfoushi O, Ghatasheh N. Predicting customer churn in telecom industry using multilayer preceptron neural networks: Modeling and analysis. Life Science Journal. 2014 Jan;11(3):75-81.

[6] Yu W, Jutla DN, Sivakumar SC. A churn-strategy alignment model for managers in mobile telecom. In3rd Annual Communication Networks and Services Research Conference (CNSR'05) 2005 May 16 (pp. 48-53). IEEE.

[7] Umayaparvathi V, Iyakutti K. A survey on customer churn prediction in telecom industry: Datasets, methods and metrics. International Research Journal of Engineering and Technology (IRJET). 2016 Apr;3(04).

[8] Available online at https://www.kaggle.com/blastchar / telco-customer churn?select=WA_Fn-UseC_-Telco-Customer-Churn.csv.

[9] Available online at https://www.kaggle.com/c/customer-churnprediction- 2020/data?select=train.csv.

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

[11] Brânduşoiu I, Toderean G, Beleiu H. Methods for churn prediction in the pre-paid mobile telecommunications industry. In2016 International conference on communications (COMM) 2016 Jun 9 (pp. 97-100). IEEE.

[12] He Y, He Z, Zhang D. A study on prediction of customer churn in fixed communication network based on data mining. In2009 sixth international conference on fuzzy systems and knowledge discovery 2009 Aug 14 (Vol. 1, pp. 92-94). IEEE.

[13] Idris A, Khan A, Lee YS. Genetic programming and adaboosting based churn prediction for telecom. In2012 IEEE international conference on Systems, Man, and Cybernetics (SMC) 2012 Oct 14 (pp. 1328-1332). IEEE.

[14] Makhtar M, Nafis S, Mohamed MA, Awang MK, Rahman MN, Deris MM. Churn classification model for local

telecommunication company based on rough set theory. Journal of Fundamental and Applied Sciences. 2017;9(6S):854-68.

[15] Alboukaey N, Joukhadar A, Ghneim N. Dynamic behavior based churn prediction in mobile telecom. Expert Systems with Applications. 2020 Dec 30; 162:113779.

[16] Pustokhina IV, Pustokhin DA, Aswathy RH, Jayasankar T, Jeyalakshmi C, Díaz VG, Shankar K. Dynamic customer churn prediction strategy for business intelligence using text analytics with evolutionary optimization algorithms. Information Processing & Management. 2021 Nov 1;58(6):102706.

[17] Vo NN, Liu S, Li X, Xu G. Leveraging unstructured call log data for customer churn prediction. Knowledge-Based Systems. 2021 Jan 5; 212:106586.

[18] Pustokhina IV, Pustokhin DA, Nguyen PT, Elhoseny M, Shankar K. Multi-objective rain optimization algorithm with WELM model for customer churn prediction in telecommunication sector. Complex & Intelligent Systems. 2021 Apr 5:1-3.