



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 9, September 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



A Dynamic Churn Prediction Model using Machine Learning Approach

Prasad Bhosale, Gaurav Jadhav, Abhishek Dhane, Prof. Nitin Pise

UG Student, School of Computer Engineering & Technology, MITWPU Pune, India

Professor, Department of Computer Engineering, SCET Pune, India

ABSTRACT: Customer retention is a significant issue and one of the most pressing issues for big businesses. Companies are working to create methods to anticipate future customer churn since it has such a direct impact on their revenues, particularly in the telecom industry. As a result, identifying factors that contribute to customer turnover is critical in order to take the required steps to minimise churn. Our work's primary contribution is the development of a churn prediction model that helps telecom carriers forecast which customers are most likely to churn. The model created in this paper employs machine learning methods on a large data platform to provide a novel approach to feature engineering and selection. This study also established churn variables that are critical in identifying the underlying causes of churn in order to evaluate the model's success. CRM may increase productivity, suggest appropriate promotions to a group of probable churn customers based on similar behaviour patterns, and vastly improve the company's marketing efforts by understanding the major churn causes from customer data. The accuracy, precision, recall, f-measure, and receiving operating characteristics (ROC) area of the suggested churn prediction model are all assessed. Furthermore, via the rules produced by the attribute-selected classifier algorithm, it offers reasons behind the churning of churn consumers.

KEYWORDS: Receiving Operating Characteristics, Deep learning, Convolution Neural Network, churn prediction, Feature selection.

I. INTRODUCTION

Before subscribing to any of the many Telecom service choices available today, consumers go through a lengthy decision-making process. Telecom providers' services aren't very distinctive, and number portability is widespread. Customer loyalty is a problem. As a result, it's becoming more essential for telecoms firms to detect reasons that cause consumers to unsubscribe and take preventative steps to keep them. Start with the number of users who churned that month to determine your likely monthly churn. Then divide the total number of user days in that month by the number of churns per user day to obtain the number of churns per user day. Then multiply the figure by the number of days in the month to obtain the monthly churn rate. According to studies done over the last several years, data mining methods are more successful in forecasting customer turnover. Developing an effective churn prediction model is a time-consuming process that includes everything from identifying acceptable predictor variables (features) from a huge amount of accessible customer data to selecting an appropriate predictive data mining method for the feature set. In addition to the network data that they produce, telecom industries gather a significant quantity of client-related data such as consumer profiling, calling patterns, and demographic data. It is possible to categorise a customer's attitude of going away or not going away based on their history of calling behaviour and behaviour. According to studies conducted over the last decade, data mining methods are more successful in forecasting turnover. Churn prediction methods that use predictive modelling are also thought to be more accurate. Churn prediction systems and sentiment analysis use classification and clustering methods to identify churn customers and the reasons for their departure. Because we produce huge amounts of data on a regular basis in the telecom sector, mining such data using specialised data mining methods is a time-consuming job, and interpreting predictions using traditional approaches is difficult. Various academics have reported efforts to remove churn from big data sets using both static and dynamic methods, but such systems still face numerous difficulties in identifying churn. Occasionally, such telecommunication data may include churn, making it critical to detect search issues. Customer relationship management must be excellent in order to successfully identify churn from big data (CRM).

Using Natural Language Processing (NLP) and machine learning methods, we suggested churn detection and prediction from a large scale telecommunication data set in this study. The first system is concerned with the strategic NLP process, which includes data pre-processing, data normalisation, feature extraction, and feature selection. TF-IDF,



Stanford NLP, and occurrence correlation methods have all been suggested as feature extraction strategies. The whole curriculum was trained and tested using machine learning classification methods.

II. LITERATURE SURVEY

Telecommunications businesses aren't always the most popular among customers, and In the telecom business, customer loyalty is critical to profitability. Whether it's complex billing, spam marketing emails, poor customer support, internet speed, connection, or expensive plans, people often express dissatisfaction with service providers' performance. As a consequence, it should come as no surprise that telecommunications firms have a high percentage of client turnover. Customer turnover (attrition) is especially troublesome in the telecom sector since it manages huge fixed infrastructures that must be compensated by income. Client acquisition is typically prioritised by businesses, with customer retention being a distant second. However, attracting a new client may cost five times as much as keeping a current one. According to study conducted by Bain & Company, increasing client retention rates by 5% may boost earnings by 25% to 95%. Customer attrition, often known as churn, is a statistic that indicates customers who quit doing business with a firm or a certain service. Most companies may use this statistic to attempt to identify the causes of high churn rates and develop reactive action plans to address those causes. But what if you knew ahead of time that a particular client was on the verge of leaving your company, and you could take proactive steps to prevent it?

Customers may cancel for a variety of reasons, including poor service quality, customer service delays, pricing changes, new rivals joining the market, and so on. Typically, there is no one cause, but rather a chain of events that resulted in consumer dissatisfaction. There is no going back if your business is unable to recognise these signs and take action before the cancel button is pressed; your client has already left. However, you still have something useful in the form of data. Your client provided plenty of hints as to where you fell short. It may be a useful tool for gaining important information and training customer churn models. Machine learning is all about learning from the past and having crucial knowledge on hand to enhance future experiences. When it comes to the telecommunications industry, there is a lot of potential for growth. The richness and volume of consumer data that carriers gather may help carriers move from a reactive to a proactive posture. The development of advanced artificial intelligence and data analytics methods has further aided in the efficient use of this rich data to combat churn.

Churn prediction employs a variety of techniques, including machine learning and data mining. The decision-tree algorithm is a dependable churn prediction technique [1]. In addition, for churn prediction, a neural network approach [7], data certainty [8], and particle swarm optimization are employed. A current collection of software to improve the quality of identifying potential churners, according to system [2]. The roles are categorised as deal, request pattern, and call pattern adjustments overview functions and are retrieved from request information and client accounts. The features are assessed using two probabilistic data mining methods, Nave Bayes and Bayesian Network, and the results are compared to those produced using the C4.5 decision tree, a commonly used approach for classification and prediction. For a variety of reasons, this has resulted in the potential that customers may rapidly switch to rivals. Improve churn prediction from huge amounts of data using extraction in the near future is one method that may be utilised to accomplish this. According to [3,] the formalisation of the collecting process's time-window, as well as a literature study. Second, this study examines the increase in churn model accuracy by extending the length of customer events from one to seventeen years using logistic regression, classification trees, and bagging combined with classification trees. As a consequence, researchers will be able to significantly decrease data-related demands such as data collecting, preparation, and analysis. The cost of a subscription is determined by the duration and promotional nature of the subscription. The newspaper industry is sending them a letter informing them that their service will be discontinued. Then ask whether they want to renew their membership and provide instructions on how to do so. Customers cannot cancel their subscriptions, although they do have a four-week grace period once their membership has expired.

According to [4,] the most effective consumer interaction techniques may be utilised to effectively increase customer satisfaction. In one of Malaysia's top telecoms companies, the researchers used a Multilayer Perceptron (MLP) neural network technique to predict customer churn. The findings were compared to the most used churn prediction methods, including Multiple Regression Analysis and Analyzing Logistic Regression. With the Levenberg Marquardt learning method, the maximum neural network design has 14 input nodes, 1 hidden node, and 1 output node (LM). When compared to the most popular churn prediction methods, such as Multiple Regression Analysis and Logistic Regression Analysis, a Multilayer Perceptron (MLP) neural network methodology was used to forecast customer churn at one of Malaysia's top telecoms firms. Using a Partial Least Square (PLS) method, system [5] focused on highly correlated intervals in data sets to create an efficient and descriptive statistical churn model. According to early findings, the

suggested approach produces more accurate results than traditional prediction models and identifies key factors to better explain churning patterns. In addition, network management, overage administration, and problem handling methods are presented and explored in the context of a few basic marketing campaigns. Burez and Van den Poel [6] compare the performance of random sampling, Advanced Under-Sampling, Gradient Boosting Method, and Weighted Random Forest in churn prediction models. Metrics were used to assess the idea (AUC, Lift). The research concludes that the sampling approach is superior than the other methods considered. Gavril et al. [7] provide a new data mining technique for explaining customer churn detection over a wide dataset type. On the basis of incoming numbers as well as outbound input calls and texts, about 3500 customer information are evaluated. For training categorization and study, certain machine learning algorithms were employed. For the whole dataset, the system's estimated average accuracy is about 90%. He et al. [8] created a prediction model based on the Neural Network technique to solve the problem of customer churn for a large Chinese telecoms company with about 5.23 million members. The average degree of accuracy was 91.1 percent, indicating a high level of predictability.

To simulate AdaBoost-churning telecommunications issues, Idris [9] proposed a genetic engineering approach. The series was confirmed by two Standard Data Sets. One from Orange Telecom and the other from cell2cell, both with an accuracy of 89 percent, and 63 percent for the other. On the big data platform, Huang et al. [10] investigated customer attrition. The researchers wanted to demonstrate that depending on the amount, diversity, and speed of data, big data substantially increases the cycle of churn prediction. Data from China's largest telecoms company's Project Support and Business Support Department was anticipated to be stored in a large data repository for fracture engineering. The forest algorithm was applied at random by AUC and evaluated. According to [11], clustered input characteristics are used using k-means and fuzzy c-means clustering algorithms to divide subscribers into distinct groups. These classes are used to build the Adaptive Neuro Fuzzy Inference System (ANFIS), which is a prediction model for active churn control. Neuro fuzzy parallel categorization is the initial stage in the prediction process. The outputs of the Neuro fuzzy classifier are then used by FIS to determine the actions of the churners. In order to identify inefficiency issues, success metrics may be utilised. Customer service network services, operations, and efficiency are all linked to churn management measures. The flexibility of GSM numbers is an important factor for churner selection. A new collection of applications has been added to System [12] to enhance the detection of probable churners. The characteristics are generated from call data and client profiles and are classified as contract, call pattern, and call pattern changes description features. The characteristics are evaluated using two probabilistic data mining methods, Naive Bayes and Bayesian Network, and the results are compared to those produced using a C4.5 decision tree, which is widely employed in many classification and prediction applications. These have led to the danger that consumers would readily move to rivals, among other things. Improve churn prediction from huge amounts of data using extraction in the near future is one method that may be utilised to accomplish this. [13] states that Formalization of the time window selection technique, as well as literature review. Second, this research examines the rise in churn model consistency by utilising logistic regression, classification trees, and bagging in conjunction with classification trees to extend the history of customer events from one to seventeen years. As a result, researchers may substantially decrease data-related difficulties in areas such as data storage, planning, and research. The amount that customers must pay is determined on the length of the subscription and the promotional value. The newspaper business sends a letter to notify them that their subscription is about to expire. Then ask whether they want to renew their membership and provide instructions on how to do so. Customers cannot cancel their subscriptions, although they do have a four-week grace period once their membership has expired.

To successfully decrease customer turnover rates, the most effective customer retention methods should be utilised, according to [14]. In one of Malaysia's top telecommunications companies, the study proposes a neural network method for Multilayer Perceptron (MLP) to forecast customer turnover. The results were compared to the most popular churn prediction methods, including Multiple Regression Analysis and Logistic Regression Analysis. With the Levenberg Marquardt (LM) learning method, the optimum design of the neural network includes 14 input nodes, 1 hidden node, and 1 output node. When compared to the most popular churn prediction methods, such as Multiple Regression Analysis and Logistic Regression Analysis, a Multilayer Perceptron (MLP) neural network methodology was used to forecast customer churn at one of Malaysia's top telecoms firms. On the [15] system Using a Partial Least Square (PLS) approach based on highly linked data sets across variables, create an accurate and succinct predictive churn model. A pilot experiment demonstrates that the model provided outperforms conventional prediction models and highlights important factors for better understanding churning patterns. A variety of fundamental churn marketing techniques are also given and addressed, including system management, overage management, and complaint management.

Burez and Van den Poel [16] evaluated the performance of Random Sampling, Advanced Under-Sampling, Gradient Boosting Model, and Weighted Random Forests in churn prediction models using unbalanced datasets. The model was

evaluated using (AUC, Lift) metrics. The under sampling method outperformed the other strategies evaluated, according to the results.

Gavril et al. [17] used a dataset for call details of 3333 consumers with 21 characteristics and a dependent churn parameter with two values: Yes/No to propose an advanced data mining technique for predicting churn for prepaid clients. Information regarding the amount of incoming and outgoing messages, as well as voicemail for each client, are among the features. To decrease data dimensionality, the author used the principal component analysis (PCA) method. To estimate the churn factor, three machine learning methods were used: Neural Networks, Support Vector Machines, and Bayes Networks. The author utilised AUC to assess the algorithms' performance. For Bayes Networks, Neural Networks, and Support Vector Machine, the AUC values were 99.10 percent, 99.55 percent, and 99.70 percent, respectively. There were no missing values in the dataset utilised in this research since it was tiny. At order to address the issue of customer turnover in a major Chinese telecom business with around 5.23 million subscribers, He et al. [18] developed a model for prediction based on the Neural Network algorithm. The total accuracy rate was used as the forecast accuracy criterion, and it came in at 91.1 percent. To simulate the churn issue in telecoms, Idris [19] presented a method based on genetic programming with AdaBoost. Two standard data sets were used to evaluate the model. One by Orange Telecom and the other by cell2cell, with the cell2cell dataset having 89 percent accuracy and the other having 63 percent. The issue of client turnover in the big data platform was investigated by Huang et al. [20]. The researchers wanted to show that, depending on the amount, diversity, and velocity of the data, big data may significantly improve the process of forecasting churn. To create the cracks in data from the Operation Support and Business Support departments at China's biggest telecoms business, a big data platform was required. AUC was used to assess the Random Forest method.

III. PROPOSED SYSTEM DESIGN

The goal of the proposed study is to use text analysis and a machine learning classifier to detect churn. During prediction, detect the customer's shifting behaviour pattern. To determine which factors have the greatest impact on churn forecast accuracy. To analyse and compute the churn rate month by month and day by day, which is helpful for improving the system's service quality. To improve system accuracy, the proposed research effort will design and build a method for churn prediction utilising NLP and machine learning techniques. Then, during prediction, we detect the customer's shifting behaviour pattern. We also analyse the factors that most affect the accuracy of churn prediction, and we ultimately review and compute churn rate month by month and day by day, which is helpful for improving the system's service quality. In this paper, we present a method for predicting churn from big data sets. The system starts with a telecoms synthetic data set that includes some imbalance meta data. To do data preparation, data normalisation, feature extraction, and feature selection, as needed. During this execution, various optimization techniques were utilised to remove duplicate features that may cause a high error rate during the execution. The suggested system's training and testing execution. After all stages are completed, the system describes the categorization accuracy for the whole data set.

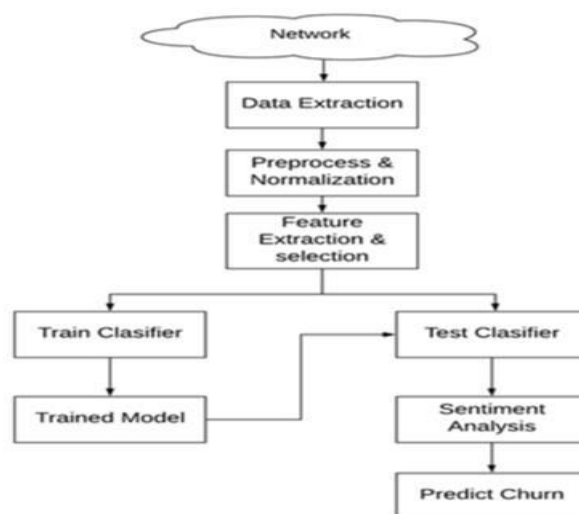


Figure 1 : Proposed system design



The aim of this such kind of research in the telecommunications industry is to help businesses make more profit. Telecom companies have become known to forecast turnover as one of the most important sources of income. Therefore, this research was aimed at building a system in the Telecom Company that predicts customer churn. Such prediction models will achieve high AUC values. The sample data was divided into 70% for training and 30% for testing to evaluate and develop the model. We chose 10-fold cross-validation for evaluating and optimizing hyper parameters. We used engineering tools, effective function transformation and selection approach. Making the interface fit for machine learning algorithms. Another concern was also found: the data was not balanced. Only about 5% of the entries are customers' churn. A problem has been solved by under-sampling or using trees algorithms that are not affected by this issue. In detecting the churn in large data and providing accurate prediction, our different classifiers can be more accurate. This work contributes to suggesting a supervised approach to the extraction of dimensional categories, selecting suitable characteristics and avoiding duplication by measuring correlation between characteristics. The results obtained show that there is a comparatively higher f-score in the weighted frequency of the term with the correlation process. In this regard, selecting features using weighted word frequency is more important. The overlap between features in a category of aspect is avoided by measuring the association.

1. Data Acquisition: First of all the information for different Telecom Sector Customer based on certain parameters is extracted data.
2. Preprocessing: Then we will apply various preprocessing steps such as lexical analysis, stop word removal, stemming (Porters algorithm), index term selection and data cleaning in order to make our dataset proper.
3. Lexical analysis: Lexical analysis separates the input alphabet into,
 - 1) Word characters (e.g. the letters a-z) and 2) Word separators (e.g space, newline, tab).
4. Stop word removal: Stop word removal refers to the removal of words that occur most frequently in documents.
5. Stemming: Stemming replaces all the variants of a word with a single stem word. Variants include plurals, gerund forms (ing forms), third person suffixes, past tense suffixes, etc.).
6. Data Training: We compile artificial as well as real time using online news data and provide training with any machine learning classifier.
7. Testing with machine learning: We predict online news using any machine learning classifier, weight calculator for real time or synthetic input data accordingly.
8. Analysis: We demonstrate the accuracy of proposed system and evaluate with other existing systems

Algorithm Design

1 Bagging Classifier

Input: inp 1.....n all input parameters, Desired Threshold Th.

Output: Executed for output as lable.

Step 1 : Read all records from database (R into DB)

Step 2: Parts [] \leftarrow Split(R)

Step 3:
$$CVal = \sum_{k=0}^n Parts[k]$$

Step 4: check (Cval with Respective threshold)

Step 5: T \leftarrow get current state with timestamp



Step 6 : if(T.time > Defined Time)

Read all measure of for TP and FN

Else continue. Tot++

Step 7: calculate score = (TP *100 / Tot)

Step 8 : if (score >= Th)

Generate event

end for

2 Decision Tree Classifier

Input : Selected feature of all test instances D_1, \dots, D_n , Training database policies $\{T_1, \dots, T_n\}$

Output: No. of probable classified trees with weight and label.

Step 1: Read (D into $D[i]$)

$V \leftarrow \text{Extract features (D)}$

Step 2: $N \leftarrow \text{Count_Features(D)}$

Step 3: for each(c into TrainDB)

Step 4: $Nc[i] = \text{Ext_Features(c)}$

Step 5: select relevant features of $w = \{Nc[i], N\}$

Step 6: Statement ($w > t$)

Step 9: Return Tree Instance $\{Nc[i], N, w, \text{label}\}$

3 Knearest neighbour Classifier

Input: Train_DatasetF TrF[], Test_DatasetF TsF[], Threshold T.

Output: Classified label

Step 1: Read R {All attributes} from current parameters.

Step 2: Map with train features with each sample.

Step 3 : Calculate distance of train DB with same evidences

$$distance = \sum_{k=0}^n (TrF, TsF)$$

Step 4 : evaluate distance > threshold

Step 5: Return the predicted label

4 Random Forest Classifier

Input: Training Rules Tr[], Test Instances Ts[], Threshold T.

Output: Weight $w \in \{0-1\}$

Step 1 : Read each test instance from (TsInstnace from Ts)

Step 1 : Read each test instance from (TsInstnace from Ts)

Step 2 : $TsIns = \sum_{k=0}^n \{Ak \dots An\}$

Step 3 : Read each train instance from (TrInstnace from Tr)

Step 4 : $TrIns = \sum_{j=0}^n \{Aj \dots Am\}$

Step 5 : $w = WeightCalc(TsIns, TrIns)$

Step 6 : if ($w \geq T$)

Step 7 : Forward feed layer to input layer for output $OutLayer[] \leftarrow \{Tsf, w\}$

Step 8 : optimized feed layer weight, $Cweight \leftarrow OutLayer[0]$

Step 9 : Return Cweight

Results and Discussions

The below analysis is the system classification graph. The graphs display how system classify the overall inputs into classification various instances. The proposed system is implemented with RNN combination, which gives all results with satisfactory level. For performance evaluation 5000 instances given for training and 1500 reviews given for testing with different cross validation. Here system compares the proposed results with two different existing systems.

Table 1 :Comparative analysis of various classification algorithms

No.	Method	Accuracy	Precision	Recall	F-1 score
1	Random Forest	0.95	0.95	1	0.97
2	DT	0.89	0.92	0.96	0.94
3	BaggingClassifier	0.94	0.96	0.94	0.97
4	Knearest neighbors	0.81	0.86	0.92	0.89

The above Table 5.3 depicts the comparative analysis of various classification algorithms evaluation for proposed churn prediction module. The KNeighbors provides lowest accuracy thus Random forest classification gives highest accuracy with 95% on various cross validation. The similar results has been demonstrated in below Figure 2.

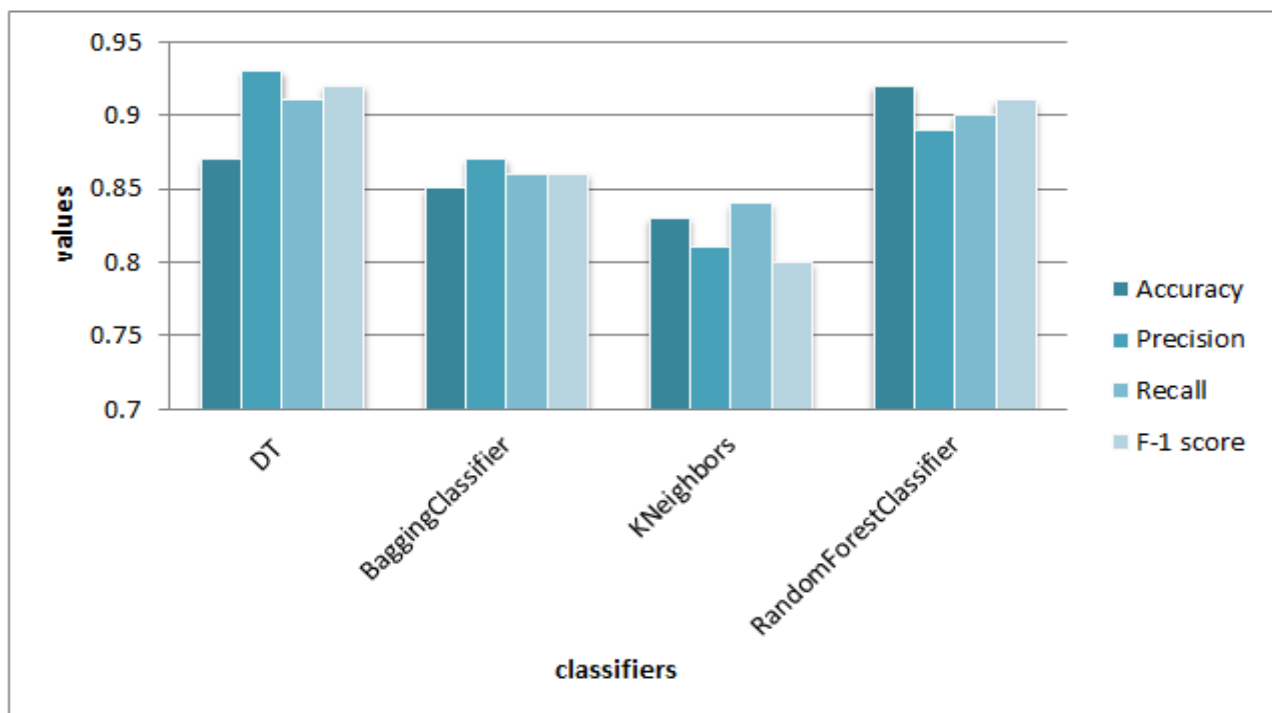


Figure 2 : Comparative analysis of various classification algorithms

IV. CONCLUSION FUTURE WORK

This research primarily focuses on finding and detecting churn customers from large telecoms data sets, as well as churn prediction systems generated by various algorithms. Some systems still have issues with linguistic data conversion, which may result in a high rate of errors during execution. Many academics have proposed combining Natural Language Processing (NLP) methods with different machine learning algorithms in the hopes of achieving excellent data structuring results. Customer turnover is a significant issue and one of the most pressing issues for big businesses. Companies are developing machine learning algorithms to anticipate future customer churn since it has such a direct impact on their profits, particularly in the telecom industry. To determine the churn prediction rate, Random Forest, Decision Tree, Bagging Classifier, and K-nearest neighbour classifiers are used. In comparison to decision trees, Bagging classifiers, and K-nearest neighbour classifiers, random forest classifier has the greatest accuracy of 95%, while KNN classifier has the lowest accuracy of 81 percent.

REFERENCES

1. Karahoca, Adem, and Dilek Karahoca. "GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system." *Expert Systems with Applications* 38.3 (2011): 1814-1822.
2. K.Shanumgapiya, R.Sugumar, An effective encryption algorithm for multi-keyword-based top-K retrieval on cloud data, *International Journal of Business Intelligence and Data Mining*, Volume 12, Issue 1, pp.126-134, January 2017.
3. Kirui, Clement, et al. "Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining." *International Journal of Computer Science Issues (IJCSI)* 10.2 Part 1 (2013): 165.
4. Ballings, Michel, and Dirk Van den Poel. "Customer event history for churn prediction: How long is long enough?." *Expert Systems with Applications* 39.18 (2012): 13517-13522.
5. Ismail, Mohammad Ridwan, et al. "A multi-layer perceptron approach for customer churn prediction." *International Journal of Multimedia and Ubiquitous Engineering* 10.7 (2015): 213-222.
6. Patel, Z., Senjaliya, N., & Tejani, A. (2019). AI-enhanced optimization of heat pump sizing and design for specific applications. *International Journal of Mechanical Engineering and Technology (IJMET)*, 10(11), 447-460.
7. Lee, Hyeseon, et al. "Mining churning behaviors and developing retention strategies based on a partial least squares (PLS) model." *Decision Support Systems* 52.1 (2011): 207-216.



8. Burez D, den Poel V. Handling class imbalance in customer churn prediction. *Expert Syst Appl.* 2009;36(3):4626–36.
9. Brandusoiu I, Todorean G, Ha B. Methods for churn prediction in the prepaid mobile telecommunications industry. In: *International conference on communications.* 2016. p. 97–100.
10. He Y, He Z, Zhang D. A study on prediction of customer churn in fixed communication network based on data mining. In: *Sixth international conference on fuzzy systems and knowledge discovery*, vol. 1. 2009. p. 92–4.
11. Idris A, Khan A, Lee YS. Genetic programming and adaboosting based churn prediction for telecom. In: *IEEE international conference on systems, man, and cybernetics (SMC).* 2012. p. 1328–32.
12. Huang F, Zhu M, Yuan K, Deng EO. Telco churn prediction with big data. In: *ACM SIGMOD international conference on management of data.* 2015. p. 607–18
13. Karahoca, Adem, and Dilek Karahoca. "GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system." *Expert Systems with Applications* 38.3 (2011): 1814-1822.
14. Kirui, Clement, et al. "Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining." *International Journal of Computer Science Issues (IJCSI)* 10.2 Part 1 (2013): 165.
15. Ballings, Michel, and Dirk Van den Poel. "Customer event history for churn prediction: How long is long enough?." *Expert Systems with Applications* 39.18 (2012): 13517-13522.
16. Ismail, Mohammad Ridwan, et al. "A multi-layer perceptron approach for customer churn prediction." *International Journal of Multimedia and Ubiquitous Engineering* 10.7 (2015): 213-222.
17. Lee, Hyeseon, et al. "Mining churning behaviors and developing retention strategies based on a partial least squares (PLS) model." *Decision Support Systems* 52.1 (2011): 207-216.
18. Burez J, Van den Poel D. Handling class imbalance in customer churn prediction. *Expert Systems with Applications.* 2009 Apr 1;36(3):4626-36.
19. Senjaliya, N., & Tejani, A. (2020). Artificial intelligence-powered autonomous energy management system for hybrid heat pump and solar thermal integration in residential buildings. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, 11(7), 1025-1037.
20. 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.
21. Li L, Jiang P, Xu H, Lin G, Guo D, Wu H. Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River, China. *Environmental Science and Pollution Research.* 2019 Jul;26(19):19879-96.
22. R.Sugumar, A.Rengarajan, C.Jayakumar, Trust based authentication technique for cluster based vehicular ad hoc networks, *The Journal of Mobile Communication, Computation and Information [Wireless Networks]*, Volume 22, Issue 6, pp.615-624, August 2016.
23. Idris A, Khan A, Lee YS. Genetic programming and adaboosting based churn prediction for telecom. In: *2012 IEEE international conference on Systems, Man, and Cybernetics (SMC) 2012 Oct 14 (pp. 1328-1332).* IEEE.
24. Huang F, Zhu M, Yuan K, Deng EO. Telco churn prediction with big data. In: *ACM SIGMOD international conference on management of data.* 2015. p. 607–18



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details