**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 35, No. 3, September 2024, pp. 1610~1617
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v35.i3.pp1610-1617 

# Predicting customer churn in telecommunication sector using Naïve Bayes algorithm

**Biswa Ranjan Agasti, Susanta Satpathy**
Department of Computer Science and Engineering, Vignan's Foundation for Science, Technology and Research, Guntur, India

**Article Info**

***Article history:***

Received Dec 17, 2023
Revised Apr 20, 2024
Accepted May 6, 2024

***Keywords:***

Customer churn prediction
Customer relationship
Machine learning
Naïve Bayes algorithm
Telecommunications

**ABSTRACT**

The telecom sector creates huge amounts of information every day as a result of its large customer base. Business professionals and decision-makers emphasized that maintaining existing clients is less expensive than recruiting new ones. Business analysts and customer relationship management (CRM) need to know the reasons why customers leave and the behavior patterns from earlier churn consumer's data. Today, there is a problem with customer churn examination and prediction in the telecom industry since it is crucial for the sector to examine customer behavior to identify those who are going to stop their subscriptions. Customer retention could be increased by utilizing detection system to detect consumer behavior. Recent advancements in machine learning (ML) have made churn prediction more precise and practical. It is essential for identifying customers ready to leave using company's products and services in the early stage. Hence in this work, predicting customers churn in telecommunication sector using Naïve Bayes (NB) model is presented. The performance of presented NB algorithm is evaluated using the parameters accuracy, precision, and sensitivity. The NB algorithm will have better performance than pervious approaches.



***Corresponding Author:***

Biswa Ranjan Agasti
Department of Computer Science and Engineering
Vignan's Foundation for Science, Technology and Research
Guntur, Andhra Pradesh, India
Email: biswaranjanagasti112@gmail.com

## 1. INTRODUCTION

In recent times, telecommunication industry is the rapidly growing industry due to the technological advancements and new life style adoptions to ubiquitous internet and communications. In the current world, telecom companies produce large amounts of information at extremely fast rates. By providing better services, many telecom service providers contend for market share. The oversaturated market and client turnover are two significant difficulties that telecom businesses must contend with as a result of the rapid improvements in global information services, as customers have several choices for better and less expensive services. The major goals of telecom companies are to maximize profits and survive in a competitive market [1]. Customers are growing with high interesting to current quality of service (QoS) produced by businesses. Today, there is more contention to provide clients with cutting-edge QoS. Effective customer relationship management (CRM) methods, on other hand, can profit the company by assisting it in attracting in more clients, retaining those it already has, and growing client retention while creating more income for its operations [2].

The customer's percentage that stops utilizing company's products or services is called customer churn. It might happen as a result of quality and price of service competition with other companies, a business's inability to build strong connections with its clients, or aggressive marketing by competing companies. Customer churn happens gradually over time; it doesn't happen suddenly [3]. These possibilities will be extracted by a data mining-based work called the customer churn analysis. Due to today's strong competition, numerous companies are now providing the same product at essentially the same level of service and quality [4]. Customer churn has a very high impact on business development because it is related directly to a company's core business practices. Reduced customer turnover is always the main objective of a business when the customer is its most valuable asset. Therefore, maintaining current customers who have been loyal to them for a long time depends on customer happiness [5].

The prediction of churn is gaining the popularity day by day by the research community and it is a powerful paradigm that supports the data-driven operational decisions [6]. The churn is predicted to identify the customers who may terminate their service contracts with the service provider. The main intention of churn customer is voluntary or involuntary in nature. The growing number of operators and scientific progress increased the opposition levels. The companies are trying hard to survive in this aggressive competitive market depending on different complicated strategies. The churn customers cause a considerable amount of loss in telecom services and become a severe issue in telecom industry [7].

The customer churn is one of the most considerable concerns in service sectors with huge competitive services. Prediction of churn is the essential and foremost to solve these issues. The prediction of customer churn can serve as a potentially big revenue source when it is predicted in the early stage [8]. The main reason that the customers churn is recognized as the biggest problem in any industry is due to the cost of obtaining new customers is whole lot more than keeping the existing customers [9]. The prediction of customer churn was done using different methods include machine learning (ML), data mining, and hybrid techniques. These kinds of methods support and enable the companies to identify, predict and retain their customers [10]. However, most of the researches confirmed that ML techniques are highly effective for predicting the churn customers [11].

Hence in this work, predicting customer churn in telecommunication sector using Naïve Bayes (NB) algorithm is presented. The rest of the work is organized as follows: the section 2 demonstrates the predicting customer churn in telecommunication sector using NB algorithm. The result analysis of developed application is described in section 3. Section 4 is where the work is finally finished.

## 2. LITERATURE SURVEY

In this section, stacking is contrasted with various heterogeneous ensembles for telecom churn prediction. In this analysis, four various telecom datasets were used to evaluate and compare the performance of five grid-searched optimized base classifiers: decision trees (DT), logistic regression (LR), multilayer perceptron's, support vector machine (SVM), K-nearest neighbours (KNN) as well as their heterogeneous ensembles: grading, majority voting, stacking, soft voting, and weighted majority voting. The results demonstrate that heterogeneous ensembles outperformed single classifiers by a massive margin, with stacked being the most successful ensemble. More investigations will be conducted to examine benefits of ensembles for trade application in greater detail [12].

SVM cloud-based genetic method for prediction customer churn (CCP) is described. The optimal genetic algorithm (OGA) with SVM model for CCP being presented in this analysis. First, the double chain quantum genetic algorithm is used to produce the OGA. The C and SVM parameters are then optimized using the produced OGA. The effectiveness of the OGA-SVM system is evaluated using benchmark dataset from the telecom sector [13].

How to anticipate telecom client churn using ML is discussed. Deep learning as well as ML approaches are being explored in order to forecast telecom client attrition. To predict whether a client will stay or leave, they compare widely used methods like random forest (RF) classifiers and SVM by relatively more modern buildings such as XGBoost and deep neural networks (DNN). By running the models through a grid search, the efficiency of the models is further evaluated. The RF system performs the best in individual case, according to the experiment's results, with such a prediction accuracy of 90.96% on test dataset before to search algorithm [14]. Enhanced churn prediction for the telecommunication sector is discussed. To decrease churn while using this well-known categorization technique, considerable thought should be given to the feature sets that will be used. SVM should also be utilised to analyse customer turnover patterns. The features will calculate in view of observational outputs, which indicated that presented model outperforms the earliest systems of ROC, sensitivity, specificity, accuracy, and processing speed [15].

Authors discussed how to predict customer churn in the telecommunication business using ML algorithms. ML methods can be used to determine users are most likely to stop their subscriptions. This plan will generate them best assistances while reducing churn. These business plans support telecom assistance become successful. In order to create this system, XGBoost, RF, and DT are used [16].

The reason of supervised learning methods to detect customer churn in telecom industry is discussed. To detect customer churn, data mining methods are implemented. In the end, it will help with studying consumer behaviour and identifying clients who are likely to churn. To predict consumer turnover behaviour to use an online set of data from Kaggle, the study’s authors used a variety of classifiers. The results demonstrate that this model, using bagging methods and provides a high level of accuracy [17].

The use of ensemble-based classifiers in a comparative evaluation of telecom industries capability to predict customer’s income is discussed. For churn detection in telecom sector, ensemble-based classifiers RF, Bagging and Boosting were employed. We compared ensemble-based classifiers to familiar classifier. In comparison to other methods, the experimental results demonstrated that RF has low specific, failure rate, high sensitivity as well as greater accuracy of 91.66% [18].

Use of weighted selective ensembles-based customer churn prediction is demonstrated. To address the customer turnover issue of CRM, this study built a system depended on properties of amount and imbalanced information and confirmed it using actual telecom data. The influence of the ensemble is definitely beneficial when the base models are SVM, which have a greater hit rate, lift coefficient, and accuracy rate. It is a reliable predictor of client attrition [19].

Customer churn prediction in the telecommunication utilizing LR analysis and DT is described. A framework is designed for the prediction of churn for telecom companies that use computational modeling methods like DT and regression models. Depending on how well these algorithms perform given the dataset, and a comparison is made. The user can identify which algorithm works the best for the data that use this comparative study, saving time that would otherwise have been required to create an accurate model [20].

A new churn prediction process is described while determining the communication patterns between subscribers and considered a propagation procedure in the network based on detail call records which can transfer the churn information from churners to non-churners. An accurate and quick propagation procedure can be possible through the detection of communication and setting the initial churners energy differently in centrality or churn date. This process was validated based on prediction model performance which is trained with social network features and personal features [21].

This approach performed churn classification very effectively from non-churn customers and in addition it predicts the customers who will become the churn possibly in the future. The performance of presented approach is evaluated using 4 rule-generation techniques like genetic algorithm (GA), covering algorithm (CA), learning from examples (LEM2) algorithm and exhaustive algorithm (EA). The results indicate that presented approach based on GA was best model to extract the knowledge in decision rules form from the benchmark telecom dataset [22].

The effects of collective inference techniques and relational classifiers are evaluated on predictive relational power learners and the performance of these models where the relational learner was combined with traditional customer churn prediction models in telecom industry. Eventually the network construction effect over model performance is investigated. The results indicated that the weights and edges definition in network have an impact on predictive model results. Thereby the best configuration was a non-relational learner with network variables without any collective inference using undirected networks and binary weights [23]. Three ML classifiers are used for the churn prediction using two popular datasets namely IBM Watson dataset which contains 21 attributes, 7033 observation and second one is cell2cell dataset which have 57 attributes and 71,047 observations. The performance of this approach is measured in terms of area under the curve (AUC). This model achieved better accuracy than previous methods [24].

Different learning strategy types were investigated to design a prediction approach. Different ML algorithms such as RF, gradient boosting (GBM), extreme randomized tree (ERT), KNN and artificial neural network (ANN) were tested to determine the best model for making a customer churn prediction technique. Here two publicly available datasets namely American Telecom Market and Southeast Asian Telecom Industry are used. The ANN have better results compared to other classifiers [25].

In order to keep customers and develop a strategy for creating a successful model that can be targeted to keep a given group of customers, several studies on the subject of CRM have been carried out in a variety of industries. Numerous data mining and statistical methods have been applied to predict churn. DT, neural, regression models, networks, clustering, SVM, and other well-known methods are a few examples. However, they take a lot of time and are not reliable. As a result, this study offers a precise way for predicting consumer attrition in the telecom business in order to address these issues.

## 3. PREDICTING CUSTOMER CHURN IN TELECOMMUNICATION SECTOR

In this section, predicting customer churn in telecommunication sector using NB algorithm is presented. The necessary information for this operation is available on the Kaggle website, which is run by a telecom business.

Customers' specifications are included in this dataset, along with data on how frequently they use different company services like calls, short message service (SMS), throughput, voice mail, download/upload speed, packet loss, latency, line of sight (LoS) and packet drop during both daytime and nighttime hours. Each attribute is given a name because the dataset obtained cannot be utilized to directly develop churn prediction models. The primary objectives of data preparation process are reducing dimension of the dataset attributes and removing unnecessary data. Data preparation is the most important phase of predictive models since unclean data contains redundancies, ambiguities, transformations and errors that need to be eliminated.

In Figure 1, block diagram is shown and the NB algorithm is used to predict future customer churn. Preprocessing is needed before a dataset is given into the models. The following is how data is collected and processed: each row having null or missing values was removed after searching the dataset for these values. The dataset's categorical data are transformed into numerical data for analysis model. Data has been normalized because some models are vulnerable to data dispersion in dataset features. All feature values are now adjusted to range from 0 to 1 as a result of doing this.

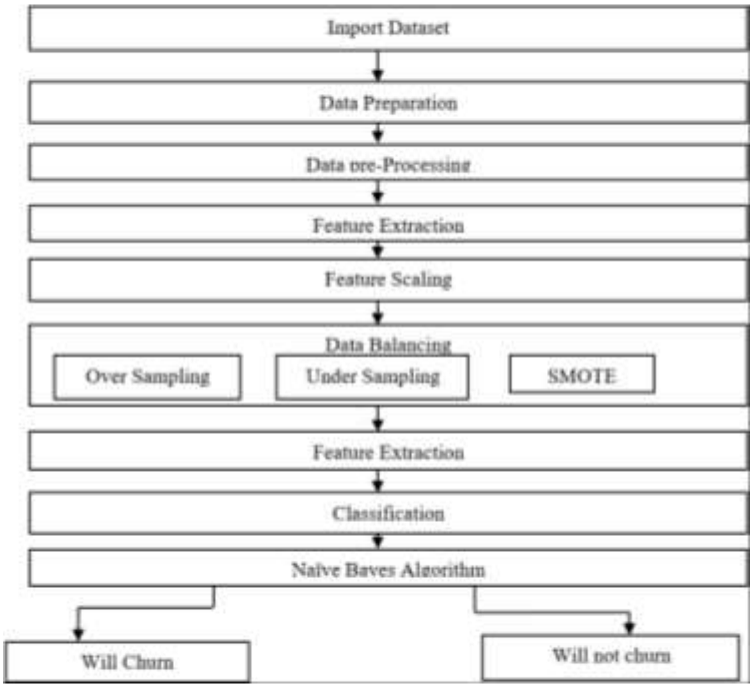


Figure 1. Block diagram for predicting customer churn

Prediction performance improves if the dataset contains highly predictive and useful variables. Therefore, increasing the effort put into selecting critical traits while decreasing the number of irrelevant ones is necessary to improve prediction performance. Finding important traits for analysis is a process known as feature extraction and selection. Using this strategy, it is feasible to extract features from the data that include redundant data and improve computational performance of ML model by identifying the useful characteristics. Currently, a few features have been disabled. In methods, it is expected that the number of instances will be distributed equally among the various groups. In case data is unequally distributed among classes, then the model tends to ignore minority class and only forecast important class. However, the data set in many actual issues is unbalanced and the data distribution between classes is uneven.

Unbalanced datasets are a topic that has been studied, and a few solutions have been raised. Through the use of excessive and under-sampling approaches, these solutions are categorized. The oversampling methods balance the data by lowering the minority to majority ratio and increasing the proportion of minority samples. Repeating minority cases is the easiest way to do this, but you may also use a tougher technique called synthetic minority over-sampling technique (SMOTE).

The neighbors of the minority class are taken into account when creating synthetic samples using this method. SMOTE is a method that uses random oversampling while allowing for the traits of the minority class. To bridge the gap between the ratios of the various class samples, under sampling approaches reduces majority group instances. It is normal, because a portion of the data is destroyed, it is best suited for situations where there is a reasonable number of information and the loss does not impact the model's accuracy. In this method, the dataset has been attempted to be balanced by three algorithms: random under-sampling, random over-sampling, and SMOTE algorithm. After the data balancing phase is finished, the attributes are chosen for the classification procedure.

A Bayesian network shows the causal probabilistic link between two random variables as well as their interdependent dependencies and joint probability distribution. A directed acyclic graph and a collection of conditional probability distributions comprise two main components. The NB mechanism is based on joint probability distributions. For instance, consider some events, such as a voice-only customer or a customer who has already been disconnected.

They can operate as Bayesian network node. At present, if node is dependent on other node in any way, an arc or array is drawn between them. It is considered to signify that child node's experience is seen with the experience of parent's node. Given that the NB explains directed acyclic network, it is possible to use conditional probability as well as the chain rule expressed in (1) below to determine complete joint distribution, or probability of last match (customer churn), provide all different incidents occurs.

$$P(x_1, x_2, x_3 \dots x_n) = P\ (x_1).P(x_2|x_1).P(x_3|x_2, x_1) \dots P(x_n|x_1, x_2, \dots x_{n)} = \pi(x_i|Parents(x_i)) \tag{1}$$

The telecom dataset has two different kinds of clients. Customers who remain with the business, do not even change, and are almost always impacted by competing companies generate the first type. Churning customers provide another category. The developed system identifies churning customers and the causes of their leaving. In order to determine the quality of services offered by a specific business, the services of multiple telecom corporations are evaluated as part of the classification process. The NB algorithm determines whether or not the customer is a churner according to the classification results.

## 4. RESULT ANALYSIS

In this section, predicting customer churn in telecommunication sector using NB algorithm is presented. The dataset used in this approach includes many characteristics, including voice service, SMS, response time, latency, packet loss, download/upload speed, and line of sight, which are defined as follows: Throughput: it is total of information which is transported from one site to other at particular period, and is typically measured in such as megabits per second (Mbps), bits per second (bps), gigabits per second (Gbps).

A packet of data's latency is the time it takes to go from one place to another. In normal speech, network latency and lag are closely related. Good user experience (UX) is correlated with low latency, while poor UX is correlated with high latency. The objective is to get a latency that is as near to 0. The round-trip time (RTT) needed for data packet to go from its origin to destination as well as back can be used to assess network latency. At initial phase, the dataset is pre-processed. Rows with empty values are removed, the data is normalized, and various table features are extracted during this step. Furthermore, the data format needs to be changed to one that algorithms can recognize. Algorithms for balancing data are used to balance the data. The NB method uses data extraction and classification models to identify whether a client is a churner or not. Here is an example of the outcome analysis of the developed model. The attribute values and extracted features will be used to generate a confusion matrix, which will be used to evaluate how well the presented NB adequately assess client turnover in the telecom industry.

According to this model, out of two classes, one is “positive” or “churns/yes” and the other is “negative” or “not churns/no.” There are four potential results, which are classified as true negative (TN), true positive (TP), and false negative (FN), false positive (FP): total instances that were both accurately predicted as positive and really positive (true positive, or TP).

- TN: total number of events that are both actually negative and correctly predicted as negative.
- FP instances are all times when a result is actually negative but was predicted as positive.
- FN: total number of times that positive events are mistakenly detected as negative.

The accuracy, sensitivity, and precision of the recommended NB classifier's performance are evaluated.

Accuracy: accuracy is defined as the ratio of correctly predicted instances to total instances predicted and it is expressed as (2).

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \tag{2}$$

Where TP, TN, FP, and FN are confusion matrix parameters and are described above.

Sensitivity: the number of records accurately classified as positive out of its total positive records is counted. It is expressed in (3) as a percentage and is also known as the true positive rate (TPR) or recall.

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (3)$$

In (3), TP, FN are true positive and false negative instances.

Precision: precision measures the percentage of data with a positive categorization that are actually positive data and is expressed as (4).

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (4)$$

Where, TP means true positive and FP means false positive instances which are used to measure the performance of the classifier.

Table 1 shows performance evaluation of various approaches for predicting customer churn in telecommunication sector. The NB algorithm performs better than the DT classifier in terms of precision, recall and accuracy. When compared to the DT algorithm, the NB algorithm does better. Figure 2 shows comparison graph for precision. In Figure 2, y-axis shows the precision values conveyed as percentages, while x-axis represents different methods used for predicting customer churn. Compared to the DT method, the NB algorithm has higher precision.

Table 1. Performance metrics evaluation

| Churn predicting approaches | Accuracy (%) | Precision (%) | Sensitivity (%) |
|---|---|---|---|
| DT based churn prediction approach | 89.2% | 87.65% | 82.34% |
| Presented NB algorithm | 98.6% | 97.8% | 97.42% |

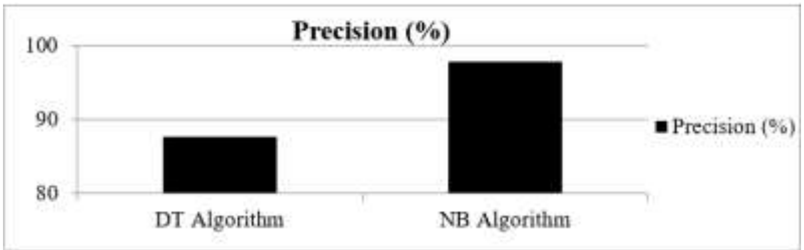


Figure 2. Comparative graph for precision

In Figure 3, y-axis shows the precision values conveyed as percentages, while x-axis represents different methods used for predicting customer churn. Compared to the DT method, the NB algorithm has higher sensitivity. The accuracy comparison between the DT and NB algorithms is shown in Figure 4.

In Figure 4, the y-axis shows accuracy as a percentage and the x-axis shows different algorithms for predicting consumer churn. The NB algorithm has high accuracy them the DT algorithm for predicting customer churn. Thus, compared to past methods, the NB algorithm that is being presented performs better and predicts churn in the telecom sector.

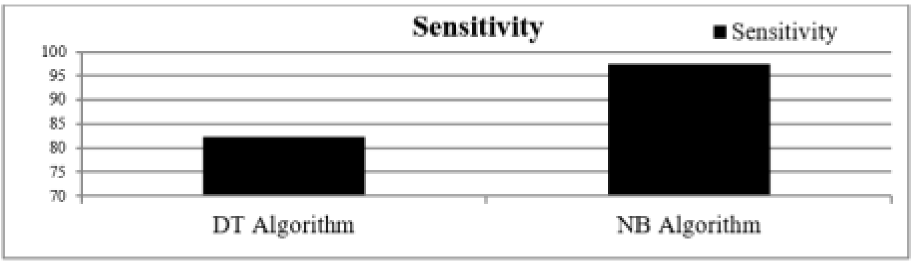


Figure 3. Comparative graph for sensitivity

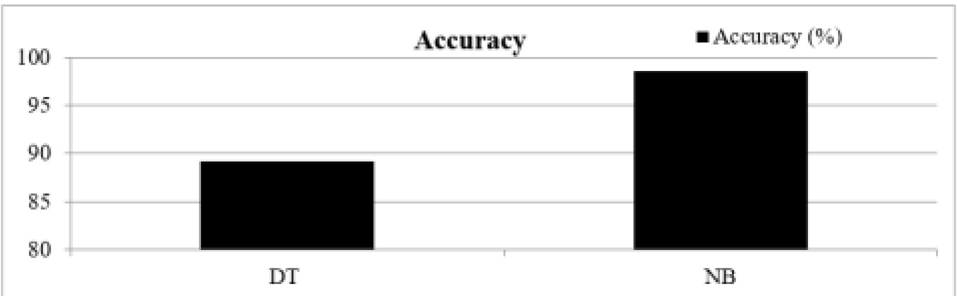


Figure 4. Accuracy comparative graph

## 5. CONCLUSION

In this work, predicting customer churn in telecommunication sector using NB algorithm is implemented. To retain its current customers, telecom service providers must be aware of causes churn. The knowledge needed to obtain this information that can be obtained from telecom data. Among the many features of this approach are latency, packet loss, throughput, voicemail, SMS, download/upload speed, line of sight and calls. First, the data is prepared by reducing the dimensions of the dataset's attributes and eliminating any unnecessary ones. Duplicate attributes are removed and special features are identified during the feature selection process. The data is balanced using three methods: SMOTE, Over sampling and under-sampling. The NB algorithm is applied to explain if particular customer is churner (i.e., whether or not it is going to remain with the business). The performance is evaluated in terms of accuracy, precision, and sensitivity. The NB algorithm has better results than earlier algorithms for customer churn prediction.

## REFERENCES


[1] S. Wu, W. C. Yau, T. S. Ong, and S. C. Chong, "Integrated churn prediction and customer segmentation framework for telco business," *IEEE Access*, vol. 9, pp. 62118–62136, 2021, doi: 10.1109/ACCESS.2021.3073776.

[2] N. Jajam, N. P. Challa, K. S. L. Prasanna, and C. H. V. S. Deepthi, "Arithmetic optimization with ensemble deep learning SBLSTM-RNN-IGSA model for customer churn prediction," *IEEE Access*, vol. 11, pp. 93111–93128, 2023, doi: 10.1109/ACCESS.2023.3304669.

[3] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," *Journal of Big Data*, vol. 6, no. 1, p. 28, Dec. 2019, doi: 10.1186/s40537-019-0191-6.

[4] V. C. Nwaogu and K. Dimililer, "Customer churn prediction for business intelligence using machine learning," in *HORA 2021 - 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*, Jun. 2021, pp. 1–7, doi: 10.1109/HORA52670.2021.9461303.

[5] S. De and P. Prabu, "A sampling-based stack framework for imbalanced learning in churn prediction," *IEEE Access*, vol. 10, pp. 68017–68028, 2022, doi: 10.1109/ACCESS.2022.3185227.

[6] K. Rothmeier, N. Pflanzl, J. A. Hullmann, and M. Preuss, "Prediction of player churn and disengagement based on user activity data of a freemium online strategy game," *IEEE Transactions on Games*, vol. 13, no. 1, pp. 78–88, Mar. 2021, doi: 10.1109/TG.2020.2992282.

[7] K. Ljubičić, A. Merćep, and Z. Kostanjčar, "Churn prediction methods based on mutual customer interdependence," *Journal of Computational Science*, vol. 67, p. 101940, Mar. 2023, doi: 10.1016/j.jocs.2022.101940.

[8] L. Calzada-Infante, M. Óskarsdóttir, and B. Baesens, "Evaluation of customer behavior with temporal centrality metrics for churn prediction of prepaid contracts," *Expert Systems with Applications*, vol. 160, p. 113553, Dec. 2020, doi: 10.1016/j.eswa.2020.113553.

[9] T. J. Shen, A. Samad, and B. Shibghatullah, "Customer churn prediction model for telecommunication industry," *Journal of Advances in Artificial Life Robotics* , vol. 3, no. 2, pp. 85–91, 2022, [Online]. Available: https://alife-robotics.org/jallr.html.

[10] Y. Suh, "Machine learning based customer churn prediction in home appliance rental business," *Journal of Big Data*, vol. 10, no. 1, p. 41, Apr. 2023, doi: 10.1186/s40537-023-00721-8.

[11] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, "A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector," *IEEE Access*, vol. 7, pp. 60134–60149, 2019, doi: 10.1109/ACCESS.2019.2914999.

[12] J. Kunnen, M. Duchateau, Z. Van Veldhoven, and J. Vanthienen, "Benchmarking stacking against other heterogeneous ensembles in telecom churn prediction," in *2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020*, Dec. 2020, pp. 1234–1240, doi: 10.1109/SSCI47803.2020.9308188.

[13] S. Venkatesh and M. Jeyakarthic, "An optimal genetic algorithm with support vector machine for cloud based customer churn prediction," in *2020 International Conference on System, Computation, Automation and Networking, ICSCAN 2020*, Jul. 2020, pp. 1–6, doi: 10.1109/ICSCAN49426.2020.9262443.

[14] P. Bhuse, A. Gandhi, P. Meswani, R. Muni, and N. Katre, "Machine learning based telecom-customer churn prediction," in *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems, ICISS 2020*, Dec. 2020, pp. 1297–1301, doi: 10.1109/ICISS49785.2020.9315951.

[15] A. Ben, "Enhanced churn prediction in the telecommunication industry," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3577712.

[16] V. Kavitha, G. H. Kumar, S. V M. Kumar, and M. Harish, "Churn prediction of customer in telecom industry using machine learning algorithms," *International Journal of Engineering Research and*, vol. V9, no. 05, May 2020, doi: 10.17577/ijertv9is050022.
[17] Z. Ullah, "LH-CAM: logic-based higher performance binary CAM architecture on FPGA," *IEEE Embedded Systems Letters*, vol. 9, no. 2, pp. 29–32, Jun. 2017, doi: 10.1109/LES.2017.2664378.
[18] A. Mishra and U. S. Reddy, "A comparative study of customer churn prediction in telecom industry using ensemble based classifiers," in *Proceedings of the International Conference on Inventive Computing and Informatics, ICICI 2017*, Nov. 2018, pp. 721–725, doi: 10.1109/ICICI.2017.8365230.
[19] G. E. Xia, H. Wang, and Y. Jiang, "Application of customer churn prediction based on weighted selective ensembles," in *2016 3rd International Conference on Systems and Informatics, ICSAI 2016*, Nov. 2017, pp. 513–519, doi: 10.1109/ICSAI.2016.7811009.
[20] P. K. Dalvi, S. K. Khandge, A. Deomore, A. Bankar, and V. A. Kanade, "Analysis of customer churn prediction in telecom industry using decision trees and logistic regression," in *2016 Symposium on Colossal Data Analysis and Networking, CDAN 2016*, Mar. 2016, pp. 1–4, doi: 10.1109/CDAN.2016.7570883.
[21] K. Kim, C. H. Jun, and J. Lee, "Improved churn prediction in telecommunication industry by analyzing a large network," *Expert Systems with Applications*, vol. 41, no. 15, pp. 6575–6584, Nov. 2014, doi: 10.1016/j.eswa.2014.05.014.
[22] A. Amin *et al.*, "Customer churn prediction in the telecommunication sector using a rough set approach," *Neurocomputing*, vol. 237, pp. 242–254, May 2017, doi: 10.1016/j.neucom.2016.12.009.
[23] M. Óskarsdóttir, C. Bravo, W. Verbeke, C. Sarraute, B. Baesens, and J. Vanthienen, "Social network analytics for churn prediction in telco: Model building, evaluation and network architecture," *Expert Systems with Applications*, vol. 85, pp. 204–220, Nov. 2017, doi: 10.1016/j.eswa.2017.05.028.
[24] K. Ebrah and S. Elnasir, "Churn prediction using machine learning and recommendations plans for telecoms," *Journal of Computer and Communications*, vol. 07, no. 11, pp. 33–53, 2019, doi: 10.4236/jcc.2019.711003.
[25] L. Saha, H. K. Tripathy, T. Gaber, H. El-Gohary, and E. S. M. El-kenawy, "Deep churn prediction method for telecommunication industry," *Sustainability (Switzerland)*, vol. 15, no. 5, p. 4543, Mar. 2023, doi: 10.3390/su15054543.


## BIOGRAPHIES OF AUTHORS

**Biswa Ranjan Agasti** completed B.Tech. at Gandhi Institute of Engineering and Technologies, Berhampur University, Odisha and M.Tech. at Keshav Memorial Inst of Technology JNTU, Hyderabad. At present he is Research Scholar at CSE Dept. in VFSTR University, Guntur, AP. He can be contacted at email: biswaranjanagasti112@gmail.com.

**Dr. Susanta Kumar Satpathy** completed B.Tech at Dr. B A Marthwada University, Maharashtra, M.Tech. at NIT, Rourkela and Ph.D. at CMJ University, Meghalaya. He has 27 Years of teaching experience. His research area is AI,ML and Image Processing. At present he is working as a Associate Professor in CSE Department at VFSTR University, Guntur, AP. He can be contacted at email: drsks_cse@vignan.ac.in.