Pastoralist Optimization Algorithm Approach For Improved Customer Churn Prediction in the Telecom Industry.

* Samuel, A. I[1], David, M[1], Salihu, B. A[1], Usman, A. U[1] & Abdullahi, I. M[2]

[1] Department of Telecommunication Engineering, Federal University of Technology, Minna, Niger State, Nigeria

[2] Department of Computer Engineering, Federal University of Technology, Minna, Niger State, Nigeria

*remmiest@gmail.com +2347032639739

ABSTRACT

In recent times, Telecom Industry customer churn has been a serious problem making it difficult to survive in the fierce competition of the industry. Survival in the industry and retention of the existing customers has become very important. Practitioners and academicians are now faced with the challenge of getting to predict likely customer churn, through predictive modeling techniques to predict potential customers who are likely to churn. Customer churn affects the revenue of the company because it cost more to acquire new customer than retaining old ones. When a company allocates its dedicated resources to retain these customers, it greatly controls the rate at which dissatisfied customers leave the company. Several techniques have been studied and we present an overview of resent works on churn prediction .Our work uses Artificial Neural Network approach for prediction of customers intending to switch over to other operators. This study uses Pastoralist Optimization Algorithm (POA) to enhance the Artificial Neural Network (ANN) by working on multiple attributes from Telecom Company's dataset with sample results. The results obtained showed that the proposed POA algorithm selected fewer attributes of ten out of fifteen for the telecom churn prediction and had a prediction accuracy of 97.0% compared to the ordinary unenhanced ANN which used the entire 15 attributes but had a prediction accuracy of 93.6%

Keywords: ANN; churn management, Neural Network, Pastoralist, prediction;

1 INTRODUCTION

In today's technologically advanced and congested marketplaces, Customers may more easily compare options and select the best one from a large number of service providers. So, businesses must not only win clients' first favor but also keep them once they have them by retaining their current clientele. In the context of the telecom sector, "churn" is the proportion of subscribers who switch from one service provider or another within a certain period of time. All industries suffer with "Churn" (Plaksij, 2022). Companies therefore need to work very hard to survive in the competitive market by generating new policies and strategies for acquiring new customers. Most organizations today like Satellite TV, telephone and internet providers heavily rely on keeping their existing customer base because it generate more income to the companies (Sandhya, 2021). Churn is a major problem for many businesses since it reveals how successful they are at retaining consumers. If companies lose their customers it can affect its finance and revenue. Businesses lose $1.6 trillion annually as a result of client attrition. (Plaksij, 2022).

Figure 1: Illustration from Forrester showing cost of attracting a new customer Plaksij (2022).

Figure 1 above isan illustration by the Forrester study demonstrating that acquiring new consumers is 5 times more expensive than retaining existing ones.. It has been demonstrated that a company's revenue increases with the number of clients it keeps. Referencing a study from the Harvard Business School that states that, on average, a 5% improvement in customer retention rates leads to a 25%–95% boost in earnings. Also, 65% of a business's revenue comes from existing clients. According to the research, keeping current customers happy is just as crucial as finding new ones, if not more so. After several studies it can be stated that Several factors influence customer churn in Telecom industry, some of which are; network coverage, tariff, voice quality, customer service, innovation e.t.c. Some of these factors are summarized in Figure 2.

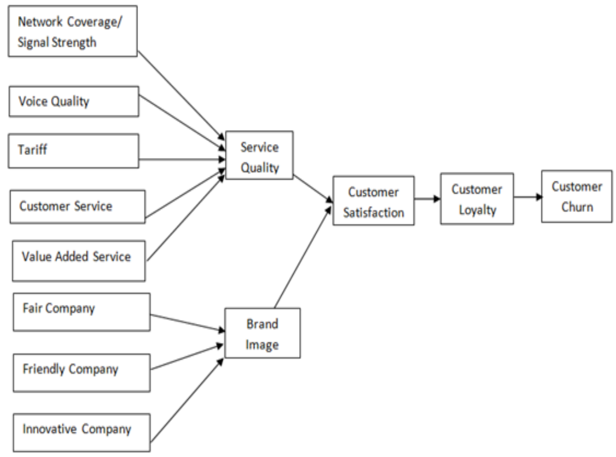

Figure 2: Factors influencing customer churn in Telecom industry Mahajan *et al.* (2017)

The figure 2 above shows the different factors affecting customer churn in telecom industry ranging from network coverage to voice quality, tariff, customer service, value added service which can all be categorized into service quality. Other factors include fairness, friendliness and innovation which are also broadly categorized as brand image. It also shows that a combination of service quality and brand image determines customer satisfaction. Customer satisfaction determines customer loyalty and customer loyalty determines customer churn.

Therefore,spending a lot of money on marketing and cam paigns to draw in new customers is not as profitable as co ncentrating on longterm relationships and tracking custom er behavior. Churn control is crucial since it has an impact on the profitability of a business. Additionally, the more you grasp your customer and their conduct, the more you'll comprehend the predicted revenue in the future. According to Marketing Metrics, the likelihood of selling to an existing customer is 60–70%, while the likelihood of selling to a new prospect is only 2-5%. So, it seems sensible that top priority should be given to lowering the rate of client attrition because keeping existing customers is profitable! But too many companies don't understand this and continue to struggle to implement a successful churn prevention tactics. It has become important for academicians and Telecom industries through predictive modeling techniques, to help policy makers come up with successful churn prevention or customer retention strategies for profit maximization

The remaining aspect of this paper is section as follows: s*ection 2* provides a few reviewed literatures, An overview of the approach is provided in ***Section 3***, the results are presented and discussed in ***Section 4***, and the conclusion is made in ***Section 5***.

2. LITERATURE REVIEW

The challenge of predicting customer churn in the telecom industry has been addressed using a variety of techniques. The majority of this field's research focuses on using various Machine Learning algorithms to create predictions and compare the outcomes subsequently. While some academics have gone as far as to create fresh algorithms to apply to this problem, others use current methods with improved parameter tweaking. Below are some of the existing predictive models as used by various authors:

In (Ngurah *et al.*, 2020). IBM customer dataset, which comprised 7040 rows and 21 fields, Deep Neural Network (DNN), Random Forest, and Extreme Gradient Boosting (XGBoost) techniques were proposed to address the issue of customer churn. The DNN model's structure is given, and it is compared to previously researched methods like Random Forest and Extreme Gradient Boosting (XGboost). It was discovered that although techniques like Random Forest and XGBoost had processing times of 529 seconds and 175 seconds, respectively, and accuracy of 77.87% and 76.45%, respectively, the suggested DNN model had a perdiction accuracy of 80.62%. In (Ahmad *et al.*, 2019). A study on the prediction of customer churn was carried out utilizing the Spark environment to create and test models using data from SyriaTel Telecom. In their work, they discussed how to prepare features for machine learning algorithms through feature engineering, feature transformation, and a selective method. The four tree-based algorithms used were the XGBoost method, Gradient Boosted Machine Tree (GBM), Random Forest, and Decision Tree. With an AUC score of 93.301%, the XGBoost model outperformed the others, followed by the GBM algorithm in second place, Random Forest in third place, and Decision Tree in fourth place. The authors discovered that applying Social Network Analysis elements improved their ability to forecast churn in the telecom industry. In (Gaur and Dubey, 2018). The researchers applied a variety of classification methods, including Logistic Regression, Support Vector Machines (SVM), Random Forest, and Gradient Boosted Tree, and then compared the results. The authors conducted a study of prior research, and they have included flowcharts for the Churn Prediction Framework and Analysis Stages in addition to information on the steps involved in developing the model. Gradient Boosting outperformed the other three models in the study, while the results from Logistic Regression and Random Forest were around average. SVM fared the worst. To assess the accuracy of their model, they employed the AUC (Area under the ROC Curve) statistic. In (Ullah *et al*., 2019). Research was conducted with the aim of resolving the issue of predicting customer attrition. The dataset underwent feature selection and noise removal during preprocessing. An 88% accuracy rate for machine learning algorithms like Random Forest was gotton. By highlighting the key dataset variables that are utilized to forecast customer churn, clustering techniques like K-Means have been applied. The "receiving operating characteristics" (ROC) area was used, along with metrics including accuracy, precision, recall, and f-measure. Using the RF method and customer profiling by K-means clustering, the findings demonstrated that their suggested churn prediction algorithm achieved better churn categorization. In (Idris *et al*., 2012). The problem of customer attrition was modelled using a Genetic Programming (GP)-based technique. Individual classifiers were thought to perform worse than an ensemble approach. Random Forest ensemble would be problematic because their dataset is unbalanced and contains fewer instances of the minority class. The flexibility and unique properties provided by GP, however, make it more ideal for categorization and subsequently churn prediction. Adaboost, on the other hand, is a boosting strategy that combines several weak classifiers to produce a powerful one. The "Orange Telecom" dataset and the "Cell2Cell" dataset were used in this instance. Here, the performance of the predictors is assessed using the area under the curve (AUC), sensitivity, and specificity metrics. On the Cell2Cell dataset, the highest churn prediction accuracy of 89% AUC is reported. In (Joolfoo and Jugumauth, 2020). A hybrid approach for predicting telecom churn, ANNs (Artificial Neural Networks) and KNNs (K-Nearest Neighbors) were employed. A comparison of earlier works has been made, and several prior methodologies have been reviewed. This research covered a total of 15 papers between 2014 and 2020. using the use of Artificial Neural Networks and the KNN machine learning model, they have attempted to develop an innovative and hybrid technique. The suggested performance indicator for this model was accuracy.

In (Abrandusoiu *et al*., 2016). The authors applied advanced data mining methodology for churn prediction using a dataset of 3333 customer call details with 21 attributes and a yes/no as a dependent parameter. The customer dataset includes information on incoming, outgoing, and voicemail calls. We saw that the author had combined Bayes Network, Support Vector Machine, and Neural Network with PCA (principal component analysis) to determine the algorithm's performance, the accuracy of the Bayes Network, neural network, and support vector machine utilized by the author was 99.10%, 99.55%, and 99.70%, respectively. In (Khan *et al*., 2015). The issue of predicting client attrition was investigated in big data platforms. Their objective was to demonstrate how big data significantly improves the process of churn prediction based on the amount, diversity, and velocity of the data. China's largest telecoms firm needed a big data platform to engineer the fissures in the data coming from the Operation Support department and the Business Support department. They applied the Random Forest algorithm and utilized AUC to assess its performance. There were 5 million active consumers at one point, and the algorithm was able to produce a draft of prepaid customers who will cancel their subscriptions in the following month with an accuracy of 0.96 for the top 50,000 customers. In (Mohanty and Rani, 2015). A work titled "Behavioral Modeling for Churn Prediction," was presented and researchers looked at the early signs of churn and created a churn score so that businesses could identify clients who were going to discontinue services. The authors moved closer to feature engineering using brute force, which might result in a significant number of features from customer data, including calls and logs, that are overlapping with one another, then determines the characteristics and indicators that are most predictive of customer attrition using two related methodologies. Features are fed into a number of supervised learning algorithms in order to forecast subscriber attrition. The method was tested using terabytes of data from a South Asian mobile phone operator, in 83.9% of cases, the authors' prediction was accurate when they classified subscribers who were idle on more than 76 percent of days during the training period as churners. As 76.6% of their sample did not churn, they had an unbalanced sample; therefore, a very simple model that just predicted the majority class (indicating "not churn") for all customers would achieve 76.6% accuracy. This simple linear discriminant performs well in this situation. They achieved accuracy rates of approximately 88.5-89.5% depending on the method employed to anticipate churn. Although utilizing a single method to forecast Churn can yield positive results, it is frequently seen that using numerous or a combination of these algorithms is the key to achieving excellent outcomes.

In (Höppner *et al*., 2020). A decision tree was utilized, and he claimed that in a benchmark analysis utilizing actual data from different telecommunication service providers, ProfTree outperforms more conventional accuracy-driven tree-based approaches. The evolutionary algorithm (EA) training times for this technique are slower than those for other classification algorithms. In (Hu *et al.,* 2020). Decision tree, a neural network, and a combination of the two were employed in a study on the prediction of customer turnover. After preprocessing, a total of 2681 entries from a supermarket's customer database between June 2018 and April 2019 were used. The decision tree model's prediction accuracy was

93.47%, while the neural network model's prediction accuracy was 96.42%. The combined model accuracy was found to be 98.87% after removing 21 customers whose churn probability is between 0.4 to 0.5. It shows that high accuracy is difficult to achieve by a single technique.

In Summary, most of the technique presented in the literature either single or hybrid techniques for customer churn prediction all tried to solving the problem of customer churn by proposing predictive models giving relatively high accuracy but none of the papers talked about finding the most important attributes within the dataset. In contrast, this paper contributes to the literature from different perspectives. It focused on finding the most important attributes from a given dataset, what to look out for in a customer and using these attributes to make churn prediction with very high accuracy. Pastoralist Optimization Algorithm used in this work was developed by (Abdullahi *et al.*, 2018) The pastoralists' herding techniques served as inspiration for the algorithm. Scouting, camp selection, camping, herd splitting, and merging are the strategies. The POA was created using mathematical models of these tactics. 10 unimodal and multimodal test benchmark functions were used to test the algorithm's performance. This is done to evaluate the algorithm's exploratory, exploitative, convergence, and escape from a local optimum solution capabilities. In order to determine the statistical significance level of the algorithm outcomes, a nonparametric statistical test (Wilcoxon rank sum tests) was also performed. The experimental findings produced with the method demonstrate that the algorithm outperformed some state-of-the-art, nature-inspired metaheuristic optimization algorithms in the majority of circumstances.

3. METHODOLOGY

This section introduces the ideas and context used as well as the technical advances made by this article. This phenomenon involved processing starting with attribute selection and ending with categorization or prediction. Effective categorization or prediction must meet three key conditions. They include having access to a lot of data to train the model, having a powerful enough machine, and using an advanced algorithm to speed up and improve the process.

3.1 DATASET DESCRIPTION

Each sample or row of data, representing a client and containing attributes for 15 columns, totaled 3150. Call failures, SMS frequency, number of complaints, number of unique calls, subscription term, age group, fee amount, service type, usage duration in seconds, status, and SMS frequency are the attributes included in this dataset. The dataset underwent feature selection and noise removal during preprocessing. The ANN is trained using the back-propagation learning algorithm, and synaptic weights were modified using gradient descent to reduce error using the transformation function. In the training mode, extra care was taken to avoid overtraining the model and to keep error levels as low as feasible. All of the covariates were normalized with values between 0 and 1 prior to the start of training. The dataset is partitioned into two sets at random: Test set 945 customers and train set 2205 customers, respectively, make up 30% and 70% of the total dataset, respectively.

3.2 MODEL DESCRIPTION

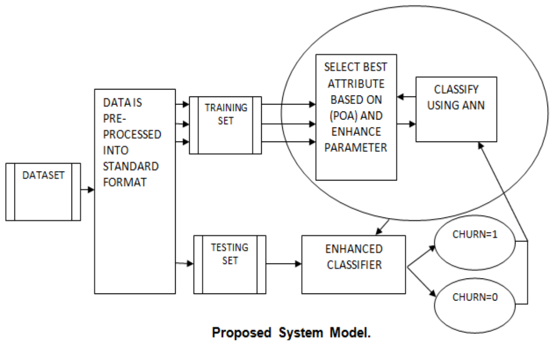

Figure 3: System Model.

In the model shown in Figure 3 above, the POA attribute selector initializes all its parameters and selects the scout pastoralists randomly, the scout location is initialized and the fitness of each scout is evaluated and the location is updated. The scouts location within the search space are also normalized until the maximum scouting rate is reached the best camping location which represents the best attribute for prediction is then selected using the fitness function and then the classification is done with the same POA to get the best pastoralist where an output of 1 represents a subscriber who is likely to churn and 0 represents a subscriber who is not likely to churn.

In this section, we describe the POA algorithm according to Abdullahi *et al*. (2018) which is the basis technique for use in our work. Each step of the algorithm is described as follows:

i. Start
ii. Initialize all POA parameters
iii. Select scout pastoralist randomly from number of Pastoralists and initialize scout location
iv. Evaluate the fitness of each scout, update scout Locations and normalize scouts' locations within the Search space until maximum scouting rate is reached
v. Select best camping location based and move pastoralist and herds to camp.
vi. Evaluate fitness of pastoralist and determine best Pastoralist within a camp Pbest
vii. Split pastoralist to different locations within camp and Evaluate fitness of each pastoralist
viii. Repeat step vii until maximum splitting rate is reached. For each split, divide the current camp radius by the number of pastoralist
ix. Update the best camp pastoralist Cbest
x. If all regions within the search space have not been explored (maximum iteration not reached), update scout location and repeat steps iv to ix and update the global camp best pastoralist Gcbest.
xi. Else, return the global best-found pastoralist Gbest,
xii. Stop

Figure 6: Pastoralist Optimization Algorithm (Abdullahi *et al*., 2018)

4 RESULTS

The main purpose of this research was to develop a model that can first determine the best attributes for the prediction of telecom churn and then make prediction based on these attributes by processing of data obtained from telecom industry. In this section, the paper presents the confusion matrix, the selected CDR attributes for customer churn prediction, optimal prediction values and a comparative analysis of optimal POA values with un optimized values . The confusion matrix is employed because it provides insight into how well categorization models perform given a certain set of test data. The only way to know is if test data's real values are knownThe matrix itself is simple to comprehend, but some of the terminology used in connection with it could be unclear. The projected values and actual values, along with the total number of outcomes, are separated into two dimensions in the matrix. Actual values are the real values for the provided data, whereas predicted values are the values that the model predicts. Table 1: Attributes selected for customer churn prediction.

S/N	ATTRIBUTE	POA_ANN	ANN
1	Call Failure	*	*
2	Complain	*	*
3	Subscription Length	*	*
4	Charge Amount	*	*
5	Seconds of Use	*	*
6	Frequency of Use	*	*
7	Frequency of Sms	*	*
8	Distinct call numbrs	*	*
9	Age Group		*
10	Tariff Plan		*
11	Status	*	*
12	Age	*	*
13	Customer Value		*
14	FN		*
15	FP		*

The table 1 above also clearly shows the attributes selected for churn prediction and this was done by an evaluation of each attribute using the fitness function. It was noticed from the table that while ANN used the entire fifteen attributes from the given dataset for prediction, POA_ANN selected fewer attributes and gave a higher prediction accuracy. The attributes selected by POA_ANN and used for the prediction were Call Failure, Complain, Subscription Length, Charge Amount, Seconds of Use, Frequency of Use, Frequency of Sms, Distinct call numbers, status, age and FN.

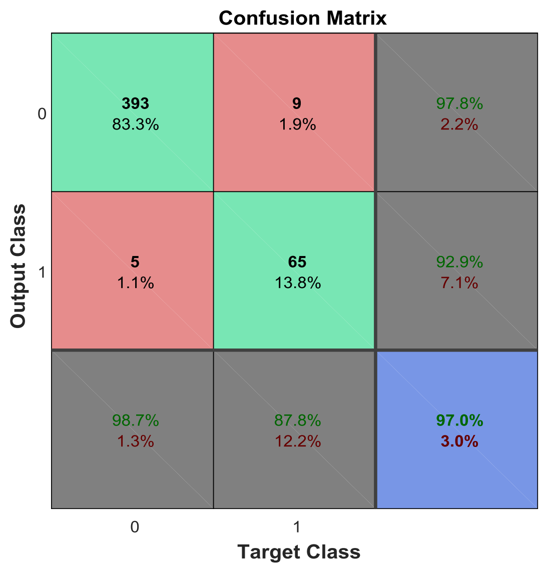

Figure 4: ANN_POA Confusion Matrix

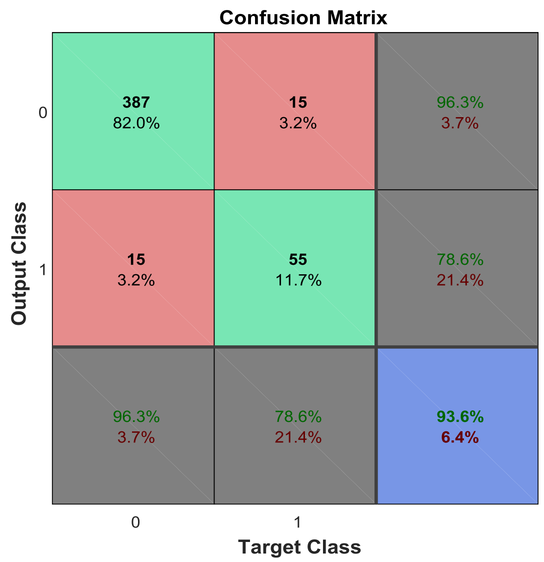

Figure 5: ANN Confusion Matrix

The confusion matrix can be used to calculate the model's properties, including accuracy, precision, sensitivity, etc. In this work, prediction accuracy is used as metric for performance evaluation for the model. Accuracy is used as the measure of classification performance because it is simple to compute among others and easy to interpret.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Where:

T_P is True positive rate

T_N is True Negative rate

F_P is false positive rate

F_N is False negative rate

Figure 4 shows the confusion matrix of POA_ANN and the area with 65 samples colored green indicates the true positive rate, which represents the actual churners correctly classified. The area with 393 samples colored green indicates the true negative rate, which represents the non churners that were correctly classified. The area with 5 samples colored red indicates the false negative rate representing the non churners incorrectly classified, the area with 9 samples colored red indicates the false positive rate representing actual churners incorrectly classified. Finally, the blue area with 97.0% represents the percentage accuracy of the technique.

Figure 5 shows the confusion matrix of ANN and the area with 55 samples colored green indicates the true positive rate, which represents the actual churners correctly classified. The area with 387 samples colored green indicates the true negative rate, which represents the non churners that were correctly classified. The area with 15 samples (row 2, column 1) colored red indicates the false negative rate representing the non churners incorrectly classified, the area with 15 samples (row 1, column 2) colored red indicates the false positive rate representing actual churners incorrectly classified. Finally, the blue area with 93.6% represents the percentage accuracy of the technique.

5 CONCLUSION

In this paper, an optimized subscriber churn prediction algorithm for the telecommunication industry is presented and its performance evaluated by comparing with unoptimized ANN technique in predicting churn. The study computed their true positive rates (TPR), true negative rates (TNR), false positive rates (FPR), false negative rates (FNR) and their prediction accuracies. The outcome showed that the POA_ANN has a high degree of prediction accuracy as it outperformed the prediction of ordinary ANN. POA_ANN only selected 10 attributes from the dataset to give an accuracy of 97.0% while the ordinary ANN selected all 15 attributes to give a prediction accuracy of 93.6% giving a 4.27% improvement in accuracy and a 50% improvement in attribute selection. This clearly shows the most important attributes for churn prediction in the Telecom industry based on the dataset. The paper also recommends that more evaluation of the performance of the algorithm be done by making comparism with other algorithms such as Particle Swarm Optimization (PSO), Cuckoo Search Optimization (CSO), Ant Colony Optimization (ACO) and Genetic algorithm (GA).

REFERENCES

Abdullahi I. M. Mu'azu M. B. Olaniyi O. M. & Agajo J. Pastoralist Optimization Algorithm (POA): A Novel Nature-Inspired Metaheuristic Optimization *Proc. of Int. Conf. on Global and Emerging Trends, (ICGET) 2018,* Base University, Abuja, Nigeria, May 2nd 2018. *pp. 101-105.*

Abrandusoiu, Ionut, Toderean, G. & Beleiu, Horia. (2016). Methods for Churn Prediction in the Prepaid MobileTelecommunicationsIndustry.97-100.10.1109/ ICComm. 2016.7528311.

Ahmad, A.K., Jafar, A. & Aljoumaa, K. Customer churn prediction in telecom using machine learning in big dataplatforms.*JBigData*6,28(2019). https://doi.org/10.1186/s40537-019-0191-6.

Chakure A. Belkhede A. Shrimali S. Garud S. & Patil S. An Overview Of Customer Churn Prediction In Telecom Industry *A Literature Survey,* Journal of Emerging Technologies and Innovative Research (JETIR) 2021 JETIR March 2021, Volume 8, Issue 3 pp. 2227-2230. 2021.

Eria K, Marikannan BP: Systematic Review of Customer Churn Prediction in the Telecom Sector. ResearchGate. 2018, September

Gaur A & Dubey R, "Predicting Customer Churn Prediction In Telecom Sector Using Various Machine Learning Techniques," 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 2018, pp. 1-5, doi: 10.1109/ICACAT.2018.8933783.

Höppner S, Stripling E & Baesens B: Profit-driven decision trees for churn prediction. Eur. J. Oper. Res. 2020; 284

Hu X. Yang Y. Chen L. & Zhu S. "Research on a Customer Churn Combination Prediction Model Based on Decision Tree and Neural Network," 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), Chengdu, China, 2020, pp. 129-132, doi: 10.1109/ICCCBDA49378.2020.9095611.

Idris A. Khan A. and Lee Y. S. "Genetic Programming and Adaboosting based churn prediction for Telecom," 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Seoul, Korea (South), 2012, pp. 1328-1332, doi: 10.1109/ICSMC.2012.6377917

Joolfoo M. B. A. & Jugumauth R. A. "A Systematic Review of Algorithms applied for Telecom Churn Prediction," 2020 3rd International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM), Balaclava, Mauritius,2020,pp.136-140,doi:10.1109/ELECOM49001. 2020 .9296999.

Khan, M., Manoj, J., Singh, A., & Blumenstock, J. (2015). Behavioral Modeling for Churn Prediction: Early Indicators and Accurate Predictors of Custom Defection and Loyalty*2015 IEEE International Congress on Big Data*.

Mahajan V. Misra R. & Mahajan R. Review on factors affecting customer churn in telecom sector. International Journal of Data Analysis Techniques and Strategies 9(2):122 DOI :10.1504 /IJDATS . 2017.085898, pp. 123-144. 2017

Mohanty R. & Rani K. J. "Application of Computational Intelligence to Predict Churn and Non-Churn of Customers in Indian Telecommunication," 2015 International Conference on Computational Intelligence and Communication Networks (CICN), Jabalpur, India, 2015, pp. 598-603, doi: 10.1109/CICN.2015.123.

Ngurah P, Oka H & Setyo A, "Telecommunication Service Subscriber Churn Likelihood Prediction Analysis Using Diverse Machine Learning Model," 2020 3rd International Conference on Mechanical, Electronics Computer, and Industrial Technology (MECnIT),Medan,Indonesia,2020,pp.24-29,doi: 10.1109/MECnIT48290.2020.9166584.

PLAKSIJ Z. CUSTOMER CHURN: 12 WAYS TO STOP CHURN IMMEDIATELY LAST UPDATED: 15 AUGUST, 2022 [ONLINE].AVAILABLE:HTTPS://WWW.SUPEROFFICE.COM/BLOG/REDUCE-CUSTOMER-CHURN/

Sandhya k Thaslima S Prasanna N.G.L. Vindhya R. and Srilakshmi P.. Analysis of Customer Churn Prediction in Telecom Industry Using Logistic Regression.
International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume-9, Issue-4, July 2021

SWIFTADMIN. HOW TO STOP CHURN IMMEDIATELY AUG11,2021[ONLINE].AVAILABLE:HTTPS://WWW.SWIFTERM.COM/HOW-TO-STOP-CHURN-IMMEDIATELY/

Ullah I. Raza B. Malik A. K. Imran M. Islam S. U. and Kim S. W. "A Churn Prediction Model Using Random Forest:Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," in IEEE Access, vol. 7, pp. 60134-60149, 2019, doi: 10.1109/ACCESS.2019.2914999.