Hybrid ML Classification Approach for Customer Churn Prediction in Telecom Industry

[1]Biswa Ranjan Agasti, [2]Dr. Susanta Satpathy
1Ph.D. Scholar, Dept of CSE, Vignan's Foundation for Science, Technology & Research, Guntur, Andhra Pradesh, India
2Associate Professor, CSE Dept, Vignan's Foundation for Science, Technology & Research, Guntur, Andhra Pradesh, India

Article Info
Page Number: 10359-10368
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: *12 August 2022*
Revised: *16 September 2022*
Accepted: *20 October 2022*
Publication: *25 November 2022*

Abstract - The rate at which customers abandon a product or service is known as the churn rate. The telecom industry must identify churn-risk customers in order to retain existing customers and maintain a higher competitive advantage. One of the main problems and major concerns for big companies is customer churn. Companies seeking to develop means to predict future customer churn especially in the telecom business because it directly affects their revenues. Customer lifetime value is greatly affected by churn rate because it affects the both company's future revenue and the length of service. Many companies have used different methods to predict churn rates and allow the development of effective customer retention plans because it cost much higher to get a new client rather than to retain an existing one. In order to predict customer churn in the telecom industry a Hybrid ML classification strategy is described in this research. Machine learning techniques are used to create the model in this paper. Customers who are likely to cancel their subscription can be predicted using machine learning algorithms. To predict churn, a combination of Support Vector Machine (SVM) and Naive Bayes (NB) algorithms is utilised. The main contribution of this work is the development of a churn prediction model that helps telecom operators to identify customers who are most likely to experience churn and better churn accuracy of the prediction.
Keywords – Telecommunication, Churn Prediction, customer, Machine Learning

I. INTRODUCTION

One of the main industries in developed countries is now telecommunications. The level of competition has increased due to technical progress and an improve in operators. Customers are the most important source of revenue for every company. In a telecoms company churn describes a person who has opted to leave a company's service. The churn model defines the individual who is most probably intending to set the business soon. Churn is the momentary movement of clients from one service provider to another for a variety of reasons, which lowers the revenue for the first service provider. Consumers have taken advantage of their choices to change to better and less cheap products as a result of greater competition in the telecom industry. The telecom industry observes a lot of customers leaving the company's services. Depending on the industry either the client voluntarily decides to stop using the services because they are disappointed with the current company plans and changes job and moving to a different location where they can no longer access the services, leaving the country, getting a job that requires them to use a specific service provider, or the service provider involuntarily terminates their services for reasons such as lack of payment of bills or the use of the services for illegal activities. [1]. Customer churn is a major concern in the service sector because products and services are very competitive. On the other hand the clients most likely to leave the company could represent in a large additional revenue source. Customer retention in telecommunication is a major concern in Customer Relationship Management (CRM). As it is difficult to acquire new consumers the primary focus of CRM is

on existing customers. In this competitive market companies are pulling hard to survive and depend on complex methods. Customer churn grows into a significant problem and causes a significant loss of telecom services [4].
In the competitive telecom industry, where customers may easily transfer providers, telecom providers are worried about their clients and how to retain them but by analyzing their behavior who will move to another provider previously. They can keep them by providing offers and services based on their past achievement thus the study's goal is to predict churn and determine the primary causes that may lead the user to switch to another telecoms provider. [10].
From service providers, every customer expects excellent service or reward points. It is more difficult to provide prompt services to valid customers because it is difficult to predict the company's genuine customers. By predicting customer behaviour, early churn prediction can save a company. As a common use for data mining methods, many industries create a model similar to churn. Every mobile phone provider in the world is near to developing their own churn model. Additionally, churn results can be effectively used for a variety of different purposes in order to retain customers [9]. To classify customers into churn and non-churners is the main goal of churn prediction.
Various techniques, like as data mining, machine learning and hybrid technologies, have been used to predict customer churn. These techniques enable enterprises to identify,predict and continue to churn customers. Additionally they help companies with CRM and decision-making. For consumer prediction and historical research data mining techniques are occasionally employed. The process of learning from data and then making judgements and predictions is described by the term "machine learning" on the other hand. [3]. When there is a large amount of data that is highly comparable and a prediction or insight is expected so that the machine learning works best. Previous research has shown how well machine learning technology predicts this situation. Hence in this work, hybrid ML classification approach for churn prediction in telecom industry is presented.
The proposed method which combines SVM and NB will make single model predictions that include customer views from many companies and it was important for the most profitable customer segment for many telecom companies when it tends to the outright irregularity issue.
The remainder of the work is structured as follows: In Section II, the literature search is covered. The hybrid ML classification strategy for churn prediction in the telecom industry is illustrated in Section III. The section IV describes the result analysis of presented approach. Finally the work is concluded in section V.

II. LITEARTURE SURVEY

Darin Ramadhanti, Aisyah Larasat, Abdul MuidYuh Wen Chen, et. al., [2] presents Deep Learning ANN Model Optimization for Customer Churn Prediction. The goal of this study is to create a classification model for predicting customer churn rates based on an improved deep learning ANN algorithm. The adaptability of ANN is beneficial. The accuracy value for the ANN model is 76.35%, according to the findings. The model uses tanh as the activation function and has an epoch parameter of 30 and a hidden layer of 50.the type of services, IPTV and contract type are three most influential variables at PT. XYZ in customer churn. 4386 consumers are likely to be loyal according to the prediction findings from the optimised deep learning-ANN model whereas 2567 customers are more likely to leave.
Aayushi Gandhi, Pushkar Bhuse, Neha Katre , Parth Meswani, Riya Muni,et. al., [7] Prediction of Telecom-Customer Churn Using Machine Learning is provided. A critical predictor of lengthy strategic success or failure is churn prediction. In this study, methods for predicting telecom customer churn using machine learning and deep learning are examined.

To predict if a customer will depart, we compare commonly used techniques such as Random Forest Classifiers and SVMs to more recent architectures such as XGBoost and Deep Neural Networks. To further evaluate their efficiency, these models are put to a grid search. The experiment's findings show that the Random Forest model performs well in this particular use case with a prediction accuracy of 90.96% on the testing data acquired before to grid search.

Malak Fraihat, Mahmoud Almomani, Ahmad Hammoudeh et. al., [11] The Selective Ensemble Model for Predicting Telecom Churn is presented. The Selective Ensemble Model (SEM) is presented in this study as an effective tool for predicting churn. SEM dynamically selects a set of machine learning models from which to form the final conclusion. According to experiment results SEM outperforms both the averaging ensemble model and its component models.

Aditya Das, Sanket Agrawal, Sudhir Dhage, Amit Gaikwad et. al., [15] Deep Learning is used to predict customer churn based on behavioural patterns analysis. This paper uses a Deep Learning Approach to predict churn on a Telco dataset. To create a non-linear classification model a multilayered Neural Network was created. Customer features that support and make use of features are used by the churn prediction model together with contextual features. Both the determining factors and the risk of churn are predicted. The trained model then predicts client churn by integrating these features with the final weights. The accuracy rate was 80.03%. Companies may use the model to look into the causes of these problems and take steps to dispose of them since it also contains churn variables.

Abinash Mishra, U. Srinivasulu Reddy et. al., [18] In order to explain a novel technique of churn prediction, deep learning is applied. Deep learning using convolutional neural networks (CNN) works well in terms of accuracy for churn prediction. The dataset is separated into training and test sets using a (60:40) ratio with 40% of the data used to test and 60% of the data used to train the mode when it is ready to make predictions. To categorize customers into non-churners and churners is the main goal of churn prediction. The predictive model for churn prediction outperforms previous models based on experimental data, with an rate error of 13.15%, accuracy of 86.85%, recall of 93.08%, F-score of 92.06% and precision of 91.08.

Lam Hai Shuan, Tan Yi Fei, Guo Xiaoning, Lai Jie Yan ,Soo Wooi King et. al., [19] Describes Discretization and Naive Bayes Classifier for Customer Churn Prediction in the Telecommunications Sector. Two experiments are conducted using the data processing methods K-Means and Equal-Width Discretization (EWD) in conjunction with Naive Bayes to compare the techniques used to identify probable churn activities. The data produced is often very large which has several dimensions. The experimental results shows that this can improve the model in identifying the key factors in churn prediction by using the correlation between attributes.

III. HYBRID ML CLASSIFICATION APPROACH FOR CUSTOMER CHURN PREDICTION

This section will present a hybrid machine learning classification method for churn prediction in the telecom industry. The block diagram of a hybrid machine learning (ML) classification approach for predicting customer churn is shown in Fig. 1. Telco customer churn data set is used in this analysis. Each column has a set of customer attributes that are described in the column metadata, and each row represents a customer. The raw data consists of 21 columns and 7043(customers) rows (features).The data set includes columns called churn that provide information about customers who have left within the past month. Each customer has signed up for the following services such as numerous lines of communication, online backup, Internet, phone, technical support, device protection ,movies and streaming TV.

Information on the customer's account, including the length of the customer's contract, the method of payment, if paperless billing is being used and the total amount charged. Demographic info about customers' gender, age and if they have partners or dependents. A number of circumstances including attributes with missing data, attributes with no values and attributes with values have been addressed in the pre-processing of the dataset. For each customer, a representation of each service and item is available. Missing values may occur because not all consumers have the same subscription. Some of them may provide a number of different services while others may have something completely different. Other columns are associated with system configurations, and the values for these columns are all null for each customer.
Dataset is a collection of feathers and N number of rows. There are numerous formats for many values. A dataset could have duplicate or null values which would cause some loss in accuracy and dependability. To represent a single value, such as gender, which is indicated by M/F or Male/Female, each data source uses a different format. An image in 3-dimension data should be converted to a 2-dimension format like to data show since the machine can only understand 0 and 1 in order to prevent noisy data, null values and incorrect size.

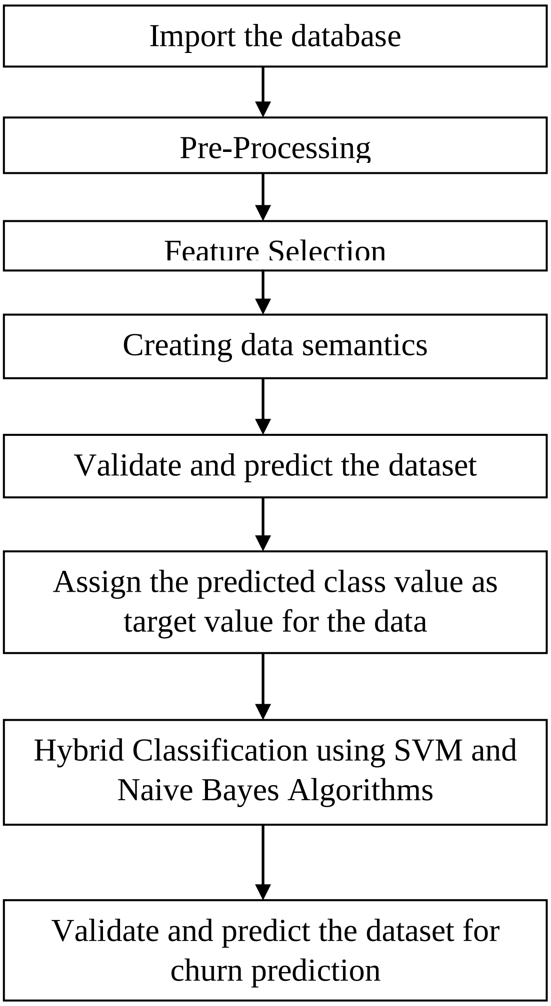

Fig. 1: Block diagram of hybrid ML classification approach for customer Churn Prediction

Making the data useful is crucial because unwanted or null values may result in results that are less accurate or unsatisfactory. The data collection includes a missing values and incorrect for small number. Only the most helpful features are listed after the original dataset has been

analysed. The list of features only covers functional features and can help with accuracy. In order to choose the necessary components from the data set based on experience.
From the data set the required elements are selected that is known as the Feature selection which is the crucial step. The dataset used in this study has a number of features but only the required features have been chosen. These features help us improve performance measurement and are helpful for making decisions while the remaining are given less importance. If the dataset contains only valuable and highly predictable variables the classification performance improves. Therefore improving classification performance requires having fewer irrelevant attributes and more significant features. As input to validate and predict the dataset are feed to develop the data semantics
After the validation and prediction of dataset, the predicted data is assigned as a target value. This stage prepares the model for strong model development for accurate prediction. The combination of SVM and NB algorithms are used as hybrid ML classification for churn prediction of telecom industry.
SVM: A supervised learning technique that can deal classification and regression issues is the support vector machine. SVM methods transfer data points from low-dimensional to high-dimensional space. Outliers and noise are problems for support vector machines. Kernel type is the primary hyperparameter in SVM to be adjusted.Sigmoid kernels, Polynomial kernels, Radial Basis Functions (RBF) and linear kernels are examples of common SVM kernel types. The conditional hyperparameter of the "kernel type" coefficient y is represented in sklearn by the symbol "gamma."When it is set to polynomial, RBF or sigmoid in sklearn the hyper parameter r is the conditional hyper parameter of the sigmoid and polynomial kernels.
The polynomial kernel function's conditional hyperparameter d also defines the polynomial kernel function's "degree."The most useful kernel for this study's analysis of customer churn in the telecom industry is one that considers all four types of kernels. The Bayes rule and a few conditional independence assumptions serve as a foundation of the Naive Bayes algorithm, which classifies data. For each class Ci, P (X | Ci)P (Ci) is evaluated in order to predict future the class label of X. If and only if according to the classifier the class Ci will be used as the label for the tuple X.

$$P(X\,|C_i)P(C_i) > P\big(X\,\big|C_j\big)P\big(C_j\big)\,for 1 \leq j \leq m, j \neq 1 \quad (1)$$

In other words, the class Ci for which P (X | Ci)P (Ci) is maximal is the class Ci the predicted class label is the class. Models posterior probability in accord with the Bayes rule, i.e., for all k = 1,...,K

$$\widehat{P}\left(Y=k\middle|X_{1,}\ldots.X_p\right)=\frac{\pi\,(Y=k)\prod_{j=1}^{p}P(X_j|Y=k)}{\sum_{k=1}^{K}\pi\,(Y=k)\prod_{j=1}^{p}P(X_j|Y=k)}\ (2)$$

Where the random variable Y represents an observation's churn class index. The predictors of an observation are X 1...X p, where (Y = k) is the prior probability that a class index is equal to the number k. Within the each class the predictors are distributed based on mean and standard deviation. In Classification using Naive Bayes in order to predict a probability distribution's parameters, assumed that predictors are conditionally independent given the class. For any unseen test data the method estimates the posterior probability of a sample belonging to each class. The approach with the highest posterior probability is then used to classify the test data.
In order to obtain better features for prediction, the final phase validates the results and forwards to the selection method. As a result, it has been shown that the system is effective in

every step of accomplishing the work of objectives. This approach predicts the churn of a telecom industry. This approach predicts the customers as churn or non-churn.

IV. RESULT ANALYSIS

In this section, hybrid ML classification approach for churn prediction in telecom industry is implemented. In this analysis, Telco customer churn dataset is used. The Table 1 shows the attributes of Telco customer churn.

Table 1: Dataset Attributes and Their Description

Attributes	Description
Customer ID	For each Customer Unique ID
Gender	Whether the customer is a male or a female
SeniorCitizen	Whether the customer is a senior citizen or not (1, 0)
Partner	Whether the customer has a partner or not (Yes, No)
Dependents	Whether the customer has dependents or not (Yes, No)
tenure	Number of months the customer has stayed with the company
PhoneService	Whether the customer has a phone service or not (Yes, No)
MultipleLines	Whether the customer has multiple lines or not (Yes, No, No phone service)
InternetService	Customer’s internet service provider (DSL, Fiber optic, No)
OnlineSecurity	Whether the customer has online security or not (Yes, No, No internet service)

After the collection of dataset, data is preprocessed and essential features are extracted. Using the combination of SVM and NB algorithms presented approach predicts and classifies the customer as churn or non-churn. The result analysis of presented approach is evaluated using confusion matrix parameters namely: True Negative (TN), True Positive (TP) and False Negative (FN), False Positive (FP) and are defined as follows:

True Positive (TP): True positive is an instance which is predicted correctly as churn and is actually churn.

True Negative (TN): If an instance is predicted correctly as non-churn and is actually non-churn.

False Positive (FP): if an instance is predicted incorrectly as churn but actually it is non-churn.

False Negative (FN): if an instance is predicted incorrectly as non-churn but actually churn.

Specificity: The ability of an algorithm or model to predict a real negative for each of the available categories is known as specificity. It is known as the True Negative Rate (TNR).Formally it can be calculated by the equation as

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (3)$$

Accuracy: It is described as being the proportion of correctly identified occurrences to all instances, and it is given as

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \quad (4)$$

Sensitivity: As the proportion of true positive cases to actual positive instances, it is also known as True Positive Rate (TPR) or recall and is mathematically stated as

$$TPR = \frac{TP}{TP + FN} \times 100 \quad (5)$$

The table 2 shows the performance evaluation of different methods for the churn prediction in telecom industry.

Table 2: Performance Metrics Evaluation

Methodologies	**True Positive Rate (%)**	**Specificity (%)**
A Novel Approach for Churn Prediction Using Deep Learning	93.8	90.8
Presented hybrid ML classification Approach	95.65	94.3

Compared to previous approach (A Novel Approach for Churn Prediction Using Deep Learning), presented hybrid ML classification Approach has better results in terms of TPR and specificity. The Fig. 2 represents the graphical representation of TPR and TNR of different Approaches.

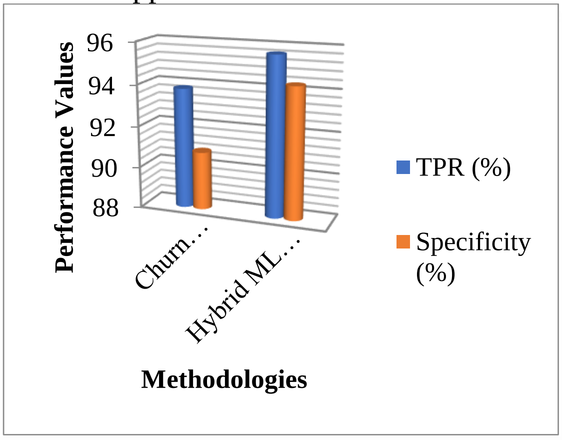

Fig. 2: Performance Comparison

In Fig.2 the y-axis represents the performance values and x-axis represents different methods. The hybrid ML classification approach has better results than customer churn prediction modeling based on Behavioural Patterns Analysis using DL. The table 3 shows the accuracy of different approaches.

Table 3: Accuracy Comparison

Methodologies	Accuracy (%)
Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning (DL)	80.3
A Novel Approach for Churn Prediction Using Deep Learning	86.85
Presented Hybrid Ml classification Approach	95.67

Compared to different approaches (A Novel Approach for Churn Prediction Using DL and Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using DL), Presented approach has high accuracy for churn prediction in telecom industry. The Fig. 3 shows the comparative graph of accuracy.

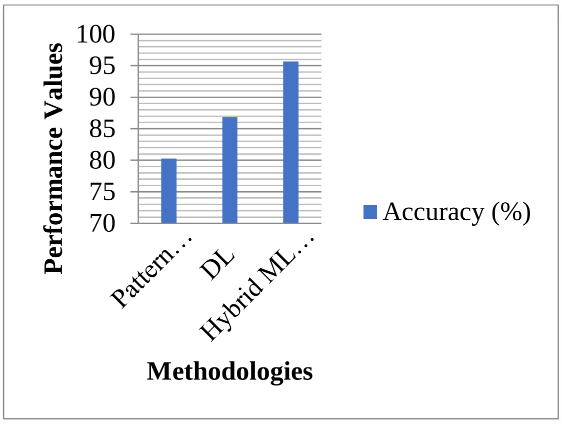

Fig. 3: Comparative Graph for Accuracy

In Fig.3 the y-axis represents the performance values and x-axis represents different methods, Presented hybrid ML classification approach has high accuracy than earlier methods. Hence, presented approach has effectively predicted the tele-customer as churn or non-churn.

V. CONCLUSION

In this work, Hybrid ML classification approach for churn prediction in telecom industry is presented. In this analysis, Telco customer churn dataset is used which contains7043 (customers) rows and 21 columns (features). The development of a churn prediction model that supports telecom operators in more accurately identifying consumers who are most likely to churn is the primary contribution of this work. The dataset has been pre-processed to include attributes with missing data attributes with no values and attributes with values. The combination of Support Vector Machine and Naive Bayes algorithm are used as a hybrid ML classification approach to predict the telecom customer as churn or non-churn. The performance of presented approach is evaluated in terms of Accuracy, True Positive Rate (TPR) and Specificity. The effectiveness of presented approach is compared with previous methods. Compared to previous methods, presented hybrid ML classification approach has better performance in terms of accuracy, TPR and specificity. Hence presented approach has effectively predicted the telecom customers as churn or non-churn.

VI. REFERENCES

[1] Nabahirwa Edwine, Wenjuan Wang, Wei Song, and Denis Ssebuggwawo, “Detecting the Risk of Customer Churn in Telecom Sector: A Comparative Study”, Hindawi, Mathematical Problems in Engineering, Volume 2022, Article ID 8534739, 16 pages, doi:10.1155/2022/8534739

[2] Aisyah Larasati, Darin Ramadhanti, Yuh Wen Chen, Abdul Muid, “Optimizing Deep Learning ANN Model to Predict Customer Churn”, 2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE),DOI: 10.1109/ICEEIE52663.2021.9616714

[3] Laurie Butgereit, “Big Data and Machine Learning for Forestalling Customer Churn Using Hybrid Software”, 2020 Conference on Information Communications Technology and Society (ICTAS) 978-1-7281-3770-4/20, 2020 IEEE, DOI: 10.1109/ICTAS47918.2020.233972

[4] V. Kavitha, S. V Mohan Kumar, G. Hemanth Kumar, M. Harish, “Churn Prediction of Customer in Telecom Industry using Machine Learning Algorithms”, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, IJERTV9IS050022, Vol. 9 Issue 05, May-2020

[5] Awodele Oludele, Adeniyi Ben, Ogbonna A.C., Kuyoro S.O., Ebiesuwa Seun, “Enhanced Churn Prediction in the Telecommunication Industry”, International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume-8, Issue-2, March 2020

[6] Buddhika Livera, D D M Ranasinghe, “Naive Bayes Based Approach on Telecommunication Churn Prediction Using Financial Data”, Sri Lanka Association for Artificial Intelligence, 2020

[7] Pushkar Bhuse, Aayushi Gandhi, Parth Meswani, Riya Muni, Neha Katre, “Machine Learning Based Telecom-Customer Churn Prediction”, 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), DOI: 10.1109/ICISS49785.2020.9315951

[8] Abdelrahim Kasem Ahmad, Assef Jafar and Kadan Aljoumaa, “Customer churn prediction in telecom using machine learning in big data platform”, Big Data (2019) 6:28, Springer, doi:10.1186/s40537-019-0191-6

[9] B.N.Krishna Sai, T. Sasikala, “Predictive Analysis and Modeling of Customer Churn in Telecom using Machine Learning Technique”, Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019), IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8

[10] Khulood Ebrah, Selma Elnasir, “Churn Prediction Using Machine Learning and Recommendations Plans for Telecoms”, Journal of Computer and Communications, 2019, 7, 33-53, DOI: 10.4236/jcc.2019.711003

[11] Ahmad Hammoudeh, Malak Fraihat, Mahmoud Almomani, “Selective Ensemble Model for Telecom Churn Prediction”, 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 978-1-5386-7942-5/19

[12] Nijhawan, V. K., Madan, M., & Dave, M, “An Analytical Implementation of CART Using RStudio for Churn Prediction”, In Information and Communication Technology for Competitive Strategies, 2019, Springer, Singapore, pp. 109-120

[13] saifil Momin, Tanuj Bohra and Purva Raut , “Prediction of Customer Churn Using Machine Learning”, Conference paper, 2019, Springer, doi:10.1007/978-3-030-19562-5_20

[14] Andrews, R, “Churn Prediction in Telecom Sector Using Machine Learning”, International Journal of Information Systems and Computer Sciences, 8, 132-134, 2019, doi:10.30534/ijiscs/2019/31822019

[15] Sanket Agrawal, Aditya Das, Amit Gaikwad, Sudhir Dhage, “Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning”, 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), DOI: 10.1109/ICSCEE.2018.8538420

[16] Abhishek Gaur, Ratnesh Dubey, “Predicting Customer Churn Prediction In Telecom Sector Using Various Machine Learning Techniques”, 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), DOI: 10.1109/ICACAT.2018.8933783

[17] Zhu, B., Baesens, B., & Vandenbroucke, S. K, “An empirical comparison of techniques for the class imbalance problem in churn prediction”, Information sciences, Vol.408, 2017, pp.84-99

[18] Abinash Mishra, U. Srinivasulu Reddy, “A Novel Approach for Churn Prediction Using Deep Learning”, 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), DOI: 10.1109/ICCIC.2017.8524551

[19] Tan Yi Fei, Lam Hai Shuan, Lai Jie Yan, Guo Xiaoning, Soo Wooi King, “Prediction on Customer Churn in the Telecommunications Sector Using Discretization and Naïve Bayes Classifier”, Int. J. Advance Soft Compu. Appl, Vol.9, No. 3, Nov 2017, ISSN 2074-8523

[20] S. A. Qureshi, A. S. Rehman, A. M. Qamar, A. Kamal, and A. Rehman, ‘‘Telecommunication subscribers’ churn prediction model using machine learning,’’ in Proc. 8th Int. Conf. Digit. Inf. Manage., Sep. 2013, pp. 131–136