International Journal of Innovation Engineering and Science Research
(ISSN: 2581-4591)
www.ijiesr.com

Marketing Strategy through Machine Learning Techniques: A Case Study at Telecom Industry

Sanjib Kumar Routray,
Faculty of Computer Science & IT, Telecom Training Center,
Bhubaneswar, Odisha, India.

ABSTRACT

The goal of machine learning is uncovering interesting information and decision making from big data in various domains of applications, which can redefine and enhance customer relationships. In the telecom Industry, on a daily basis huge volume of customer detailed record data is being generated due to a large client base. Telecom management stressed upon the decision making about finding new customers is costlier than retaining the existing ones. Accordingly, marketing analyst and customer relationship management (CRM) professionals need to analyze on the reason for churn customers, as well as, study the interesting patterns from the existing churn customers' behaviour.This paper proposes a segmentation of the market and prediction of the customer's behaviour (Churn, Choice of PLAN) model by Clustering and Classifying the customers based on their attributes. Since customer grouping is the main part of Customer Relationship Management (CRM) on predictive analysis on CRM data of telecommunication industry by using classification, as well as, clustering techniques, identification of churn customers and provisioning of suitable affordable choice PLAN are essential factors behind the churning of customers along with making of decision in sales and marketing in the telecom sector. The best data mining strategies are proposed to be used for classifying the CRM datasets for predicting the choice of PLAN and predicting the churn for retention of customers for efficient managerial decision and making project plan to reach the related goals. Accordingly, this paper proposed “Random Forest” Algorithm for predicting the churn output and “Decision Tree” is proposed for predicting selection of choice of mobile PLAN for predicting profitable mobile customers of telecommunication industry by using Rapid-Miner-9.07-tool-kit. For creating effective retention policies under CRM, to prevent churners, categorization of churn customers is highly essential. Accordingly, after classification, “k-means” clustering algorithm is proposed for providing group-based retention offers. The proposed prediction model is evaluated using no. of metrics, like accuracy, precision, recall, f-measure, and receiving operating characteristics (ROC) area.

Keywords-*Customer Relationship Management, Churn prediction, CDR,*

I. INTRODUCTION

Telecommunication industry every time offering quality products and services with new technology for attracting new customers and retaining existing customers while competing with private sectors to defend their positions in the market. In the present scenario, a huge volume of data sets is being generated by telecom companies at an exceedingly fast rate, as there is a range of telecom service providers competing in the market promotion to increase their customer share. Also the target of telecom companies is to maximize their profit and stay on competitive market [1].But customers have multiple options in the form of quality with less expensive services. When a more no. of customers not satisfied with the services of Telecom Company, service migration or switching found from one service provider to another and churn takes place. Churning takes place due to several reasons like service dissatisfaction, binding limit of post-paid customers and saturated competitive market. On continuous churning, the reputation of a company will affect and loyal customers also get affected. Thus customer's population will be decreased with a high loss of revenue [3].For building intimate long term relationships with commercial important customers or individual customer in marketing strategy, achieving quality of service or sustained customer patronage are highly essential [1, 2]. Accordingly, customer relationship management is the efficient tool for achieving the company's goal. The challenging job of CRM administration is the churn prediction for retaining their valuable customers and enhancing the CRM mechanism [5], [6]. For acquiring new customers, requirement of advertisement, involvement of the workforce and promotional marketing is more expensive than retaining existing customers [3]. With immediate attention, identification of existing churning customers can be stopped. Accordingly a high performance model is needed for identifying churn customers and in future also. Many machine learning and data mining solutions can be used to analyze such Call Data Record (CDR) data on which it can identify reasons behind customer churning. For maximizing the profit, *CRM* can employ to design retention strategies for reducing the percentage of churning customers as well as offering of suitable plans to retain customers [2]. Data mining is now popular by different names such as knowledge discovery, machine learning, business intelligence, predictive analysis, and predictive analytics. From the extracted pattern of data, in the data mining process, the behaviour of churn customers easily identified and decision maker's analysis for taking right decisions through CRM. The development of customers depends on service quality, network coverage, load errors, high technology, billing and rewards where customers can compare the service quality and benefits between the service providers [5]. Researchers focus on global prediction rate of churn customers. In general 2% of normal churn customers, create an annual loss of 100 billion dollars and predicting churn customers is 16 times cheaper than attracting new customers [8]. Ensuring of high level churning, many companies already implemented various data mining techniques on differentiating customers due to technological improvements [9] [10]. Telecom companies also need to understand about customer demands and fulfilling their needs for ultimate escaping from other competitor [11] [12].Accordingly CRM controls the business of customers as per need of services and promotions.Existing models adopted pre-processing methods initially for removing noise for better classifying with improving performance of the models on benchmark CDR data sets [13].But for true representation, multiple algorithms are chosen on these data sets for helping correct decision making of retention and churn prediction [14]. A various machine learning algorithm is proposed in this study and it is validated on CDR of a BSNL, as an Indian telecom company. The performance of a classifier is measured by using the TP rate, FP rate, Precision, Recall, F-measure and ROC-area to prove the best model for churn prediction with the offering of suitable plan to retain customers by achieving high accuracy. For churn and no-churn classification, number of machine learning algorithm we used, where shown better performance compared to other algorithms. Further for market segmentation and study the behaviour of customers, we used k-means clustering algorithm by using attribute selection measures.The remaining part of the paper is structured as follows that provides related work, proposed customer churn and choice of PLAN prediction model, experimental evaluation with results and finally conclusion with future scope.

II. RELATED WORK

Data mining has more importance in analyzing the data and generating interesting, useful patterns and relationships. In the research paper of Clifton (2010), Author stated that originally data being stored in data warehouses, where data mining application needed to analyze the data. As storage became cheaper during 1980s, many organizations develop their storage warehouses and store transaction data for extracting patterns or relationships by adapting Artificial Intelligence methods in the area of KDD. Through applying of data mining techniques, marketing aids of customers of Telecom Company will be fulfilled by proper decision making by applying data mining techniques. According to Author Pradnya, A Shirsath, Vidya Kumar Verma (2013), data mining technology needed for transformation in many fields such as banks, Airlines, railways and so many public/private sector service provider companies to manage and make a decision of such types of huge amount of data. V .Jaaraj, J. Lavanya, J. JagatheshAmairaj, M. Rajkumar (2013) focuses on the existing data mining techniques, i.e. Association rule, Clustering, Classification and described how these techniques were applied to improve the customer relationship with the company for generating considerable profit. Ayman Alawin, Mohmmad Al. ma aitah, Al. Balqua (2014) stated to identify the profitable customers, churn one and develop two models, i.e. one is a physical model (OLAP) (Continuous mining of data sets wherever it resides) and the other is logical model, i.e. adoption of "Classification " on research point of view. From the customer churn prediction point of view in Telecommunication Industry, supervised data mining technique, i.e. "Classification" is used to model the churn prediction with reference to Telecom Industry [25] and in the computer science, application of Telecom big data is being researched from online classification from huge collected data pool [26]. Similarly for finding the targeted high value customers of Telecom Industry, market segmentation is essential through clustering data mining techniques and for increasing the market share, company growth and profitability assessment of data mining based CRM technique, i.e. "Classification" is essential.The problems associated with churning of customers through the most commonly recognized machine learning, data mining techniques that actually supporting the telecommunication company for predicting the churn customers and ultimate retaining the customers under CRM [15]. In the decision tree, there is a limitation in nonlinear complex connections between attributes, but on purning of decision tree, the accuracy of decision tree improves [16]. Decision tree algorithms have many advantages like easy visualizations, use of nonparametric method and processing of numerical and categorical data [17]. *Naive Bayes* is also a guided learning module that predicts invisible data based on the position of Bayesian, which is used to predict churn customer [18]. The novel model presented in [13] by using the classical rule inductive technique (FOIL) shows a hybrid approach linking the adapted *k-means* clustering algorithm to predict churn customer behavior. The control of a large volume of data in today's world provides an opportunity to improve the quality of service to the users. This data includes information about customers' behavior, usage pattern and network operations. Over the past decade, the telecommunications service sector has undergone a major change due to new services, state-of-the-art upgrades [19]–[24] and intensified competition due to deregulation [4]. There is a need to secure important customers, strengthen connection management and retention of CRM. In this paper, in the first phase, customers are segmented using decision rules and in the second phase, a model is developed for every leaf of the tree. Performance of random forests algorithm is best for different types of datasets on selected attributes compared to j48-decision tree, Naïve Bayes for measuring student's performance [27].This hybrid approach is compared with decision trees, random forests, Naïve Bayes decision rule and random tree [28] for churn prediction. Finally, random forest Classification Algorithm and K-Means Clustering Algorithm is proposed in the article as reviewed in this research paper for maximizing organization's satisfaction for increasing loyalty, retaining customer business over their lifetimes [28]

III. PROBLEM DEFINITION

Telecom management seriously thinks to come out of monopolistic mode and planned for enhancing customer relations, strengthening technology and upgrade infrastructure. Accordingly the present study of CRM module of marketing management with different issues analysed through machine learning techniques. Using supervised and unsupervised learning techniques, prediction of performance of machine learning algorithms through monitoring various performance parameters on different telecom datasets will be possible. Churn prediction and profit prediction on suitable PLAN are mainly two important issue in marketing management.Dataset 1 & 2, which is acquired from Telecom Company records numerous attributes act input for a classifier models predicts appropriate churn class with profit class on suitable PLAN.Similarly, the behaviour of customer information with their relationship, clustering model is used by partitioning the complete customer data sets into groups. Through clustering technique, Low, Medium & Risky customer is determined under churn class category of datasets, which will help the management to take suitable action by offering suitable PLAN to the esteemed customers.Different DM and ML techniques are applied throughout the research for suitable accuracy prediction. Algorithms used are Naïve Bayes, Decision Tree, Random Tree and Random Forest on both the datasets.

IV. PROPOSED MODEL

This section presents the proposed customer churn and profit prediction model. Fig. 1 shows the proposed churn prediction model and profit prediction model indicates number of steps. In the first step, data pre-processing is performed which includes data filtering for noise removal, removal of imbalanced data features with normalization of the data. In the second step, different classification algorithms like Decision Tree, Random Tree, Random Forest and Naïve Bayes applied for categorizing the customers into the churn and non-churn customers of one part. In the second part same algorithms applied for categorizing profit and no profit from customers on CDR data sets. In the third step, customer segmentation according to churn and profit is performed using *k-means* clustering techniques. Based on patterns of customer transactional behaviour from data sets, cluster analysis is developed. Finally the model recommends retention strategies for each churn category and individual category on a suitable PLAN of customers.

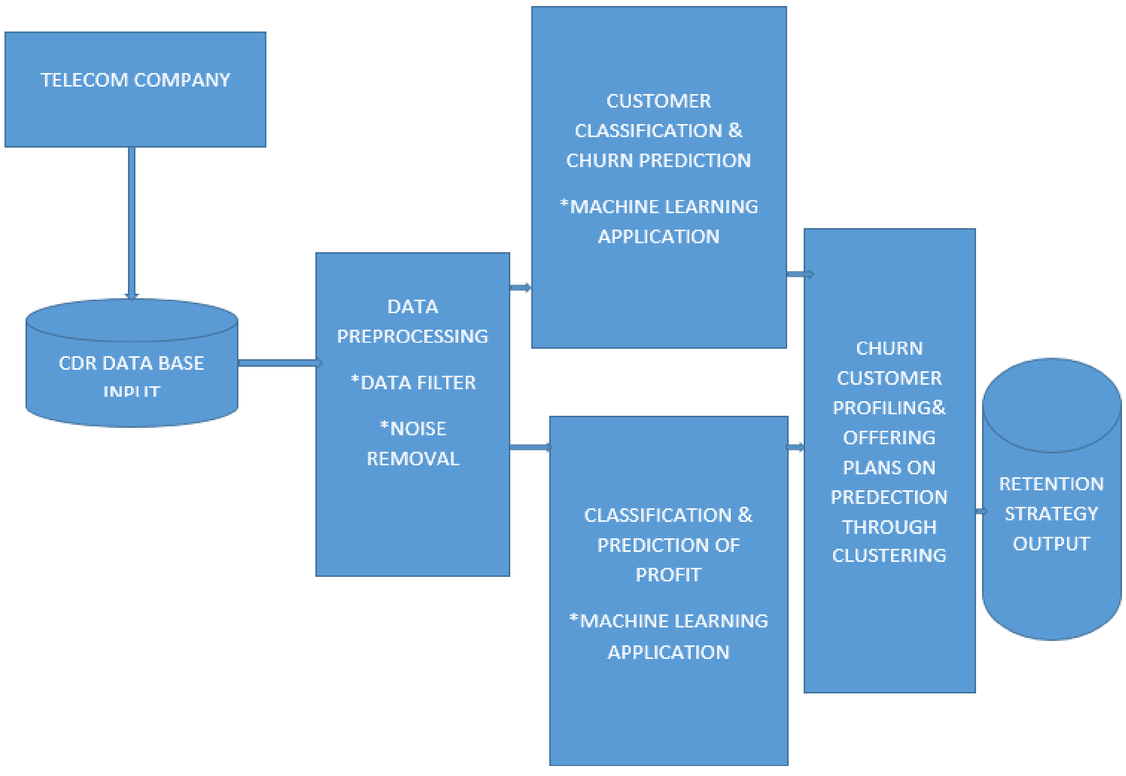

Fig.1 Proposed model for customer churn and profit prediction.

A. Data Pre-processing

1) In telecom CDR (Call Detailed Records) data set-1 as mentioned in Table-1, number of missing values, "Null" value first removed and further filtered imbalance attributes, which is the first step of noise removal under data pre-processing. In this telecom dataset, the number of attributes is 10, with useful features for using information gain and correlation attributes ranking filter techniques for feature selection on RapidMiner-9.7 tool kit to get high ranking values. Out of 10 attributes, some attributes improve performance measures and are useful for churn-decision-making process. Same procedures adopted for CDR data set-2 with 5 no. of equal important attributes for predicting profit-decision making process through performance measurement. Accordingly the performance of classification increases, if the dataset contains highly predictive and valuable variables.

.

Table-1 Number of Filtered Features of CDR-data set-1 & 2.

Sl no.	#Features	Sl no	#Features
1	State	1	GSM NO
2	Area Code	2	Category
3	Phone No.	3	Group
4	Int. Plan	4	FRC PLAN
5	Voice Mail Plan	5	PROFIT
6	Voice Mail No.		
7	Total Day Minute		
8	Total Day Call		
9	Total Day Charge		
10	Churn		

B. Customer Category and Prediction

There are three categories of customers like individual, business and others. Out of which three groups of customers divided like lower income group, middle income group and higher income group exist. According to choice of service of different customer groups, some of the customers are loyal and they are declared as non-churn customers. But some are churn customers, who is not satisfied with the services of the telecom company. The proposed model targets those churn customers by applying different machine learning classification techniques like Decision Tree, Random Tree, Random Forest & Naïve Bayes and devise retention strategies to overcome the churning problems created by different customers. Further from performance measurements of different algorithms, identification of best classification is analysed between churn and non-churn customers. Also above algorithms correctly classified the profitable customers to the company from non-profitable one between different customer groups.

V. EXPERIMENTS and RESULTS

A. Dataset Description

In this study, datasets are obtained from one GSM telecom service provider for studying the customer churn prediction problem. The data are extracted from the customer service usage pattern Call Detail Record (CDR), which consists of labeled data with two classes where 15% data is labeled as "T" (true customers) that represents churner and 85% data is labeled as "F" (false customers) that represent non-churners. It has three main types of' attributes that include call behaviour, or usage attributes, financial information attributes and marketing related attributes, selected on feature selection techniques that allows identifying the most relevant, useful and effective attributes for customer churn prediction. Another part of transactional datasets are also taken with four main attributes like lower income group (LIG), Middle Income Group (MIG) and higher income group (HIG) selecting suitable PLAN for selecting two categories of customers like profitable and no profitable (Table-2).

Table-2 Dataset Description

CRM Data sets	Instances	Selected Attributes	Models Used
1	49	5	Classification for Churning of Customers

			Clustering for Market Segmentation
2	120	4	Classification for Profitable PLAN Prediction

On the two CDR data sets, we performed a number of experiments on the proposed machine learning model of by using a Rapid Miner-9.7 toolkit for providing the factors behind customer churn. Different techniques like Random Forest, Decision Tree, Random Tree and Naïve Bayes tested. Out of which, Random forest is a useful technique for classification and shows better accuracy (85.71%) and less error in comparison to other on correctly classifying of churn and non-churn customers with 10 fold cross validation as mentioned in Table-3. Similarly for selecting the suitable PLAN, for future decision making by the management, above machine learning techniques also used. The Decision Tree technique shows excellent (100%) classification accuracy with no error with 10-fold cross validation, as mentioned in Table-4 among individual, business and other groups of people under Lower income group (LIG), Middle Income Group (MIG) and higher income group (HIG) selecting suitable PLAN. The performance vector output of Rapid-Miner tool-kit in case of above two best techniques. Further, in this study, Attribute Selected Classifier algorithm is used for identification of factors, which clearly indicates the finding of churn customers. Also in this paper, k-means clustering algorithm is used for creating retention policies by decision makers considering between of churn and no-churn customers with profit and no-profit customers of different groups.

B. Performance Evaluation Matrix

The performance of different classifiers is measured for predicting of churn and profit on the basis of some parameters i.e. (1) True positive rate (2) False positive rate(3) Precision (4) Recall (5) F-Measure(6) ROC (Receiver Operating Characteristics). Confusion matrix is used to evaluate the classifier quality for a two class problem, i.e. True Positive, True Negative, and False Positive & False Negative. Confusion matrix is a useful tool for analyzing the performance of classifiers by recognizing tuples of different classes. Similarity measurement of classifier's performance as mentioned in Table-5 includes accuracy, sensitivity (recall), and specificity, precision etc. Where further construction and evaluation requires partitioning of training set and test set. Holdout, Random Sampling and Cross validation are important steps of partitioning.

Table-3 Classification Accuracy & Errors of Various Algorithms on Own Churn-Balance Dataset for churn prediction

Method Used	Classification Accuracy (%)	Classification Error (%)
Random Forest	85.71	14.29
Random Tree	78.57	21.43
Decision Tree	57.14	42.86
Naïve Bayes	50.00	50.00

Table-4 Classification Accuracy & Errors of Various Algorithms on Own Profit Measurement Dataset for Profit prediction

Method Used	Classification Accuracy (%)	Classification Error (%)
Decision Tree	100	0.00
Random Forest	97.22	2.78
Random Tree	97.22	2.78
Naïve Bayes	91.67	8.33

Table-5: Performance measurements of various Algorithms of Classification for Churn and Profit

Method	TP	FP	Precision	Recall	F-measure

Used for Churn	**Rate**	**Rate**			
Random Forest	**1**	**0**	**0.77**	**1**	**0.87**
Random Tree	**0.87**	**0.2**	**0.77**	**0.87**	**0.40**
Decision Tree	**0.8**	**0.2**	**0.44**	**0.8**	**0.56**
Naïve Bayes	**0.62**	**0.6**	**0.55**	**0.62**	**0.58**
Method Used for Profit	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-measure**
Decision Tree	**1**	**0**	**1**	**1**	**1**
Random Forest	**0.97**	**1**	**1**	**0.97**	**0.98**
Random Tree	**0.97**	**1**	**1**	**0.97**	**0.98**
Naïve Bayes	**0.97**	**1**	**0.94**	**0.97**	**0.96**

To further validate our findings, the performance of algorithms in Table-5 show that TP and F-measure is higher for Random Forest classifier and Decision Tree classifier as compared to others in case of churning and profit. The Area under Curve (AUC) is a selective performance measure which is used by many researchers in the prediction model for measuring the accuracy.

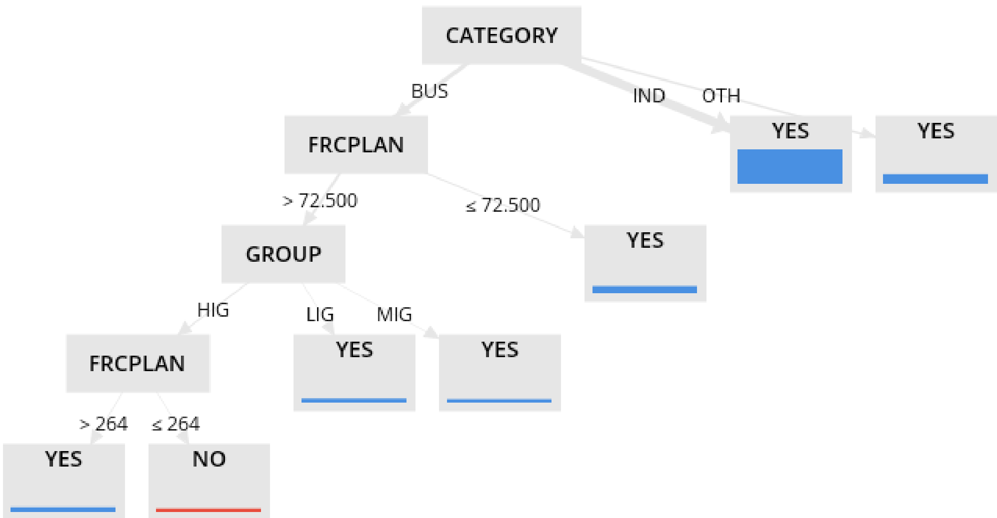

Fig. 2 Decision Tree output of Rapid-Miner tool-kit for PROFIT prediction

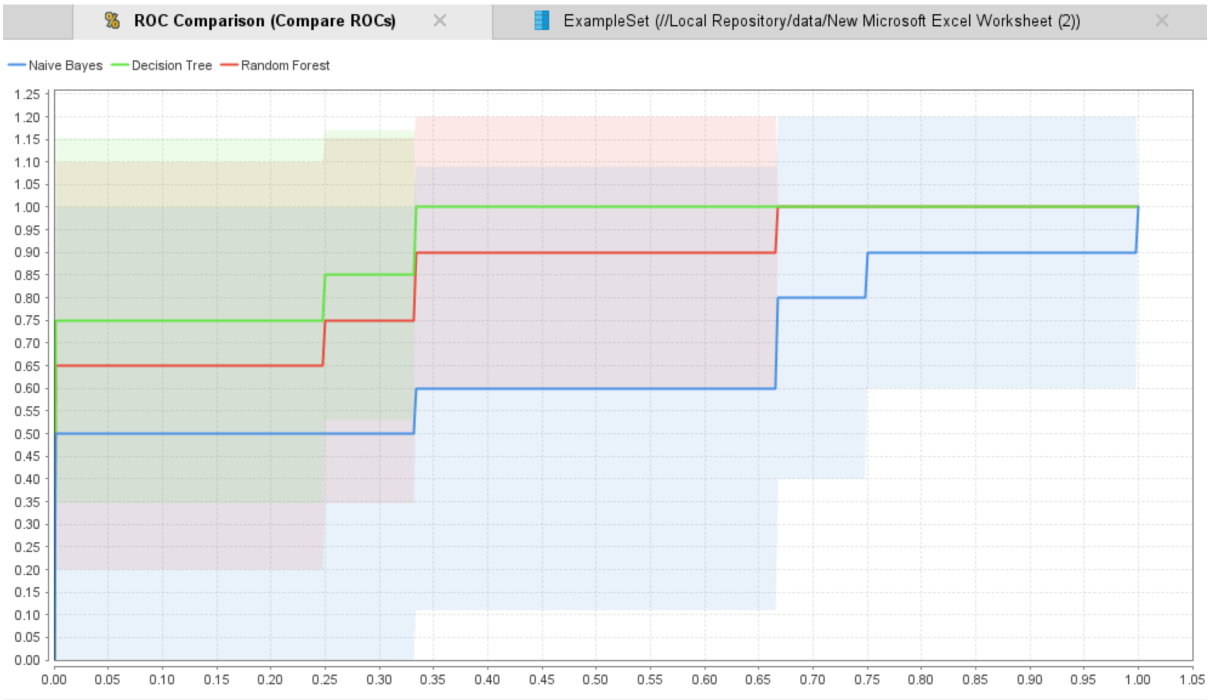

Fig. 3 ROC outputs in Rapid-Miner tool-kit for PROFIT

The complete Decision Tree output of a Rapid-Miner in Fig. 2 indicates more profit in case of "IND" head. ROC area denotes the average performance against all possible cost ratios between FP and FN. If the ROC area value is equal to 1.0, this is a perfect prediction. Similarly, the values 0.5, 0.6, 0.7, 0.8 and 0.9 represent a random prediction, bad, moderate, good and superior respectively. Accordingly ROC output shows better output in case of Decision Tree as per Fig.3.

C. Customer Segmentation and Retention

Based on the behaviour of customer information with their relationship, segmentation or clustering is used by partitioning the complete customer data sets into groups. Out of number of clustering algorithms, we used k-means clustering algorithm. K-means clustering is the iterative approach, where data can be segmented into different groups from complex heterogeneous large data sets. Arithmetic mean value of real valued data is the representative of a cluster, where we can find the hidden pattern that represents one class or cluster. In this study, the *k-means* algorithm segments the data into three groups due to the nature of the data. The three groups represent cluster-0, cluster-1 & cluster-2 among Low, Medium and Risky customers as mentioned in Fig. 4. Fig. 5, shows the number of customers in each segment, according to the value of k in *k-means* to 3 can lead to better segmentation. During n testing, it is found that, three clusters (0, 1 & 2) along with 31 numbers of customers are under no-churn category and 10 under churn category of sample data sets. Therefore, for retaining the churned customers, suitable action to be taken by the decision makers for making profit by offering suitable PLAN to the individual customers instead of business and other groups as per above classification technique.By offering suitable PLAN to a specific group of customers, Telecom Company easily understands the behaviour of customers and ultimately enhances retention and marketing performance.

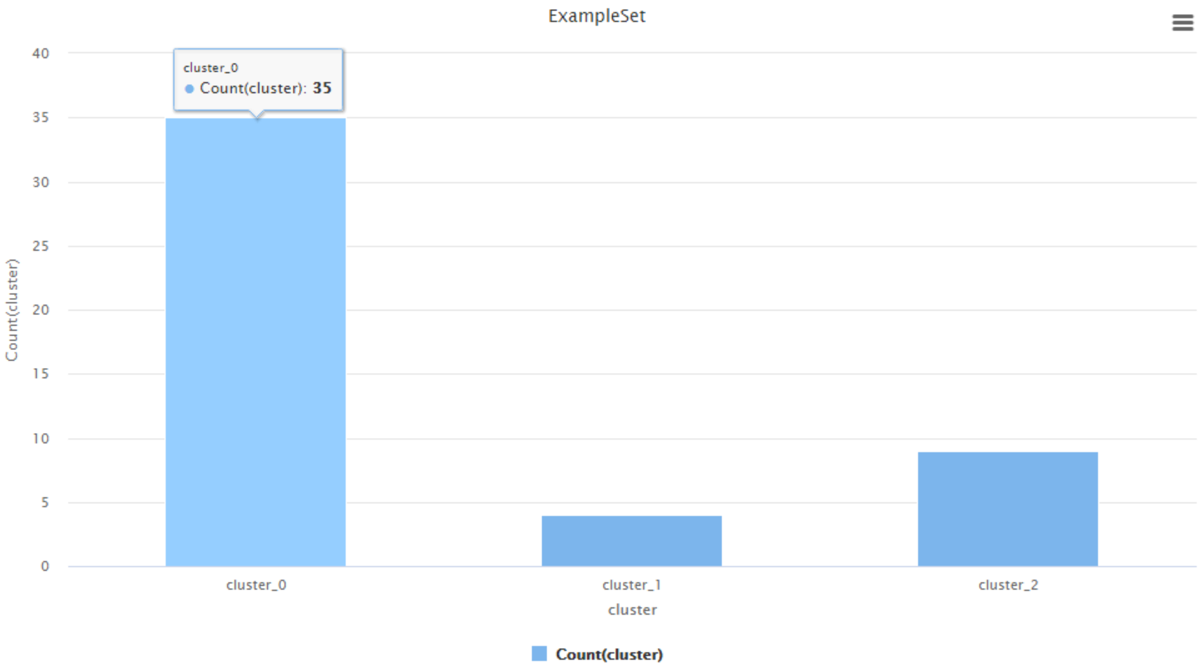

Fig. 4 Cluster count output of k-means Algorithm

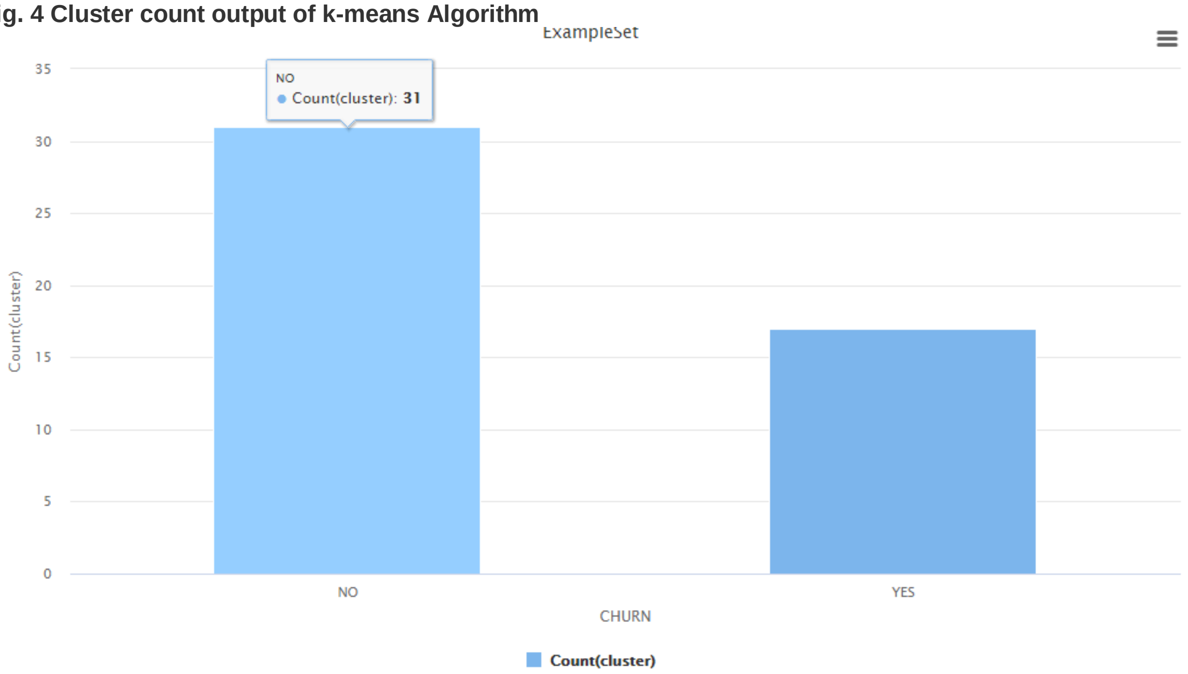

Fig. 5 Cluster count output churn and no-churn customers

VI. CONCLUSION

The challenging issue in the present competitive telecom market is churning prediction of the CRM to retain the liable customers by identifying a similar groups of customers and offering competitive PLAN (Tariff of mobile customers) to the respective groups for making more profit in organization. Accordingly, researchers finding the key factors of churning by developing certain machine learning models for retaining of customers in telecom organization. In this paper, through standard performance metrics, Random Forest, Random Tree, Decision Tree & Naïve Bayes classification techniques proposed and evaluated for the purpose of churning. Out of which Random Forest produced a better result that is 85.17% (Accuracy). Similarly for the purpose of profit calculation by decision makers, Decision Tree produced a better result that is 100% (Accuracy). Finally from the dataset, churn factors identified and performed un-supervised k-means clustering technique according to their risk of churning also recommended certain guidelines on customer retention under CRM for efficient decision making by decision makers.

In the scope of the study, applying of deep learning under the Artificial Intelligence for predictions and pattern analysis investigation is highly essential according to the frequent changing behaviour of churn customers in future.

REFERENCES

[1] S. Babu, D.N. Ananthanarayan and V. Ramesh, "A Survey on factors impacting churn in telecommunication using data mining techniques", int. j. Eng. Ras. Techmol.,vol 3, no.3, pp.1743-1748, Mar 2014.

[2] C. Geppert, "Customer Churn Management: Retaining high-margin customers with relationship management techniques, KPMG & Associates YarhandsDissou Arthur/KwakuAhenkrah/David Asamoath, 2002

[3] W.Verbeke, D.Martens, C.Mues, and B.Basens, "Building Comprehensible Customer Churn prediction models with advanced rule induction techniques", Expert system, Appl., vol.38, no. 3, pp. 2354-2364, Mar.2011

[4] Y. Huang, and M . T. Kechadi, "A rule based method for customer Churn Prediction in telecommunication services", in Proc. Pacific-Asia Conf: knowl. Discovery Data mining Berlin, Germany: springer, 2011, pp. 411-422

[5] Bhave and Ashish, "Customer Satisfaction Measurement", Quality & productive journal, pp. 562-572, 2002

[6] M. Kaur, K. Sing, and N. Sharma, "Data mining s a tool to predict the churn behaviour among Indian Bank Customers," Int. j. Recent Innov. Trends Comp. Commun., vol.1, no. 9, pp. 720-725, sept. 2013

[7] V.L. Migueis, D. VanmdanPoel, A. S Camanho and j.F.ecunha, "Modelling Partial Customer Churn: On the value of first product-category purchase sequences," Expert syst. Appl., vol.12, no.12, pp. 11250-11256, Sept. 2012.

[8] D .Manzano-Machob, "The architecture of a churn prediction system based on stream mining ", in Proc. ArtifIntell. Res. Develop, 16th int. conf. catalan Assoc. Artif. Intell., vol.256, oct 2013, p.157.

[9] P.T. Kotler, Marketing Management Analysis, Planning, Implementation and control, London, U.K: Prentice-Hall, 1994.

[10] F.F. Reichheld and W.E. Sasser, jr., "Zero defections:, Quality comes to services", Harvard Bus Rev., vol 68, no. 5, pp. 105-111,1990

[11] NematSheereen, S, "A Study on Customer Satisfaction of BSNL Services in Kerala", Intercontinental Journal of Marketing Research Review, ISSN: 2321-0346 – online, ISSN: 2347-1670,Vol. 2, Issue 9 pp.137- 140, 2014

[12] H.-S. Kim and C.-H. Yoon, "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market", Telecommunication. Policy, vol.28, nos. 9-10, pp.751-765, Nov.2004.

[13] Y. Huang and T. Kechadi, "An effective hybrid learning system for telecommunication churn prediction," Expert System. Appl., vol.40, no.14, pp.5635-5647, Oct.2013

[14] O. G. Ali and U. Ariturk, "Dynamic churn prediction framework with more effective use of rare event data: The case of private banking", Expert syst. Appl., vol.41, no.17, pp. 7889-7903, Dec. 2014.

[15] V. Lazarov and M. Capota, "Churn prediction", Bus. Anal. Course, TUM Compt. Sci, Technische Univ. Munchen, Tech. Rep., 2007. [online]. Available :http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.462.7201 & rep=rep1& type=pdf

[16] R .Vadakattu, B.Panda, S. Narayan, and H. Godhia, "Enterprise subscription - churn prediction", Proc. IEEE int. conf. Big Data pp.1317-1321, Nov-2015

[17] M. Hassouna, A. Tarhini, T. Elyas and M.S. AbouTrab, "Customer churn in a mobile markets a comparision of techniques, jan. 2016 [Online] Available http://arxiv.org/abs/1607.07792.

[18] V. Umayaparvathi and K. Lyakutti, "Applications of data mining techniques in telecom churn prediction", int. j. Compt. Appl. Vol. 42, pp. 5-9, Mar. 2012

[19] A.Amin et al., "Cross company customer churn prediction in telecommunications: A comparision of data transformation methods", Int. j. inf. Manage., vol. 46, pp. 304-319, jun.2019.

[20] A. Amin, B. Shah, A. M. Khattak, T. Baker, H. Ur. Rahman Durani, and S. Anwar, "just-in-time customer churn prediction: with and without data transformation", in Proc. IEEE Congr. Evol. Comput. Jul.2018, pp.1-6.

[21] A. Amin et al., "customer churn prediction in the telecommunication sector using a rough set approach", Neurocomputing, vol. 237, pp. 242-254. May-2017.

[22] A. Amin et al., "Comparing oversampling frequencies to handle the class imbalance problem: A customer churn prediction case study",IEEE Access ,vol.4, pp.7940-7957, 2016.

[23] M. Ahmed, H. Afzal, A. Majeed, and B. Khan, "A Survey of evolution in predictive models and impacting factors in customer churn", Adv. Data Sci. Adapt. Anal., vol. 9, no. 3, jul. 2017, Art. No. 1750007.

[24] R. Rajamohamed and j. Monokaran, "Improved credit card churn prediction based on rough clustering and supervised learning techniques", Cluster comput., vol. 21, no. 1, pp. 65-77, Mar.2018

[25] Kiran Dahiya, KanikaTalwar, "Customer Churn Prediction in Telecommunication Industry using Data mining technique-A Review", International Journal of Advanced Research in Computer Science and Software Engineering, Vol-5, Issue-4, 2015.

[26] Jiang Heling, An Yang, Fengyun Yan, Hang Miao, "Research on pattern Analysis and data Classification methodology of Data mining and knowledge Recovery", International journal of Hybrid Information Technology, IEEE, Vol-9, pp.179-186, 2016.

[27] Madhuri T. Sathe, Amol C. Adamuthe, "Comparitive study of supervised algorithms for prediction of student's performance", I. J .Modern Education and Computer Science, Published Online in MECS (http://www.mecs-press.org/) vol-1,pp.1-21, Feb – 2021.DOI: 10.5815/ijmecs.2021.01.01

[28] Irfan Ullah, Basit Raza, Ahmad Kamran Malik, Muhammad Imran, SaifUl Islam & Sung Won Kim, "A churn prediction model using Random Forest Analysis of Machine Learning Techniques for churn prediction & factor Identification in Telecom sector", Open Access Journal, IEEE Access, volume-7, pp.60134-60149,2019.