Classification Methods in Telecommunication for Churn Prediction of Customers: A Novel Approach Using Data Mining

Sativa, Research Scholar (SET), Ansal University, Gurugram.

Abstract-The telecom sector is vigorous with recent technologies and current resources in today's time. In telecommunication, the customers have different types of options. Customers have a lot of complaints like customers call drop-in particular area, issues related with customer bills, disconnection complaints, issues with promotions, call timing issues, complaint related to resolve duration call delayed problems, quality of services problems and coverage related complaints, churn prediction problem. But now a day's churn prediction for customers is an   in telecommunication. In this review paper complete focus is on, how to develop a classification model, to handle the customer complaints in telecommunication. This work introduced churn prediction model for customers in telecommunication using data mining classification methods. Various classifiers have used in telecommunication for churn prediction of customers. Classification algorithms comparison depends upon efficiency of the available dataset. Different data mining-based classifiers namely decision tree, KNN, logistic regression, SVM, naïve Bayes and rule based and random forest are used to handle customer complaints. The performance of classification model is measured with different performance measures like accuracy, recall and precision measures with confusion matrices.

Keywords: Telecommunication, Churn Prediction, CRM, Classification techniques, Discussion and performance, Conclusion .

Introduction

These daysenormous dataare produced by different industrieswhich is growingat fast pace.In telecommunicationsector, customer is considered as an important resource by service providers [1]. Thetelecommunication sector faced many critical problems and customer churn is one of these.The marketwhere customers have several choices of different service providers, they can easily switch the provideror plan within same provider [2]. Acquiring new customers in telecommunication company is costly instead of retaining existing customer.There are some reasons for churning customerlike customer are not satisfied with qualities and services,also customers are not satisfied with features and higher cost.Therefore, there is little difficulty to calculate customers churn rate. Another factor is customer loyaltyas the loyal customer increase company revenue. The loyal customers increase the profit of company, by referring some retainingstrategies plan to the friends, family members. The company profit may be determined by quality of services and quality of productprovided to customers.

Slow response to customer complaints related with billing, high prices and insufficient features are the factors that cause customers defecting to the competition. Customers analyzingtheir services with another service providers services and change to those providers who provides better services plan [3]. In this survey paper various data mining-based classification prediction methods are studied for churning.

The ample amount of data, such as network related data, demographic data, call and short messages data, billing record data and other type of data that described behavior of customer,helpstoprovide the accuracy of data mining classification algorithm in telecommunication sector. Due to all this ample amount of data, there is a probability for classifiers to develop predictive modeling to handle the churn prediction in telecommunication. Data mining-based classifiers have completely used as solution to predict customer churn by identifying the factors.

In telecom, prediction for churn customers is a demanding activity. In recent years, the incredible and rapidly growing research problem is churn prediction for customers[4]. Most of the studies have analyzed customer churning problem from different perspectives to design and suggested the best solution to telecommunication field. Therefore, various probabilistic and non-probabilistic techniques namely ensemble techniques [5], probabilistic method [6], Support Vector machine [7], k-NN, Rough Set Theory (RST) [8], Naive Bayes, Fuzzy Logic, neural network [9] have been used to differentiate customer with ratio to churn. Hence there is trial of which one datamining classification method can be used for churn prediction model. Customer complaints like churned problems in telecom sector can be resolved by classification methods involving unknown records or data.The decision-making prediction methodsare used to handle customer complaints. Some relevant complaint can be differentiated from irrelevant complaints [10], automatically by using complaints handling prediction model or decision-making systems.Mapping the customer complaints and to measure customer satisfaction, different type of decision-making prediction system like (hybrid and web based) are used [11]. The churned customers are recognized with many development techniques such as data analysis tools and data mining classification methods.whenthecomplaintsare relatedwithincomplete and incompleteinformation due to customers limited knowledge about service or productfailure, some classification faults commonlyexist.

Customer Churn

Churn prediction for customer helps the CRM to avoid customers who may be churned in near future by suggesting retention plans and provides better packages to attract the possible churn customers that the telecommunication company retain them.

There are two type of churn customer 1. Direct churn customer 2. Indirect churn customers.

Indirect churn customers are those customers that the telecommunication company itself delete them because of non-payment status, no usage of phone. These indirect churned customers are easily identified as compared to direct.The direct churn customers are the customers that itself decide to stop their services with the particular telecommunication service provider like (airtel, orange, Vodafone etc.)[12].
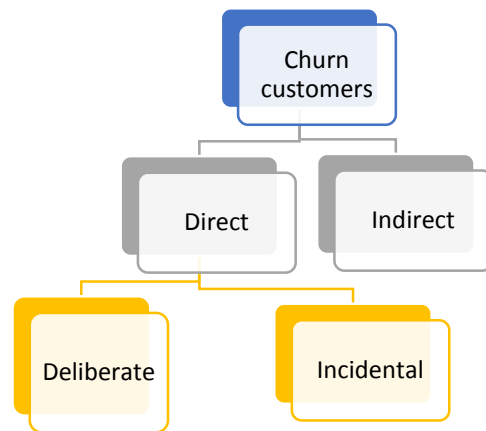
Figure1: churn customers representation

This study introduced that the direct churn customers are difficult to identify because of some varying parameters such as economy factors, technology factors and poor customer services factors.

CRM

It helps in business, to communicate with client or customers properly. The function of CRM is to enable an organization to understand customers need and behavior and provide better quality of service.The new task of CRM is to retain existing customers and provides retention strategy plan for organization and customers. CRM can analyze data and generate reports. CRM helps the companies to maximize their profit [13]. CRM can be used for retention strategies in telecommunication to reduce customers churn ratio. By using data mining classification techniques, the CRM can carry the customer retention objective of a company by recognizing exact customer needs.The goal of researchers is to distinguished customerswhoidentifythosewho are likely to churn and who are not churned from service provider [14].Nowadays the customer has more than one option to fulfil their requirements due to the deregulation of telecommunication companies.Therefore, the telecom companies should put in more efforts to understand and then meet the customer needs, whilst taking into consideration their competitor's offerings in the same market [15]. To identify business unit and customers, it provides retention plan such as which itemsservicesand promotions are suitable for which customers,as they need. So, it is easy to handle customer complaints with the help of different CRM systems like operational, collaborative, Analytical in telecom sector.

Data mining

It is process of extracting relevant pattern from existing large databases, in order to discover structured information from the data [9]. This process of knowledge discovery from database is called as KDD process.
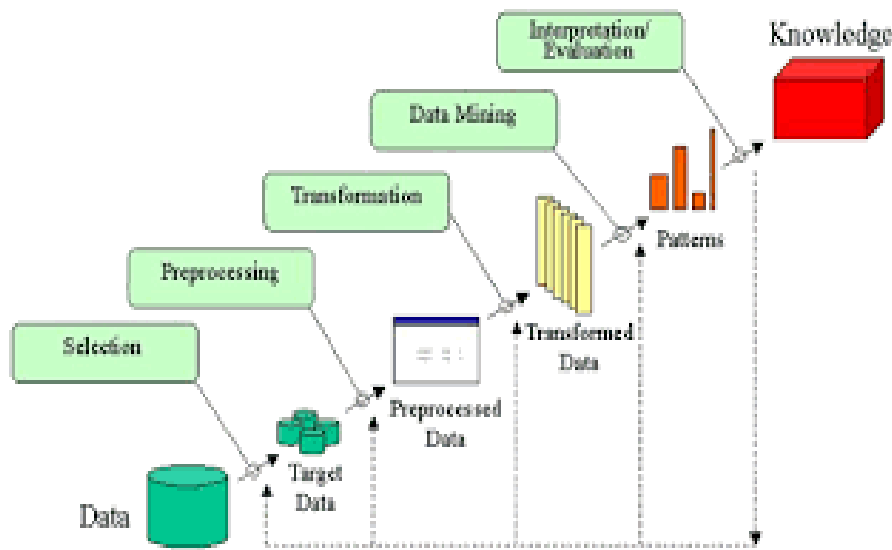
Figure 2: KDD-datamining approach.jpeg (tan & Kumar,2009)

All the steps implemented in this approach determines the successive levels to extract information from large dataset and documents the knowledge discovery for further detection. Different data mining classification algorithms are suggested by different researchers to predict the churn model.

Methodology

The below given figure showed the prediction model steps for churn customers. Data mining with process of data preprocessing helps to recognize and extraction of patterns for the behavior of churn customers.Data analysis can be done after processing the collected data. The function of preprocessing stage is to turn data into suitable state [16].
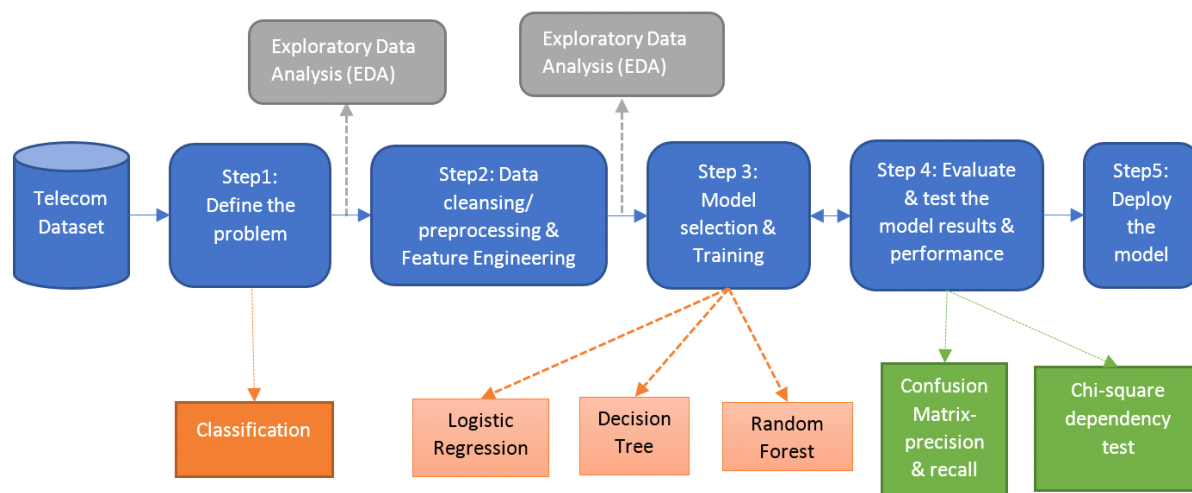

Figure3: prediction model in telecom [12]

Data preprocessing: Different data processing techniques are suggested for filtering the data such as by removing noise, removal of imbalance data and under sampling of data.

Feature Selection: Only the relevant attributes are selected from large dataset that is based on domain knowledge. Features can be selected with two different types of methods; one is information gain and attribute ranking filter[17].

Noise removal: Missing data, outliers and unique characters that existed in telecom dataset are included in noisy type data. Removal of noise is required because noisy data can cause poor performance and poor analysis results.

Under Sampling: It is a technique to resolve class imbalance issues that are faced by telcom sector. This problem arises when data of class is overrepresented the majority class, then technique of under-sampling is used to modify majority classes with minority classes.

Classification Techniques:

Decision Trees (DT)

It is a tree structured classifier.In this tree, attributes are represented by internal nodes and class label corresponds to a leaf node. It is predictive modeling method used in data mining. Decision classifier are useful in concept of transparency. These are easy to understand, are inexpensive to build, decision classifiers can be performed on both type of, nominal and categorical datasets. Gini index and information gain are the criteria, to find impurity and can be calculated from the formula given below

Information gain= -(p/p+n) log(p/p+n)-(n/p+n) log(n/p+n)

Gini index=$1-\sum Pi^2$

Entropy= $-\sum P_{i+}N_i/P+N$

Gain =Entropy Class-Entropy of attribute

Naïve Bayes (NB)

This classifier usessupervisedlearningmethod. A classifier that predicts class conditional probability and classifies input data set based on Bayes theorem. It computes probability whether input dataset belongs to specified class or not. As [18] and [19] indicates that it's based on the Bayes rule which is defined as follows:

$$P(B/D) = \frac{P(D/B) \cdot P(B)}{P(D)}$$

Rule induction (RI): The rule-based classifierextracts formal rules from observations. These rules represent patterns in data. RI provides different ways to generate rules by specifying concentration on class at a time [8]. The if-Then rule and rules generating from decision tree are used for pattern specification.

SVM: It is supervised learning method. The purpose of SVM is to find a hyperplan that separately classify the data points. The two classes of data points are separated by different hyperplanes. We could choose any type of hyperplan. This method helps to find the maximum margin hyperplan between m1, m2 (the maximum distance) between data points of both classes. Maximum the margin better will be the classification of datapoints between two classes [7].

Linear Regression: Regression is technique used to computetarget value based on independent variables. Regression helps to predict independent service, when dependent variables are given.

K-Nearest Neighbors: KNN is instance-based learning approaches. KNN can be called as lazy algorithm, because all training data are used at testing phase. It can be applicable for classification by finding the distance or similarity measures between records. KNN suggested that the datapoints must be scalar, so that the distance can be easily measured. Euclidean distance is best method used in KNN [8].

Random Forest: It is an effective ensemble learning technique for classification and can handle largedata efficiently. In random forest, number of different decision trees are used for prediction. We got better results on churn customer prediction [8]. These classifier use Gini index to determine final tree. To make final classifier, different random trees are voted by weighted values.

Ensemble Learning methods (Ada Boosting): It is non probabilistic method that consists of multiple weak learners that are individually trained. In this approach divide the training dataset into number of subset of data, from which a proposed model was created. It is a very powerful classification technique. The popular ensemble algorithms are as: boosting, Ada-boost, random forest [8].

Related Work

This literature work showsvarious data mining classification techniques, are used for Churn prediction.Classification technique helpscompanies to identify, predict and retain churning customers. Classifiers are also used for decision making and CRM. The different kind of research studies wereintroduced by different researchers for problems of prediction with the customer churn [18][19].

In telecommunication industry an important issue is to predict the behavior of churn customers using datamining methodologies. A review of churn prediction analysis is conducted with dataset from data mining survey techniques. The datasets consist of training data and test data.so, from these datasets only essential variables were selected for different algorithmimplementation[22].Neural network and genetic algorithm are two classification techniques for the churn prediction in telecommunication customers[23].In this approach GA is used for feature selection and the neural network was used for prediction model. The finding of this study was, not to recommended for business profit maximization. Another approach is hybrid neural network for churn prediction is provided in [24]. This approach is combination of neural network and self-organized map. The combination of NN & SOM performs data reduction and filtering. So, it was found that it leads to missing of some data variables from training set.In this literature review comparison of churn prediction methods are provided to conclude better accuracy results. After comparing NNand decision trees, the performanceand correctness of decision tree is better than the neural network [20]. It was found that, the decision tree method has some limitation on nonlinear attributes. This study results that accuracy of decision tree can be improved with pruning method [21]. For churn prediction two step classifiers algorithm is introduced. First stepin classifier algorithm describe how the attributes aresplitted into different clusters, and in second step implement the classifier algorithm to test churn data. In another study, authors proposed a new approach of data certainty for churn prediction [25]. In data certainty, different distance factors based on clustering method are used for churn prediction. In this study, enhanced concept of data certainty for data mining is used.Noisy data has bad

impact on performance of classifiers. Customer churn prediction is also performed by another research based on random forest with factor identification in telecom sector [26]. This work recognized the factors that are required to determine the churn root cause. This proposed working model was calculated by confusion matrix that help to evaluate accuracy, recall and precision. It was found that Random forest algorithm produced better accuracy. In some another studies, random sampling technique and SVM classifierwere suggested for customer churn prediction. The under-samplingtechnique are used to modify the class imbalance with balanced dataset. The performance of model cannot improve, if there exist class imbalance problem [27].

An ensemble learning algorithm is performed to achieve the better predictive performance of churn prediction model. It is provided in study [28]. After designing a number of decision trees at initial stage and display the status of the classification of the individual decision trees. An ensemble learning is an aggregation of several classifiers into one prediction model. Rotation forest and Rot boost are two classification methods. The new induced algorithm Adaboost to improve performance of prediction model. This study found the drawback of interpretability and understanding of churn factors. In literature a rough set theory approach is used to classify churn customer [29]. Another study on Naive Bayes classifier, that is used to evaluate performance of model that involves some extra variable. The estimated result is compared with C4.5 decision tree algorithm. This decision tree was showed better accuracy on dataset. So, it was found that it was not address the imbalance problem of initial datasets.

In this literature, there is existence of various hybrid techniques for churn prediction. With this research some techniques, named KNN, SVM, DT and ANN are compared to evaluate the accuracy performance.Thepresence of highly predictive variable will improve the performanceofclassification. But this study also introduced a set of variables that also introduced a new technique of dimensionality reduction in order to find out the essential set of variables involved in datamining process [31].

As we have reviewed all the classifiers algorithms, we concluded that random forest performs best results and then other algorithms like Decision tree, SVM, ANN, KNN.Thereisa lot of research workindatamining and applicationsof datamining focuses on somechurnprocess. But from this literaturesurvey,itisprovedthatthereexists some workbased onRule Induced algorithm for prediction of customerchurn. Ourfutureworkwill beevidential based Rule Induced techniques for multiple classontelecomdatasetsandcomparetheperformancewithsome probabilistic and non-probabilistic techniquesforchurnpredictionbecause theyareveryefficienttoolsfordataminingapplications.

Discussion: The confusion metrics is designed to compare accuracy, recall and precision for all the classifiers. This matrix helps to conclude the performance of different classification-based datamining algorithms used in churn prediction.

Performance matrix		Predicted record	
		Churn	Non churn
Actual record	Churn	TP	FN
	Non churn	FP	TN

Confusion matrix

Performance Measures: By applying performance measures on different classifiers weconcluded the accuracy, recall and precision [16].From above confusion matrix, the TP =true positive, TN= true negative, FP= false positive, FN= false

Using all terms, we evaluate different measures:

Accuracy**:** It evaluate the correctly classified records.

$Accuracy$: (TP+TN)/ $(TP + FP + TN + FN)$

Precision:It identified that which part of instances/record is positive.It was dealing with replacing the errors which can be computed with the following equation:

$TPPrecision = TP$ /(TP+FP)

Recall: It is the probability that all the required instances are selected as positive. It deals with completeness means instances predicted positive and correctly identified. It deals with substitution and deletion errors, [16] had stated that it can be calculated by the following formula:

$TPRecall$ = TP/ $(TP + FN)$

F Performance Measure: In FMeasure computed value is placed between correctly predicted records and finding data points ofevery class belongs to same class. It is computedas:

F-measure=2*(PRECISION*RECALL/PRECISION+RECALL)

Conclusion

Inthis review paper,classification model isdevelopedfor handling customerchurn in telecommunications.The proposed random forest classifier is used to handle customer complaints. The classification accuracy of non-probabilistic models (SVM and DT and KNN, Random Forest) is better than probabilistic methods. The performance of the suggested classifier (Random Forest) is better than some other ensemble classifiers (bagging and ad boosting). So,to improve customer satisfaction it is helpful to handle customer complaints efficiently and effectively. Limitation of algorithms used in this study are only binary classification is considered. In future, we will investigate multiple class classification methods for churn prediction. Moreover, this work can be explained to describe the behavior of churn customer by applying evidential reasoning based multiple class classification technique. We will use text mining using some framework.

References

[1] S. Babu, D. N.Ananthanarayananand V. Ramesh,"A survey on factorsimpacting churn in telecommunication using datamining techniques,"*Inc. Eng. Res. Technol*., vol. 3, pp. 1745-1748, Mar. 2014.

[2] M.Óskarsdóttir, C.Bravo, W.Verbeke, C.Sarraute, B.Baesens andJ. 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.

[3] W. Verbeke, D. Martens, C. Muse, and B. Baesens, ''Building comprehensible customer churn prediction models with advanced rule induction techniques,'' *Expert Syst. Appl.*, vol. 38, pp. 2354–2364, Mar. 2011.

[4] A. Amin et al., ''Customer churn prediction in the telecommunication sector using a rough set approach,'' *Neurocomputing*, vol. 237, pp. 242–254, May 2017.

[5] A.Idris and K.Asifullah,"Ensemble based efficient churn prediction model for telecom,"In 12th Confr. Frontiers of Information Technology, pp. 1–7.

[6] C.Kirui, L.Hong, W.Cheruiyot andH.Kirui, "Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining," *IJCSI International Journal of Computer Science Issues,* vol. 10, no. 2, pp. 165–172, Jan. 2013.

[7]M.Farquaad, V.Ravi and S. B. "Churn prediction using comprehensible support vector machine: An analytical CRM application.*" Applied Soft Computing*,pp. 19, 31–40, Sep. 2014.

[8]A. Ahmed and D.Maheshwari, "Churn prediction on huge telecom data using hybrid firefly-based classification,"*Egyptian Informatics Journal*, vol. 18 no. 3, pp. 215–220, Nov.2017, Available: http://www.sciencedirect.com/science/article/pii/S1110866517300403https://doi. org/10.1016/j.eij.2017.02.002.

[9]G. Nie, W. Rowe, L. Zhang, Y. Tian, and Y. Shi, ''Credit card churn forecasting by logistic regression and decision tree,'' *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15273–15285, Dec.2011.

[10] K. Coussement, S.Lessmann,and G.Verstraeten, "A comparative analysis of datapreparation algorithms for customer churn prediction: A case study in the telecommunicationindustry," *Decision Support Systems*, vol. 95, pp. 27–36, Available: https://doi.org/10. 1016/j.dss.2016.11.007.

[11] M.Farquad, V.Ravi and S. B.Raju, "Churn prediction using comprehensiblesupport vector machine: An analytical CRM application,"*Applied Soft Computing*, vol. 19, pp. 31–40.Jun. 2014.

[12] E. Shaaban, Y. Helmy, A. Khedr and M. Nasr, "A Proposed Churn Prediction Model," *International Journal of Engineering Research and Applications*. Vol.2, no. 4, pp.693-697.

[13] H.S. Kim and C. H. Yoon, ''Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market,'' *Telecommun. Policy*, vol. 28, no. 9–10, pp. 751–765, Nov.2004.

[14] T. Gairn-Muñoz, T. Pérez-Amaral, C. Gijon et al., "Consumer complaint behavior in telecommunications: The case of mobile phone users in Spain[J],"*Telecommunications Policy*, vol. 40, no. 8, pp. 804-820, Aug. 2016.

[15] H. Yu *et al.*, ''Feature engineering and classifier ensemble for KDD cup 2010,"Dept.Computer Sci. Inf. Eng., National Taiwan Univ., Taipei, Taiwan, Tech. Rep., 2010, vol. 1, pp.1-16.

[16] O.R.Zaïane, "IntroductiontoDataMining," CMPUT690PrinciplesofKnowledgeDiscoveryinDatabasesChapter,pp.1–15,Mar. 1999.

[17] A. Sharma and P. K. Kumar,"Neural network-based approach for predicting customer churn in cellular network services,"*International journal of computer application*, vol. 27, no. 11, [Online].Available:https://arxiv.org/abs/1309.3945.

[18] A.Amin,B.Shah,A.M.Khattak,T.Baker,H.RahmanDurani,andS. Anwar, ''Just-in-time customer churn prediction: With and without data transformation,'' in *Proc. IEEE Congr. Evol. Comput*. pp. 1–6, Jul.2018.

[19]Q.Zhao, X.Gao, A. Dongand X. Dong, "K-local maximum margin feature extraction algorithm for churn prediction in telecom,"*Cluster Compu*ting, vol. 20, no. 2, pp. 1401–1409, Jun.2017.

[20] A. Fade, E. Chang, M.Sabariet al., "Intelligent customer complaint handling utilizing principal component and data envelopment analysis," *Applied Soft Computing*, vol. 47, pp. 614-630.

[21] V. Umayaparvathi, K. Iyakutti, "Attribute Selection and Customer Churn Prediction in Telecom Industry," In confr. Data Mining and Advanced Computing, Mar. 2016.

[22]V. Umayaparvathi and K. Iyakutti, "Applications of Data Mining Techniques in Telecom Churn Prediction," *International Journal of Computer Applications*, vol. 42, no. 20, Mar.2016.

[23]S.Maldonado, Á.Flores, T.Verbraken, B.Baesens and R.Weber, "Profit-based feature selection using support vector machines - General framework and an application for customer retention,"*Applied Soft Computing Journal*, vol. 35, pp. 740–748. Oct. 2015.

[24] S.A.Qureshi, A.S.Rehman, A.M.Qamar, A.KamalandA.Rehman, "Telecommunications Subscribers Churn Prediction Model Using Machine Learning." *In Digital Information Management (ICDIM),* Eighth International Conference, 2013, pp. 131-136.

[25]A. Amin, F. AlObeidat, B. Shah, A. Adnan, J. Loo, and S. Anwar, "Customer churn prediction in telecommunication industry using data certainty," *J. Bus. Res.*, vol. 94, pp. 290–301, Jan.2019.

[26] Y.Xie,X. Li,E.W. T.Ngai& W.Ying, "Customer churn prediction using improved balanced random forests,*" Expert Systems with Applications,*vol. 36, no. 3, pp. 5445–5449. Available:https://doi.org/10.1016/j.eswa.2008.06.121.

[27] J.Burezand D. Van den Poel,"Handling class imbalance in customer churn prediction,"*Expert Systems with Applications*, vol. 36, no. 3, pp. 4626–4636, Apr. 2013.

[28]K.W. De Bock and D. Van Den Poel,"Reconciling performance and interpretability in customer churn prediction using ensemble learning based on generalized additive models,"*Expert Systems with Applications*, vol 39, Issue 8, pp. 6816–6826. Jun. 2014.

[29]A. Amin et al.*,* "Customer churn prediction in the telecommunicationsector using a rough set approach," Neurocomputing, vol. 237, pp. 242–254, May 2017.

[30] C.Kirui, L.Hong, W.Cheruiyotand H.Kirui,"Predicting Customer Churn in Mobile Telephony Industry Using Probabilistic Classifiers in Data Mining," *International Journal of Computer Science Issues.*vol. 10, no. 2, pp.165-172, Jan. 2013.

[31] A. Keramati, R. Jafari, M. Alinejad, I. Ahmadian, M. Mozaffari, and U. Abbasi, "Improved Churn Prediction in Telecommunications Industry Using Datamining Techniques," *Applied Soft Computing*. vol.24, pp .994-1012, Nov. 2017.