ISSN No. 0976-5697

Volume 8, No. 5, May – June 2017

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

# Customer Churn Prediction:A Survey

Ms. Chinnu P Johny
P.G Scholar
Department of Computer Science & Engineering
FISAT, Mookkannoor, Kerala, India

Mr. Paul P. Mathai
Assistant Professor
Department of Computer Science & Engineering
FISAT, Mookkannoor, Kerala, India

*Abstract:* Customer churn prediction is a core research topic in recent years. Churners are persons who quit a company's service for some reasons. Companies should be able to predict the behavior of customer correctly in order to reduce customer churn rate. Customer churn has emerged as one of the major issues in every Industry. Researches indicates that it is more expensive to gain a new customer than to retain an existing one. In order to retain existing customers, service providers need to know the reasons of churn, which can be realized through the knowledge extracted from the data. To prevent the customer churn, many different prediction techniques are used .The commonly used techniques are neural networks, statistical based techniques, decision trees, covering algorithms, regression analysis, kmeans etc. This paper surveys the commonly used techniques to identify customer churn patterns.

*Keywords:* - Customer churn; Customer Churn Prediction ; Data Mining Methods; Prediction Models;

## I. INTRODUCTION

In this competitive world, companies have to put much effort not only to convince customers but also to keep hold of existing customers. Churners are persons who quit a company's service for some reasons. Companies should be able to predict the behavior of customer correctly in order to reduce customer churn rate. Churners and non-churners can be differentiated using churn prediction which is a binary classification task. Obviously, customer churn figures the life time value of a customer in a company. This is done by analyzing the customer's lifetime profit to a company. Normally, it is easy to find from the analysis that most of the company's profits are contributed by frequent customers and the more expense is for attracting new customers than retaining the old ones. Therefore, finding the churners can help companies retain their customers and retaining the relationships with existing customers is of greater importance [1]. Churn can be both voluntary and involuntary. When existing customer leaves the company and joins competing company, voluntary churn happens. While in involuntary churn customer is asked by the company to leave due to reasons like non-payments etc. Voluntary churn can be sub-divided into: incidental churn and deliberate churn Incidental churn occurs, not because the customers had already planned but because of something happened in their lives e.g. a change in financial conditions, change in current location etc. Deliberate churn occurs for reasons of technology (customers wanting a newer or better technology, price , service quality factors, social or psychological factors and convenience reasons) The main reasons for customer churn are no satisfaction with the customer service, no understanding of the service plan ,high costs, unattractive plans, bad support. This is an expensive problem since acquiring new customer costs five to six times more expensive than retaining existing ones .The aim of customer churn prediction is to detect customers with high tendency to leave a company.

In general, Customer Churn means, the customers who are about to move their usage of service to a competing service provider. Different Churn prediction methods gives the prediction about customers who likely to churn in the near future and churn management help to identify such churners and give them some positive offers in order to reduce churn effect .These customers can be identified using their behavior , demographic details etc. [2]

Customer behavior can be recognized by the exact components of the services they are utilizing and how often it is employed. In telecommunication industry, the number and length of calls, period between calls, the usage of data, loyalty points etc, could be taken in order to determine the customer behavior. Customer demographic includes education, social status, geographical data, age, sex are also used for churn calculation.

In the last few years, there are revolutionary things have happened in the telecommunications industry, such as, new services, technologies and the liberalizations of the market opening up to competition in the market. Since the customer is the major source of profit, a method to promptly manage customer churn gains vital significance for the survival and development of any telecommunication company [3]. For many telecoms companies, figuring out how to deal with churn is turning out to be the key for continued existence of their organizations.

The remainder of this paper is organized as follows: Section II introduces the different churn prediction methods briefly. Related works in the literature of churn prediction are analyzed in Section III. Finally, a brief conclusion is given in Section IV.

## II. BACKGROUND DETAILS

To successfully manage the churn prediction challenge, different data mining methods are put forward by different researchers. The main data mining methods are based on neural networks, statistical based techniques, decision trees, covering algorithms, Regression Analysis, K-means etc. Decision tree, neural network, and k-means are selected by [4] as main techniques to build predictive models for telecom customer churn prediction

### A. *Neural Networks*

The use of neural networks in churn prediction has a big asset that is the likelihood of each classification made can be determined. The neural networks outperform decision trees for prediction of churn[6]. They state the biggest disadvantage of neural networks they do not uncover patterns in an easily understandable form, categorizing them as a 'black box' model. The basic idea behind neural networks is that each attribute is associated with a weight and combinations of weighted attributes participate in the prediction task. During learning the weights are constantly updated, thus correcting the 'effect' which an attribute has. Given a customer data set and the set of predictor variables the neural network tries to calculate a combination of the inputs and to output the probability that the customer is a churner.

.

### B. *Decision Trees*

For predictions and classification of future events, decision trees are the commonly used. In order to develop such trees, two major steps are considered: building and pruning. During the first phase ie building, the data set is partitioned recursively until most of the records in each partition contain identical value. The second phase ie pruning, then removes some branches which contain noisy data (those with the largest estimated error rate)

Each node in a decision tree is a test condition and the branching is based on the value of the attribute being tested. The tree is representing a collection of multiple rule sets. When evaluating a customer data set the classification is done by traversing through the tree until a leaf node is reached. The label of this leaf node (Churner or Non Churner) is assigned to the customer record under evaluation.

### C. *Regression Analysis*

Regression is considered to be a good technique for identifying and predicting customer satisfaction. For each of the variables in a regression model the standard error rate is calculated using SPSS. Then the variables with the most significance in respect to linear regressions for churn prediction are obtained and a regression model is constructed.

Since the prediction task in churn prognosis is to identify a customer as a churner or non churner and therefore the prediction attribute is associated with only two values logistic regression techniques are suitable. While linear regression models are useful for prediction of continuous valued attributes, logistic regression models are suitable for binary attributes.

### D. *K-means*

K-means clustering is an algorithm to classify or to group the different objects based on attributes or features into K number of group. K-means is one of the simplest unsupervised learning algorithms. K centroids are defined for K clusters which are generally far away from each other. Then the elements are grouped into clusters which are nearer to the centroid of that cluster. After this first step, again the new centroid for each cluster is calculated based on the elements of that cluster. The same method is followed, and the element is grouped based on new centroid. In every step, the centroid changes and elements move from one cluster to another. The same process is followed till no element is moving from one cluster to another.

## III. RELATED WORKS

This section gives an analysis on the various works that have been proposed in the area of churn prediction, stating both their merits and demerits.

Chih-Ping Wei [5] proposes a churn-prediction technique that predicts the churn rate from subscriber contractual information and call pattern changes extracted from call details. For a specific prediction time-period, the proposed technique is capable of identifying potential churners at the contract level. In addition, the proposed technique incorporates the multi-classifier class- combiner approach to address the challenge of a highly skewed class distribution between churners and non-churners. The proposed method exploit the use of call pattern changes and contractual data for developing a churn-prediction technique that identifies potential churners at the contract level .This method outperforms the single classifier approach. But it only considers few variables and cannot process large inputs.

Wai-Ho Au [6] proposes a new data mining algorithm, called data mining by evolutionary learning (DMEL), to solve classification problems. DMEL algorithm searches through the possible rule space using an evolutionary approach which has got the following steps:

1) The process starts with the generation of an initial set of first-order rules using a probabilistic induction technique and based on these rules obtained, rules of higher order (two or more conjuncts) are found iteratively.
2) An objective interestingness measure is used for identifying interesting rules
3) The fitness can be defined in terms of the probability that the attribute values of a record can be correctly determined using the rules it encodes.
4) The likelihood of predictions are estimated so that customers can be ranked according to their likelihood to churn.

This method provides accurate classification and is capable of finding both positive and negative relationships among attributes without user input. But this method is also incapable of handle data of higher dimension.

Bart Larivie're [7] propose a random forests techniques in order to predict customer's profitability evolution and the customer's next buy and partial-defection decisions. Two types of random forests techniques are introduced to analyze the data: random forests which are used for binary classification and the regression forests that are applied for the models with linear dependent variables. Research findings of this paper demonstrate that both random forests techniques provide better fit for the estimation and validation sample compared to ordinary linear regression and logistic regression models. But this method does not consider the correlation between the variables.

Kristof Coussement [8] proposes two aspects in which churn prediction models could be improved by (i) relying on customer information type diversity and (ii) choosing the best performing classification technique. This study contributes to the literature by finding evidence that adding emotions expressed in client/company emails increases the predictive performance of an extended RFM churn model. An in-depth study of the impact of the emotionality indicators on churn behavior is done as a substantive contribution.The studies shows that the prediction performance is improved. But the service recovery measures not considered.

Chih-Fong Tsai [9] proposes the important processes of developing MOD customer churn prediction models by data mining techniques. This method contain the preprocessing stage which helps to select important variables by association rules. The model construction stage is constructed by neural networks(NN) and decision trees (DT)and four evaluation measures including prediction accuracy, precision, recall, and F-measure, are used to examine the model performance. This method shows that decision trees performs better than NN model when association rules are used. Here time series and trend analysis is not considered

Adem Karahoca [10] proposes an integrated diagnostic system for the churn management application presented is based on a multiple Adaptive Neuro Fuzzy Inference System with fuzzy c-means. By usage of series of ANFIS units helps in reducing the scale and complexity of the system and speeds up the training of the network. This system can be applied to a range of telecom applications where monitoring and management is required continuously. The results of this paper proves that, ANFIS method combines both precision of fuzzy based classification system and adaptability (back propagation) feature of neural networks in classification of data.

Based on the accuracy of the results shown in this paper, it can be stated that the ANFIS models can be used as an alternative to current CRM churn management mechanism (detection techniques currently in use). This method can be applied to different telecom networks or other industries .And once it is trained, it can then be used during operation to provide instant detection results to the task. One disadvantage of the ANFIS method is that the complexity of the algorithm is high when there are more than a number of inputs fed into the system. However, it can be used efficiently against large datasets when the system reaches an optimal configuration of membership functions.

Yi-Fan Wang [11] proposes a recommender system for wireless network companies to understand and avoid customer churn. To ensure the accuracy of the analysis, the decision tree algorithm is used to analyze data of over 60,000 transactions and of more than 4000 members, over a period of three months. The training data is the first nine weeks data, and testing data is the data of last month. The experimental results are found to be very useful for making strategical recommendations to avoid customer churn. system. The proposed system, users can gain recommendations from the system. The process is application-oriented, so different applications may need different classification approaches as appropriate. Here the Decision Tree classification approach is used. This method helps in maintaining congenial customer relationships. But this approach cannot process inputs with higher dimension.

Ning Lu [12] proposes churn prediction model in the telecommunication industry using a boosting algorithm which is believed to be very robust and has demonstrated success in churn prediction in the banking industry . The established literature only uses boosting as a general method to boost accuracy, and few researchers have ever tried to take advantage of the weight assigned by boosting algorithms. The weight also provides important information, specifically, outliers. The testing results show that boosting provides a good separation of the customer base, which also leads to a better overall performance. Here it provides a distinct prediction model for each cluster. But the reason for customer churn is not identified and large inputs cannot be processed.

Emiliano G. Castro [13] proposes a frequency analysis approach for feature representation from login records for churn prediction modeling. These records (from real data) were converted into fixed-length data arrays using four different methods, and then these were given as input for training probabilistic classifiers with the nearest neighbors machine learning algorithm.Using predictive performance metrics, the classifiers were then evaluated and compared. Probabilistic classifiers based on the -nearest neighbors ( -NN) algorithm using four distinct feature representation methods for the login records:

1) RFM-based
2) Time domain
3) Frequency domain
4) Time frequency plane domain

The studies shows that this method provides significant economic gain and the predictive performance is very reasonable. But here this method focus solely on login records which limits its predictive performance.

Yiqing Huang [14] makes two main contributions. The primary contribution is an empirical demonstration that indeed churn prediction performance can be significantly improved with telco big data by integrating both BSS and OSS data. Although BSS data have been utilized in churn prediction very well in the past decade, we have shown that it is worthwhile collecting, storing and mining OSS data, which takes around 97 This paper presents a churn prediction system which is one of the important components for the deployed telco big data platform in one of the biggest operators, that is China. The second contribution is the integration of churn prediction with retention campaign systems as a closed loop. After each campaign, identifies which potential churners accept the retention offers, which can be used as class labels to build a multi-class classifier automatically matching proper offers with churners. This means that reasonable campaign cost to make the most profit. This method can process large input but not able to identify the reason for customer churn.

Hui Li [15] proposes to leverage the power of big data to mitigate the problem of subscriber churn and enhance the service quality of telecom operators. The telecom operators have collected a huge volume of valuable data on customer/subscriber behaviors, service usage, and network operations. For efficient big data processing, first build a dedicated distributed cloud infrastructure that integrates both online and online processing capabilities; Second, a complete churn analysis model based on deep data mining techniques, is developed and utilize inter-subscriber influence to improve prediction accuracy. Here also large data can be processed. But information from complex attributes are not integrated.

Wenjie Bi [16] proposes a new clustering algorithm called semantic-drive subtractive clustering method (SDSCM). Then, a parallel SDSCM algorithm is implemented through a Hadoop MapReduce framework. The three main contributions of this paper are as follows. First, a new algorithm called SDSCM is proposed, which improves clustering accuracy of SCM and k-means. And, this algorithm decreases the risk of imprecise operations management using AFS. Second, to deal with industrial big data, proposes a parallel SDSCM algorithm through a Hadoop MapReduce framework. Third, the results show that the parallel SDSCM and parallel k-means have high performance, when compared with other traditional methods. This method improves clustering accuracy and reduces the risk of imprecise operations management. But doesn't give any retention offers.

## IV. CONCLUSION

To successfully manage the churn prediction challenge, different data mining methods are put forward by different researchers. The main data mining methods are based on neural networks, statistical based techniques, decision trees, covering algorithms, Regression Analysis, Kmeans etc. This paper provides a detailed study on the methods used for the process of customer churn prediction. Each of the above churn prediction models has its own advantages and disadvantages. Hence a good prediction model is required in order to avoid the customer churn problem. This can be achieved by considering a method which can process large inputs with higher dimension and complex attributes for future work for Churn prediction. Good prediction models have to be constantly developed and a combination of the proposed methods has to be used.

## V. REFERENCES

[1] T.Vafeiadis, K.I. Diamantaras, G. Sarigiannidis, K. Chatzisavvas" Customer churn prediction intelecommunications", Simulation Modelling: Practiceand Theory 55 (2015) 1-9

[2] V. Lazarov and M. Capota, "Churn Prediction," TUM Comput. Sci., 2007

[3] W. Verbeke, D. Martens, C. Mues, and B. Baesens, "Building comprehensible customer churn prediction models with advanced rule induction techniques," Expert Syst. Appl., vol. 38, no. 3, pp. 2354-2364,2011

[4] S. Y. Hung, D. C. Yen, and H. Y.Wang, "Applying data mining to telecom churn management," Expert Syst. Appl., vol. 31, no. 3, pp. 515-524, Oct. 2006.

[5] Wei C., Chiu I., "Turning telecommunications call details to churn prediction: a data mining approach", ExpertSystems with Applications, Elsevier, pp. 103-112, vol. 23, no. 2, Aug.2002

[6] W. H. Au, K. C. C. Chan, and Y. Xin, "A novel evolutionary data mining algorithm with applications to churn prediction," IEEE Trans. Evol. Comput., vol. 7, no. 6, pp. 532-545, Dec. 2003.

[7] Bart Larivie re, Dirk Van den Poel, "Predicting customer retention and profitability by using random forests and regression forests techniques" Expert Systems with Applications 29 (2005) 472-484

[8] Kristof Coussement, Dirk Van den Poel "Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers" Expert Systems with Applications 36 (2009) 6127-6134

[9] Chih-Fong Tsai, Mao-Yuan Chen, "Variable selection by association rules for customer churn prediction of multimedia on demand", Expert Systems with Applications 37 (2010) 2006-2015.

[10] Adem Karahoca, Dilek Karahoca,"GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system", Expert Systems with Applications 38 (2011) 1814-1822

[11] Yi-Fan Wang a, Ding-An Chiang b, Mei-Hua Hsu c,*, Cheng-Jung Lin b, I-Long Lin d "A recommender system to avoid customer churn: A case study", Expert Systems with Applications 36 (2009) 8071-8075

[12] N. Lu, H. Lin, J. Lu, and G. Zhang, "A customer churn prediction model in telecom industry using boosting," IEEE Trans. Ind. Informat., vol. 10, no. 2, pp. 1659-1665, May 2014

[13] E. G. Castro and M.S.G. Tsuzuki, "Churn prediction in online games using players' login records: A frequency analysis approach," IEEE Trans. Comput. Intell. AI Games, vol. 7, no. 3, Sep. 2015

[14] Y. Huang et al., "Telco churn prediction with big data," in Proc. ACM SIGMOD Int. Conf. Manage. Data, San Francisco, CA, USA, 2015, pp. 607-618

[15] H. Li, D. Wu, and G. X. Li, "Enhancing telco service quality with big data enabled churn analysis: Infrastructure, model, and deployment,"J. Comput. Sci. Technol., vol. 30, no. 6, pp. 1201-1214, Nov. 2015

[16] Wenjie Bi, Meili Cai, Mengqi Liu, and Guo Li"A Big Data Clustering Algorithm for Mitigating the Risk of Customer Churn" IEEE Transactions on Industrial Informatics, Vol. 12, No. 3, June 2016