STM JOURNALS

Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720 (Online)
Volume 6, Issue 2
www.stmjournals.com

# Predicting Churners from the User Base of a Business: A Review

***Saima Jan[1,*], Afaq Alam Khan[2]***
[1]M.Tech. Student, Department of Information Technology, Central University of Kashmir, Ganderbal, Jammu and Kashmir, India
[2]Department of Information Technology, Central University of Kashmir, Ganderbal, Jammu and Kashmir, India

***Abstract***
*Customer churn prediction and management is a foremost job aimed by enterprises to maintain cost-effective customers and thus many customer churn prediction models are developed which aim to categorize consumers as non-churners or churners and then retention actions are taken to prevent churners from churning. The data from different companies like telecommunication companies, financial companies, retail companies etc. can be used to discover valuable patterns which predict the customers' relationship with the company. From this data, customers' behavior can be used to predict his loyalty with the company. In this study, different techniques and approaches are discussed given by different researchers to predict churners, so as to apply retention strategies on them. This study also presents the results that were recorded by various researchers in their research.*

***Keywords**: Churn prediction, data mining, customer lifetime value, customer relationship management*

***Author for Correspondence** E-mail: jansaima119@gmail.com

## INTRODUCTION

In the present world of high competition when a person gets so many offers from different companies to use their product or service, it becomes necessary for companies to focus on their customer base. This can be done by preventing the existing customers from leaving their company as well as by attaining new customers. Customer churn is a term used when a buyer ceases his or her connection with a corporation. It is the movement of the customer commencing one company to a new one in the quest of improved facilities or rates or the remunerations provided by the another company. So, for the company to survive in such a competitive environment, it is essential for the company to identify these customers who may churn and then taking proper retention actions like providing incentives or some other offers to prevent them from churning. The churn rate can be calculated as the percentage of customers leaving the company over some specified period of time. Losing only few customers is inevitable but a high rate of churn is costly. The cost of attaining new customers by different ways like advertising, promotional costs etc. is more as compared to retaining existing customers, as making trade from new customers involves functioning all the way over the sales, exploiting your promoting and sales assets all through the procedure. Customer holding is mostly more worthwhile as by this time we have received the faith and fidelity of prevailing customers. Thus companies attempt to win back these churning customers, because recovering long term customers can be profitable than attaining new ones.

### Churning Types

Churn can be voluntary or involuntary.

- In case of voluntary churn, a customer makes a decision towards switching to some other company.
- While an involuntary churn occurs due to different reasons like relocation of customer to some long term care facility, death of customer or relocation to some other far away location etc.

Thus the companies need to identify the type of churn and thus should work on only voluntary churners as it occurs due to the relationship between company and customer and the company can work on improving such kind of relationship. Thus the primary focus should be finding the reason behind parting of customer or identifying the customers with high risk of leaving.

Customer churn hinders the growth of company; thus companies should have some defined method to detect the churners. By this they not only keep check on their customers i.e. whether they may churn or not but also identify the reasons of the churn and determine their customer retention strategies as well as success rates and hence identify methods for improvement.

**Methods Employed by the Companies to Determine the Churners**

Some of the methods which can be employed by the companies to determine the churners include:

***Customer Lifetime Value***

It can be demarcated as the current worth of forthcoming currency movements generated by the purchaser's product usage, without captivating into account formerly consumed costs. Value of a purchaser is the estimated figure of reduced impending incomes where a purchaser produces a boundary for each interval. All the transactions a purchaser will render on behalf of the various merchandises the corporation markets, is referred to as CLV. CLV of a purchaser is the totality of net cash flows generated by the transaction on some product say p after it is discounted at some rate say r. CLV can thus be used as a tool for customer selection i.e. it can be used to determine whether a customer is churner or not.

***Customer Relationship Management***

It is a method of gaining, holding and associating with discerning consumers to generate greater worth for the enterprise and the consumer. Since enterprises mainly rely on the profits earned from customers, so CRM should be an important focus for the enterprises to make decisions about their customers i.e. most profitable customers can be recognized with the help of CRM. It is used to determine the potential value of customer's risks associated which hinders him from paying his bills and foresee their future needs. CRM can be analytical or operational. Operational CRM is the automation of industry method, and analytical CRM means to analyze the behavior of consumer so as to help the company in effective allocation of resources.

***Data Mining Techniques***

Since the data available today is very huge, there must be some method to extract valuable patterns from this huge data. This can be done with the help of mining techniques. Data mining is a method to unearth formerly unfamiliar, permitted and action taking patterns or valuable information from big databases for critical commercial judgment support. The objective of mining the data is description plus prediction. Description is used to describe the data i.e. to discover patterns which are easily interpretable by humans from the data; and prediction is employed to predict the hidden or forthcoming prices of entities of concern. Prediction and description can be attained by various mining methods like classification, clustering, regression, dependency analysis, etc.

In this study, we are trying to briefly describe a number of related states of art methodologies used in different but correlated studies composed into a single location.

## RELATED WORK

Customer is an important entity for enterprises and should be of primary focus for them. The data available to the companies about the customers can be used to study the behavior of the customers. The companies should collect the data about their customers. The data can be collected by constantly monitoring the customers in order to reveal information about them. The information gathered is required to find the buying behavior of the customers. From this information the profitable customers can be detected. Also the customers whose loyalty is decreasing can be determined. For such customers i.e. churners, retention action is needed. The retention action should be taken for only those customers who are profitable to the company. Thus in order to determine the profitable churners and then taking retention action for them, many strategies are followed by different enterprises and much work is done in this regard. If we observe the available literature, many studies have revealed different strategies to predict the churning customers, some of them are presented below:

**Hoppner [1]**

Hoppner employed the telecom operator consumer churn dataset from southern region of Korea that included 889 operators and 10 features [1]. Among them, 277 operators came out as churners. They took profit into consideration and built a model for prediction of would be churners. After detecting churners, retention action would be taken accordingly. If we talk of machine learning techniques to classify operators into two categories i.e. non-churners or churners, it is simply a binary classification problem. Based on this, many predictive classifiers are built and then the one with better performance is selected by taking into account measures like Area under receiver operating curve (AUROC). But these models do not consider misclassification costs or the benefits obtained by the correct classification of the same. So to achieve this, an expected maximum profit measure for customer churn (EMPC) is developed recently. In this paper, a new classifier is presented that combines the EMPC metric for model construction directly. That classifier is ProfTree. The result of this study shows that the performance of ProfTree is much enhanced than those of classic methods.

**Glady *et al*. [2]**

Glady *et al*. used customer lifetime value (CLV) for predicting churning behavior of customers [2]. They defined a loss function in order to predict the loss/profit suffered by a misclassification. They first gave a product centric definition of churn by using product usage of a customer and from that they arrived at customer centric one in which overall product usage of customer is used to define churner. According to them, a churner is thus any person whose customer lifetime value is diminishing with time. They determined that to miss a small number of cost-effective customers incurs greater loss for a company than losing many non-profitable ones. They used the dataset of a financial service company located in Belgian in which they took a section of 10,000 consumers for a time period of 9 months from Jan 2004 to September 2004. They considered two diverse product practices i.e., the overall debit transactions and the overall sum debited in month. They divided the total data into two sets i.e. 66% training set which comprises of item trades from January 2004 to the month of June in the same year for designing classifiers and test set of identical consumers from July 2004 to the month of September, 2004 for checking the performance. They measured the performance of classifiers i.e. logistic regression, neural networks, decision trees, Ada cost and cost-sensitive tree by measuring different parameters i.e. AUPROC, percentage of correctly classified observations (PCC), AUROC, cumulative profit percentage and percentage of real positives. The result of their research was that cost sensitive tactics accomplish much better outcomes for not only classification but also correctly classifying the profitable churners.

**Jahromi *et al*. [3]**

Jahromi *et al*. used CRM as a tool to determine the possible churners [3]. They used the call record data i.e. date, time, type and duration of call and fee of 34504 customers of an Iranian mobile telecommunications service provider company when observed from the time period of 2007 November to 2008 April. This is a non-contractual company as customers do not have any contract with the company which can terminate or expire. In such conditions, predicting churn is a complex problem. For constructing a model in such a case they found that if persons are divided into groups on the basis of whether they use their cellphone for short period of time, then the persons who use their cellphones occasionally will be wrongly predicted as churners and if longer time period is used, then the model may not be able to predict real churners. So they divided the process of constructing the model into two parts i.e., clustering part for arriving at the meaning for concept of consumer churn and second part is classification for predicting churners. They extracted the features from the available data which are RFM related features for dividing customers into four individual clusters based on different maximum distances giving four groups of observation length, retaining length and estimation length used for these four clusters. From these clusters, some new features like minutes of incoming and outgoing calls, frequency of use of incoming and outgoing calls etc. were extracted. By utilizing these features, a predictive model was constructed using decision tree technique for each cluster. Then by using gain measures, performance of predictive model was confirmed as satisfactory. Also in order to deal with imbalanced dataset the learning methods which are cost sensitive (weiss) were also verified and their effectiveness was confirmed.

**Khan *et al*. [4]**

Khan *et al*. used the dataset provided by ISP company (Sepanta Co.) of Iran [4]. Three tables containing records related to subscribers, their usage and their cards were used. These tables were reduced to 9148, 2093120 and 61993 records respectively. Customer who was recorded on some date was observed after a time period of 6 months that whether he is still present there or not. If he is not present, then he was considered as a churner. The features employed for this study were either billing related or demography based or usage based. Also three classifiers were employed for testing of these features i.e. decision tree, logistic regression and neural network. Results show that neural network performs slightly better than others. Also a dependency network showing the importance of features was formed using a decision tree. The main focus of their study was to find the importance of features used. The result shows that billing and usage related features are important for churn prediction. However, demographic features have little effect on predicting churners.

**Liu *et al.***

Liu *et al.* [6] emphasized on churn prediction in case of mobile games. They used two large real datasets which were obtained by collecting data from Samsung game launcher platform independently for consecutive 4 months from users in Korea and United States of America. From this data, labels and features for training model were extracted. The motive was to detect user behavior to predict whether they are churners. In this regard, five solutions to predict churn were presented. Taking into account limitations of traditional methodologies, a new inductive and semi-supervised model was devised. This model learns embedding and prediction function for application relationships. The embedding function and prediction function are modeled using deep neural networks. A unique technique of these deep networks is used to get the relationship dynamics and contextual information. By considering similarity between attributes and adjacency in topology, a new random walk technique was designed. Then the performance was measured by using the collected data. It was observed that this method is superior to other existing prior methods.

**Burez *et al.***

Burez *et al*. [5] worked on handling imbalanced data as nowadays, complex factual problems occurring in the world are considered, the difficulty of imbalanced data occurs. When we talk of rarity, six data mining problems occur as given by Weiss when mining imbalanced data, and there are 10 methods given by Weiss to address them. For churn datasets three of the six problems were addressed by Burez *et al*. i.e. improper evaluation metrics, relative lack of data and noise. Six real life proprietary churn modeling datasets of Europe were used. Of those six datasets, five are contractual and one is non-contractual. Effect of sampling, cost sensitive learner and boosting were investigated distinctly for logistic regression and also for random forests. For both of them, averages over 5×2 CV of lift and area under curve are calculated for number of samples and number of cases. The results show that among them, the good valuation metrics is lift and area under curve. Predictive performance is not increased by the under-sampling method CUBE. Among weighted random forests and random forests, weighted one performs better

ISSN: 2395-6720 (Online)

and is thus recommended. But it is better to compare it with logistic regression. Also, among these classifiers, boosting is strong but not on any occasion it outperforms any other method.

**Tsai *et al.* [7]**

Tsai *et al.* used a CRM dataset of a telecom companies of America [7]. It mainly emphasizes on the duty of consumer churn prediction. This dataset consists of 51306 subscribers, in which, 34761 are churners and 16545 are non-churners in the time period of July 2001 to January 2002. The subscribers included those customers who are for at least 6 months with this telecom company. Churn prediction was done during the period of 31–60 days, i.e. whether for that period customer left the company or not. The paper divides the dataset into five equal parts, in which for model training, 80% of dataset is used and rest for testing that model by using fivefold cross validation. The paper started by first using the real dataset for training a multi-layer perceptron neural network to make it function like a baseline ANN model for associations. By using four different learning epochs and five distinct hidden layer nodes, 20 distinct ANN representations are developed for comparing them. First hybrid design consists of two ANN models; among them, one model is used for reduction of data and the 2nd one for prediction of churn. Here by using the real dataset for training, the first ANN model is tested. The result consists of both incorrectly and correctly foretold data from that training set. The data which is incorrectly predicted is viewed as outliers and appropriately foretold data by first ANN model is employed to train 2nd ANN model as prediction archetypal for advanced guess. The result expresses that hybrid ANN model performs better than a single baseline ANN model. For construction of 2nd hybrid model, they combined SOM with ANN. Firstly, [2×2], [3×3], [4×4], and [5×5] SOMs are used for clustering of data. Among them, 3×3 SOM performs best by providing the maximum rate of accurateness for the two given clusters viz. churners and non-churners. The result shows SOM and ANN hybrid models act is not superior to baseline ANN model. Thus the final result of their research was that ANN + ANN model achieves superior results than both baseline ANN model as well as SOM + ANN hybrid model.

**Verbeke *et al*. [8]**

Verbeke *et al*. used 11 datasets from wireless telecom operators across the globe [8]. The performance of various data mining classification algorithms is studied which comprises of rule based classifiers, decision tree approaches, neural networks, nearest neighbor, ensemble methods and classic statistical methods. Also, SVM and LSSVM are applied using radial and linear basis function kernels. The choice of classification technique has significant impact on the analytical control of subsequent model. Performance of classification archetypal is measured by means of AUROC and top decile lift. Also, a new profit centric metric for measuring performance is used. It is seen that top decile and lift measures can give suboptimal model choice and there may be loss in earnings. The result also shows that generally, decision trees yield the best performance. The preprocessing techniques for data like oversampling and variable selection are tested in the benchmarking experiments. It is found that outcome of oversampling is influenced by classification system used and dataset and can be positive or negative. Thus an empirical approach should be adopted and effect of oversampling should be consistently tested to check whether it is beneficial or not. In this paper, a model is developed and evaluated by measuring the total earnings that can be produced by using guesses of this model together with optimum section of consumers in a retention operation. The result of wide benchmarking experimentation indicates that both, use of highest profit measure and enhancing the involved portion of consumers to pick a classification model produce worthy cost saving. For developing the customer churn prediction model, a process model was formed consisting of four steps viz. data gathering and picking applicant descriptive variables. The consequential data is scrubbed and preprocessed. 2nd step encompasses actual building of a model. The model training is encompassed in this step. In the 3rd step, implementation of model is done by giving it to a professional expert to check the correctness of model and whether it is consistent with business facts. After that, an archetype is build and finally once final model with satisfactory performance is built, it is reviewed regularly with the intention of assessing whether that model still accomplishes good. The datasets used are divided into two parts 2/3rd part is used for training and 1/3rd for testing. The number of variables used for each dataset is reduced to a maximum of 20. Four categories of features are recognized in data set i.e. social and demographic data, call activities measurements or usage features financial information and marketing related variables. Input selection is important criterion for achieving noble performance and generally 6 to 8 features are sufficient to guess results with greater accuracy, selected variables will predict usage attributes of data in a better way. It can be also seen that for designing effective retention campaigns and strategies, clarity or interpretability is reflected as a significant feature of a classifier.

**Wei *et al*. [9]**

Wei *et al*. depicted that for thought-provoking issues like prospect profiling, fraud detection, churn prediction and management, several corporations are revolving around data mining techniques [9]. They used the data set of mobile telecommunications company of Taiwan with 21,000,000 subscribers. The variables used were payment type, length of service and contract type. To describe call patterns, they used frequency of use, minutes of use and sphere of influence parameters. The number of features used for the construction of a model depends on observation period which is divided into different sub periods to represent changes in call pattern. It can be depicted that if classification algorithm uses more number of sub periods, more will exist the chance towards detecting churners but the sub periods should not be that much that it may lead to impairment in the generalization and prediction effectiveness. The different classifiers used include decision rule, CN2, neural networks and decision tree. Also, a multi classifier class combiner approach is used to state the greatly skewed distribution problem. Here several experts who are not perfect are involved in decision making practice to make discrete recommendations. These recommendations are then jointly combined by alternative expert to make a concluding decision. It was found that multi-classifier method outperforms the single classifier method. Also, the planned churn prediction system which was based on call behavior gave acceptable predictive efficiency if more fresh call details would be given to the prediction model. Also, by using churn prediction systems based on past demographics as reference, evaluated lift factors were also reflected acceptable.

**Au *et al*. [10]**

Au *et al*. used DMEL (data mining by evolutionary learning) to predict churn [10]. A database of 100,000 subscribers is used to reduce churn rate. Among them, some had already churned and for others they need to predict churners by mining this database to prevent further subscribers from churning. Also, probability of churning is also calculated so as to take retention action on high probability churners first and then on low probability ones. Thus DMEL is used to extract rules which represent patterns of churning and to determine probability that does a subscriber churns in coming future or not. It is also used to handle missing values effectively. For evaluating the result of DMEL, it is given to numerous databases existing in real world. Those databases are allocated to two datasets. The records in both datasets are arbitrarily selected. The mining rules were accomplished on training set and testing set was set aside for testing. In every single test set, the results of one of the features were detected and then calculated by using the rules mined from training set. These two values are then compared to see whether they are same. The accuracy is measured by using DMEL, SCS, C4.5 decision tree, Michigan-style model and GABL, a Pittsburgh style learner. When test result is not matched to its removed value by any of the above classifiers, the record is assigned to majority class in training set and if it is in step with more than single rule of diverse classes, again it is assigned to majority class among them. Among all the databases, performance of different classifiers shows that DMEL performs better than other approaches i.e. DMEL is a robust way for determining churners. Also a lift curve was plotted which indicated that if some portion of all subscribers are communicated, then all churners can be trapped irrespective of restricted resources.

**Gupta *et al.* [11]**

Gupta *et al.* reviewed various implementable customer lifetime value models [11]. These CLV models are beneficial for segmenting the marketing besides allocating resources to gain customers, holding them and for cross-selling. Many companies use CLV as a device to measure and achieve the feat of their business. For companies to do effective marketing, traditional promoting metrics like awareness about brand, outlooks, share and sales are not enough, neither aggregate financial metrics like stock price or collective profit of corporation are enough. There should be some disaggregate metric which can be used at customer level to determine the nature of each individual i.e. whether they are profitable or not and hence whether should be kept or fired. Such a disaggregate metric is CLV which is used for determining these profitable customers and allocating resources accordingly. From CLV, CE, i.e., customer equity can be determined which is a decent proxy of complete firm's value. In this paper, a conceptual framework was given for modeling CLV which shows that by the marketing programs of a firm, customer base is affected i.e. customer acquisition, retention and cross-selling effects is influenced by marketing actions which in turn affects CLV and CE and finally the CLV and CE forms a proxy for firms value. Different models are highlighted in this study.

- First category of models is RFM models which create clusters of consumers based on frequency, recency and monetary values of their past purchases. It can be seen that response rates of consumers vary maximum by recency and then by frequency of purchase, followed by monetary values. This model is used to guess the future behavior of a purchaser and is thus directly related to CLV. However, they predict behavior for only consecutive period not for 3rd, 4th and so on periods. But RFM model is easy to implement and so is preferred more. CLV models perform better than RFM models; however RFM variable can be used to build CLV models.
- Second category is probability model. The behavior of a customer varies according to some probability distribution. It requires only recency and frequency of use variables. The probability can be given that whether a customer is yet active and probable number of transactions in some observed period. The model can be used as an input to CLV or to derive expression for CLV.
- Third model is econometric model. Econometric models share the same philosophy as that of probability models. Customer acquisition, retention and expansion are modeled and then combined for CLV estimation.
- Fourth model is persistence model. These models concentrate on modeling behavior of its constituents i.e. acquiring, holding and cross-selling. Long time duration is considered and in that time period, it is determined how one variable affects the other. Different methods for increasing customer base are examined to examine difference in CLV.
- Fifth model is computer science model. Many computer science models perform better when we use them for studying customer churn. These models can be used for determining CLV.
- Sixth model is diffusion/growth model. CLV can be used to determine long run profitability of a customer. However, it can also be combined to give a metric which can prove beneficial for senior managers. That metric is CE which gives CLV of existing and forthcoming customers.

The generalized system implementation for predicting churners can be shown as in Figure 1.

Summarization of the previous work is done in Table 1:

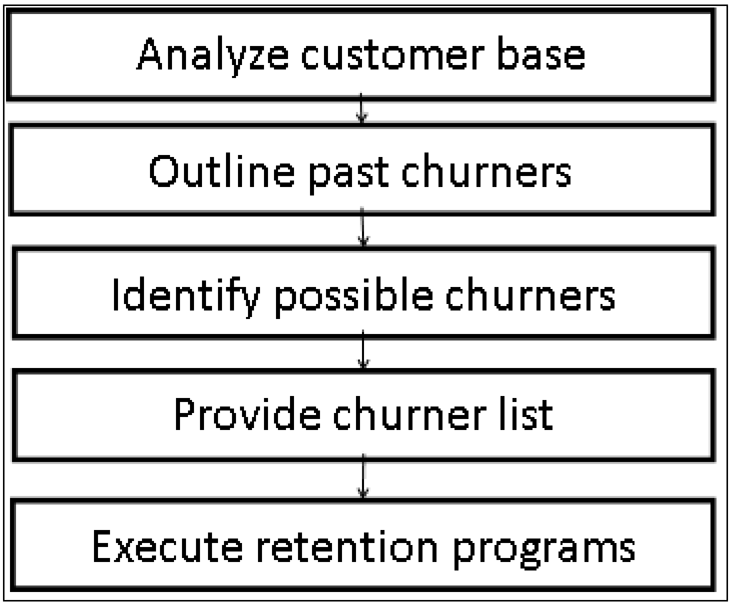


***Fig. 1.*** *Overview of Churn Prediction.*

***Table 1:*** *Summarization of Results Obtained.*

| Prior Work | Year | Dataset | Approaches Used | Results |
|---|---|---|---|---|
| Hoppner *et al*. [1] | 2017 | South Korean Telecom operator | ProfTree using EMPC | Enhanced performance of ProfTree from traditional methods |
| Glady *et al.* [2] | 2009 | Belgian financial service company | Cost sensitive classifier i.e. Ada cost, decision trees, Logistic regression, neural networks | Cost sensitive methods attain great results as compared to traditional ones |
| Jahromi *et al*. [3] | 2010 | Iran's mobile telecommunication company | Two step clustering technique to develop Predictive model | A predictive model to evaluate churners |
| Khan *et al*. [4] | 2010 | Iran's ISP company | Decision tree, logistic regression and neural networks | Billing and usage based features are important for churn prediction but not demographic related features |
| Xi-Liu et al. [6] | 2018 | Two Datasets collected from Samsung game launcher | New inductive and semi-supervised random walk approach | The approach used proved superior than traditional ones |
| J. Burez et al. [5] | 2009 | Six real world consumer attrition guessing data sets | Sampling, boosting, cost-sensitive learner, logistic regression, Random forests | Lift, AUC are good estimate metrics. Also increased performance of cost-sensitive techniques over standard ones |
| Tsai *et al.* [7] | 2009 | American telecom companies CRM dataset | Artificial neural networks (ANN) and self-organizing maps (SOM) | ANN + ANN outperforms other baseline and hybrid models |
| Verbeke *et al*. [8] | 2012 | Wireless telecom operators | Decision trees, ensemble methods, nearest neighbor methods, neural networks, rule induction techniques and statistical classifiers | Decision trees perform better overall performance than other models |
| Wei *et al.* [9] | 2002 | Taiwan's mobile telecommunication company | Decision trees, C4.5, decision rule, neural networks, multi-classifier class combiner | Multi-classifier class combiner method performed better when compared to single classifier method |
| Au *et al.* [10] | 2003 | Different real world databases | DMEL | DMEL performs better than traditional approaches like decision trees, logit regression, neural networks etc. |
| Gupta *et al*. [11] | 2006 | Real world data | Comparison of different models for predicting CLV and hence churners | CLV is good evaluation metric for churn prediction |

## CONCLUSION

In this study, we discussed how data from different real world companies like mobile telecommunications, financial service companies, retail companies, etc. can be used to predict the churners. This field is important to consider for the growth of businesses. It helps in better understanding of the customers and finding effective ways to study customer behavior and the strength of their relationship with the company. For increasing customer base, customer acquisition, retention and expansion techniques can be implemented. Among them, the companies mostly need to focus more on retention techniques as they are better than acquisition ones as they involve less cost and more profit. So, different methodologies are implemented to detect churners and hence applying retention actions to prevent them. In this study, we have given a brief summary of important state of art techniques and approaches necessary to predict churners. A summary of the results obtained and the data sets used by different researchers are also reported in Table 1.

## REFERENCES

1. Höppner S, Stripling E, Baesens B, *et al.* Profit Driven Decision Trees for Churn Prediction. *Eur J Oper Res*. Nov 30, 2018.
2. Glady N, Baesens B, Croux C. Modeling Churn Using Customer Lifetime Value. *Eur J Oper Res*. Aug 16, 2009; 197(1): 402-11p.
3. Tamaddoni Jahromi A, Sepehri MM, Teimourpour B, *et al.* Modeling Customer Churn in a Non-Contractual Setting: The Case of Telecommunications Service Providers. *J Strat Market*. Dec 1, 2010; 18(7): 587-98p.
4. Khan AA, Jamwal S, Sepehri MM. Applying Data Mining to Customer Churn Prediction in an Internet Service Provider. *International Journal of Computer Applications (IJCA)*. Nov 21, 2010; 9(7): 8-14p.
5. Burez J, Van den Poel D. Handling Class Imbalance in Customer Churn Prediction. *Expert Syst Appl*. Apr 1, 2009; 36(3): 4626-36p.
6. Liu X, Xie M, Wen X, *et al.* A Semi-Supervised and Inductive Embedding Model for Churn Prediction of Large-Scale Mobile Games. In *2018 IEEE International Conference on Data Mining (ICDM)*. Nov 17. 2018; 277-286p.
7. Tsai CF, Lu YH. Customer Churn Prediction by Hybrid Neural Networks. *Expert Syst Appl*. Dec 1, 2009; 36(10): 12547-53p.
8. Verbeke W, Dejaeger K, Martens D, *et al.* New Insights into Churn Prediction in the Telecommunication Sector: A Profit Driven Data Mining Approach. *Eur J Oper Res*. Apr 1, 2012; 218(1): 211-29p.
9. Wei CP, Chiu IT. Turning Telecommunications Call Details to Churn Prediction: A Data Mining Approach. *Expert Syst Appl*. Aug 1, 2002; 23(2): 103-12p.
10. Au WH, Chan KC, Yao X. A Novel Evolutionary Data Mining Algorithm with Applications to Churn Prediction. *IEEE Trans Evolut Comput*. Dec 2003; 7(6): 532-45p.
11. Gupta S, Hanssens D, Hardie B, *et al.* Modeling Customer Lifetime Value. *J Serv Res*. Nov 2006; 9(2): 139-55p.

**Cite this Article**
Saima Jan, Afaq Alam Khan. Predicting Churners from the User Base of a Business: A Review. *Journal of Artificial Intelligence Research & Advances.* 2019; 6(2): 8–16p.