CUSTOMER CHURN PREDITION IN TELECOM SECTOR USING SUPPORT VECTOR MACHINE

[1]**K.SOWMYA**,MASTER OF ENGINEERING, Department of Computer Science and Engineering MAHENDRA ENGINEERING COLLEGE (Autonomous) Mahendhirapuri, Mallasamudram, Namakkal DT -637 503

[2]**DR.R.VIJAYARAJESWARI**, Associate Professor, Department of Computer Science and Engineering MAHENDRA ENGINEERING COLLEGE (Autonomous) Mahendhirapuri, Mallasamudram, Namakkal DT -637 503

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 7 June2021

ABSTRACT

Acquisition and retention of new clients are one of the most significant concerns of businesses. While recipient companies concentrate on acquiring new customers, mature ones try to focus on retention of the existing ones in order to provide themselves with the opportunity of cross – selling. According to Freeman (1999) one of the most significant ways of increasing customers' value is to keep them for longer period of time.

Keywords: Acquisition, Freeman, Retention.

1. INTRODUCTION

In the new era emergence of electronic commerce has boosted the available information, and as Peppard (2000) believes, the internet channel has empowered the customers who are no longer stuck with the decisions of a single company and has led to exacerbation of the competition, while competitors are only one "click away", customer empowerment is likely to amplify the attrition rate of a company's customers (Lejeune, 2001). Facing with this threat companies should be equipped and armed with the most efficient and effective methods of examining their client's behavior predicting their possible future failure.

In accordance with (Lejeune, 2001) churn management consists of developing techniques that enable firms in keeping their profitable customers. The study at your disposal aims at finding an efficient and accurate predictive model for customer churn in pre-paid mobile telephony market segment by utilizing machine learning technique. With the intention of making you more familiar with the research's realm and its importance we start the report by providing you with statistics regarding the customer churn magnification in telecommunications industry and afterwards we address our problem definition and the question of our research.

1.1. Definition of Churn

It was discovered by Berson et al. (2000) noted that 'customer churn' is defined as the

process of subscribers (either prepaid or post paid) switching from one service provider. Churn can be active / deliberate, rotational / incidental, passive / nonvoluntary . With proper management of customers, we can minimize the susceptibility to churn and maximize the profitability of the company. A mechanism needs to be established to analyze the attributed of profitability. Churn Prediction can also be described as a method which helps in identifying possible churners in advance.

2. LITERATURE SURVEY

Swift (2001) defined CRM as an" enterprise approach to understanding and influencing customer behavior through meaningful communications in order to improve customer acquisition, customer retention, customer loyalty, and customer profitability. Kotler and Keller (2006) have defined Customer relationship management (CRM) as the process of managing detailed information about individual customers and carefully managing all customer "touch points" to maximize customer loyalty. Kincaid (2003) viewed CRM as "the strategic use of information, processes, technology, and people to manage the customer's relationship with your company (Marketing, Sales, Services, and Support) across the whole customer life cycle". Bose (2002) viewed CRM as an integration of technologies and business processes used to satisfy the needs of a customer during any given interaction more specifically from his point of view Customer relationship management (CRM) involves acquisition, analysis and use of knowledge about customers in order to sell goods or services and to do it more efficiently. Richards and Jones (2008) have defined CRM as "a set of business activities supported by both technology and processes that is directed by strategy and is designed to improve business performance in an area of customer management".

As it is presented in the above figure, a prospect that responds to the marketing campaigns of the company in acquisition phase, becomes a customer and this "New Customer" becomes a established one once the relationship between him/her and the company has been established and this is the point that in which the company can benefit from its established customers by revenue that comes from cross – selling and up – selling, but the peril that threatens the company in this stage is that at some point established customers stop being customers (Churn) (Olafsson, Li, & Wu, 2008).

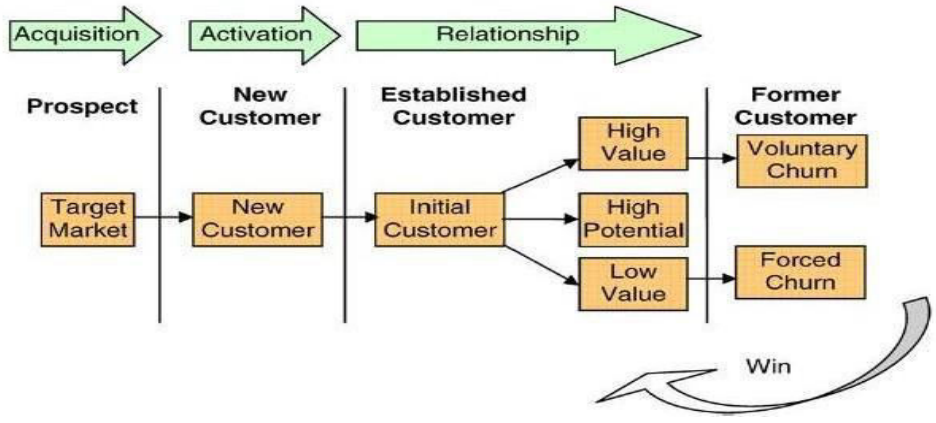

Fig. 1 Illustration of a customer life-cycle (source: Olafsson, Li, & Wu, 2008)

Thus, in simple words, the main goal of customer relationship management is to create satisfaction and delight among customers in order to prevent customer churn which is the most important threat that threatens all companies. It has been shown that a small change in retention rate can result in significant changes in contribution (Van den Poel & Larivie're, 2004).

3. PROPOSED MODEL

This section presents the proposed customer churn prediction model. Fig. 2 shows the proposed churn prediction model and describes its steps. In the first step, data preprocessing is performed which includes data filtering for noise removal, removal of imbalanced data features and normalization of the data. Important features are extracted from data using information gain attributes ranking filter and correlation attributes ranking filter. In the second step, Exponential Data Analysis is done on our Dataset to Understand the Demographics, Customer Account information, Distribution of services, Relation between monthly and total Charges, Churn rate w.r.t. all Attributes. In the third step, different classification algorithms are applied for categorizing the customers into the churn and non-churn customers. The classification algorithms include Logistic Regression (LR), Support Vector Machine(SVM), Random Forest (RF), ADABoost, Artificial Neural Network(ANN). In the final step, the all models are compared and Best fit model is suggested.

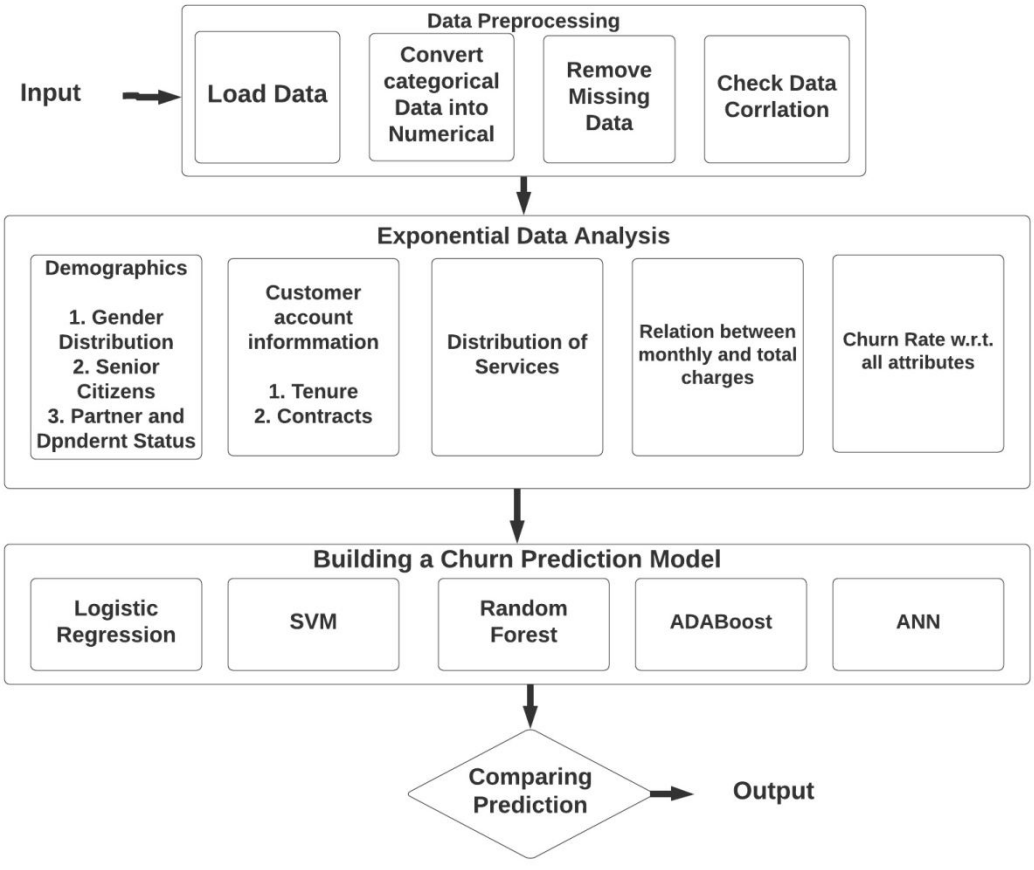

Fig. 2 Proposed Model

We can see that some variables have a negative relation to our predicted variable (Churn), while some have positive relation. Negative relation means that likeliness of churn decreases with that variable. Let us summarize some of the interesting features below: As we saw in our EDA, having a 2 month contract reduces chances of churn. 2 month contract along with tenure have the most negative relation with Churn as predicted by logistic regressions Having DSL internet service also reduces the proability of Churn
Lastly, total charges, monthly contracts, fibre optic internet services and seniority can lead to higher churn rates. This is interesting because although fibre optic services are faster, customers are likely to churn because of it. I think we need to explore more to better understad why this is happening.

"Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well (look at the below snapshot).

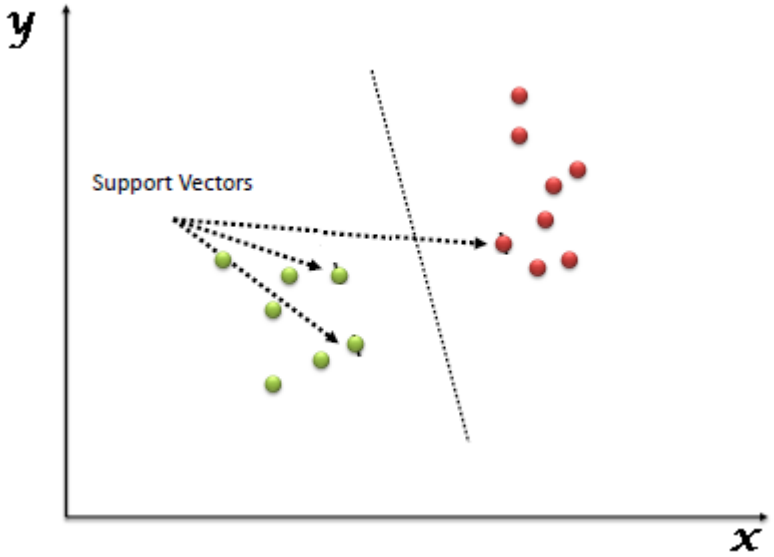

Fig. 3 Hyperplane Classifying the objects
Support Vectors are simply the co-ordinates of individual observation. The SVM classifier is a frontier which best segregates the two classes (hyper-plane/ line).

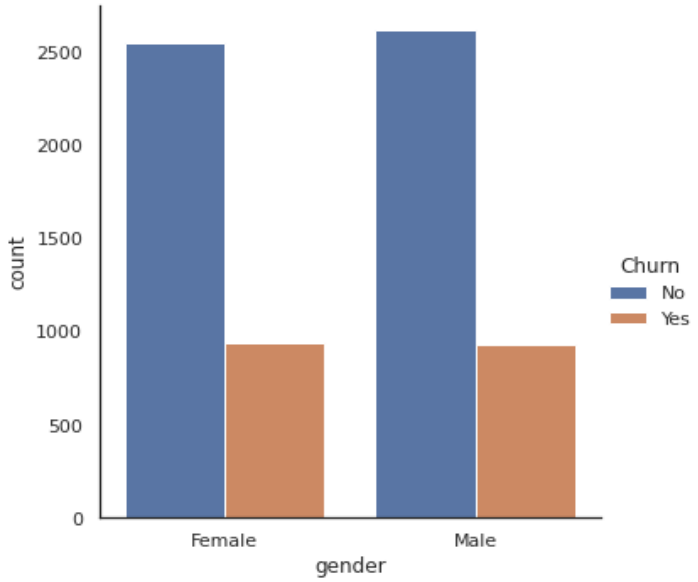

Fig.4 churn w.r.t male vs female

Wth SVM was able to increase the accuracy to upto 81%. However, we need to take a deeper look at the true positive and true negative rates, including the Area Under the Curve (AUC) for a better prediction.

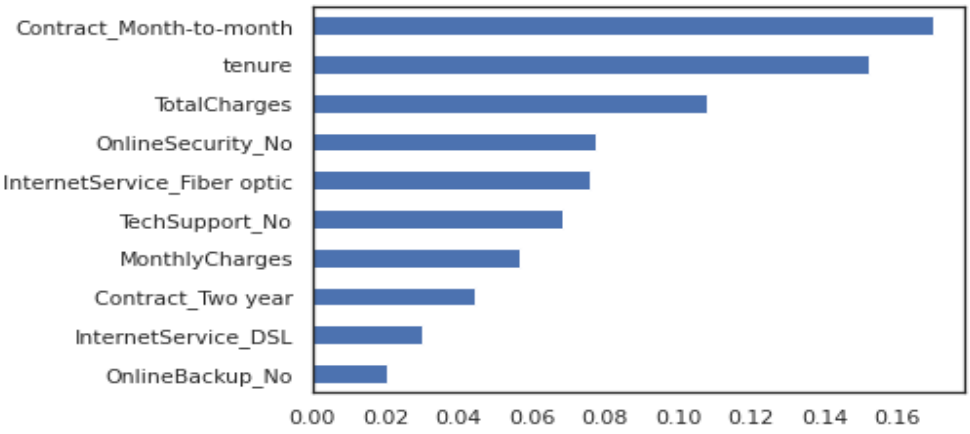

Fig. 5 Churn w.r.t. Services

The results from random forest are very similar to that of the logistic regression and in line to what we had expected from our EDA

AdaBoost algorithm, short for Adaptive Boosting, is a Boosting technique that is used as an Ensemble Method in Machine Learning. It is called Adaptive Boosting as the weights are re-assigned to each instance, with higher weights to incorrectly classified instances. Boosting is used to reduce bias as well as the variance for supervised learning. It works on the principle where learners are grown sequentially. Except for the first, each subsequent learner is grown from previously grown learners. In simple words, weak learners are converted into strong ones. Adaboost algorithm also works on the same principle as boosting

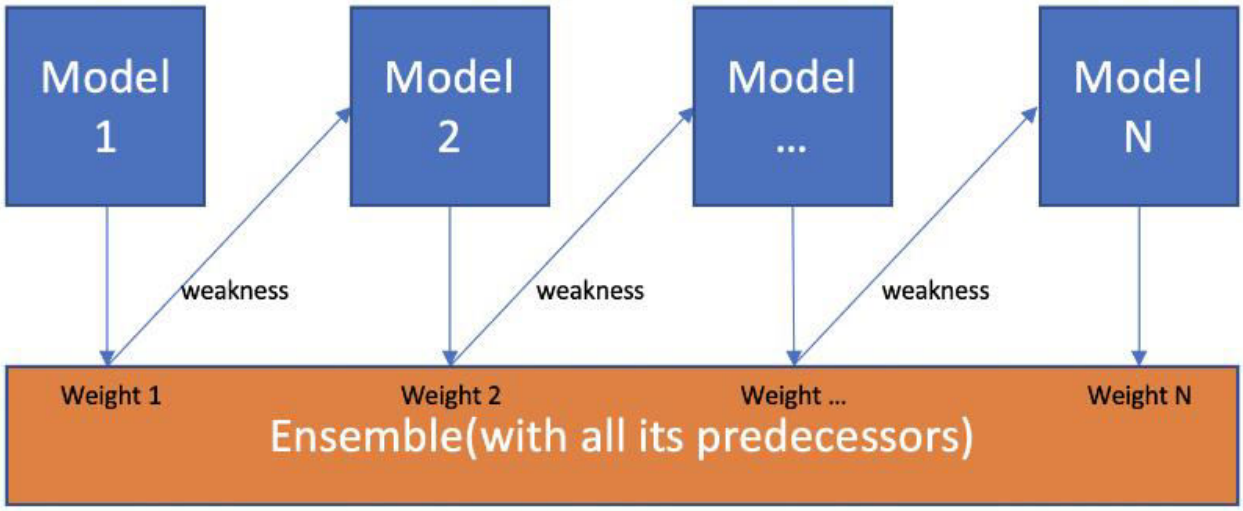

Fig.6 Adaptive Boosting

Artificial Neaural Network

A neural network is a machine learning algorithm based on the model of a human neuron. The human brain consists of millions of neurons. It sends and process signals in the form of electrical and chemical signals. These neurons are connected with a special structure known as synapses. Synapses allow neurons to pass signals. From large numbers of simulated neurons neural networks forms.

An Artificial Neural Network is an information processing technique. It works like the way human brain processes information. ANN includes a large number of connected processing units that work together to process information. They also generate meaningful results from it.

We can apply Neural network not only for classification. It can also apply for regression of continuous target attributes.

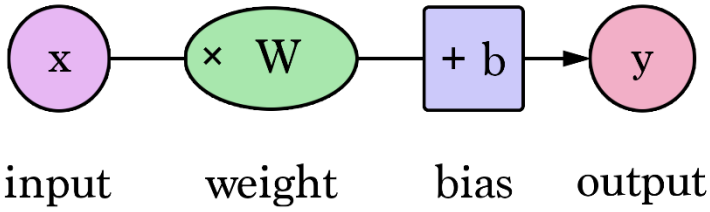

Fig.7 Basic Block Diagram of ANN

Performance Evaluation Matrix

In this study, the proposed churn prediction model is evaluated using accuracy, precision, recall, f-measure, and ROC area. Equation below calculates the accuracy metric. It identifies a number of instances that were correctly classified.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

Here ‘‘TN’’ stands for True Negative, ‘‘TP’’ stands for True Positive, ‘‘FN’’ stands for False Negative and ‘‘FP’’ stands for False Positive. TP Rate is also known as sensitivity. It tells us what portion of the data is correctly classified as positive.

The F-measure value is a trade-off between correctly classifying all the data points and ensuring that each class contains points of only one class. It is calculated by using above Equation. 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 random prediction, bad, moderate, good and superior respectively. Values of ROC areas other than these indicate something is wrong. Using Confusion Matrix, Performance evaluation matrix and ROC Curve our Model is measured and Best Model for Churn Prediction is Validated.

CONCLUSION

Telecommunication industry has suffered from high churn rates and immense churning loss. Although the business loss is unavoidable, but still churn can be managed and kept in an acceptable level. Good methods need to be developed and existing methods have to be enhanced to prevent telecommunication industry to face challenges. The importance of this project to the telecom market is continuously growing. Data collection is becoming an everyday task to all companies, and the value of that data can come from multiple sources. Churn prediction is becoming one of those sources that create revenue to the company. Being able to prevent when clients are going to cease their contract with the company opens the possibility of renegotiating that contract in order to retain the costumer. Although multiple studies have been conducted in this area, most of them use only one or two algorithms.

REFERENCES

[1] S. Babu, D. N. Ananthanarayanan, and V. Ramesh, ‘‘A survey on factors impacting churn in telecommunication using datamininig techniques,’’ Int. J. Eng. Res. Technol., vol. 3, no. 3, pp. 1745–1748, Mar. 2014.

[2] C. Geppert, ‘‘Customer churn management: Retaining high-margin customers with customer relationship management techniques,’’ KPMG & Associates Yarhands Dissou Arthur/Kwaku Ahenkrah/David Asamoah, 2002.

[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, Mar. 2011.

[4] Y. Huang, B. 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] A. Idris and A. Khan, ‘‘Customer churn prediction for telecommunication:Employing various various features selection techniques and tree based ensemble classifiers,’’ in Proc. 15th Int. Multitopic Conf., Dec. 2012,pp. 23–27.

[6] M. Kaur, K. Singh, and N. Sharma, ‘‘Data mining as a tool to predict the churn behaviour among Indian bank customers,’’ Int. J. Recent Innov. Trends Comput. Commun., vol. 1, no. 9, pp. 720–725, Sep. 2013.

[7] V. L. MiguØis, D. van den Poel, A. S. Camanho, and J. F. e Cunha, ‘‘Modeling partial customer churn: On the value of first product-category purchase sequences,’’ Expert Syst. Appl., vol. 12, no. 12, pp. 11250–11256, Sep. 2012.

[8] D. Manzano-Machob, ‘‘The architecture of a churn prediction system based on stream mining,’’ in Proc. Artif. Intell. 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