

Analysis of Customer Churn Prediction in Telecom Industry Using Logistic Regression

K. Sandhya Rani¹, Shaik Thaslima², N.G.L. Prasanna³, R.Vindhya⁴, and P. Srilakshmi⁵

¹ Assistant Professor, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India

^{2,3,4,5} Students, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India

Correspondence should be addressed to Shaik Thaslima; thaslimashaik221@gmail.com

Copyright © 2021 Made K. Sandhya Rani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Customers play an important role in industry to run industry. Churn of the customer may lead many consequences. Customer churn prediction must be the important aspect of any company. This helps in the detection of customers who are likely to cancel a subscription to a service. Recently, the mobile telecommunication market has changed from a rapidly growing market into a state of saturation. The focus of telecommunication companies is to shift from growing of large customer into keeping customers in house. For that reason, it is valuable to know which customers are likely to switch to a competitor in future. The model is proposed for churn prediction for telecommunication companies using machine learning techniques namely logistic regression. A comparison is done on the efficiency of the algorithm on the available dataset.

KEYWORDS- Machine learning, logistic regression, variance reduction, Bayesian Models, CRM (Customer Relationship Management)

I. INTRODUCTION

With the aspects of development and the advancements around the world communication is one of the main required entities in the development. Communication needs to be fast and reliable for the people for their works and necessities to be done. In developed countries telecom industries play a major role and have become a part of necessities required for people to live. The developments and technological progress and the steady increase in the operating network the competition between them has peaked. Companies need to thrive to survive in the competitive market by implementing new strategies and policies for acquiring the customer base around the world. The main theme behind these strategies is to develop and generate more income to the companies. There are different strategies followed by the marketing team in the companies to attract new customers, making the existing customers to buy or upgrade to new services within the same company and finally to keep the customer base for longer period. However, a comparison of these strategies taking into account the Return on Investment (RoI) of each showed that the third strategy is the most profitable [3], showing that retaining existing customers has a much lower cost than

acquiring a new one. A new customer, and is also considered much easier than an upselling strategy. To adopt a third strategy, companies must minimize the possibility of customer turmoil, known as movement of customers from one supplier to another.

Customers who are likely to churn is the considerable issue in service sectors with the high competition in the market. On the other side, analysing the customers who are likely to leave the company will represent potentially large additional revenue source if it is done in the early phase. Many researchers confirmed that machine learning techniques are highly useful and efficient to predict the churning and non-churning events by learning from the previous company data. The data used in this contains all customer information throughout a period of time. In this experiment, it mainly focuses on tree-based and regression-based machine learning methods and algorithms for prediction of churn in telecom industries; we mainly focussed on Decision tree algorithm and logistic regression for efficient prediction model for customer churn. There are different companies to build their own customer churn prediction models because acquiring a new customer costs six times higher than the cost of retaining the customer likely to churn.

II. EXISTING SYSTEM

In the proposed system R programming will be used to build the model for churn prediction [10]. It is widely used among statisticians and data miners for developing [2] statistical software and data analysis. R is freely available and a powerful statistical analysis tool which has not yet been explored for building model for churn prediction [5]. The system will have three main options namely View performance analysis [1] – which displays the results obtained by applying logistic regression and decision tree on the available dataset, Testing [7] – to construct a list of customers which have a high probability to churn from the input, given that the attributes of the input data are same as the available dataset used for training, Training and testing [4] – which builds a model along with generating a churn list if any other type of dataset is provided [3]. In performance analysis the results after using logistic regression and decision trees on the available dataset is illustrated using confusion matrix analysis [6]. In the next

operation the user can provide data for testing the system provided the features of the data are same as that used for training using the publicly available dataset.

Disadvantages:

- Prediction is done in R programming, which has not yet been explored for building model for churn prediction.

In existing system, the system can only predict by using limited features

III. PROPOSED SYSTEM

This technology involves deep studying and analysis, it is extremely difficult for people to predict the percentage of that likely to churn. Our system helps understand the people who want to churn and it will easily represent the data in the form of bar plots, pie charts. Our system can easily identify the risk factor by using the previous data. Logistic regression algorithms are used for the efficacy in prediction results.

Figure 1 describes about the work flow of the system and the steps involved in the description of the system. This shows the work flow of the system and management of the system

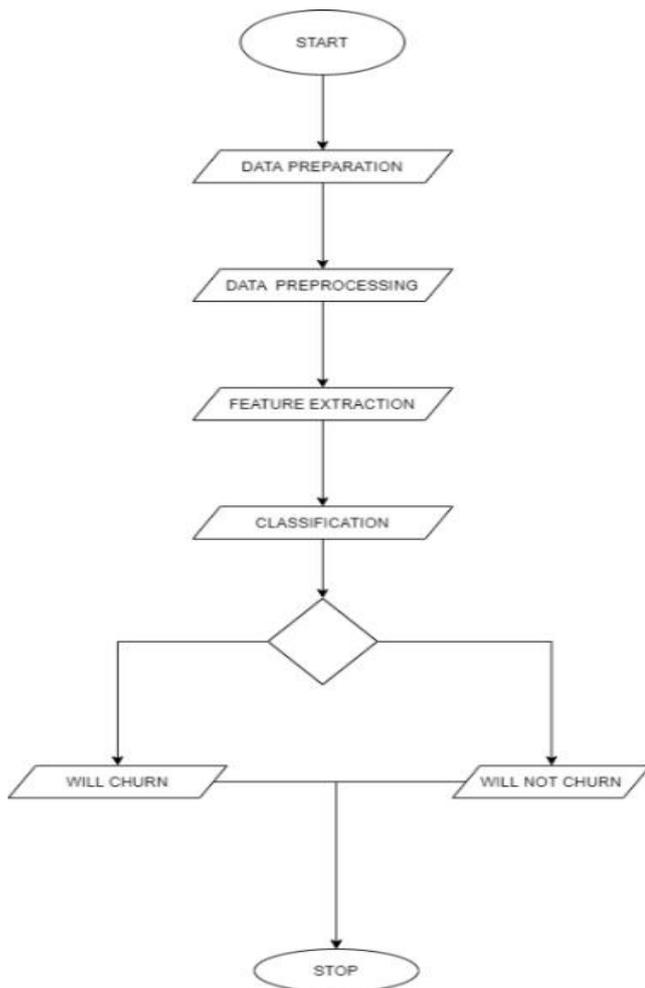


Fig. 1: Working process

A. Functional Requirements

Dataset must be collected from a trusted source or a telecom company which will be useful for real time validation

B. Non- Functional Requirements

System with minimum requirements mentioned must be used for better prediction in the time frame. Basic knowledge on the market and risks involved in customer base and how the churn works. Reading and understanding the graphical data that are drawn in the application.

IV. RESULTS

In the result, we have shown in Figure no. 2, the result in the form of ROC Curves for Voting Classifier which shows the accuracy of the result by AUC. Stacked model had improved precision on a good amount, the blended still got a better F1 and it is saved as the best model. This model is saved and its ROC Curves are plotted below. Using the techniques by plotting the true positive rate (TPR) against the false positive rate (FPR)

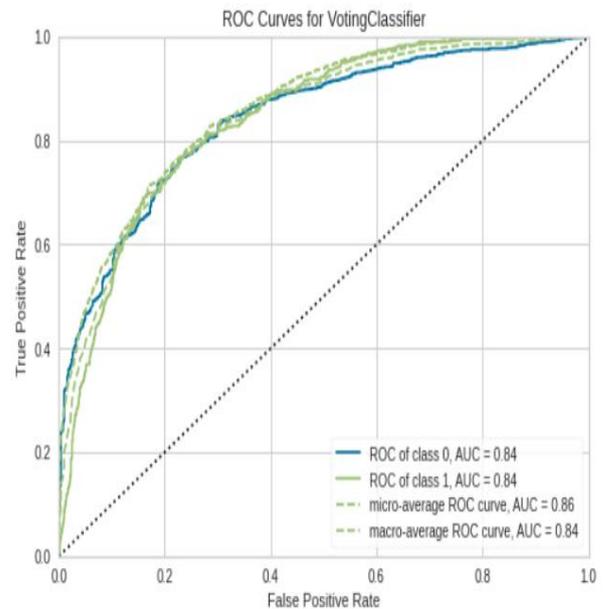


Fig. 2: Analysis of Customer Churn prediction using logistic Regression

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

V. CONCLUSION

The main objective of the project is to predict the customers that are likely to churn in a telecom industry, analysis of the companies with their previous data for a period of time which helps them to understand that day's situation of the company and helps them to assess churn factor of the company so that they can implement new plans to bring new customers and the customers already in their company. This system is used to reduce the time and burden of company by going through the historical data and news of the company to assess and react to the situations.

REFERENCES

- [1] A Hybrid Churn Prediction Model in Mobile Telecommunication Industry ,Georges D. Olle Olle and Shuqin Cai ,International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 4, No. 1, February 2014.
- [2] Telecommunication Subscribers' Churn Prediction Model Using Machine Learning, Saad Ahmed Qureshi, Ammar Saleem Rehman, Ali Mustafa Qamar, Aatif Kamal, Ahsan Rehman, IEEE International Conference on Digital Information Management (ICDIM), 2013 Eighth International Conference on, 2013, pp. 131–136.
- [3] Customer Churn Prediction in Telecommunication Industries using Data Mining Techniques- A Review, Kiran Dahiya and Kanika Talwar, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4, 2015.
- [4] Churn Prediction in Telecommunication Using Classification Techniques Based on Data Mining: A Survey, Nisha Saini and Monika, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 3, March 2015.
- [5] Churn Prediction In Mobile Telecom System Using Data Mining Techniques, Dr. M. Balasubramaniam, M.Selvarani, International Journal of Scientific and Research Publications, Volume 4, Issue 4, April 2014.
- [6] Predicting Customer Churn in Mobile Telephony Industry Using Probabilistic Classifiers in Data Mining, Clement Kirui1, Li Hong, Wilson Cheruiyot and Hillary Kirui, IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013.
- [7] Applying Data Mining Techniques in Telecom Churn Prediction, N.Kamalraj and Dr.A.Malathi, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 10, October 2013.
- [8] Churn Prediction in Telecommunication Using Data Mining Technology, Rahul J. Jadhav and Usharani T. Pawar, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No.2, February 2011.
- [9] Analysis of Customer Churn in Mobile Industry using Data Mining, Aishwarya Churi, Mayuri Divekar, Sonal Dashpute, Prajakta Kamble, International Journal of Emerging Technology and Advanced Engineering, Volume 5, Issue 3, March 2015.
- [10] A Proposed Churn Prediction Model, Essam Shaaban, Yehia Helmy, Ayman Khedr, Mona Nasr / International Journal of Engineering Research and Applications (IJERA) ,Vol. 2, Issue 4, June-July 2012.

ABOUT THE AUTHORS



Mrs. K. Sandhya Rani, M.Tech, Assistant Professor, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India



Shaik Thaslima, Student of B.Tech Final Year, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India



N.G.L. Prasanna, Student of B.Tech Final Year, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India



R. Vindhya, Student of B.Tech Final Year, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India



P. Srilakshmi, Student of B.Tech Final Year, Department of Computer Science and Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India