International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Churn Prediction and Prevention for OTT/Tele- Communication

Aniket Shinde[1], Ajinkya Pawar[2], Pratik Dhamne[3], Krushnakant Shaha [4] ,Prof. Sharad M.Rokade[5]

BE Students, Dept. Of Computer Engineering, SVIT, Nashik

Abstract -Customers are becoming more drawn to the quality of service (QoS) offered by businesses in the present. However, the present day shows greater rivalry in offering clients technologically cutting-edge QoS. However, effective customer relationship management systems may help the organization attract new clients, preserve client connections, and enhance client retention by generating more revenue for the company's operations. Churners have always been a major problem for every business that offers services. Churning drives up a company's expenses while also lowering its profit margin. However, it is possible to forecast if a consumer wishes to cancel service using predictive analysis based on historical service usage, service performance, expenditure, and other behavioral patterns. The problem with churn analysis is that it can reveal unnecessary information when used on databases that are combined by a company that owns confidential information.

Keywords—*customer relationship management, customer retention, machine learning, churn analysis*

1.INTRODUCTION

The customer gain is the important factor in the business. The product of any company is not good if it doesn't have any customers or they don't buy their product. So, company make their product best to attract more customers and their business grows. But the difficult task for the business growth is to make the customer stick to their product, because over the period of time customer start losing their interest in companies particular product and it results great growing company in the initial starts going backward after some times.

To avoid this problem company should know why their customers are losing interest from their product. So, to handle this situation we have made the solution that predicts whether the customer will leave the product or not, if leaving then what are the factors that made them loose the interest in product. When company will know that they will improve that particular thing which affecting the company's growth.

We have come with machine learning model called "Churn Prediction". The technical meaning of churning is whether customer is start losing interest in the product or not. This ML model will predict the churning based on the dataset used to train the model. The dataset will be all the basic information of the customer with respect to with product.

2. LITERATURE REVIEW

- M.A.H. Farquad et.at. [6] proposed a hybrid approach to overcome the drawbacks of general SVM model which generates a black box model (i.e., it does not reveal the knowledge gained during training in human understandable form). The hybrid approach contains three phases: In the first phase, SVM-RFE (SVM-recursive feature elimination) is employed to reduce the feature set. In the second phase, dataset with reduced features is then used to obtain SVM model and support vectors are extracted. In the final phase, rules are then generated using Naive Bayes Tree (NBTree which is combination of Decision tree with naive Bayesian Classifier). The dataset used here is bank credit card customer dataset (Business Intelligence Cup 2004) which is highly unbalanced with 93.24% loyal and 6.76% churned customers. The experimental showed that the model does not scalable to large datsets.
- Wouter Verbeke et.at.[7] proposed the application of Ant-Miner+ and ALBA algorithms on a publicly available churn prediction dataset in order to build accurate as well as comprehensible classification rule-sets churn prediction models. Ant-Miner+ is a high performing data mining method based on the principles of Ant Colony Optimization which allows to include domain knowledge by imposing monotonicity constraints on the final rule-set. The advantages of Ant-Miner+ are high accuracy, comprehensibility of the generated models and the possibility to demand intuitive predictive models. Active Learning Based Approach (ALBA) for SVM rule extraction is a rule extraction algorithm, which

combines the high predictive accuracy of a non-linear support vector machine model with the comprehensibility of the ruleset format.

- Ning Lu et.at.[8] proposed the use of boosting algorithms to enhance a customer churn prediction model in which customers are separated into two clusters based on the weight assigned by the boosting algorithm. As a result, a high risky customer cluster has been found. Logistic regression is used as a basis learner, and a churn prediction model is built on each cluster, respectively. The experimental results showed that boosting algorithm provides a good separation of churn data when compared with a single logistic regression model.
- Ssu-Han Chen et.at.[9] used a novel mechanism based on the gamma Cumulative SUM (CUSUM) chart in which the gamma CUSUM chart monitors individual customer's Inter Arrival Time (IAT) by introducing a finite mixture model to design the reference value and decision interval of the chart and used a hierarchical Bayesian model to capture the heterogeneity of customers. Recency, another time interval variable which is complementary to IAT, is combined into the model and tracks the recent status of the login behavior. In addition, benefits from the basic nature of control charts, the graphical interface for each customer is an additional advantage of the proposed method. The results showed that the accuracy rate (ACC) for gamma CUSUM chart is 5.2% higher than exponential CUSUM and the Average Time to Signal (ATS) is about two days longer than required for exponential CUSUM.

3. PROBLEM STATEMENT

One of the main growth pillars for products with a subscription-based business model is customer retention. The SaaS industry is very competitive since clients have a wide range of options for providers, even within a single product category. A consumer may stop buying from you after one or more negative experiences. And if hordes of dissatisfied consumers leave at a rapid clip, there would be significant financial losses as well as reputational harm. We will learn methods and recommended procedures for resolving this issue. We'll go into gathering information regarding client relationships with brands, customer behaviour traits that are most associated with churn, and the reasoning behind choosing the top machine learning models.

4. METHODOLOGY

Data analysis is one of the most convenient ways of perceiving instances of customer's deprival in business. for example, in an E-commerce business there will be database that contains the basic information of the customer as well as the data that shows the how is the customers response towards the product that customer uses. All those data are used for analysis and to predict the churning rate of the customer.

There are various data analysis technique in the machine learning. We will be using following techniques.

1. KNN (K-Nearest Neighbor)
2. Random Forest

- Random Forest:

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

Steps involved in random forest algorithm:

i. In Random Forest n number of random records are taken from the data set having k number of records.
ii. Individual decision trees are constructed for each sample.
iii. Each decision tree will generate an output.

Final output is considered based on Majority Voting or Averaging for Classification and regression respectively

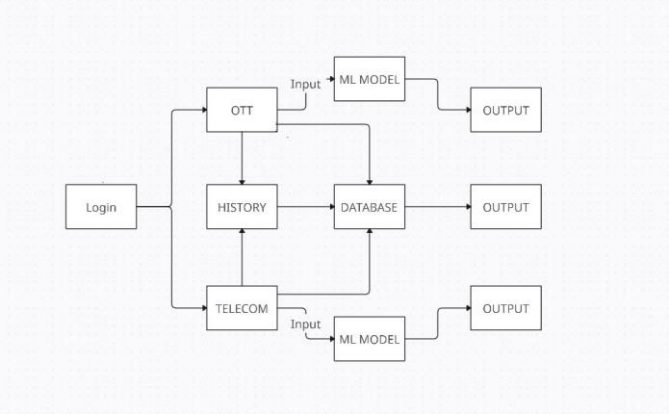
5. PROPOSED SYSTEM

Customer churn is a concept that is quickly gaining significance, particularly in the most influential and competitive organizations worldwide. The term "customer churn" refers to a group of users who stop using a service over an extended period of time.

The revenue received by the business for a particular service or product is adversely affected by the in-fluence of customers who left an organization in the past to those who are willing to continue an ex-change with it. Therefore, in order to maintain business growth while maintaining customer satisfaction, we have suggested this system using a churn prediction model. The dataset will be entered into the machine learning model using a GUI.

Model for machine learning: In order to use the raw data in this model, it must first be cleaned. The company itself will provide the raw data. Data will train in the machine learning model after being cleaned. The model will then forecast both customer behaviour and the rate of churn.

6. SYSTEM ARCHITECTURE



7. FUTURE SCOPE

1.Focus on ways to achieve better results such as using different hyper-parameter optimization techniques for the same algorithms in a smaller time frame.

2.Different combinations of attributes can be used in the future to determine the customer retention policies.

3.Also, the performance factors can be improved using different deep learning approaches.

4.Behavioral analysis can be extended to association between prediction of churn and lifetime customer value.

5.This can help shape retention policies as well, which will further boost the revenue of the company.

6.A churn retention model can also be designed to retain customers. These retention models can further be studied to analyse the cost of adding a new customer to the base against the cost of implementing retention policies to retain existing customers.

8. RESULTS

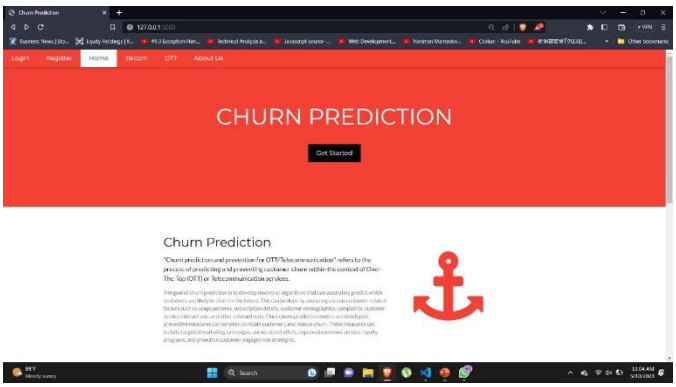


Fig 1.1 Homepage of the Website

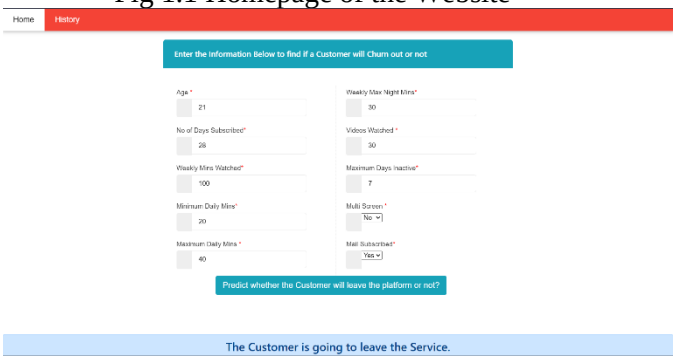


Fig 1.2 Churn Prediction for OTT

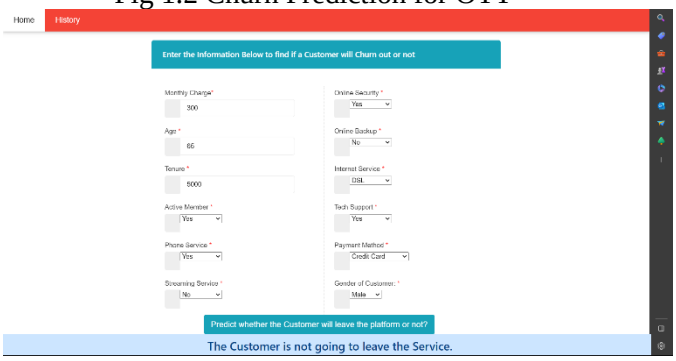


Fig 1.3 Churn Prediction Telecom

9. CONCLUSION

Therefore, from the above discussion, it can be concluded that, regardless of the type of organization, every organization need to concern about the customer churn. Customer retention is the process of maintaining the loyalty of the customer by understanding the customer demand and serving them accordingly. Powerful churn prediction model will help the organizational management to predict the customer churn. Depending on the complex data of the industry, K-nearest neighbor can turn out advantageous for predicting the churn rate. The above report has focused on the concept of customer retention along with the churn prediction. Apart from that, the use of random forest in order to enhance the

churn prediction process has been discussed here along with the algorithm.

10. REFERENCES

1] RajaGopal Kesiraju VLN, P. Deeplakshmi, "Dynamic Churn Prediction using Machine Learning Algorithms - Predict your customer through customer behavior" (ICCCI -2021) Jan. 27 – 29, 2021.

2] Pushkar Bhuse, Aayushi Gandhi, Parth Meswani, Riya Muni, Neha Katre, "Machine Learning Based Telecom-Customer Churn Prediction" [ICISS 2020]

3] Sanket Agrawal, Aditya Das, Amit Gaikwad, "Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning" (ICSCEE2018)

4] Saran Kumar A, Chandrakala D, "A Survey on Customer Churn Prediction using Machine Learning Techniques" International Journal of Computer Applications (0975 – 8887) Volume 154 – No.10, November 2016.

5] Ssu-Han Chen, "The gamma CUSUM chart method for online customer churn prediction", Electronic Commerce Research and Applications, 17 (2016) 99–111.

6] M.A.H. Farquad, Vadlamani Ravi, S. Bapi Raju "Churn prediction using comprehensible support vector machine: An analytical CRM application", Applied Soft Computing 19 (2014) 31–40.

7] Ning Lu, Hua Lin, Jie Lu, Guangquan Zhang "A Customer Churn Prediction Model in Telecom Industry Using Boosting", IEEE Transactions on Industrial Informatics, vol. 10, no. 2, may 2014.

8] Navid Forhad, Md. Shahriar Hussain, Rashedur M Rahman "Churn Analysis: Predicting Churners" from https://ieeexplore.ieee.org/document/6991433.

9] Wouter Verbeke, David Martens, Christophe Mues, Bart Baesens "Building comprehensible customer churn prediction models with advanced rule induction techniques", Expert Systems with Applications 38 (2011) 2354–2364.