

Telecom churn prediction using machine learning algorithm

Pallavi Mogare, Vaishnavi kadam, Mohini Pawar, Gayatri Kadam

Department of Computer Engineering
Matoshri College of Engineering and Research Centre,
Nashik-422105

Abstract: Customers always play vital role in increasing profit and revenue of every organization; hence, to gain customer satisfaction it is important for the organizational managers to maintain one efficient customer relationship management system by selecting the target customers and maintaining effective relationship with them. In current days, the customers are getting more attracted towards the quality of service (QoS) provided by the organizations. However, the current era is evidencing higher competition in providing technologically advanced QoS to the customers. Nevertheless, efficient customer relationship management systems can be advantageous for the organization for gaining more customers, maintaining customer relationships and improve customer retention by adding more profit to the organizational business. Furthermore, the machine learning models such as support vector machine algorithms can add more value to the customer retention strategies.

Keywords: customer relationship management, customer retention, machine learning, support vector algorithm.

INTRODUCTION

The CRM system will be helpful for the organization in identifying the most prominent group of customers and their behavior; which will become beneficial for the organization in understanding the retention strategies in a better way. Additionally, higher the customer loyalty, lesser is the customer churn rate; hence using machine learning algorithm such as support vector algorithm can add value in preventing the customer churn. This report will focus on the customer retention with the usage of support vector machine learning in gaining customer loyalty and increasing retention..

MOTIVATION

The main Motive of implementing this system is to overcome the Churning manifests either voluntarily which means the consumer willingly decides to terminate relations with a given seller or involuntarily where the consumer's action renders them to their evacuation from undertaking any form of business relations with a given entity..

LITRATURE SURVEY

This chapter discuss brief literature regarding the project. Literature survey is mainly used to identify information relevant to the project work and know impact of it within the project area. It defines as till yet how many surveys have been done knowledge of latest technology and implementation designs.

| Sr.no | Title | Details |
|-------|---|---|
| 1 | A Churn Prediction Model Using Random Forest | This project is to manage all Business analysts and customer relationship management (CRM) analyzers need to know the reasons for churn customers Google cloud. |
| 2 | Research on a Customer Churn Combination Prediction Model | SMost of the current customer churn predictions use a single prediction model, which makes it difficult to accurately predict customer churn |

LIMITATION OF EXISTING SYSTEM

- Internet: It is an important factor to which training the dataset.

EXPERIMENTAL SETUP

• MySQL Database MySQL is an open source relational database management system(RDBMS).It uses a standard form of the well known SQL data language and works quickly and works will even with huge dataset

• Language: Python Python is a popular programming language. It was created by Guido van Rossum, and released in 1991. It is used for:

- (a) web development (server-side).
- (b) software development.

• Operating System: Windows 8 and above Windows is the most widely used operating system for desktop and laptop computers. Develop by Microsoft, Windows preliminary runs on x86 based computers. Windows provides Graphical User Interface and desktop Environment in which application displayed in resizable,movable windows on screen.

SCOPE:

Understanding a problem and a final goal

It's important to understand what insights one needs to get from the analysis. In short, you must decide what question to ask and consequently what type of machine learning problem to solve: classification or regression. Sounds complicated, but bear with us.

PROBLEM STATEMENT:

Examining the customer attrition rate in an organization implies the process of churn analysis. In the telecommunication industries, the churn can be identified as the number customers who had discontinued their subscription in a certain time period [4]. A typical churn rate measures the number of customers moving in and out within a given time period. Moreover, for the telecommunication industry, the movement of the customers from one company to another

SYSTEM ARCHITECTURE

Aspects of marketing mix theory give better explanations on why customers initiate the action of moving away from a given business hub. Product features and mainly the quality and price topples the perceptions held by customers towards a given company. For instance, in a study in an Australian telecommunications industry, it is proved that the quality of the network used by the customers was an underpinning feature to the customers' decision of not using the local networks [1]. Poor network quality nearly affected most of the users typically due to the large number of small-scale users who only call and surf the internet. Pricing and promotions also dramatically determine consumer behavior in any market. Once a firm tag favorable prices to its customers, it will, in return, increase and maintain their customers' loyaltySS

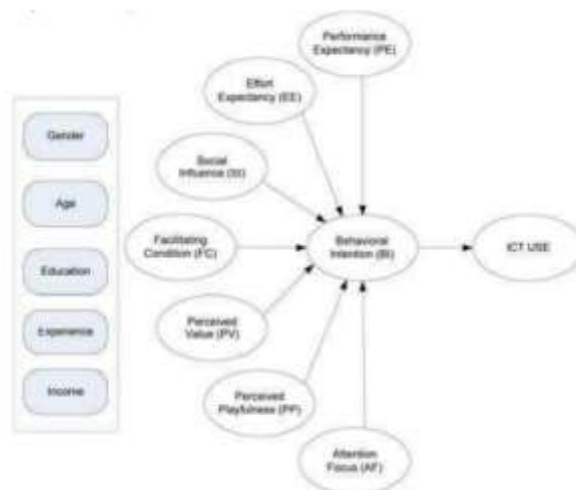


Fig -1: System Architecture Diagram

ADVANTAGES

1. Identify at-risk customers
2. Optimization products and services
3. Increased revenue

METHODOLOGY

Head of ScienceSoft data analytics department Alex Bekker notes that such methods as permutation importance, ELI5 Python package, and SHAP (SHapley Additive exPlanations) can be used to define the most relevant and useful features.

The principle of work of all methods lies in explaining how models make their predictions (based on what features a model made a particular conclusion.) Model interpretability is one of the high-priority issues in the field, and data scientists keep developing solutions to solve it. You can read more about interpretability in our article on AI and data science advances and trends.

Permutation importance is one of the ways of defining feature importance – an impact a feature makes on the predictions. It's calculated on models that have already been trained. This is how permutation importance is done: A data scientist changes the order of data points in a single column, feeds the model with the resulting dataset, and defines to what extent that change decreases its accuracy. Features that have the biggest influence on results are the most important.

Another way to do permutation importance is to remove a feature from a dataset and retrain the model.

Permutation importance can be done with ELI5 – an open source Python library that allows for visualizing, debugging ML classifiers (algorithms), and interpreting their outputs.

According to the ELI5 documentation, this method works best on datasets that don't contain a large number of columns (features).

Using the SHAP (SHapley Additive exPlanations) framework, specialists can interpret decisions of “any machine learning model.” SHAP also assigns each feature an importance value for a particular prediction.

5. 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 telecommunication industry, support vector machine can turned 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 support vector machine in order to enhance the churn prediction process has been discussed here along with the algorithm

REFERENCES

- Siu NY, Zhang TJ, Yau CY. The roles of justice and customer satisfaction in customer retention: A lesson from service recovery. *Journal of business ethics*. 2013 Jun 1;114(4):675-86
- Hossain MM, Suchy NJ. Influence of customer satisfaction on loyalty: A study on mobile telecommunication industry. *Journal of Social Sciences*. 2013;9(2):73- 80.
- Maldonado S, Flores A, Verbraken T, Baesens B, Weber R. Profitbased feature selection using support vector machines—General framework and an application for customer retention. *Applied Soft Computing*. 2015 Oct 1;35:740-8.
- Maga M, Canale P, Bohe A, inventors; Accenture Global Services Ltd, assignee. Churn prediction and management system. United States patent US 8,712,828. 2014 Apr 29
- Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC. A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*. 2015 Jun 1;55:1-9.
- Haenlein M. Social interactions in customer churn decisions: The impact of relationship directionality. *International Journal of Research in Marketing*. 2013 Sep 1;30(3):236-48.
- Farquard MA, Ravi V, Raju SB. Churn prediction using comprehensible support vector machine: An analytical CRM application. *Applied Soft Computing*. 2014 Jun 1;19:31-40.
- Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC. A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*. 2015 Jun 1;55:1-9.