



Predict Customer Churn through Customer Behaviour using Machine Learning Algorithms

Harini T¹, Hari Krishna T², Sushma G¹, Charankumar Reddy P¹, Mohammad Thahir S¹

¹Department of Computer Science and Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, Andhra Pradesh, India.

²Associate Professor & Head, Department of Artificial Intelligence and Machine Learning, Annamacharya Institute of Technology and Sciences, Rajampet, Andhra Pradesh, India.

Email: thimmarreddyharini2006@gmail.com

Article History

Received: 1 March 2023

Accepted: 22 March 2023

Keywords:

Customer Relationship Management(CRM) system; Machine Learning Algorithms; Decision Tree; Naïve-Bayes Classification; Logistic Regression algorithms; Customer Churn

Abstract

Customers are becoming more concerned to the quality of service (QoS) offered by organizations in the present. However, the present day shows greater rivalry in offering the clients with technologically innovative QoS. However, an organization may benefit from effective customer relationship management systems in order to increase sales, maintain relationships with existing customers and improve customer retention. The customer retention strategies can benefit greatly by the use of machine learning models like Decision Tree, Naïve-Bayes Classification, Logistic Regression algorithms.

1. Introduction

Every organization's profit and revenue are always increased by its customers, so it is crucial for organisational managers to maintain an effective customer relationship management system by identifying their target clients and building strong relationships with them. (Brandusoiu, Todorean, et al.) Additionally, the company will benefit from the CRM system by being able to recognise the most significant client group and their behaviour which helps to better understand its retention strategy. Lower customer churn rates are associated with stronger customer loyalty, therefore applying machine learning algorithms like the Decision Tree, Naïve-Bayes Classification, Logistic Regression algorithms can help to prevent the customer churn. (Maldonado et al.)

2. Background

Churning can occur voluntarily where a customer chooses to end their relationship with a specific seller or involuntarily where the customer's act forces them to avoid engaging in any kind of business relationship with a specific entity. (Tsai, Lu, et al.) Forceful churning occurs when a customer violates the traditional rules for conducting business within a specific location. In this situation, it may be the result of illegal activities like theft or a failure to make a timely payment. A voluntary kind of churning is sparked by issues related to consumer satisfaction or the loss of an organizations competitive edge. (Burez and Van Den Poel) Any company that deals with a variety of consumers should regularly rotate its evaluation practices.

3. Causes of Churning

The influence by former customers on current consumers who are willing to continue doing business with a company has a negative effect on the amount of money a company receives for a certain service or good. Customers may have unrealistic expectations that are not met when they buy products from a certain company, as evidenced by their preference for newer products. For the corporations that manufacture machinery is difficult to persuade customers about a certain machine when they test it and discover it underperformed. (Hossain, Suchy, *et al.*) If the device meets additional requirements, the consumer will choose a different seller. Particularly in these situations, consumer management strategies pay off, prompting the organisation to increase the quality of its production and expand their brand portfolio. (Vafeiadis *et al.*) Price is one of the primary factors influencing consumer decisions to buy a particular good. More than any other factor, customers feel the pressure on their sale of premium services and goods when they have choices to less costly products that can perform the intended duty. Price discrepancies in the product market destroy a company's ability to retain customers in comparison to that of their rivals. According to the hedonic pricing theory, customers assess the value of a product by comparing its price to that of competing businesses. (Kaya *et al.*) A corporation's goods are at danger of being rejected by potential customers if the pricing plan is thought to be flawed by the general public. Customer happiness is the primary determinant of whether a certain consumer would adhere to the terms of a particular business or reject the offer. (Idris, Rizwan, Khan, *et al.*) Customers are satisfied if their expectations are met or exceeded by the performance of the specific good or service they are about to purchase. In most circumstances customers prefer to stick with the company that offer top-notch customer care services. Once a consumer develops the habit of purchasing goods from a certain seller, the company has won the customer's loyalty, and it is up to the company to do all reasonable efforts to keep the customer. (Baghla, Gupta, *et al.*)

4. Conceptual Framework

Aspects of the marketing mix theory provide stronger justifications for why clients decide to leave a certain company hub. Customers' impressions

of a certain corporation are overturned by product attributes, particularly the worth and pricing. For instance, a research in the Australian telecommunications sector established that the consumers' choice to decline utilising local networks was supported by the quality of the network they utilised. Due to the vast number of infrequent small scale customers who just phone and use the internet, poor network quality almost always affects the majority of patrons. Selling prices and advertising have a tremendous influence on consumer behaviour in any industry. When a business offers its clients advantageous rates, their loyalty will grow and stay constant. (Farquad, Ravi, Raju, *et al.*)

The anticipation and confirmation hypothesis frequently offers a crystal-clear explanation for why consumers act in particular ways. The broad strokes of typical customer behaviour are seen in Figure 1. As was previously stated, a customer's expectations greatly affect whether they would purchase a particular product, and it is the obligation of a brand's marketing team to ensure that the pricing are reasonable. When customers use a service and determine that it complies with standards, their loyalty to a particular company always grows. (Siu, Zhang, Yau, *et al.*)

5. Churn Prediction

When monitoring a company's sustainability turnover rate, churn analysis is implicit. Churn is a term used to describe the number of customers who cancelled their subscriptions in a given period of time in the telecom industry. The amount of customers who arrive and exit over a specific time period is measured as the churn rate. Additionally, the shifting of consumers from one business to another is referred to as churn in the telecommunications sector. As the particular industry works hard to keep more profitable clients, the current situation is showing a bigger quantity of churning customers. Figure 2 depicts the algorithm used to train your data set and model. The churn can also be divided into two categories. Involuntary churn is the term used when an industry decides to lose customers due to bill paying issues, fraud, or another similar issue. Contrarily, the specific behaviour turns into a voluntary model of potential customer behaviour. (Ullah *et al.*)

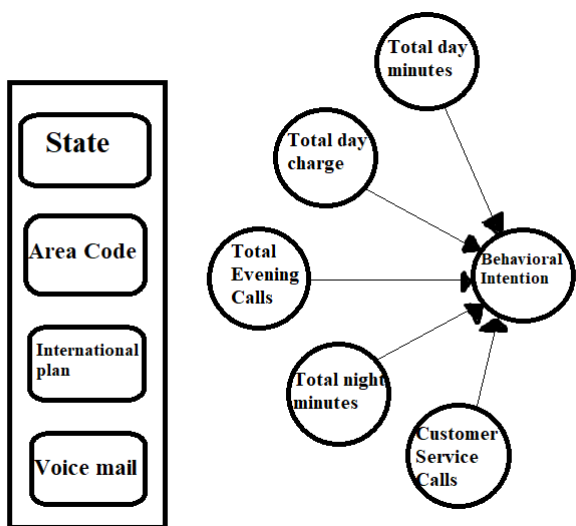


FIGURE 1. The broad strokes of typical customer behaviour

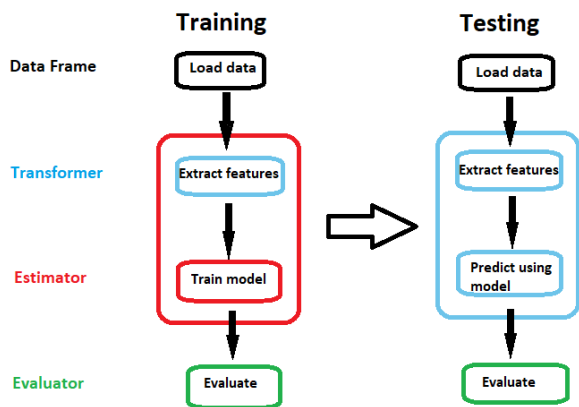


FIGURE 2. An algorithmic flow chart

6. Prediction of the churning through a machine learning algorithm

The quickest way to spot client turnover in a firm is through data analysis. When deciding on the appropriate course of action to be followed by a corporation to sustain a regular trade environment in the product market, for example, using consumer databases is made easier by employing E-commerce. The knowledge of consumer behaviour in relation to their data pre-processing facilities is highlighted for a producer. For instance, it is much preferable to use a machine learning algorithm when conducting a regression study that will show how different businesses and goods affect consumer behaviour. The managing team should be persuaded to take action against customer fraud by

using graphical presentations of the results, it is advised. Tools for analysing analytical data should be used, particularly when providing the marketing team with an instant snapshot of ongoing usual market behaviour. When analysing a company’s trading behaviour, categorization is crucial in determining its strong point and weak point. Additionally, the variable distribution tables will clearly demonstrate the mean and mode features of the purchasing teams in a larger context.

7. Decision Tree

A decision tree is a tree structure that resembles a flowchart, in which each leaf node symbolises the result and each interior node represents a feature (or attribute). A decision tree’s root node is the first node from the top. Based on the attribute value, it learns to partition. It is known as recursive partitioning and it divides the tree in this way. Your decision-making is aided by the flowchart. It is an image that, like a flowchart diagram, faithfully represents how people think. It is easy to grasp and interpret decision trees.

8. Logistical Regression

In the early 20th century, the biological sciences employed logistic regression. Later, it was put to use in numerous social science contexts. When the target variable (dependent variable) is categorical, logistic regression is employed. To identify spam in an email, either (0). whether the tumour is malignant(1) or not(0). Think of a situation where we must categorise if an email is spam or not. Setting up a threshold on which classification can be performed is necessary if we utilise linear regression to solve this problem. If the actual class is malicious, the estimated constant value is 0.4, and the minimum value is 0.5, the data point will be categorised as not malicious, which could have significant effects in real time. This example demonstrates that categorization issues do not lend itself to linear regression. Due of the limitless nature of linear regression, logistic regression enters the picture. They just range from 0 to 1, and nothing else.

9. Naive Bayes Classifier

Naive Bayes classifiers are a subset of Bayes Theorem-based classification methods. It is a collection of algorithms rather than a single method, all of which follow the same essential tenet: that each

pair of characteristics being categorised is independently of the others. To get things going, let's consider a dataset. Consider a fictitious dataset that describes the optimum conditions for a round of golf. Based on the present conditions, each tuple classifies whether the weather is suitable for playing golf as "Yes" or "No".

10. Conclusion

As a result of the discussion above, it can be said that any organisation, whatever of its form, needs to be concerned about customer turnover. Customer retention is the process of preserving a customer's loyalty through comprehending their needs and meeting them appropriately. The organisational management will be helped by a strong churn prediction model to predict consumer churn.

Decision Tree, Naïve-Bayes Classification, Logistic Regression may be useful for estimating the turnover rate depending on the complicated data of the telecommunication sector. The concept of client retention as well as the churn forecast were the main topics of the mentioned paper. Along with the technique, the usage of Decision Tree, Naïve Bayes Classification, Logistic Regression is to improve the churn prediction process has also been covered here.

References

- Baghla, Seema, Gaurav Gupta, et al. "Performance Evaluation of Various Classification Techniques for Customer Churn Prediction in E-commerce". *Microprocessors and Microsystems* 94 (2022): 104680–104680.
- Brandusoiu, Ionut, Gavril Todorean, et al. "CHURN PREDICTION IN THE TELECOMMUNICATIONS SECTOR USING SUPPORT VECTOR MACHINES". *ANNALS OF THE ORADEA UNIVERSITY. Fascicle of Management and Technological Engineering*. XXII (XII), 2013/1.1 (2013): 1–1.
- Burez, J and D Van Den Poel. "Handling class imbalance in customer churn prediction". *Expert Systems with Applications* 36.3 (2009): 4626–4636.
- Farquad, M A H, Vadlamani Ravi, S Bapi Raju, et al. "Churn prediction using comprehensible support vector machine: An analytical CRM application". *Applied Soft Computing* 19 (2014): 31–40.
- Hossain, M M, N J Suchy, et al. "INFLUENCE OF CUSTOMER SATISFACTION ON LOYALTY: A STUDY ON MOBILE TELECOMMUNICATION INDUSTRY". *Journal of Social Sciences* 9.2 (2013): 73–80.
- Idris, Adnan, Muhammad Rizwan, Asifullah Khan, et al. "Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies". *Computers & Electrical Engineering* 38.6 (2012): 1808–1819.
- Kaya, Erdem, et al. "Behavioral attributes and financial churn prediction". *EPJ Data Science* 7.1 (2018).
- Maldonado, Sebastián, et al. "Profit-based feature selection using support vector machines – General framework and an application for customer retention". *Applied Soft Computing* 35 (2015): 740–748.
- Siu, Noel Yee-Man Y, Tracy Jun-Feng J Zhang, Cheuk-Ying Jackie Y Yau, et al. "The Roles of Justice and Customer Satisfaction in Customer Retention: A Lesson from Service Recovery". *Journal of Business Ethics* 114.4 (2013): 675–686.
- Tsai, Chih-Fong, Yu-Hsin Lu, et al. "Customer churn prediction by hybrid neural networks". *Expert Systems with Applications* 36.10 (2009): 12547–12553.
- Ullah, Irfan, et al. "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector". *IEEE Access* 7 (2019): 60134–60149.
- Vafeiadis, T, et al. "A comparison of machine learning techniques for customer churn prediction". *Simulation Modelling Practice and Theory* 55 (2015): 1–9.



© Harini T et al. 2023 Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Embargo period: The article has no embargo period.

To cite this Article: , Harini T, Hari Krishna T , Sushma G , Charankumar Reddy P , and Mohammad Thahir S . “**Predict Customer Churn through Customer Behaviour using Machine Learning Algorithms.**” International Research Journal on Advanced Science Hub 05.05S May (2023): 333–337. <http://dx.doi.org/10.47392/irjash.2023.S045>