


Will the Customer Survive or Not in the Organization? A Perspective of Churn Prediction Using Supervised Learning

Neelamadhab Padhy, Computer Science & Engineering, School of Engineering & Technology, GIET University, Odisha, India*

 <https://orcid.org/0000-0002-8512-3469>

Sanskriti Panda, Computer Science & Engineering, School of Engineering & Technology, GIET University, Odisha, India

Jigyashu Suraj, Computer Science & Engineering, School of Engineering & Technology, GIET University, Odisha, India

ABSTRACT

The technology of machine learning and data science is gradually evolving and improving. In this process, we feel the importance of data science to solve a problem. In this article the main objective is to predict the customer churn (i.e., whether the customer will leave the telecom service or they will continue with the service). In this paper, the authors have also followed some statistical measures. They have computed the mean, standard deviation, min, max, 25%, 50%, 75% values of the data. Mean is the average value of the data values. The standard deviation is a measure of the amount of variation or dispersion of a set of values. The authors have done an extensive data pre-processing and built machine learning models and found that among all the models, logistic regression gives the best performance (i.e., 81.5%). Hence, they chose that as their final model to indicate the churn prediction.

KEYWORDS

Churn Prediction, Extensive Data Pre-Processing Technique, Machine Learning Algorithms

1. INTRODUCTION

Machine learning and data science technology are gradually evolving and improving. Because of this process in many practice fields, we feel the importance of data science to solve a problem. Here we will discuss a real word telecom dataset. This research focuses on employees between the ages of 20 and 39, who are thought to be the key contributors to any organization's highest turnover rate. We have to predict whether the customer will leave the telecom service or continue with the service using a machine learning algorithm. In this dataset, we will apply several algorithms and techniques. We have

DOI: 10.4018/IJOSSP.300753

*Corresponding Author

Copyright © 2022, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

to observe that which of the applied technique gives the best result on the dataset. Accordingly, we have to find the best fit model for the dataset. Employees that are motivated to boost their productivity are more likely to be more productive overall. This means that managers should devote more time to this area of their jobs to obtain a deeper understanding of it. Employees may be aware of their goals and the significance of those goals. The technology of machine learning and data science is gradually evolving and improving. In this process in many of the practice fields, we feel the importance of data science to solve a problem. One of the reasons why data science is important is that we can make high-valued predictions that can guide better decisions and smart actions in real-time without human intervention through data science. Here we will discuss an accurate word telecom dataset. We have to predict whether the customer will leave the telecom service or continue using a machine learning algorithm. In this dataset, we will apply several algorithms and techniques. We have to observe which of the applied technique gives the best result on the dataset. Accordingly, we have to find the best fit model for the dataset. Khan, Y., Shafiq, S., Naeem, A., Hussain, S., Ahmed, S., & Safwan, N. (2019). In this article, telecom data of Pakistan is given the job is to build a model that can predict whether the customer is switching to a different service provider or is sticking to the same service provider. Here Deep learning ANN model is used, and the accuracy is 79%. In this article, Brandusoiu, I., & Todorean, G. (2013). We are given a dataset of 3333 records that contains 21 attributes. Here Support Vector Machine Algorithm is used to predict the customer churn so that the telecom service can get improvised. The model uses the Polynomial kernel function to predict churners and non-churners, and its accuracy is 88.56%. This article analyzes how work stress factors, quality of leadership, work motivation, organizational culture, and job satisfaction affect turnover intention. Prasada, P. P. B., & Sawitri, N. N. (2019), this research Structural Equation Model (SEM) analysis method is used, and Partial Least Square (PLS) analysis tool is used. The study found that work stress and job satisfaction have an impact on the likelihood of leaving. Qualities of leadership, motivation at work, and organizational culture, on the other hand, have little bearing on the possibility of turnover. The empirical data of 350 Kuwaiti mobile telephone subscribers are examined in this study.

Rizomyliotis, I., Poulis, A., Apostolos, G., Konstantoulaki, K., & Kostopoulos, I. (2020), The objective is to investigate the reciprocal effects of customer loyalty and its antecedents. This study models a fuzzy cognitive map (FCM) to analyze the facts. This study shows that the model reaches the equilibrium when brand equity and customer loyalty are increased and reach an optimal level. Balmer, R. E., Levin, S. L., & Schmidt, S. (2020). This paper analyses and explores the location and use of telecommunication and other network businesses for artificial intelligence. This article shows that Artificial Intelligence has a significant role in these network Industries, and due to this, people can get a satisfactory service. Alaskar, L., Crane, M., & Alduailij, M. (2019). This research identifies the best predictors using the Select Best, RFE, and Random Forest (RF) models. Different ML methods like logistical regression, decision tree (DT), naïve bays, SVM, and AdaBoost with optimum hyperparameters have been developed. Several crucial indicators assessed the prediction model's performance in the final phase of the experiment. Two predictive models have been shown to have improved results: DT with Select KBest and the SVM-polynomial kernel with RF. Jae-Hyeon Ahn (2006) discussed the changes in the customer status explain the relationship between churn determinants and churn's probability, which can be explained directly or indirectly through a customer's status changer. They have discussed the retention of clients in the maturing mobile communications sector is one of the most important problems. This study studies customer churn drivers in the Korean mobile telecommunications service market using customer transaction and billing data. (FlorisDevriendt et al. 2021) developed the models for the uplift of client turnover that can be established for predictive customer churn. Adhikary, D. D., & Gupta, D. (2021). The authors examined and compare the performance of over 100 classifiers in a telecom company's churn forecast. We employed well-known classifiers from various families. Churn prediction is a moderately unbalanced group of problems that degrades classifier performance. The Regularized Random Forest classifier has the best accuracy

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/will-the-customer-survive-or-not-in-the-organization/300753

Related Content

Open Source Software in the Arab World: A Literature Survey

Manar Abu Talib (2016). *International Journal of Open Source Software and Processes* (pp. 49-64).

www.irma-international.org/article/open-source-software-in-the-arab-world/179925

A Model for the Successful Migration to Desktop OSS

Daniel Brink (2007). *Handbook of Research on Open Source Software: Technological, Economic, and Social Perspectives* (pp. 154-167).

www.irma-international.org/chapter/model-successful-migration-desktop-oss/21186

Analyzing OSS Project Health with Heterogeneous Data Sources

Wikan Damar Sunindyo, Thomas Moser, Dietmar Winkler and Stefan Biffel (2013). *Open Source Software Dynamics, Processes, and Applications* (pp. 207-230).

www.irma-international.org/chapter/analyzing-oss-project-health-heterogeneous/74670

Evaluating the Potential of Free and Open Source Software in the Developing World

Victor van Reijswoud and Emmanuel Mulo (2012). *International Journal of Open Source Software and Processes* (pp. 38-51).

www.irma-international.org/article/evaluating-the-potential-of-free-and-open-source-software-in-the-developing-world/101205

Open Source Adoption Index: Quantifying FOSS Adoption by an Organisation

Sanjeev K. Saini, C. N. Krishnan and L. N. Rajaram (2010). *International Journal of Open Source Software and Processes* (pp. 48-60).

www.irma-international.org/article/open-source-adoption-index/51586