

Directing customers to subscription through app behavior analysis

¹Rithanya P, ²Thirukkural Selvan A L, ³Sowmiya V, ⁴Suseendran S

¹ Bachelor of Engineering, Electronics and Communication Engineering, Bannari Amman Institute of Technology, Erode, India

² Bachelor of Engineering, Computer Science and Engineering, Bannari Amman Institute of Technology, Erode, India

³ Bachelor of Engineering, Electronics and Communication Engineering, Bannari Amman Institute of Technology, Erode, India

⁴ Assistant Professor Level -II, Computer Science and Engineering, Bannari Amman Institute of Technology, Erode, India

ABSTRACT

Mobile applications have become essential tools for businesses to interact with their clients in the changing environment of the modern market. Platforms like YouTube Red and Pandora Premium serve as effective examples of the shift from free to paid app memberships. The difficulty, though, is in identifying those who are less likely to subscribe, which calls for precision targeting and customized marketing initiatives. To improve ease and efficiency in financial tracking, this study explores a fintech company's entry into the mobile app market with a premium membership model. This research was motivated by the requirement for efficient targeting and conversions. This study's main objective is to strategically deliver promotions to users who are most likely to respond positively to maximize the return on investment. It starts by carefully examining data obtained from users' app usage during a free trial period of 24 hours. Predictive models are then built using cutting-edge machine learning techniques like Logistic Regression, Support Vector Machines (SVM), and XG Boost. The creation of an interactive web page is a crucial element of this process. This website's dynamic estimates of consumer registration in the premium membership are displayed, offering a clear and interesting user experience. The identification of potential subscribers and the use of customized marketing techniques are among the key findings. By making the most of available resources, streamlining marketing initiatives, and encouraging openness, the technique raises conversion rates. The predictive model encourages conversions and supports well informed decisions, aligned with the business's financial goals.

Keywords: Mobile applications, Premium membership, Machine learning, Marketing techniques, Predictive model.

1. INTRODUCTION

A distinctive strategy is being offered in response to the challenges mentioned in the existing methodology. The fundamental concept is to utilize machine learning algorithms to predict user behaviour and give an organization the ability to choose precise marketing strategies. Unlike the conventional approaches, this methodology's recommended approach makes use of the latest technologies to accomplish precise targeting and improved conversion rates. The incorporation of several machine learning techniques, such as Logistic Regression, Support Vector Machines (SVM), and XG Boost is one of the methodology's significant innovations. These algorithms provide a strong framework for predictive modelling, making it possible to recognize customers who are less likely to subscribe with more accuracy. For comparison analysis, conventional methods like Decision Trees and Naive Bayes are also investigated. In addition, the suggested methodology includes developing a website page that dynamically shows forecasts regarding

whether customers are likely to sign up for the premium membership. The procedure is made more interactive and transparent thanks to this user-friendly interface, which improves the user experience. Additionally, the system emphasizes hyperparameter tuning. The machine learning models are optimized for hyperparameters to ensure the best performance. As a result, forecasts get more precise, and marketing tactics become more successful overall. In conclusion, the suggested methodology incorporates distinctive features such as numerous sophisticated algorithms, interactive web displays, and hyperparameter adjustment to improve the accuracy and efficiency of the prediction model in addition to addressing the issues raised by the current approach.

2. LITERATURE SURVEY

The literature surveys linked to recent app behaviour analysis deployments that direct users to subscriptions are summarized in this chapter. After studying and reviewing all the relevant cases then discussed with our guide and identified the problem statement. According to Peter Foy's 2020 study, numerous companies have switched to the "freemium" pricing model in recent years, necessitating the requirement to effectively convert free app users into premium paid subscribers. Researchers and professionals have used machine learning approaches to forecast the possibility of user registration in premium subscriptions to address this difficulty. The blog post by Peter Foy offers a thorough solution to this issue. In order to target marketing efforts toward customers who are most likely to upgrade to the premium version, the study establishes a clear business objective. The study uses exploratory data analysis and correlation plots to figure out the most important determinants of subscription behaviour by examining customer activity data, such as installation time, engaged features, and screen interactions. With logistic regression and L1 regularization, the machine learning model achieves a remarkable accuracy of about 77%. Omprakash Yadav et al., 2019 have studied that organizations are becoming more aware of the value of offering their clients high-quality services as a result of increased competition. Companies are offering free trial versions of their apps and researching user behavior data to figure out client preferences in order to satisfy consumer requirements and increase customer loyalty. With this specific dataset, they generate predictions about user-paid subscriptions using several algorithms, including linear regression, Decision tree and logistic regression. The dataset includes customer age, screens opened, time of subscription, and other information. This enables them to distinguish their services based on actual customer needs. For measuring the service level and quality of service organizations, it is essential to satisfy customer service needs. According to a 2020 study by Qiang Ren, the rapid growth of mobile Internet services and the rise in the number of users have created major challenges for internet service providers and increased competition. Accurate marketing based on user behaviour monitoring has become crucial to staying ahead. However, the big data era's enormous amount of data is difficult for traditional design to handle. This article offers a way for internet service providers to adjust to the shifting environment by focusing on traffic management and performing in-depth analyses of user behaviour patterns to identify market needs. The proposed solution consists of a cloud based mobile Internet big data user behaviour analysis engine that includes system architectural design, big data warehousing and preprocessing components, as well as essential user behaviour analysis modules. Charlie C. Chen et al., 2018 studied that the music industry has strategically focused on boosting paid subscriptions to convert free listeners into paying customers in an effort to grow revenues and ensure economic sustainability. Nevertheless, reaching this goal has been difficult, mostly because there is a lack of a thorough understanding of the psychological and social buying reasons that influence consumers subscription selections. In this study, 243 customers of on-demand music streaming services have their two separate buying motivations—social influence and hedonic performance. According to our data study, social influence has a major impact on consumer perceptions toward music streaming, which in turn influences their purchase intention.

3. METHODOLOGY

METHODOLOGY PROPOSED OUTPUTS OF OTHER ML MODELS:

1) RANDOM FOREST REGRESSION:

Random forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both classification and regression problems in ml. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

ACCURACY: 0.7905

2) XGB CLASSIFIER: (XG BOOST ALGORITHM)

Xgboost is an optimized distributed gradient boosting library designed for efficient and scalable training of machine learning models. It is an ensemble learning method that combines the predictions of multiple weak models to produce a stronger prediction.

ACCURACY: 0.9811

3) ADABOOST CLASSIFIER:

Adaboost, short for adaptive boosting, is an ensemble machine learning algorithm that can be used in a wide variety of classification and regression tasks. It is a supervised learning algorithm that is used to classify data by combining multiple weak or base learners (e.g., decision trees) into a strong learner.

ACCURACY: 0.746556

4) LOGISTIC REGRESSION:

Logistic regression is much like the linear regression except that how they are used. Linear regression is used for solving regression problems, whereas logistic regression is used for solving the classification problems.

ACCURACY: 0.6215

5) SVM CLASSIFIER (SUPPORT VECTOR MACHINE) MODEL:

The goal of the svm algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. Svm chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as support vector machine.

ACCURACY: 0.479

6) KNN ALGORITHM

K-nn algorithm can be used for regression as well as for classification but mostly it is used for classification problems. K-nn is a non-parametric algorithm, which means it does not make any

assumption on underlying data. It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

ACCURACY: 0.5668

4. CONCLUSION

This study presents an innovative approach to the evolving landscape of mobile applications and premium memberships. The primary aim was to enhance financial tracking and maximize return on investment through precise targeting and tailored marketing strategies. By employing advanced machine learning techniques, including Logistic Regression, Support Vector Machines (SVM), and XG Boost, predictive models were developed to identify potential premium subscribers. An interactive web page was introduced to dynamically estimate consumer registrations for premium memberships, enhancing the overall user experience. Furthermore, hyperparameter tuning optimized machine learning models, leading to more accurate predictions and more effective marketing tactics. This methodology represents a departure from traditional approaches, incorporating advanced algorithms, interactive interfaces, and hyperparameter optimization. It addresses the challenges of predicting user behaviour in a rapidly changing mobile app market, providing businesses with a powerful tool to achieve their financial objectives and enhance customer engagement.

5. REFERENCES

- [1] Hoang Thi Thuy Dung and Tran Duc Quynh. 28-October-2021. "Prediction of Customer Behavior using Machine Learning", Proceedings of the 2nd International Conference on Human centered Artificial Intelligence
- [2] Chen, Q., Zhao, X and Zhang, M (2020), "Analyzing customer behavior in mobile app usage", Industrial Management & Data Systems, Vol. 117 No. 2, pp. 425-438. [3] Mohamed Hamadab, Mahmoud SalahEldin Kasema, Islam TajEddinc .3 feb 2023 "Customer Profiling, Segmentation, and Sales Prediction using AI in Direct Marketing", Proceedings of the
- DOI:10.48550/arXiv.2302.01786
- [4] Omprakash Yadav, Vivekkumar mishra, Abubaker khan, Tejas Nivatkar . May 2019. "Driving Customer Through Subscription Through App Behavior Analysis", Proceedings of the International Journal of Information And Computing Science Volume 6, Issue 5.
- [5] V. Shrirame, J. Sabade, H. Soneta and M. Vijayalakshmi. 2020. "Consumer Behavior Analytics using Machine Learning Algorithms," 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT),pp.16,doi10.1109/CONECCT50063.2020.9198562.
- [6] Xie, Y, Li, X, Ying, and W Ngai, E. 2009. "Customer churn prediction using improved balanced random forests, Expert Systems with Applications" 36(3):5445–5449. [7] Larivière, B., and Van den Poel, D. 2005. "Predicting customer retention by using random forests and regression forests techniques." Expert Systems with Applications 29(2):472–484.
- [8] Buckinx, W and Van den Poel, D. 2007. Predicting customer loyalty using the internal transactional database. Expert Systems with Applications 32(1):125–134. [9] C. C. H. Chan.2008. Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer, Expert systems with applications 34 (4) 2754– 2762.
- [10] M. Hosseini, M. Shabani, New approach to customer segmentation based on changes in customer value, Journal of Marketing Analytics 3 (2015) 110–121.
- [11] S.-Y. Kim, T.-S. Jung, E.-H. Suh, H.-S. Hwang, Customer segmentation and strategy development based on customer lifetime value: A case study, Expert systems with applications 31 (1) (2006) 101–107.

- [12] X. He, C. Li, The research and application of customer segmentation on e-commerce websites, in: 2016 6th International Conference on Digital Home (ICDH), IEEE, 2016, pp. 203–208.
- [13] P. Q. Brito, S. Almeida, C. Soares, A. Monte, M. Byvoet. 2015. “Customer segmentation in a large database of an online customized Robotics and Computer-Integrated Manufacturing” 36 93–100.
- [14] A. Sheshasaayee, L. Logeshwari, An efficiency analysis on the tpa clustering methods for intelligent customer segmentation, in: 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), IEEE, 2017, pp. 784–788. [15] P. Grau-Carles, M. T. Ballestar, J. Sainz. 2018. Customer segmentation in e-commerce: Applications to the cashback business model, Journal of Business Research 88 407–414.

