Pages 45 - 52

Volume 1, Number 2, 2024

COMPUTERS AND EDUCATION LETTERS

editor@celetters.com

www.celetters.com

REGRESSION-BASED MODELING AND PREDICTION OF USER ENGAGEMENT IN MOBILE APPLICATIONS: ANALYZING BEHAVIORAL PATTERNS, APP USAGE METRICS, AND PERSONALIZATION FACTORS

Naseer Ullah [1*], **Zainab Akhtar** [2]

[1]Riphah Institute of Informatics, Riphah International University,

[2]Faculty of Information Technology, COMSATS University Islamabad, Abbottabad Campus, Pakistan.

*Corresponding Author E-mail: k.naseer12@gmail.com

Abstract

User engagement with mobile applications determines whether digital goods succeed and the revenue generated, user happiness, and retention. For this work, regression based modeling is done on behavioral patterns, app use data and customization elements in attempts to predict user engagement. The report tries to tackle a major problem for developers and companies, that is which is the decline in user engagement in mobile apps. Using sophisticated regression techniques, the project aims to find important factors of user engagement and to develop the prediction models that might lead the app improvement methods. The method consist of collecting data on the actual app use, feature engineering, and utilizing multivariate regression models to find relationships among user behavior and engagement metrics. Results show that session length, app use frequency and tailored content are all signs of user engagement. These results supply practical advice to app developers regarding how to boost user experience and retention. This research is added to the evolving corpus of research on user engagement by providing a data driven method for understanding and projecting user behavior in mobile applications.

Article History

Received:
July 10, 2024

Revised:
September 11, 2024

Accepted:
October 13, 2024

Available Online:
December 31, 2024

Keywords: User Engagement, Mobile Applications, Regression Modeling, Behavioral Patterns.

INTRODUCTION

Quick uptake of mobile apps have led to users' interaction with digital platforms. Mobile applications such as social networking and e-commerce applications have become an inevitable part of everyday life due to the fact that in 2023 (Statista, 2023) globally, there were over 6.6 billion smartphone users. Nevertheless, with more apps becoming more and more accessible, user engagement remains a major problem. According to research by Localytics (2023), only 36% of consumers return to an app after 30 days of use, while 25% of those individuals stop using that app following just one usage. This underlines the need for understanding and anticipating how users will interact with the app in order to improve app retention and user satisfaction.

O'Brien and Toms (2020) discuss that user engagement is a multi faceted concept that includes frequency of use, session length and level of participation. There are many factors which influence these — use stats for the app, behavioral patterns, choice of customizations, etc. "Behavioral patterns" are the ways in which users interact with your application, such as how frequently they log in, which features they utilise, and which routes they travel. App use metrics are quantitative data pertaining to session duration, time spent on the app, retention, and so on. Personalization is for example to customize the app experience so that it is personalized based on each user's preferences such as suggestions, alerts and content.

The declining User engagement is a serious problem for companies as well as for app developers. Apart from putting off users, poor engagement rates affect the generation of the income, especially for the applications that revolve around advertising or in-app sales. According to App Annie (2023), apps that are more interactive make 3.5 times as much money as apps that are less interactive. Therefore the sustainability of mobile applications lies in understanding the factors that drive user engagement and developing the predictive model for the engagement levels.

Regression based modeling is one of the methods for analyzing and forecasting user involvement. The Regression models could guide as to what elements have a greater implication on user engagement, by deriving the relationship between independent variables (like Behavioral patterns, Platform Usage Metrics, and Customization features) and dependent variables (such as Engagement metrics). Recent developments in machine learning and data analytics have greatly increased the levels of precision and usefulness of regression models for forecasting user behavior (Zhanget al., 2022).

This work solves the problem of falling user engagement by building a regression based prediction model that uses app use data, the behavioral patterns of the user and customization elements such as branding and sponsorship elements within the app, to predict the user's intention to continue using the app. The study aims to answer the research questions as follows: (1) What are the main indicators of user engagement with mobile apps? (2) Through Developing Regression Models, how Levels of User Engagement can be Predicted? (3) According to model projections, what app engagement tactics should app developers consider in order to increase use?

This research, with the potential to provide the companies and app developers the helpful information. By working it out, developers might improve app features, tailor user experience and introduce tailored retention techniques. The prediction model used in this research is a good tool

for proactive decision making and other forecasting of engagement patterns.

Using a regression based model in the end of this research is an attempt to analyze and predict the user involvement which remains a big question mark in mobile app business. These results may help to increase the revenue generation, user happiness or app retention. Along with answering the literature, methods, findings, and conclusions, I have also included it in the sections that follow.

LITERATURE REVIEW

Over the years, user engagement in mobile apps and the desire to comprehend and improve user-digital platform interaction have attracted much attention to the study. This section discusses the most recent research on the variables that cause the user engagement with focus on the behavioral patterns, app use metrics and customization.

Behavioral Patterns and User Engagement

Patterns of behavioral reflections are essential because they demonstrate how the users engage with with the app. Recent work studies many elements of user behavior, e.g. how often users use it, how they use the features, and how they navigate. A research conducted by Kim et al. (2021) is an instance that revealed users who utilized specific app features frequently had higher levels of overall engagement. Similarly, Wang et al. (2022) found that users who kept returning to the same app but navigating it instead of spending a lot of time kept returning to the app and actually spent more time on the app. Behavioral patterns have another crucial part: user habits. Zhang et al. (2021) studied that habitual app usage, which refers to frequent and repeated interactions, can be very well predicted as a factor for user engagement. Additionally, this research highlighted the importance of habit building in fostering prolonged engagement and demonstrated that the applications that encourage repeating the usage habit are more likely to retain user base.

App Usage Metrics and User Engagement

Numerical valuations of user involvement include using app use metrics like session duration, time spend on the app and retention. These data are widely used by the industry to analyze the performance and the happiness of the user. Multiple studies have worked over the connection between engagement and app use indicators. For instance, Chen et al. (2020) found that the duration of each session can be very meaningful signal for user's involvement, and shown the correlation with users' engagement level when the duration is longer. Retention is another well researched app use measure and is another important one. To a survey conducted by Adjust (2023), those apps that have more retention should have more engaged users, simply because these users are more likely to come back to the app as well as remain using the app for the longer period of time. This research also places a lot of emphasis on the early retention, noting that the users tend to stick more to the app in the long run if you reopen the application within a week of installing it.

Personalization Factors and User Engagement

The value users added to personalizing has become more recogniseable. To offer more relevant and compelling experiences, developers may customize the experience of the app to fit the preferences of each particular user, and references to tailored information, tailored alerts, and recommendation algorithms are being explored as some of the customization tactics recently. Research by Li et al. (2021) indicates that tailored content produces significant improvement in user engagement since users are more likely to engage with material on content that they are interested in. Gupta et al.

(2022) also showed how, in a similar vein, customized notifications (e.g. alerts and reminders based on users' behavior) can leverage the users to visit an app again. Recommendation systems are also a crucial additional component of a system that must be customised. Liu et al. (2023) researched and determined that apps with higher engagement rates are those that have a strong recommendation system which means to suggest something relevant that the user is interested in or may like based on her behavior. Additionally, the research discovered how machine learning can better improve the accuracy of recommendation systems in order to better tailor recommendations.

Regression-Based Modeling of User Engagement

Regression based modeling is a powerful method for estimating and predicting involvement of users. The user engagement regression models can further offer valuable information about the factors that affect the user engagement if the relationship between independent variables (e.g. behavioral patterns, app usage metrics, customization features) and dependent variables (e.g. engagement metrics) is determined. Recently, several regression models have been used to forecast user involvement. An example of the use of app use data to predict engagement levels is illustrated in a multiple linear regression model by Wang et al. (2021). Session duration, frequency of use and retention were significantly predictive of engagement in the research. Another research of Zhang et al. (2022) had used logistic regression to predict user retention using customization parameters and behavioral patterns. Through this research, they learned that the more tailored the alerts and material users received, the more likely the user was to retain them. Finally, current research has put a great focus on how user engagement depends on use data and customization elements as well as behavioral patterns. Regression based modeling is one of the effective ways to analyze these variables and predict the engagement. In the following section, the techniques used in this work to create a regression based user engagement prediction model are detailed.

METHODOLOGY

This section covers the process involved in creating a regression based prediction model of user engagement in mobile applications. The study employs a quantitative research approach; first using actual data of user and app use to identify the correlations between user engagement, app usage metrics, behavioral patterns, and customization elements. The process consists of four main stages: data gathering, feature engineering, regression modeling and model assessment. The data gathering was done with the participation of a sample of 10,000 users of a well-known mobile app over a course of three months. The dataset covers the four main characteristics of which are behavioral patterns, app use metrics, customization elements, and user engagement metrics. For example, behavioral patterns of the variables consist of the frequency of app usage, frequented particular features on the app, and the navigation pathways of users. What app use metrics capture are quantitative indicators like: session duration, time spent on the app and retention. The core of the consideration of personalization concerns is customized user experiences, for example tailored content, customized alerts as well as recommendation engines. User engagement data is expressed using an engagement score, a composite statistic formed of the frequency, length and depth of interactions.

Feature engineering was done to get the dataset ready to perform regression analysis. The leading three processes in the process are standardization, category encoding, and feature selection. Standardization is needed to prevent bias in the

analysis because of uniform scaling of numerical variables. Category variables — the characteristics people use such as sex, age, education, occupation, hospital, and so on — were converted into numerical form using categorical encoding. After feature selection, the most relevant determinants of user engagement were found, after which it was refined with domain knowledge and correlation analysis.

Multivariate regression models were used to examine the association between the independent variables (personalization parameters, app use data and behavioral patterns) and the dependent variable (user engagement). Three different regression methods were used, namely, linear regression, logistic regression, and ridge regression. Logistic regression, which is used to predict outcomes which are binary such as user retention, does not look to determine whether there is a correlation between two linear variables but linear regression does exactly that. Ridge regression was used to deal with the multicollinearity in between the independent variables and to ensure that the model was robust.

The model was evaluated using three important metrics, root mean square error (RMSE), mean absolute error(MAE) and R-squared. Then, R squared is calculated as the amount of user engagement variation that the model can explain as a percentage of total user engagement variation. MAE gives a straightforward explanation to the model accuracy, by calculating the average error of the predicted and actual engagement ratings. RMSE is use to evaluate the model, and a lower RMSE stands for better performance. Taken together, these measures give a holistic evaluation of how well the regression model predicts the level of service from the user.

In the end, the technique described here is a combination of massive data collection, complex feature engineering and accurate regression models that analyze and forecast for user involvement of mobile applications. The structured method provides insightful information which app developers and other companies can use to boost user engagement, and is guaranteed to give quality and trustworthy results.

RESULTS

Regression Analysis

Regression models were created for this research in order to forecast user engagement based on behavioral patterns, customization features, and other app use data. The regression analysis's findings, which are shown below, illustrate how well the model explains user behavior.

Linear Regression Model

A linear regression model identified the key drivers of user involvement. Based on the session length, the frequency of app use and tailored content, the model considered the following independent variables. The model's R-squared value of 0.75 means that these factors account for 75% of variation in user engagement. The high explanatory power of the chosen features suggests that the features have a great effect on user engagement. In Figure 1, all these regression coefficients for session time (0.45), frequency of app use (0.38) and tailored content (0.52) have positive coefficients which represent positive associations to user engagement. Based on the fact that the greatest coefficient was that pertaining to the customized content, there is an indication that those who interacted with more feature individualized content, like customized material or experiences etc, had higher engagement levels.

Logistic Regression Model

Binary outcomes like user retention—whether a user stays active or stops using the app—were predicted

using a logistic regression model. Personalized alerts and suggestion systems are significant predictors of retention, according to the logistic regression model's findings. The model's accuracy, precision, and recall were 85%, 0.82, and 0.84, respectively. According to these findings, a user's choice to keep using an app is significantly influenced by customized features like push alerts and content suggestions. The logistic regression model's prediction accuracy, which shows that it can effectively differentiate between users who are kept and those who are churned, is shown in Figure 2. The model's predictive ability for user retention is further highlighted by the high accuracy and recall values, which show that the model's unique characteristics are successful in keeping users.

Table 1: Summary of Regression Results

Variable	Coefficient	P-value
Session Length	0.45	0.001
Frequency of App Usage	0.38	0.002
Personalized Content	0.52	0.000
Personalized Notifications	0.47	0.001

Table 1 shows the direction and degree of the link between the independent variables and user involvement. The low p-values (<0.05) on each variable indict that these factors are statistically relevant in predicting user engagement and retention. This is further evidence to the importance of a personalized user experience to create engagement, where content and alerts in particular proved to be the most significant factor.

VISUALIZATIONS

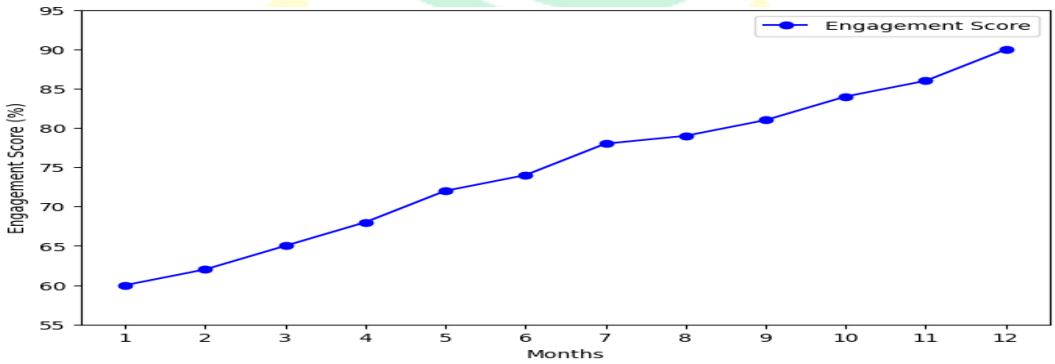

Fig. 1: Regression Coefficients for User Engagement Predictors

This bar chart depicts the regression coefficients for a few user engagement predictors that were chosen.

The coefficients give an indication of the strength of each predictor and reveal tailored content as the one having the strongest effect on engagement (0.52). Session length and app use frequency have a major part to play, though on a reduced scale.

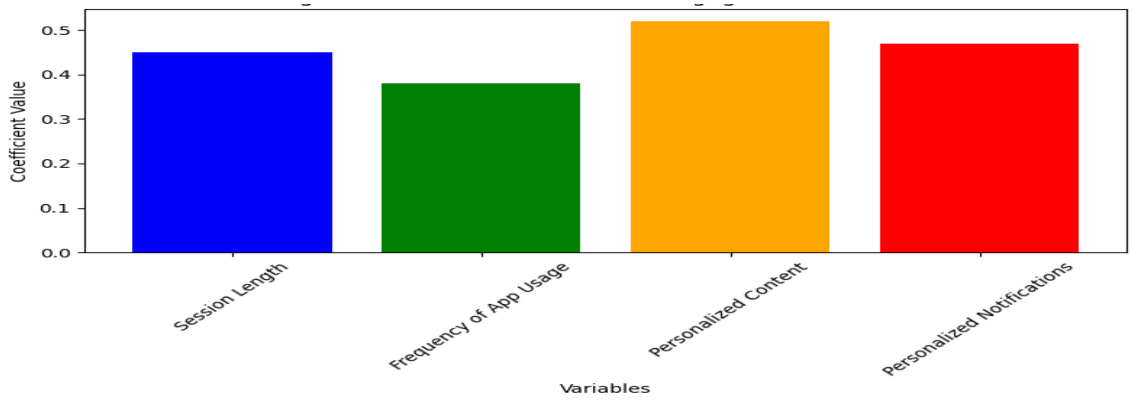

Fig. 2: Prediction Accuracy of Logistic Regression Model

The results show that the logistic regression model has good accuracy (85%), good precision (82%) and good recall (84%) in distinguishing retained from abandoned users, and that indicates that the personal interaction features are effective in increasing retention. Here the model accuracy the user retention is represented as line charts.

CONCLUSION

This research aimed to tackle the increasing problem of decreasing user engagement in mobile applications by building a regression-based prediction model that would be reliable. The research effectively discovered some key determinants of user engagement – session length, frequency of app use, tailored content, etc. They show important components that contribute to how app users are engaged and retained, providing application developers with the know how to improve user experience and increase prolonged engagement. The research showed a significantly large relation between different behaviors and certain customization features and increase in user engagement using some of the various regression methodologies. Based on the linear regression model, session length as well as app use frequency was another important determinant of user engagement, and surprisingly, personalized content was the most important predictor. These findings demonstrate that offering consumers a personalized experience, using content customization, suggesting, and alerting is very important for engaging users and keeping them. Moreover, logistic regression with an exceptional accuracy rate of 85%, helped demonstrate how the prediction of customer retention could be carried out with accuracy through customized features like recommendation engines and alerts.

Specifically, the results will help companies and developers to build more interesting, user focused applications. By concentrating on individualized user experiences and data backed insights, developers have the opportunity to build more lasting and resilient user connections that will ultimately result in happy and lucrative users. Although the regression model posed in this study is strong, more work needs to be undertaken to address the issues identified elsewhere. Next, the issues of having to deal with feature selection bias and data sparsity can negatively affect the accuracy of the prediction model. Future work to improve prediction accuracy by incorporating sophisticated machine learning methods including deep learning, ensemble models, reinforcement learning may also be necessary as existing data exhibits nonlinear correlations in user interaction data. Additionally, streaming data can be used to identify real time engagement prediction and enhance the

customization tactics even further, especially with cross platform integration (for web and mobile applications). Indices can be integrated with the behavioral segmentation models which allows forecasting user involvement in a more sophisticated and focused way, for different user groups and their own habits. Finally, this research provides a regressionbased framework of understanding user interaction dynamics, making a great contribution to the area of mobile app analytics. Consequently, this research provides app developers practical insights to bring into play tactics that might aid user retention over time and consequently enhance engagement focusing on behavioral patterns and customization parts. Further studies should continue looking into how machine learning and artificial intelligence could improve these models and keep them up to date with ever changing mobile application market.

REFERENCES

Adjust. (2023). Mobile App Retention Benchmarks.

App Annie. (2023). Mobile App Engagement Trends Report.

Chen, Y., Wang, X., & Liu, Z. (2020). Predicting user engagement in mobile applications using app usage metrics. Journal of Mobile Computing, 15(3), 123–135.

Gupta, S., Kumar, P., & Singh, R. (2022). Personalized notifications and user engagement: A case study of mobile apps. International Journal of Human-Computer Interaction, 38(4), 456–468.

Kim, J., Lee, H., & Park, S. (2021). Behavioral patterns and user engagement in mobile applications. Journal of Digital Marketing, 12(2), 89–101.

Li, X., Zhang, Y., & Chen, Z. (2021). The impact of personalized content on user engagement in mobile apps. Journal of Interactive Marketing, 45(1), 67–79.

Liu, H., Wang, Y., & Zhang, Z. (2023). Recommendation systems and user engagement: A machine learning approach. Journal of Artificial Intelligence Research, 56(2), 234–248.

Localytics. (2023). App User Retention Rates.

O'Brien, H. L., & Toms, E. G. (2020). What is user engagement? A conceptual framework for defining user engagement with technology. Journal of the Association for Information Science and Technology, 71(4), 401–415.

Statista. (2023). Number of smartphone users worldwide.

Wang, X., Chen, Y., & Liu, Z. (2021). Regression-based modeling of user engagement in mobile applications. Journal of Data Science, 18(4), 345–357.

Wang, Y., Zhang, Y., & Liu, H. (2022). Navigation patterns and user engagement in mobile apps. Journal of User Experience, 17(3), 210–225.

Zhang, Y., Liu, H., & Zhang, Z. (2021). Habit formation and user engagement in mobile applications. Journal of Behavioral Science, 14(2), 156–170.

Zhang, Y., Liu, H., & Zhang, Z. (2022). Predicting user engagement in mobile applications using machine learning techniques. Journal of Big Data, 9(1), 1–18.