Identifying the effect of Success Factors on Predicting Android Mobile App Success on Google Play Store

Dr. Anureet Kaur
School of Computer Science and Applications,
Khalsa College Amritsar, Punjab,143002, India
E-Mail: anumahal@gmail.com

Abstract—At the present time, mobile apps play an imperative role in everyone's life. Thousands of apps are being put on the Play Store market for users to install every day. But the life of the app on the Play Store depends on its success parameters. As there is substantial growth in the mobile app market, the competition among software app developers has also increased. Several attributes of an app, attributes that are either internal or external to it, can perform a substantial role in deciding whether the app will be successful or not. In this research paper, an effort has been made to find the success factors that may affect the success of the app. For this, a literature survey is first conducted to find these success factors, and then a survey is conducted through an online questionnaire from mobile app users. Findings from the survey corroborate that out of 14 identified success factors, the security of on app plays a pivotal role. After that, design quality and content quality equally play another major role in the success of an app. The mobile app reviews/input collected via the survey are further utilized as influencers/predictors in predicting the overall success of an app using Machine learning algorithms. To achieve this, various Machine Learning algorithms are applied using "Python web-based interactive Jupiter Notebook 7.0.8". The accuracy of the final selected model is 96%, which is the highest accuracy among many existing models in the literature.

Keywords: Mobile Applications, Prediction, Android, success, failure, survey, Machine Learning.

INTRODUCTION

The use of mobile technology has boomed all over the world in the last 20 years. As the demand for mobile devices has drastically increased in the last two decades, the mobile app development and deployment on respective app stores of mobile device has also amplified. Unlike many other mobile OS, Android apps play a major role in the mobile software industry. Google Play Store on Android mobile devices provides access to millions of apps to work on. Daily, this platform is flooded with an enormous number of fresh apps. Developers of these apps are in anticipation that their hard work should not be in vain and try to build an app that will be successful in this competitive market. It would be beneficial for the developers to know in advance if their app will be successful or not before uploading it to the Google Play Store. It would prove to be highly advantageous for developers to have a method that can aid in predicting the success possibility of an application.

The success of an app can be ascertained by features like ratings, number of installs, and reviews rather than the revenue it generates. Generally, many apps in the Play Store are not charged; the revenue generated by subscriptions, in-app purchases, and in-app advertisements is practically unknown. Although many apps are added to the Play Store daily, only a few apps achieve their monetary benefits and

 https://doi.org/10.22232/stj.2025.13.01.13

endure in their respective competitive marketplaces. It would be really helpful if the possibility of app success could be ascertained in advance.

This paper is divided into four sections. After the introduction in section 1, the methodology adopted in this paper is presented in section 2. Then in section 3, the entire process of forming the Research Questions (RQs), the method adopted for answering the formed RQs, and finally, the results of the RQs are presented. Section 4 presents the conclusion and future work.

MATERIAL AND METHODS

The analysis of the survey starts with defining the Research Questions and the elaboration conduct of the online survey.

Defining Research Questions and Conducting the Survey

The goal of this research is to help developers understand and predict the app's success on the Play Store. In reaching this final goal, first, there is a need to identify what forms the basis. In light of this, three Research Questions (RQs) are formed:

RQ1. What are the different mobile app success factors?

RQ2. Can success factors affect the mobile app success prediction process?

RQ3. Is there a way/model to predict an app's success on the Google Play Store using Machine Learning Algorithms?

The answer to the first RQ is achieved through a literature survey from existing sources. RQ2 is further answered with the help of an online survey. For conducting the survey Google form was created, and a total of 09 questions were designed to be asked in 2-3 minutes. The questionnaire is forwarded as a Google form link https://forms.gle/ke9DK66LTuwwtCVK9 through various social media messaging systems to around 1200 users as a volunteer from 3.07.2024 to 3.10.2024. The response rate was 67 percent. The questionnaire is added in Appendix A. For accomplishing RQ3, various Machine Learning algorithms are used. Data sets for experimental analysis are collected from various sources (mostly mobile app developers and freelancers).

RESULTS REPORTING

The results for RQ1, RQ2, and RQ3 are presented in this section.

Answering RQ1

After a thorough review of existing research in the literature, some of the success factors were identified for investigation under this research, thus answering RQ1: -

1. ***Subjective Norms:*** The intent to use mobile apps can be influenced by the opinions of one's friends, and the adherence to social norms can also impact this decision [1], [2], [3], [4], [5], [6].
2. ***Perceived Compatibility:*** Compatibility measures the level of conformity between newly released products [1], [3], [7], [8], [9].
3. ***Perceived Playfulness:*** The notion of perceived playfulness pertains to the interaction between users and information systems, stemming from their focus, inquisitiveness, and delight.[1], [10], [11], [12], [13].
4. ***Satisfaction:*** Satisfaction is obtained through practical application, which ultimately influences the users' intention to continue. [1], [8], [9], [14], [15].
5. ***Perceived Usefulness:*** This refers to the extent to which an individual perceives that the utilization of a specific product or system will improve their job performance [1], [11], [13], [14], [16], [17], [18], [19].
6. ***Perceived Ease of Use:*** This is the extent to which an individual perceives that utilizing a specific product or system will be uncomplicated. [1], [10], [13], [17], [20], [21].
7. ***Content Quality:*** Quality of data supplied by Apps [22], [23], [24], [25].
8. ***Reliability:*** Refers to App crashes or any other cause for an app not working as anticipated [9], [19], [22], [26].
9. ***Security:*** It deals with probable harm to personal data, protection, possible disclosure to deceit, and malicious action on the web [18], [22], [26], [27], [28].
10. ***Design Quality:*** It refers to the user interface, i.e., what users, when they are engaging with your app, see, touch, and experience on their phone [9], [13], [22].
11. ***Usefulness:*** Professed usefulness of the app, its functionality, and services [18], [19], [22], [29], [30], [31].
12. ***Performance:*** It indicates how the app is functioning and how responsive the app is to the end-user [18], [22], [30], [31], [32].
13. ***Social Media integration:*** Allows users to easily share content with their friends on social networks [9], [13], [33].
14. ***Regularly updated:*** It means there is a certain system of mobile app updates that arise out of necessity, such as bug fixing, new feature releases, and so on [9].

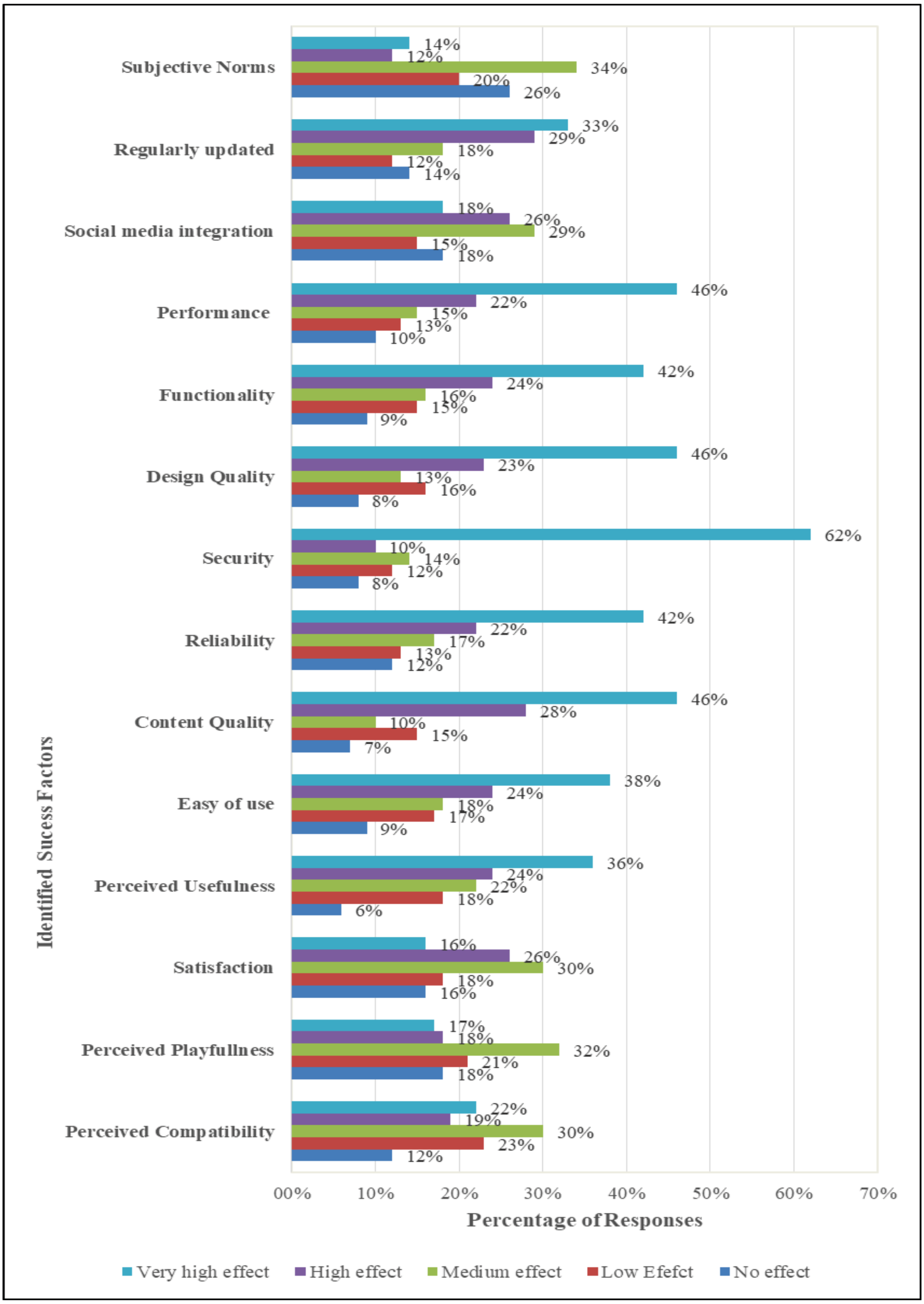

Fig 1: Influence of Success Factors on Mobile App Success on the Play Store by Respondents

Answering RQ2

The answer to RQ1 forms the basis for getting an answer to RQ2. The goal of the online survey is to identify the impact of various identified success factors on success prediction. The target audience of this survey is actual users of the mobile app. Figure 1 presents the influence of success factors on mobile app success in the Play Store by respondents from "No Effect" to "Very High Effect".

The fourteen success attributes/factors identified from the existing literature are further investigated by mobile app users via an online survey. The extent of influence of success factors on mobile app success prediction is collected on a 5-point Likert scale from "No Effect" to "Very high Effect". The result of these responses is further used for deriving weights using weighted average, which is popular and applied in many studies. [34], [35][36]. The formula for the weighted average is shown in Eq. (1).

$$W_j = \frac{\sum_{k=0}^{4}(x * w_k)}{N} \qquad (1)$$

Where W_j notes weight assigned to the jth success factor.

w_k is the rating assigned to jth success factor in the range 0 to 4, signifying no effect, low, medium, high, and very high effect. x is the no. of responses for a w_k and N is the total number of responses. The weighted average values for each success factor using eq. (1) are shown in Table 1. The motive behind identifying the success factors and their influence on the success of the app on the Play Store is to assist in the further understanding of how these factors can impact in predicting the success using a prediction algorithm. It This also be noted that the existing research studies on success prediction methods for mobile apps do on incorporate the influence of the identified 14 success factors. The author believes that this research gap can further be investigated by implementing them in prediction methods and analyzing the results.

Table 1: Success Factors with Assigned Weights

Sr. No.	Success factor	Assigned weights
1.	Subjective Norms	1.69
2.	Perceived Compatibility	2.15
3.	Perceived Playfulness	1.95
4.	Satisfaction	2.07
5.	Perceived Usefulness	2.62
6.	Perceived Ease of Use	2.61
7.	Content Quality	2.85
8.	Reliability	2.65
9.	Security	3
10.	Design Quality	2.78
11.	Usefulness	2.70
12.	Performance	2.76
13.	Social Media Integration	2.10
14.	Regularly updated	2.51

Answering RQ3

1. Data Collection

In this research work, a mobile app dataset is collected from three freelancing mobile app developers. The data providers kept account of the limited variables/features of the mobile app. However, after approaching them for the required data (to consider 14 success factors while developing the app) as per our studies, they were able to keep and share the data. The dataset had 21 columns and 230 rows (200 after removing outliers). The variables such as App Name, Category, LOC, APK Size, Avg Rating, Installs, Version, and Free/Paid are already kept by the developers. The rest of the variables identified as success factors by the author are also accounted for as requested. Each of the success factors is given a score by the developer based on the guidelines in Table 2.

2. Data Cleaning

The dataset initially had 230 rows (mobile apps). The duplicate, missing, and inappropriate rows were removed and leaving 200 rows. The valid data is selected, and the garbage values are cleaned from the dataset collected from the developers. Relevant features or attributes required for success prediction are chosen.

3. Data Pre-processing and Data Selection

Data is prepared for classification to enhance classifier accuracy, and the Avg Rating column is chosen as the target variable for success prediction. The following are the preprocessing tasks performed on the data to prepare for classification models: -

- Casting of Installs to int, Avg. Rating and Price to float (and remove $ from Price column).
  Encode features (Type, Category) by label encoding method.
- Replaced the Size column with unified size values by converting Megabytes and Kilobytes into Bytes, then normalizing App size values.
- Remove unimportant features like (App Name, Version, and Last update).

Table 2: Guidelines for Developers for Scoring the App Success Factors

Sr. No.	Variable	Description	Categories	Score
1.	App Name	Name of the App		None
2.	Category	Category of the app		None
3.	LOC	Line of Code for each app	Numeric	None
4.	APK Size	Size of an App in terms of MB	Numeric	None
5.	Avg Rating	App Avg rating on the Play Store	Numeric	None
6.	Installs	No. of downloads	Numeric	None
7.	Version	Android Version supported by the app	Numeric	None
8.	Free/Paid	If the app is free on the Play Store or not	Yes	1
			No	0
9.	Subjective Norms	Was the app sponsored or recommended	Yes	1
			No	0
10.	Perceived compatibility	Installing the new version of an app on an older version	Yes	1
			No	0
11.	Perceived Playfulness	Does the app pertain to the users stemming from their focus, inquisitiveness, and delight?	Yes	1
			No	0
12.	Satisfaction	Is the user willing, after being fully satisfied, to continuously use the app in the future	Yes	1
			No	0
13.	Perceived Usefulness	Does the app improve users' job performance	Yes	1
			No	0
14.	Perceived ease of use	Whether the usage of the app is easy or not	No	0
			Yes	1
15.	Content Quality	If the user is satisfied with the Content quality of the app or not	Yes	1
			No	0
16.	Reliability	If the app completes the Stress test or not	Success	1
			Fail	0
17.	Security	Does the app provide a shield against probable harm to personal data or any malicious action?	Yes	1
			No	0
18.	Design Quality	Whether the user is satisfied with the design quality of the app or not	Satisfied	1
			Not Satisfied	0
19.	Usefulness	Professed usefulness, functionality, and services	Yes	1
			No	0
20.	Performance	The performance of an app depends on the RAM and CPU utilization of the App	0-20 Mb Ram & < 20 % CPU	1
			0-20 Mb Ram & 20%-50% CPU	2
			0-20 Mb Ram & > 50% CPU	3
			>20 Mb Ram & < 20 % CPU	4
			>20 Mb Ram & 20 % - 50% CPU	5
			>20 Mb Ram & > 50 % CPU	6

(Table 2 Contd ....)

(.... Contd Table 2)

21.	Social Media Integration	The app has options with Facebook, Instagram, or another social network	Yes	1
			No	0
22.	Regularly Updated	Whether the app is regularly updated or not	Yes	1
			No	0

The average rating column is considered as targets/labels for classifier Models. As App success is based on the average rating, there is a need to categorize the target as follows

- The Avg rating below 2.0 is considered a failed application under category (0).
- The Avg rating between 2.0 and 3.0 is considered a limited success application under category (1).
- The Avg rating above 4.0 is considered a successful application under category (2).

There were 22 features in the dataset. Only 18 were selected after understanding the meaning and suitability for statistical analysis and modeling. The final selected features are LOC, APK Size, Avg Rating, Installs, Version, Free/Paid, Subjective Norms, Perceived compatibility, Perceived Playfulness, Satisfaction, Perceived Usefulness, Perceived ease of use, Content Quality, Reliability, Security, Design Quality, Usefulness, Performance, Social Media Integration, Regularly Updated.

4. Machine Learning Models

Five Machine Learning models were selected as our problem is a classification problem (success or failure). The selected ones are as follows:

1. **Logistic Regression Classifier** [37], LRC is used for binary classification where the sigmoid function is used, which takes input as independent variables and produces a probability value between 0 and 1. Logistic regression is a technique that is limited to classification issues. Because it uses the logistic function (Sigmoid) as the activation function to separate the data into distinct classes, it can only be taken into consideration for classification problems, although it shares characteristics with linear regression. It is a popular technique in research on classification [38], [39].

2. **Naive Bayes Classifier** [40], NBC is based on Bayes' Theorem, which describes the probability of a certain event, based on prior knowledge of conditions that might be related to this event. Using Bayes' Theorem to classify data according to the probability of various classes given the data's attributes is the fundamental concept of the Naive Bayes classifier [41]. The majority of its applications are in high-dimensional text classification. Naive Bayes classifiers use the Bayes theorem to categorize events into classes. They belong to the category of extremely effective classification algorithms due to their accuracy and simplicity. Naive Bayes classifiers classify unknown events by calculating the likelihood of each class occurring and then choosing the class with the highest probability. They are regarded as very efficient supervised classifiers because of their minimal calculation time and high degree of accuracy[42].

3. **Decision Tree classifier** [43], DTC's purpose is to generate a model that predicts a target based on input features using Tree models. In these models, leaves represent class labels, and feature conjunctions that lead to those class labels are represented by branches. Data instances and attributes are represented hierarchically in a Decision Tree (DT). Regression and classification tasks are both handled by decision trees (DTs). Decisions are made in decision nodes, and a tree shape is constructed with a root node as the starting node. A DTC's primary benefit is the minimal amount of computing time needed after it is created. However, the primary disadvantage of the DTC is the determination of the sequences of these nodes, which can be changed to create multiple DTCs by altering the root or decision nodes [44], [45] .

4. **Support Vector Machine classifier**[37], SVM is considered a robust classifier that aims to find an optimal hyperplane that divides feature space by finding how far the margin is from the hyperplane, and the traits that are closest across all classes. Following the classification of the data and the mapping of the data into hyperspace using a kernel function, support vectors—the data points of each class that are nearest to one another—are assigned in SVM. Consequently, the classification uses a subset produced by the input data's support vectors to reduce the separation between the input data points and the hyperplane [46].

5. **Gradient Boosting Classifier** [47], Another type of tree-based ensemble machine learning algorithm after

the Decision Tree classifier is gradient boosting. GBC is a powerful boosting algorithm that combines several weak learners into strong learners, in which each new model is trained to minimize the loss function. By taking into account the loss experienced by the weak learners, who are also Decision Trees, it maximizes the outputs. A freshly built or altered Decision Tree is added to use a gradient descent approach to minimize the overall loss. The gradient boosting approach works well for tasks involving classification and regression [48], [49].

5. **Prediction:** The ML models are used to fit the model and predict the values after splitting the dataset into training (75%) and testing (25%). The time (in seconds) is also recorded for each classifier to compare the performance metric.

6. **Performance Evaluation:** For performance evaluation, the F1-score accuracy data instances and attributes are represented hierarchically in a DT. Regression and classification tasks are both handled by decision trees (DTs). Decisions are made in decision nodes, and a tree shape is constructed with a root node as the starting node. A DT's primary benefit is the minimal amount of computing time needed after it is created. However, the primary disadvantage of the DT is the determination of the sequences of these nodes, which can be changed to create multiple DTs by altering the root or decision nodes [30], [50], [51] measure is used, which ranges from 0.1 to 1.0. For the reliability of the models, K-fold cross-validation [50] is used.

The results in Table 3 for F1-scores show that the best prediction accuracy is for Logistic Regression, Gaussian Naïve Bayes, and the SVC ML classifier with 96% accuracy. Performance-wise wise Gradient Boosting succeeded with 0.0082 seconds, making it the fastest training algorithm among the other four algorithms.

Table 3: Accuracy Measure of the ML Classifiers using F1-Score

Machine Learning Classifiers	F1-Scores		Time Elapsed	
	Training	Testing	Training	Testing
Logistic Regression	0.9067	**0.9600**	0.0291	0.0122
Gaussian Naïve Bayes	0.9067	**0.9600**	0.0175	0.0130
Decision Tree	1.000	0.800	0.0172	0.0093
SVC	1.000	**0.9600**	1.0140	0.0224
Gradient Boosting	1.000	0.9200	0.2539	**0.0082**

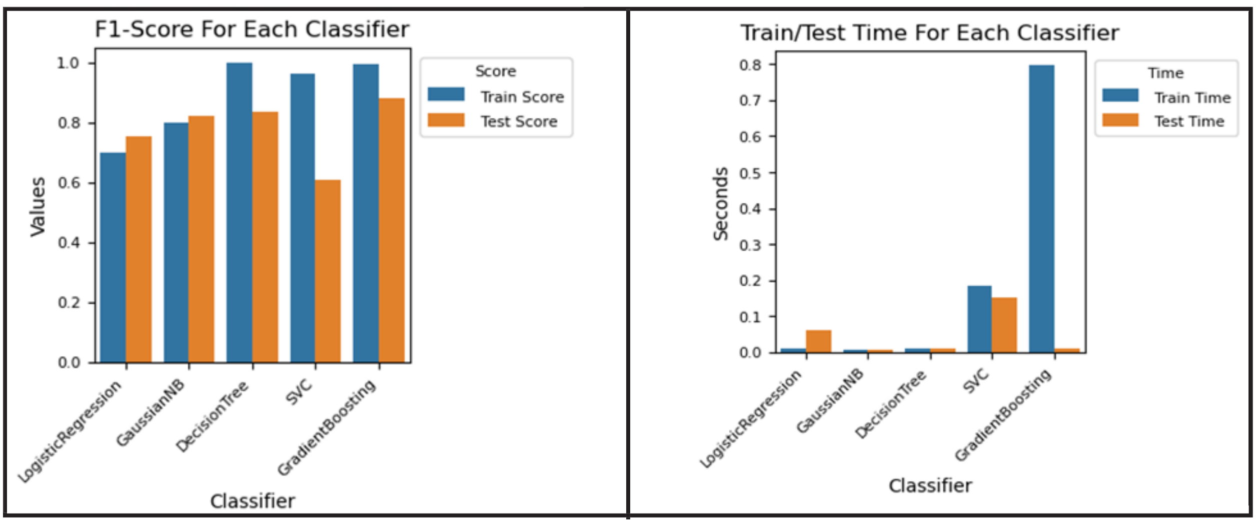

Fig. 2: Comparison of Machine Learning Classifiers

The cross-validation technique is used to test the classifier's reliability in this study. Cross-validation divides the training data into k parts (k=10 in this implementation), then uses 9 parts as a training dataset and tests with the remaining part (validation dataset). This operation is repeated 10 times, then the final result is the average score of 10 trials. The results of cross-validation are presented in Table 4. Logistic Regression and SVC, again, prove that they are the most reliable classifiers with mean scores of 0.91.

Table 4: ML Classifiers Reliability Measure using Cross-Validation

Machine Learning Classifiers	Training	Testing
Logistic Regression	0.910	**0.910**
Gaussian Naïve Bayes	0.880	0.90
Decision Tree	1.00	0.850
SVC	1.000	**0.910**
Gradient Boosting	1.000	0.90

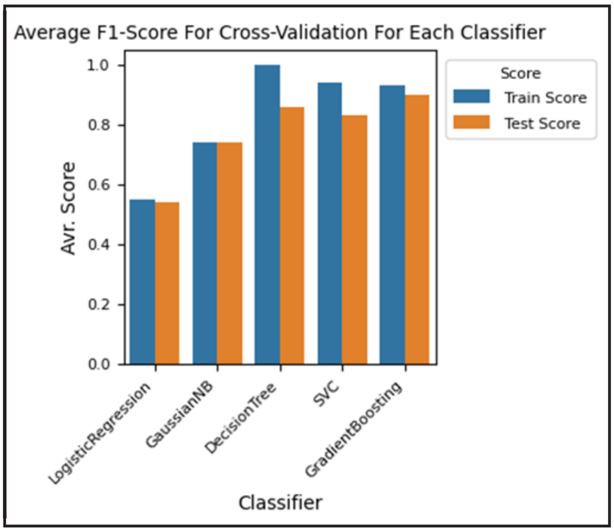

Fig. 3: Cross-Validation Score for ML Classifier

CONCLUSION AND FUTURE WORK

This study report presents the current state-of-the-art for success prediction of mobile app software before its launch on the Play Store. 14 critical success factors of mobile apps that play a very important role in their success on the Play Store are identified in the literature and are further investigated through an online survey of various mobile app users. The research gap indicates that critical success factors are currently not considered by existing approaches to predict success. There is no formal model that solely reflects these factors. This study successfully implemented machine learning models (using Python Jupyter Notebook 7.0.8) to predict the 'Average rating' of an app as a success predictor. A developer can now first predict whether an app will be successful, limited successful, or failed, with a given set of features, after being uploaded to the Google Play store. The accuracy of the final selected model is 96%, which is the highest accuracy among many existing models in the literature. In the future, to further enhance the model accuracy, one can try to tune classifier parameters instead of using default values or try a Neural Network for classification.

ACKNOWLEDGMENTS

The author is highly thankful to Khalsa College, Amritsar for offering the seed money to direct this research.

DISCLOSURE OF INTERESTS

The author has no competing interests to declare that are relevant to the content of this article.

REFERENCES

I. C. Fang and S. C. Fang, "Factors affecting consumer stickiness to continue using mobile applications," *Int. J. Mob. Commun.*, 2016, doi: 10.1504/IJMC.2016.078720.

V. Venkatesh and F. D. Davis, "Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies," *Manage. Sci.*, 2000, doi: 10.1287/mnsc. 46.2.186.11926.

C. Su-Chao and T. Feng-Cheng, "An empirical investigation of students' behavioural intentions to use the online learning course Web sites," *Br. J. Educ. Technol.*, 2008.

T. P. Liang, Y. L. Ling, Y. H. Yeh, and B. Lin, "Contextual factors and continuance intention of mobile services," *Int. J. Mob. Commun.*, 2013, doi: 10.1504/IJMC.2013.055746.

O. Al-Shamaileh and A. Sutcliffe, "Investigating a multi-faceted view of user experience," in *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, 2012. doi: 10.1145/2414536.2414538.

O. Al-Shamaileh, "I Have Issues with Facebook: But I Will Keep Using It," *IEEE Technol. Soc. Mag.*, 2018, doi: 10.1109/MTS.2018.2826078.

I. L. Wu and J. L. Chen, "An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study," *Int. J. Hum. Comput. Stud.*, 2005, doi: 10.1016/j.ijhcs.2005.03.003.

T. Zhou, "Understanding continuance usage of mobile services," *Int. J. Mob. Commun.*, 2013, doi: 10.1504/IJMC.2013.050995.

P. A. Paramartha, H. B. Santoso, and P. H. Putra, "FACTORS INFLUENCING THE USER STICKINESS OF A MOBILE NEWS APPLICATION: THE CASE OF 'LINE TODAY APP,'" *JST (Jurnal Sains dan Teknol.*, 2021, doi: 10.23887/jstundiksha.v10i2.36345.

T. Ahn, S. Ryu, and I. Han, "The impact of Web quality and playfulness on user acceptance of online retailing," *Inf. Manag.*, 2007, doi: 10.1016/j.im.2006.12.008.

D. Sledgianowski and S. Kulviwat, "Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context," *J. Comput. Inf. Syst.*, 2009, doi: 10.1080/08874417.2009.11645342.

Y. Zhang, Y. Fang, K. K. Wei, E. Ramsey, P. McCole, and H. Chen, "Repurchase intention in B2C e-commerce - A relationship quality perspective," *Inf. Manag.*, 2011, doi: 10.1016/j.im.2011.05.003.

T. H. Hsu and J. W. Tang, "Development of hierarchical structure and analytical model of key factors for mobile app stickiness," *J. Innov. Knowl.*, 2020, doi: 10.1016/j.jik.2019.01.006.

A. Bhattacherjee, "UNDERSTANDING INFORMATION SYSTEMS CONTINUANCE: AN EXPECTATION- CONFIRMATION MODEL Motivation for the Study," *Inf. Syst. Contin. MIS Q.*, 2001.

C. C. Tu, K. Fang, and C. Y. Lin, "Perceived ease of use, trust, and satisfaction as determinants of loyalty in e-auction marketplace," *J. Comput.*, 2012, doi: 10.4304/jcp.7.3.645-652.

H. Van der Heijden, "Factors influencing the usage of websites: The case of a generic portal in The Netherlands," *Inf. Manag.*, 2003, doi: 10.1016/S0378-7206(02)00079-4.

F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q. Manag. Inf. Syst.*, 1989, doi: 10.2307/249008.

L. Guerrouj and O. Baysal, "Investigating the android apps' success: An empirical study," in *IEEE International Conference on Program Comprehension*, 2016. doi: 10.1109/ICPC.2016.7503724.

M. Linares-Vásquez, G. Bavota, C. Bernal-Cárdenas, M. Di Penta, R. Oliveto, and D. Poshyvanyk, "API change and fault proneness: A threat to the success of android apps," in *2013 9th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering, ESEC/FSE 2013 - Proceedings*, 2013. doi: 10.1145/2491411.2491428.

J. Y. L. Thong, S. J. Hong, and K. Y. Tam, "The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance," *Int. J. Hum. Comput. Stud.*, 2006, doi: 10.1016/j.ijhcs.2006.05.001.

G. C. Bruner and A. Kumar, "Explaining consumer acceptance of handheld Internet devices," *J. Bus. Res.*, 2005, doi: 10.1016/j.jbusres.2003.08.002.

O. Al-Shamaileh and A. Sutcliffe, "Why people choose Apps: An evaluation of the ecology and user experience of mobile applications," *Int. J. Hum. Comput. Stud.*, 2023, doi: 10.1016/j.ijhcs.2022.102965.

C. I. Wei, P.S., Lee, S.Y., Lu, H.P., Tzou, J.C., Weng, "Why do people abandon mobile social games? Using Candy Crush Saga as an example.," *Int. J. Soc. Educ. Econ. Manag. Eng.*, vol. 9, no. 1, pp. 13–18., 2015.

J. Hartmann, A. Sutcliffe, and A. De Angeli, "Towards a theory of user judgment of aesthetics and user interface quality," *ACM Trans. Comput. Interact.*, 2008, doi: 10.1145/1460355.1460357.

M. J. Noh and K. T. Lee, "An analysis of the relationship between quality and user acceptance in smartphone apps," *Inf. Syst. E-bus. Manag.*, 2016, doi: 10.1007/s10257-015-0283-6.

R. Adhikari, D. Richards, and K. Scott, "Security and privacy issues related to the use of mobile health apps," in *Proceedings of the 25th Australasian Conference on Information Systems, ACIS 2014*, 2014.

B. H. Jones and A. G. Chin, "On the efficacy of smartphone security: A critical analysis of modifications in business students' practices over time," *Int. J. Inf. Manage.*, 2015, doi: 10.1016/j.ijinfomgt.2015.06.003.

M. A. Harris, S. Furnell, and K. Patten, "Comparing the Mobile Device Security Behavior of College Students and Information Technology Professionals," *J. Inf. Priv. Secur.*, 2014, doi: 10.1080/15536548.2014.974429.

Y. Ding and K. H. Chai, "Emotions and continued usage of mobile applications," *Ind. Manag. Data Syst.*, 2015, doi: 10.1108/IMDS-11-2014-0338.

D. W. Gefen, D., Karahanna, E., Straub, "Trust and TAM in online shopping: an integrated model.," *MIS Q.*, vol. 27, no. 1, pp. 51–90, 2003.

V. Venkatesh, J. Y. L. Thong, and X. Xu, "Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology," *MIS Q. Manag. Inf. Syst.*, 2012, doi: 10.2307/41410412.

A. Malik, S. Suresh, and S. Sharma, "Factors influencing consumers' attitude towards adoption and continuous use of mobile applications: A conceptual model," in *Procedia Computer Science*, 2017. doi: 10.1016/j.procs.2017.11.348.

M. Kapias, "10-key-characteristics-of-a-successful-mobile-app-for-every-business." [Online]. Available: https://fireart.studio/blog/10-key-characteristics-of-a-successful-mobile-app-for-every-business/

S. Amjad *et al.*, "Calculating Completeness of Agile Scope in Scaled Agile Development," *IEEE Access*, vol. 6, pp. 5822–5847, 2018, doi: 10.1109/ACCESS.2017.2765351.

A. Kaur and K. Kaur, "A COSMIC function points based test effort estimation model for mobile applications," *J. King Saud Univ. - Comput. Inf. Sci.*, 2019, doi: 10.1016/j.jksuci.2019.03.001.

G. Grela, "Does Weighted average really work," in *In Active Citizenship by Knowledge Management & Innovation: Proceedings of the Management, Knowledge and Learning International Conference*, ToKnowPress, 2013, pp. 1219–1227.

J. T. Hancock, R. A. Bauder, H. Wang, and T. M. Khoshgoftaar, "Explainable machine learning models for Medicare fraud detection," *J. Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00821-5.

S. Chalichalamala, N. Govindan, and R. Kasarapu, “Logistic Regression Ensemble Classifier for Intrusion Detection System in Internet of Things,” *Sensors*, vol. 23, no. 23, 2023, doi: 10.3390/s23239583.

A. A. Muideen, C. K. M. Lee, J. Chan, B. Pang, and H. Alaka, “Broad Embedded Logistic Regression Classifier for Prediction of Air Pressure Systems Failure,” *Mathematics*, vol. 11, no. 4, 2023, doi: 10.3390/math11041014.

W. M. Shaban, A. H. Rabie, A. I. Saleh, and M. A. Abo-Elsoud, “Accurate detection of COVID-19 patients based on distance biased Naïve Bayes (DBNB) classification strategy,” *Pattern Recognit.*, vol. 119, 2021, doi: 10.1016/j.patcog.2021.108110.

N. Memon, S. B. Patel, and D. P. Patel, “A novel approach of polsarimage classification using Naïve Bayes classifier,” in *MathematicalModeling, Computational Intelligence Techniques and RenewableEnergy: Proceedings of the First International Conference,MMCITRE 2020*, 2021, pp. 93–104.

I. Wickramasinghe and H. Kalutarage, “Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation,” *Soft Comput.*, vol. 25, no. 3, 2021, doi: 10.1007/s00500-020-05297-6.

A. Singh, “Mobile App Success Prediction,” *Int. J. Res. Appl. Sci. Eng. Technol.*, 2020, doi: 10.22214/ijraset.2020.6273.

D. T. Pham, H. N. Do, and H. T. Phan, “Rule-enriched Decision Tree Classifier for Conditional Sentence Sentiment Analysis,” *J. Tech. Educ. Sci.*, vol. 19, no. 1, 2024, doi: 10.54644/jte.2024.1530.

J. V. Shukla and A. Deshpande, “A decision tree classifier approach for predicting customer’s inclination toward use of online food delivery services,” *Multidiscip. Sci. J.*, vol. 6, no. 5, 2024, doi: 10.31893/multiscience.2024072.

T. Adugna, W. Xu, and J. Fan, “Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images,” *Remote Sens.*, vol. 14, no. 3, 2022, doi: 10.3390/rs14030574.

J. T. Hancock and T. M. Khoshgoftaar, “Gradient Boosted Decision Tree Algorithms for Medicare Fraud Detection,” *SN Comput. Sci.*, vol. 2, no. 4, 2021, doi: 10.1007/s42979-021-00655-z.

Abdullah-All-Tanvir, I. Ali Khandokar, A. K. M. Muzahidul Islam, S. Islam, and S. Shatabda, “A gradient boosting classifier for purchase intention prediction of online shoppers,” *Heliyon*, vol. 9, no. 4, 2023, doi: 10.1016/j.heliyon.2023.e15163.

A. Shakya, M. Biswas, and M. Pal, “Fusion and Classification of SAR and Optical Data Using Multi-Image Color Components with Differential Gradients,” *Remote Sens.*, vol. 15, no. 1, 2023, doi: 10.3390/rs15010274.

F. Sohil, M. U. Sohali, and J. Shabbir, “An introduction to statistical learning with applications in R,” *Stat. Theory Relat. Fields*, vol. 6, no. 1, 2022, doi: 10.1080/24754269.2021.1980261.

L. Bognár and T. Fauszt, “Factors and conditions that affect the goodness of machine learning models for predicting the success of learning,” *Comput. Educ. Artif. Intell.*, vol. 3, Jan. 2022, doi: 10.1016/j.caeai.2022.100100.