*I. J. Education and Management Engineering,* **2024, 3, 13-21**
Published Online on June 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijeme.2024.03.02

# Data Analysis and Success Prediction of Mobile Games before Publishing on Google Play Store

**Muhammad Muhtasim***
Department of Computer Science and Engineering, Bangladesh Army University of Engineering & Technology, Natore-6431, Bangladesh
E-mail: muhtasim222@gmail.com
ORCID iD: https://orcid.org/0000-0001-5741-603X
*Corresponding Author

**Md. Showrov Hossen**
Department of Computer Science and Engineering, City University, Dhaka, Bangladesh
Email: sourovhossen96@gmail.com
ORCID iD: https://orcid.org/0009-0008-8758-6273

Received: 17 July, 2023; Revised: 15 August, 2023; Accepted: 11 September, 2023; Published: 08 June, 2024

**Abstract:** The popularity of mobile games has expanded among individuals of all ages, and the mobile gaming businesses are quickly expanding day by day. The Google Play Store, one of the most well-known platforms for the distribution of Android applications and games, sees a daily influx of thousands of new mobile games. One of the biggest problems in the gaming industry is predicting a mobile game's performance. Every day, thousands of new games are released. But just a couple of them are successful, while most of them fail. The study was done with the intention of analyzing any relationship between a mobile game's success and its distinctive features. Many of the mobile game developers work independently or work in the mobile game industries to make their games successful on the digital market. Before they are released, game makers can increase the quality of their games if they are confident in their products' commercial viability. For that reason, more than 17,000 games were taken into consideration. We show that the success of a mobile game is clearly influenced by its category, number of supported languages, developer profile, and release month. Furthermore, we show that specific aesthetic features of game symbols are more frequently linked to higher rating counts. We analyzed Google Play Store mobile games data and used a variety of machine learning algorithms for predicting the performance of mobile games based on the total number of downloads and the total user rating.

**Index Terms:** Mobile games, Google Play Store, Game industries, Digital market, Downloads, User rating, Machine learning

## 1. Introduction

Today's gaming business is a large one that is steadily expanding. By 2020, the US video game market was expected to be worth more than billion, according to statistics, demonstrating the importance and scale of this industry. Presently, there are more than 300 thousand games available only on the Apple App Store. Both startups and independent developers were able to readily access this growing industry because to free distribution and the cheap cost of mobile application development. Moreover, mobile phones are now more accessible and robust due to the latest developments in mobile technology. As a result, mobile devices' memory capacity, processor speed, and graphics quality all grew, supporting more challenging and fun games.

In the game development industry success prediction of a mobile game is a major issue. Every day the game industries release thousands of mobile games where a few of them succeed and the majority fail. Google Play Store has thousands of different mobile games, and several thousands of new mobile games accumulate to its list every day [1]. In the huge competitive digital market, it is too difficult to understand for a developer if he is walking in the right direction or wrong direction [2]. In this research we address these challenges through machine learning model for data mining to prove a successful technique for the prediction of android mobile games application success. The success of a mobile game can be determined by the total number of downloads and user rating. In the digital market, there are two types of

 

games: free and paid version and there is also various category of games [3, 4].

We use data mining methods on an existing dataset of mobile games to predict this novel success indication. We use data mining to identify the most crucial characteristics affecting the value of this measure, and we offer methods for predicting it based on these variables. To discover the underlying facts and principles that determine the phenomena under study, a sequence of stages and procedures must be carefully followed. Pre-processing is typically essential to remove irrelevant data (data cleaning), change the data's format, or do some data regularization and normalization in preparation for the next steps.

In this work we show which category games are most release. We used statistical data analysis to show the top 10 popular game categories and show the most game release industry with bar chart. We proposed a machine learning technique to predict the mobile games success or failure. We used four famous machine learning algorithms: Random Forest, Naïve Bayes, K Nearest Neighbors and Support Vector Machine. With the help of these algorithms, we can predict the download number and user rating of a mobile game. At first the game developer or game industry set up a target value of total number of game downloads and user rating and then compare it to the predicted download number and user rating. If the predicted downloads number and user rating are higher or equal to the developer targeted download number and install, then our proposed machine learning model take a decision that it will success or failure.

In this work, developer can understand the condition of the game before publishing the mobile game on Google Play Store and then developer can take the necessary step. So, this model will be helpful for the game developer and game industries students. The main contribution of this paper is machine learning models to understand and predict mobile game success.

## 2. Literature Review

There are a few studies in literature that raise comparable issues to the ones posed in this study. The success definition and the analyzed qualities, however, are what set them apart. In [5], a unique success score measure that represents various objectives was explored for over 17thousand games. The future success of a game is found to be positively impacted by several game qualities, including the number of IAPs (In-App Purchases), the puzzle genre, supporting several languages, and being built by an experienced developer. employed Deep Learning (DL), Random Forest (RF), and Support Vector Machine (SVM) specifically to forecast this success score measure.

In [6] concluded that they propose the use of data mining techniques in this article to determine if the newly hosted application will be a success or not. The methodologies results can be utilized by a variety of venture capitalists and potential investors to make an educated judgment. In analyzed the income of about several characteristics for iPhone games in parallel to that work. The study's objective was to list the top 10 elements that influence game creation. They confirmed a significant correlation between a mobile game's income and the overall number of users that run or engage with it. We investigated numerous game qualities and their impact on average rating and rating count, much like these two publications. To boost the statistical significance of the results, we considered 20,000 mobile game applications as opposed to a small number of mobile games.

In [7] investigated the icon design and its relationships to mobile game downloads and purchases for 68 game applications. Realism and originality have been shown to increase downloads, clicks, and purchases. We don't take the icon design into account when working. But a thorough examination of the aesthetic aspects of the most popular and least popular games was done. Then, inferences were reached based on traits that each cluster shared.

In parallel to that work [8] with respect to they examine the notion of gaming sales statistics and sales projections briefly in this study. The different machine learning approaches and metrics that were utilized to forecast sales are displayed.

In [9] demonstrate that promoting a mobile game on TV and putting it up on several App Stores improves its likelihood of success. At the same time, they discovered that the game's effectiveness declines over time. The advertising criteria are not considered in the scope of our activity. But it may be done in a later piece. The key question is whether our results hold true for games that have been published in various App Stores.

In [10] a database of PC games is constructed, and the data is visually presented. Finally, studies using machine learning approaches are carried out to determine how effectively post-release success of PC games may be predicted before their release.

In paper [11] the authors predict the success of android apps before uploading on Google Play Store based on user rating and installs using machine learning algorithms.

Color, motion, or textual features can affect how appealing people find public screens, as demonstrated in [12]. Icon design and mobile gaming appear to be connected. Considering that the symbol itself has a significant impact on the player's first perception of the game. We investigate the color information of mobile icons in a manner like this work. However, from a more specific angle.

In [13] because desktop programs differ from mobile applications in terms of functionalities, they are not suitable to mobile apps. So far, there hasn't been a repository or even a mechanism for predicting the likelihood of an Android application's success. They utilize the library to train a neural network and other classification techniques to predict their chances of success. The greatest accuracy we attained in the experiments was 99.99 percent, compared to 96 percent in prior work on desktop platforms [14].

In [15] video games are one of our society's most popular forms of entertainment. The increasing demand from customers has resulted in high manufacturing costs, putting development businesses at greater risk. A machine learning technique to predict video game success before its first release is proposed in research [16]. The implementation involves use of Word2vec based techniques for Natural Language Processing (NLP).

Game components and ARM strategies are discussed and studied in [17] as important factors affecting the success of mobile apps. By conducting data-driven research on the mobile gaming application market, mobile gaming application features, user acquisition and retention trends, and monetization strategies, this study employs the CRISP-DM model for data mining to prove a credible strategy for forecasting the success of mobile game apps. The best game predictions were singled out and contrasted with the most typical user behaviors for further study and implications.

In our paper, based on the total number of downloads and total user rating, we evaluated Google Play Store mobile game data and utilized several machine learning algorithms to forecast the success of mobile games. This work will be helpful for those who want to upload the mobile game application to the play store. Before uploading they predict their success or failure value. Then decided, they will be uploaded or not in play store.

## 3. Methodology

In we collect our dataset from online sources (Kaggle). We collect mobile games dataset for our research. Datasets contain the garbage value which should be handled before training the model otherwise it affects the performance of the model to predict the output. Before data analysis and result finding data preprocessing is necessary. More than 20,000 mobile games from the Apple App Store are included in the dataset that was utilized. Each game has 21 unique qualities when it first starts. There are very few missing values among these. Given that it was compiled in January of 2023, the dataset is quite recent. Being a decent proxy for the real gaming market as a result. The primary attributes of the dataset are displayed with their categories.

### *3.1. Data Cleaning*

The raw data from online source contain garbage value, null value and unnecessary data that not needed for model fitting. First, we are cleaning the unnecessary, null and garbage value of the dataset.

### *3.2. Data Conversion*

Some of the data needs to be converted into an appropriate format. We convert data into similar formats to improve the efficiency of the machine learning model. We converted the categorical data into numeric values to fit the model. Dealing with missing data Games with null User Rating Count values have ratings of four or fewer. They receive five rating counts total. Simultaneously, games with average User Rating null values indicate that they did not more than 5 rating counts are gained. As a result, we allocate their general rating is 0.

### *3.3. Feature Selection*

The game name, platform, year of release, category, publisher, total ratings, total downloads, average user ratings, growth, price, worldwide sales, and developer are among the features included in the raw dataset. Among these features we select game name, category, publisher, total downloads, average user rating and price for data analysis and training model. Aspect Engineering displays the properties of the final dataset. A few more features were generated from the raw dataset to increase the amount of data related to each individual game. By tallying the number of languages that each game supports, the number of languages is determined. By tallying the number of genres that the game is categorized under, the number of genres is calculated accordingly. Then, by categorizing all developers included in the dataset into two groups—Newbie and Professional, the game's developer maturity (Developer Category) is determined.

### *3.4. Machine Learning Methods*

#### *3.4.1. Random Forest*

A technique for ensemble learning that may be used to solve a variety of classification and regression issues [8, 9] in fields including bioinformatics, remote sensing, and finance. Generally speaking, random forest is used to improve learning by merging several decision trees. To lower the total variance, these decision trees are trained on several partitions of the same training set. enhancing the trained model's generalizability in the process. They used two random forest models, which is like our research. the first to foresee if a consumer would buy a game. The other tries to predict how many purchases a certain user will make in the future.

#### *3.4.2. Na ïve Bayes*

It is an algorithm that discovers each object's probability, its characteristics, and the groups to which they belong. An alternative name for it is a probabilistic classifier. The Naive Bayes Algorithm, which is categorized as supervised learning, is mostly employed to address classification issues.

#### *3.4.3. K-Nearest Neighbors*

The k-most comparable records from the training dataset to a new record are then found when a prediction is

needed. A condensed forecast is created based on these neighbors. There are several techniques to measure how similar two recordings are. One might employ a problem- or data-specific approach. The Euclidean distance is typically a decent place to start when working with tabular data. Once the neighbors have been identified, the summary forecast may be created by either taking the average or returning the most frequent result. KNN can therefore be applied to classification or regression issues.

*3.4.4. Support Vector Machine*

The k-most comparable records from the training dataset to a new record are then found when a prediction is needed. A condensed forecast is created based on these neighbors. There are several techniques to measure how similar two recordings are. One might employ a problem- or data-specific approach. The Euclidean distance is typically a decent place to start when working with tabular data. Once the neighbors have been identified, the summary forecast may be created by either taking the average or returning the most frequent result. KNN can therefore be applied to classification or regression.

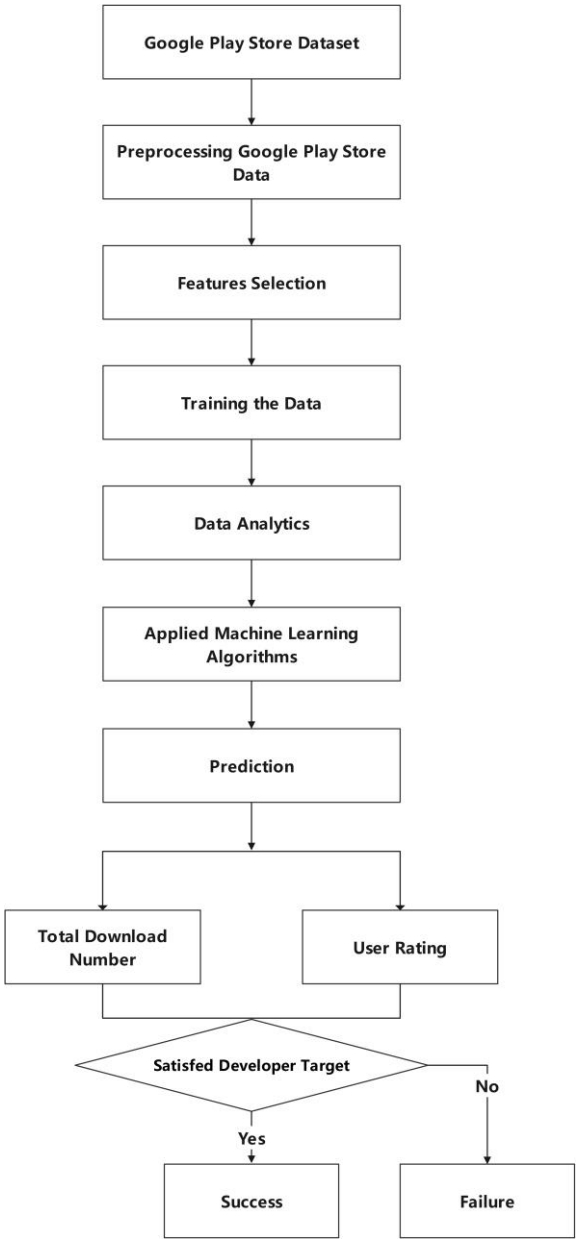


Fig. 1. System methodology flowchart

## 4. Data Analysis

*4.1. Categorical Games Total release*

The Game Category feature of the dataset has a total of 12 types of categorical games. Figure 2 shows the bar

chart of the number of categorical games total release. In this bar chart we show that the action type of games and the sports type of games are mostly release and the demand of these categories is more than the other categories. We also show the top 10 most popular games categories and their frequency in Table 1. We said these games category is most popular because this is mostly released on Google Play Store. It helps the games developer to know which types of mobile games are more demandable in the digital market.

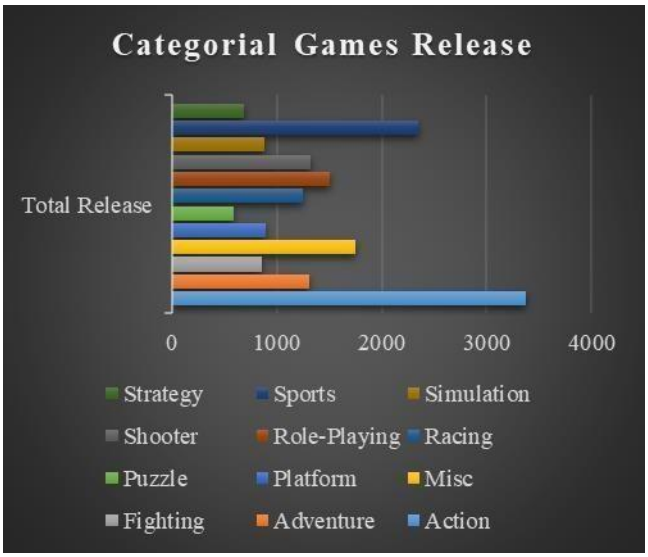


Fig. 2. Bar chart of categorical games vs total release

In the following table, we have provided a list of different video game categories along with their frequencies. This data appears to represent the number of games in each category or their popularity. Keep in mind that the mobile gaming industry is dynamic, and new trends may have emerged since my last update. Additionally, regional preferences can influence the popularity of specific categories in different parts of the world.

Table 1. Top 10 most popular mobile games category

| Place | Games Category | Frequency |
|---|---|---|
| 1 | Action | 3370 |
| 2 | Sports | 2348 |
| 3 | Role Playing | 1500 |
| 4 | Shooter | 1323 |
| 5 | Adventure | 1303 |
| 6 | Racing | 1249 |
| 7 | Simulation | 874 |
| 8 | Platform | 863 |
| 9 | Fighting | 849 |
| 10 | Puzzle | 580 |

### *4.2. Mobile Games Company Total Games release*

In tech today the gaming industry is one of the most innovative and important sectors. The game industry involves the development, monetization, and the marketing of mobile games. The gaming sector is excellently large. There are many games industry in the world. Figure 3 shows the Bar Chart of top gaming industries based on the total number of games released. The gaming sector is currently among the most cutting-edge and significant in the technological world. It covers a broad spectrum of tasks, including game creation, monetization, and marketing. It has seen enormous development and transition recently, partly because of technological improvements and shifting customer tastes. Let's examine Figure 3, a bar graph showing the top gaming sectors based on the total number of games launched. This bar chart provides insights into the diversity of gaming categories and their respective contributions to the overall gaming industry. The high number of games in the "Action" and "Sports" categories suggests their popularity and appeal to a broad audience, while categories like "Puzzle" and "Simulation" also have a significant presence in the market.The gaming industry's ability to cater to various tastes and preferences has contributed to its continued growth and innovation, making it a vital and dynamic sector within the technology and entertainment landscape.

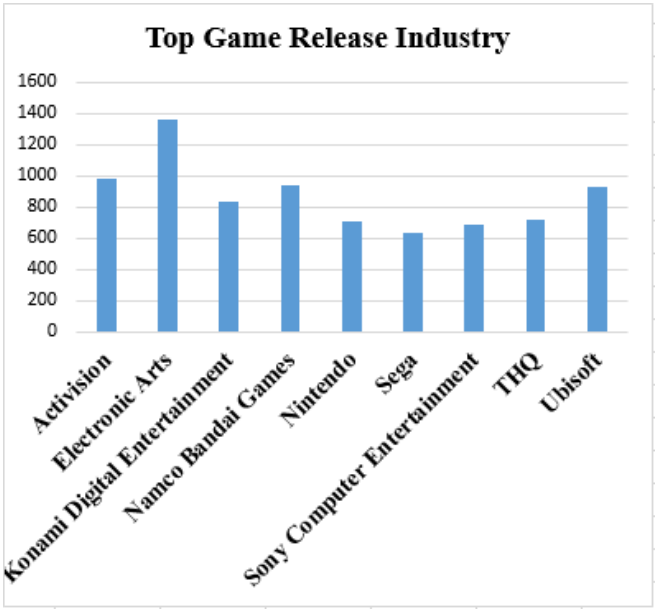


Fig. 3. Bar Chart of top gaming industry against total game release

### *4.3. Most Download Mobile Games*

In this section we show which mobile games are mostly downloaded from Google Play Store. Table 2 shows the top mobile games which are downloaded more than 500 million from Google Play.

Table 2. List of more than 500 million downloaded mobile games.

| | |
|---|---|
| 1.Garena Free Fire- World Series | 8.My Talking Tom |
| 2. PUBG MOBILE - Traverse | 9.My Talking Angela |
| 3.Sniper 3D: Fun Free Online FPS Shooting Game | 10.Hill Climb Racing |
| 4.Temple Run 2 | 11.8 Ball Pool |
| 5.Subway Surfers | 12.Clash of Clans |
| 6.Temple Run | 13.Candy Crush Saga |
| 7.Ludo King | 14.Fruit Ninja |

## 5. Result Analysis

We used some popular machine learning algorithms to predict the success of mobile games before launching on Google Play Store. We predict the number of downloads and user rating by using machine learning algorithms. Here the input features are game category, game publisher industry and game price (free or paid). If the predicted downloads number and user rating meet the target of mobile game developers or mobile game industries, we will then take a decision that the game is success or otherwise failure.

### *5.1. Random Forest*

A supervised learning machine learning algorithm is random forest. To discover a more reliable and precise forecast, it constructs numerous decision trees and blends them.

Table 3. Result prediction by Random Forest algorithm

| Game Name | Developer Target Value-User Rating | Developer Target Value-Total Downloads | Predicted Value-User Rating | Predicted Value-Total Downloads | Decision |
|---|---|---|---|---|---|
| Angry Birds | 4.5 | 100M | 4.20 | 90M | Failure |
| Ludo King | 4.2 | 500M | 4.55 | 550M | Success |
| Temple Run | 4.6 | 500M | 4.61 | 510M | Success |
| 8 Ball Pool | 4.45 | 200M | 4.34 | 130M | Failure |
| Candy Crush | 4.55 | 300M | 4.60 | 450M | Success |

### *5.2. Naïve Bayes*

The theorem is the foundation of the probabilistic machine learning method known as Naive Bayes, which is utilized in a wide range of categorization applications. It is a family of algorithms rather than a single method, and they are all united by the same fundamental idea—that each pair of characteristics being categorized is independent of the others.

Table 4. Result prediction by Naïve Bayes algorithm

| Game Name | Developer Target Value-User Rating | Developer Target Value-Total Downloads | Predicted Value-User Rating | Predicted Value-Total Downloads | Decision |
|---|---|---|---|---|---|
| Angry Birds | 4.5 | 100M | 4.52 | 115M | Success |
| Ludo King | 4.2 | 500M | 4.43 | 570M | Success |
| Temple Run | 4.6 | 500M | 4.51 | 430M | Failure |
| 8 Ball Pool | 4.45 | 200M | 4.54 | 250M | Success |
| Candy Crush | 4.55 | 300M | 4.63 | 430M | Success |

### *5.3. K-Nearest Neighbors*

Support A group of supervised learning techniques called vector machines are utilized in classification, regression, and outlier identification. ata is transformed using a method known as the kernel trick, and based on these changes, it determines the best border between the potential outputs.

Table 5. Result prediction by KNN algorithm

| Game Name | Developer Target Value-User Rating | Developer Target Value-Total Downloads | Predicted Value-User Rating | Predicted Value-Total Downloads | Decision |
|---|---|---|---|---|---|
| Angry Birds | 4.5 | 100M | 4.59 | 125M | Success |
| Ludo King | 4.2 | 500M | 4.05 | 220M | Failure |
| Temple Run | 4.6 | 500M | 4.32 | 425M | Failure |
| 8 Ball Pool | 4.45 | 200M | 4.49 | 290M | Success |
| Candy Crush | 4.55 | 300M | 4.30 | 180M | Failure |

### *5.4. Support Vector Machine*

A supervised machine learning approach called K- Nearest Neighbors can be utilized to resolve classification and regression issues. A similarity metric is used to categorize new instances and store all the existing cases.

Table 6. Result prediction by SVM algorithm

| Game Name | Developer Target Value-User Rating | Developer Target Value-Total Downloads | Predicted Value-User Rating | Predicted Value-Total Downloads | Decision |
|---|---|---|---|---|---|
| Angry Birds | 4.5 | 100M | 4.59 | 125M | Success |
| Ludo King | 4.2 | 500M | 4.14 | 380M | Failure |
| Temple Run | 4.6 | 500M | 4.70 | 600M | Success |
| 8 Ball Pool | 4.45 | 200M | 4.39 | 100M | Failure |
| Candy Crush | 4.55 | 300M | 4.63 | 500M | Success |

### *5.5. Accuracy Comparison of Different Machine Learning Algorithms*

We showed the comparison of different machine learning algorithms accuracy in Table 7. From Table 7, it is observed that Random Forest obtained the highest accuracy 98.82% than all the others machine learning algorithms for predicting the success of mobile games before publishing on Google Play Store.

Table 7. Comparison of the accuracy of different ML algorithms

| Game Name | Developer Target Value-User Rating | Developer Target Value-Total Downloads | Predicted Value-User Rating | Predicted Value-Total Downloads | Decision |
|---|---|---|---|---|---|
| Angry Birds | 4.5 | 100M | 4.59 | 125M | Success |
| Ludo King | 4.2 | 500M | 4.14 | 380M | Failure |
| Temple Run | 4.6 | 500M | 4.70 | 600M | Success |
| 8 Ball Pool | 4.45 | 200M | 4.39 | 100M | Failure |

## 6. Conclusion

Before submitting the mobile gaming applications to the Google Play Store, we applied several machine learning algorithms to assess the success or failure. Because these data are inherently dynamic, predicting the possibility of a new mobile game application's success is extremely challenging. The process for determining the possibility of success is lengthy and time-consuming. The information was gathered by examining many popular mobile gaming applications that were available across all platforms. The game developer can easily determine the success or failure by using this proposed technique. We demonstrate how certain game characteristics, such as the puzzle genre, multilingual support, and development by an experienced studio, have a significant, positive impact on a game's potential for future success. Furthermore, we demonstrate that the failure of the game appears to be strongly related to its July release and the absence of any. At the same time, using merely a game's outward qualities, we were able to accurately forecast its success. Therefore, it is crucial that game creators consider the advice provided by our statistical study to prevent failures caused by straightforward selection possibilities. In the future, we intend to explore other datasets and do comparable analyses to demonstrate the consistency of the findings. Data mining techniques may also be used to learn the prior data before applying comparison techniques. This would minimize effort and enhance efficiency even further.

## References


[1] S. Faralli, S. Rittinghaus, N. Samsami, D. Distante, and E. Rocha, "Emotional Intensity-based Success Prediction Model for Crowdfunded Campaigns," Inf Process Manag, vol. 58, no. 1, Jan. 2021, doi: 10.1016/j.ipm.2020.102394.

[2] M. R. Dehkordi, H. Seifzadeh, G. Beydoun, and M. H. Nadimi- Shahraki, "Success prediction of android applications in a novel repository using neural networks," Complex and Intelligent Systems, vol. 6, no. 3, pp. 573–590, Oct. 2020, doi: 10.1007/s40747-020-00154-3.

[3] R. M. Frey, R. Xu, and A. Ilic, "Mobile app adoption in different life stages: An empirical analysis," Pervasive Mob Comput, vol. 40, pp. 512–527, Sep. 2017, doi: 10.1016/j.pmcj.2017.01.006.

[4] A. Kerim and B. Genç "Mobile games success and failure: mining the hidden factors," Neural Comput Appl, 2022, doi: 10.1007/s00521-022-07154-z.

[5] R. Xu, R. M. Frey, E. Fleisch, and A. Ilic, "Understanding the impact of personality traits on mobile app adoption - Insights from a large-scale field study," Comput Human Behav, vol. 62, pp. 244–256, Sep. 2016, doi: 10.1016/j.chb.2016.04.011.

[6] M. Ben Hajhmida and O. Oueslati, "Predicting mobile application breakout using sentiment analysis of Facebook posts," J Inf Sci, vol. 47, no. 4, pp. 502–516, Aug. 2021, doi: 10.1177/0165551520917099.

[7] M. Kajko-Mattsson, A. Hauzenberger, and R. Fredriksson, EM 3 : Software Retirement Process Model.

[8] T. Dinh Le and B.-T. Ho Nguyen, "ATTITUDES TOWARD MOBILE ADVERTISING: A STUDY OF MOBILE WEB DISPLAY AND MOBILE APP DISPLAY ADVERTISING," 2014.

[9] K. L. Hsiao and C. C. Chen, "What drives in-app purchase intention for mobile games? An examination of perceived values and loyalty," Electron Commer Res Appl, vol. 16, pp. 18–29, Mar. 2016, doi: 10.1016/j.elerap.2016.01.001.

[10] Y. Su, "Data-driven Analysis Platform for Indie Mobile Game Publishing," 2022.

[11] Y. Su, P. Backlund, and H. Engström, "Comprehensive review and classification of game analytics," in Service Oriented Computing and Applications, Springer Science and Business Media Deutschland GmbH, Jun. 2021, pp. 141–156. doi: 10.1007/s11761-020-00303-z.

[12] M. R. Dehkordi, H. Seifzadeh, G. Beydoun, and M. H. Nadimi- Shahraki, "Success prediction of android applications in a novel repository using neural networks," Complex and Intelligent Systems, vol. 6, no. 3, pp. 573–590, Oct. 2020, doi: 10.1007/s40747-020-00154-3.

[13] B. Olabenjo, "Applying Naive Bayes Classification to Google Play Apps Categorization," Aug. 2016, [Online]. Available: http://arxiv.org/abs/1608.08574

[14] O. Rutz, A. Aravindakshan, and O. Rubel, "Measuring and forecasting mobile game app engagement," International Journal of Research in Marketing, vol. 36, no. 2, pp. 185–199, Jun. 2019, doi: 10.1016/j.ijresmar.2019.01.002.

[15] E. M. Ghazali, D. S. Mutum, and M. Y. Woon, "Multiple sequential mediation in an extended uses and gratifications model of augmented reality game Pokémon Go," Internet Research, vol. 29, no. 3, pp. 504–528, Jun. 2019, doi: 10.1108/IntR-12-2017-0505.

[16] S. C. Hung, H. Hsu, S. Y. Lien, and K. C. Chen, "Architecture harmonization between cloud radio access networks and fog networks," IEEE Access, vol. 3, pp. 3019–3034, Dec. 2015, doi: 10.1109/ACCESS.2015.2509638.

[17] V. A. Burégio, E. S. De Almeida, D. Ludrédio, and S. L. Meira, "A reuse repository system: from specification to deployment," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2008, pp. 88–99. doi: 10.1007/978-3-540- 68073-4_8.

## Authors' Profiles

**Muhammad Muhtasim** is currently working as a Lecturer in the Department of Computer Science and Engineering under the Faculty of Electrical and Computer Engineering at Bangladesh Army University of Engineering & Technology (BAUET). He has completed his bachelor's in computer science and engineering and master's in computer science and information Technology from Patuakhali Science and Technology University (PSTU)in 2020 and 2022, respectively. He is interested in research areas including Data Mining, Machine Learning, Data Analytics, Data Mining and Cyber Security.

**Md. Showrov Hossen** is currently working as a Lecturer in the Department of Computer Science and Engineering at City University. He has completed his bachelor's in computer science and engineering and master's in computer science and information Technology from Patuakhali Science and Technology University (PSTU)in 2020 and 2022, respectively. He is interested in research areas including Data Mining, Machine Learning, Data Analytics, Data Mining and Cyber Security.

**How to cite this paper:** Muhammad Muhtasim, Md. Showrov Hossen, "Data Analysis and Success Prediction of Mobile Games before Publishing on Google Play Store ", International Journal of Education and Management Engineering (IJEME), Vol.14, No.3, pp. 13-21, 2024. DOI:10.5815/ijeme.2024.03.02