

# Mobile Games Success and Failure: Mining the Hidden Factors

Abdulrahman Kerim  
Department of Computer Engineering  
Hacettepe University  
Ankara, Turkey  
ahaje@hacettepe.edu.tr

Burkay Genç  
Department of Computer Engineering  
Hacettepe University  
Ankara, Turkey  
burkay.genc@hacettepe.edu.tr

**Abstract**—Predicting the success of a mobile game is a prime issue in game industry. Thousands of games are being released each day. However, a few of them succeed while the majority fail. Towards the goal of investigating the potential correlation between the success of a mobile game and its specific attributes, this work was conducted. More than 17 thousands games were considered for that reason. We show that specific game attributes, such as number of IAPs (In-App Purchases), belonging to the puzzle genre, supporting different languages and being produced by a mature developer highly and positively affect the success of the game in the future. Moreover, we show that releasing the game in July and not including any IAPs seems to be highly associated with the game’s failure. Our second main contribution, is the proposal of a novel success score metric that reflects multiple objectives, in contrast to evaluating only revenue, average rating or rating count. We also employ different machine learning models, namely, SVM (Support Vector Machine), RF (Random Forest) and Deep Learning (DL) to predict this success score metric of a mobile game given its attributes. The trained models were able to predict this score, as well as the rating average and rating count of a mobile game with more than 70% accuracy. This prediction can help developers before releasing their game to the market to avoid any potential disappointments.

**Keywords**-data mining, mobile games, game design, game features.

## I. INTRODUCTION

In 1997, the first mobile game was released by NOKIA, called “Snake”. This mobile game initiated a new era in the world of game industry. Game industry, today, is a huge market and notably growing by time. Statistics reveal that the value of the video game market in the United States by 2017 was more than 18 billion dollars [1] which discloses the seriousness and the value of this business. In Apple app store only there is currently more than 265 thousands gaming applications [2]. Free distribution and the low cost of mobile application development made both small size companies and individuals able to enter this flourishing market easily. While the recent advancement in mobile technology made them more powerful and cheaper. This in turn, increased memory size, processing power and graphics quality of mobile devices, which supported more sophisticated and enjoyable games. Nevertheless, a perfect graphical design, an interesting story

line, appealing animations and attractive characters are expected to be the main factors behind the success of a game, we can see a lot of examples where simple and non-professional games reached the peak -in term of popularity and revenue. Apparently, following the common sense may not lead -necessarily- to a success. In contrast, studying games’ success and failure stories may help us to better comprehend the direct and indirect causes of success.

In this work, we present a thorough study where we analyze more than 17 thousand games to answer the main question of “what makes a game fail or succeed?” First major problem in mobile game industry is the great difficulty in defining a general game success measure and second issue is predicting the success of a particular game from its early development phase. We undertook this study to ask and answer the mentioned two questions following a rigorous scientific approach. We argue that number of ratings or average rating of a mobile game shown in app store are not enough measures to describe the success of a game. Thus, we introduce a new measurement to quantitatively assess the success of a mobile game.

Data mining comes as a powerful tool for highlighting the unseen patterns in large collections of data. So by performing Exploratory Data Analysis (EDA), it is possible to extract wisdom from data. A series of steps and procedures ought to be followed carefully to reach the ultimate goal of understanding the hidden facts and rules that control the phenomenon under consideration. The collected data can be subjective, biased, based on wrong assumptions, invalid or insufficient. A pre-processing step is usually required to discard unnecessary data (data cleaning), transform the data from one form to another, or perform some data regularization and normalization for later steps. Following this, the feature engineering takes place. With the help of domain knowledge, new features can be introduced, some attributes can be removed or some data entries can be discarded (outliers). Finally, suitable algorithms can be applied to extract information from the data. Under the scope of this study, we try to call attention to some unique attributes that were shown to influence the success of a mobile game.

Furthermore, this work is dedicated to predict the success of a game based on different set of attributes. In conjunction

with this, we study the most important attributes that are used in the prediction process. That is done to highlight the importance of these features so that future developers consider such attributes more reasonably.

## II. RELATED WORK

In the literature, we can see a few number of studies that asked similar questions to the one asked in this paper. However, what distinguishes each of them is the success definition and the studied attributes. In [3], the top 100 games of Google Play app store were considered. The number of downloads and revenues are studied with respect to 37 different features associated with the games. They conclude that a feature like allowing IAPs is associated with success. On the other hand, features like inviting friends and customizing the game are usually linked with the opposite. In parallel to that work, the revenue is studied in [4] with respect to 31 attributes for 50 iPhone games. The aim of the study is to identify the most 10 important features in game development. They confirm a strong relation between the total number of people who run or get engaged with a mobile game and its revenue. Similar to these two works, we examine various game’s attributes and their effect on average rating and rating count. However, instead of considering a limited number of mobile games, we consider 17 thousands mobile game apps to increase the statistical significance of the findings.

Another work [5] concluded that selecting a less competitive genre for an application and keeping the application’s quality at higher levels by updating the application frequently were shown to be positively correlated with the application staying a longer time among the top applications. In [6], the authors state that the previous history of the application releasing company was shown to be correlated with application success. Similar to these two studies, we examine the success of a mobile game given its genre and the developer maturity. However, we do not study all apps categories but we focus only on mobile application. Alternatively, the icon design and its relations to mobile game downloads and purchases were examined in [7] for 68 game applications. Uniqueness and realism are found to cause more downloads, clicks and purchases. In our work, we do not consider the icon design. However, it could be a future work to further enhance the predictive models accuracy.

In [8], they show that advertising a mobile game on TV and uploading it on many app stores increase its chance of success. While, they found that the success of the game decreases over time.

In contrast to the previous works, first we define a new success measure combining the main three objectives of game development, namely, revenue, popularity and reputation. Second, we consider a large number of mobile games spanning different genres, various developers, for free and

paid games and other unique attributes. Third, we conduct both analytical and predictive studies.

Table I: The main attributes of the dataset used in this work.

Attribute Name	Type	Attribute Name	Type
URL	textual	In-app Purchases	numeric
ID	numeric	Description	textual
Name	textual	Developer	textual
Subtitle	textual	Age Rating	categorical
Icon URL	textual	Languages	categorical
Average User Rating	numeric	Size	numeric
User Rating Count	numeric	Primary Genre	categorical
Price	numeric	Genres	categorical
Current Version	date	Original Release Date	date

## III. METHODOLOGY

In this section we outline our methodology. We first discuss our feature engineering efforts, then we describe our novel success score metric and finally we summarise the predictive models we used.

### A. Feature Engineering

The used dataset [9] contains more than 17 thousand mobile games taken from Apple App Store. Initially, each game is associated with 18 different attributes. A negligible number of these values is missing. It is relatively a recent dataset since it was collected in August of 2019. That makes it a good proxy of the actual game market. In Table I the main attributes of the dataset are shown together with their types.

**Attributes Selection** The unnecessary attributes such as the URL, ID and ICON.URL are removed. Because they are irrelevant to the success of a mobile game. Other attributes like the Subtitle and Description are discarded as well under this scope. However, they can be good candidates for our future work. Primary.Genre is neglected since more than 96% of the observations were marked as Games. We note that the actual game genre information is presented in the Genres attribute which was taken into account.

**Handling Missing Data** Games with null Rating Count means they have less than 5 ratings. We assign them 5 rating counts. At the same time, games with Rating Average of null, means they did not gain more than 5 rating counts. Because of this, we assign them Rating Average of 0. Following this, we remove games with missing values for Price and Size attributes. After that, in the Languages column, we see some cells are left with no entry, we handle this by assigning the English language for these games. We note that the English language constitutes the majority in the dataset and hence was selected as the default for imputation.

**Deriving New Attributes** The attributes of the final dataset is shown in Table II. Based on the raw dataset, few more

Table II: The main attributes of the categorized dataset.

Attribute	Categories	Attribute	Categories
Num. Languages	Single, Many	Release Month	1,2,3,4,5,6,7,8,9,10,11,12
Num. Genres	Less than 4, Equal or More than 4	Elapsed Months	Less than 17, [17, 31], [32, 45], [46, 64], More than 65
Age Rating	+4, +9	Developer Category	Newbie, Professional
Availability IAP	No, Yes	Game Size	S1, S2, S3, S4, S5
Average IAP	No, Some, High	Game Free	No, Yes
Min IAP	0\$, ~1\$, Equal or More than 1.5\$	Rating Average	Low, Moderate, High
Max IAP	0\$, Less than 2.5\$, Equal or More than 2.5\$	Success Score	Successful, Undetermined, Unsuccessful
Sum IAP	No, Some, High	Rating Count	Low, Moderate, High

attributes were derived to further enhance the level of information associated with each single game. These are the number of languages and genres, the game’s developer maturity, Number, Min, Max, Sum and Average of IAP, Price and Game Size.

### B. Success Score Measure

In this work, we propose a new success measure that encodes the three famous traditional success measures, namely, gross revenue, number of downloads and rating.

The suggested success measure shown in (1) gives importance to the following three objectives: Revenue, Reputation and Success Speed. **Revenue** represents the expected income by this game. It is calculated by multiplying the price of the game with the expected number of people who will buy the game. Then, the expected income via IAPs is added. **Reputation** is represented by the average rating of the game. **Success Speed** is encoded in the success formula as well.

$$S = \frac{(RC * G * P + RC * K * AP)_N + RA_N}{T_N + \epsilon} \quad (1)$$

where: S is the success score, RC denotes rating count, RA denotes rating average, P is the initial download and installation price, AP stands for average IAPs, T denotes elapsed time since first release, K is expected number of users who will buy in app purchase per rating count, G is the expected number of users who will buy the game per rating count, (.)\_N stands for the operation of normalizing the operand to [0, 1] interval. Epsilon is a very small number to avoid division by zero.

For a game to score high on this measure, it should maximize the revenue and rating average. At the same time, it should do that in a short period of time. It is important to note however, that K and G are estimated variables that can be changed to match the targeted audience. Parameters K and G were guesstimated as 0.05 and 0.5, respectively.

### C. Predictive Models

In the process of conducting our statistical analysis, four different models were applied. Under the scope of

understating the correlations between success and mobile game’s attributes, the Apriori algorithm [10] and [11] was utilized based on [12] and [13] implementations. At the same time, we ranked the games in terms of success score, rating average and rating count. Then, the attributes distribution in the best and worst 100 games were analysed. On the other hand, for the predictive part, SVM [14], RF [15] and DL models were used.

In Table III the used models and their associated parameters are shown.

## IV. ANALYSIS AND FINDINGS

### A. Wisdom Extraction

One main goal of this work, is to show the link between the success and the various game development attributes. For that goal in mind, Apriori algorithm was applied for mining possible associations. The support parameter was set to 0.001 and the confidence one to 0.8. Following this, the highest 5 rules were picked in terms of the confidence and support. In this sections, our main findings are highlighted in term the proposed success measure, the average rating and the rating count.

Table III: Models and their corresponding hyper parameters.

Model	Full Name	Parameters
Apriori	Mining Associations With Apriori	Support = 0.001, Confidence = 0.8
SVM	Support Vector Machine	Kernel = radial, Cost=19.5
RF	Random Forest	n tree = 800, importance = TRUE
DL	Deep Neural Network	Hidden Layers = [70, 50, 20, 10, 4] Number of Epochs = 35 L1 = 0.001, L2 = 0.01 Activation Function = Maxout Adaptive Rate = TRUE

**Success Score** Most important rules with respect to our novel success score are given in Figure 1. Some of the interesting findings appear to be being a long time developer, availability of cheap IAPs, and publishing Puzzle and Travel games help with success. In parallel to that, making your game available in many languages can increase the popularity and revenue. Thus, the probability of success. Surprisingly, releasing a game in July is expected to decrease

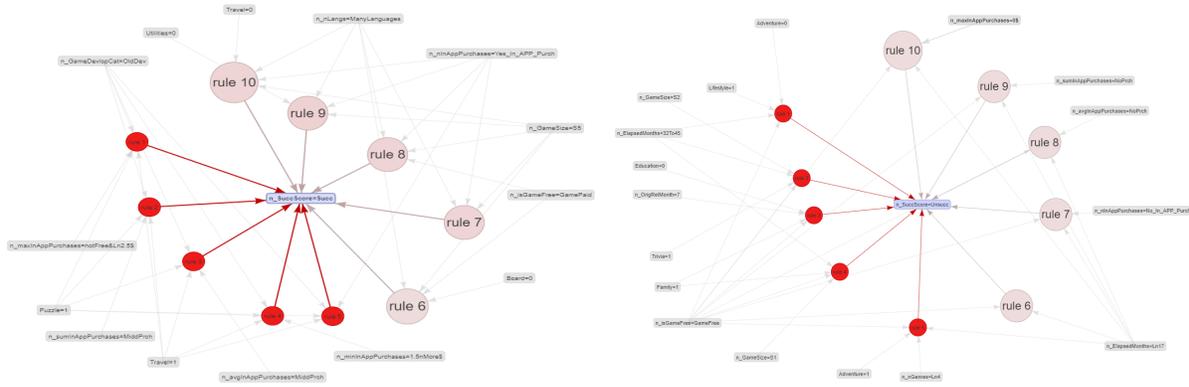


Figure 1: The strongest 10 rules associated with high (on the left) and low (on the right) success scores. Size and color represent support and confidence' respectively.

its chance in achieving a success. The reason could be that US and Europe start their holidays at that time so people do not prefer playing video games for fun.

In addition to the previous general observations, specifically we study the 100 games with the highest and lowest success score. As shown in Figure 2, in a successful game the number of game genres is relatively less, IAPs are found, and the game size is larger.

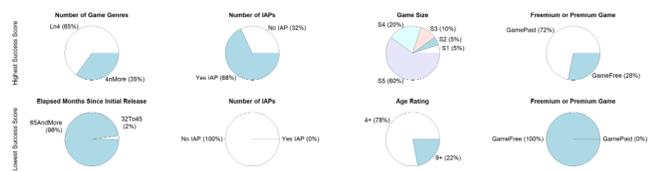


Figure 2: Attributes distribution in the best and worst 100 games in term of the success score.

**Average Rating** The rules associated to average rating and observations from Figure 3 suggest game developers who want to achieve higher rating average to think of avoiding both small size and free games. At the same time, they should aim at including IAPs and focusing on +9 audience.

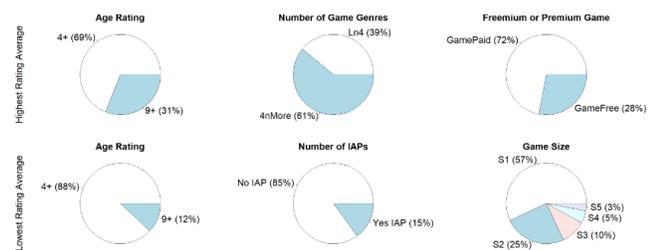


Figure 3: Attributes distribution in the best and worst 100 games in term of the rating average.

**Rating Count** Most important rules with respect to rating count show that if the goal of the game is to attract high

audience number, the developer should think of making the game available in many languages and targeting +9 audience. In parallel to that, Figure 4 reflects observations on different features. We can see that making a game free does not necessarily make it popular. However, the opposite was shown to happen.

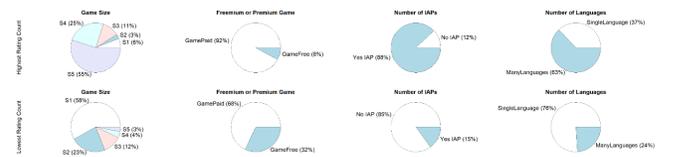


Figure 4: Attributes distribution in the best and worst 100 games in term of the rating count.

**B. Prediction**

The second major aim of this work is to predict the success of a mobile game. Firstly, the dataset is partitioned into train, validation and test splits by taking 80% for training, 10% for validation and 10% for testing. Following this, SVM, RF and DL models were trained. All the models obtain around 70% classification accuracy as shown in Figure 5. However, it can be observed that predicting the proposed success score is more accurate as compared to predicting rating average and rating count. That could be the case because our success measure is closely associated with the game development features much more than the other success criteria. In other words, rating average and rating count might be influenced by other factors that are not included in the main studied features. In parallel to that, it is clear that the best model among the four predictive models is the RF model. Thus, it is worth studying the true and false positives and negatives of the this model. For the success score prediction, it sounds that the model does the best at predicting the unsuccessful games. While the same observations were shown for the low rating and low

Table IV: Confusion matrices for the best model (RF) for (a) Success Score, (b) Rating Average and (c) Rating Count.

	Prediction		
	Succ.	Undet.	Fail
Succ.	<b>285</b>	142	139
Undet.	119	<b>435</b>	10
Fail	13	61	<b>495</b>

(a) Success Score

	Prediction		
	Low	Medium	High
Low	<b>778</b>	35	132
Medium	83	<b>30</b>	101
High	194	18	<b>328</b>

(b) Rating Average

	Prediction		
	Low	Medium	High
Low	<b>863</b>	32	85
Medium	202	<b>21</b>	81
High	120	20	<b>275</b>

(c) Rating Count

number of ratings as shown in Tables IVa, IVb and IVc. In fact this observation is critical because it sheds light on two points. First there are some features and decisions that lead to game failure whether the success was in term of rating average, rating count or the general proposed success measure. Thus, the developer should avoid these features and decisions. Oppositely, the successful games are harder to be linked to specific attributes like unsuccessful ones. It implies that there is a clear path for your game to fail. However, success path is dependent on more sophisticated factors like creativity and novelty.

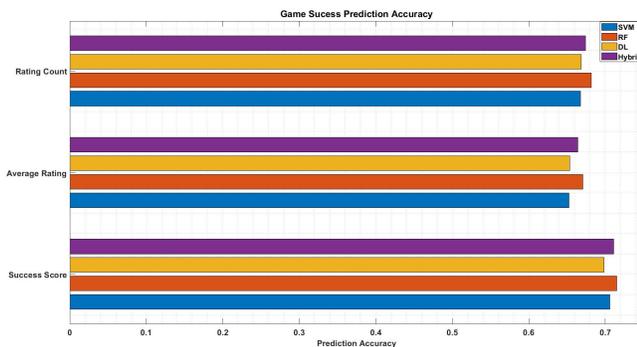


Figure 5: The prediction accuracy of game success using SVM, RF, DL and hybrid models.

## V. CONCLUSION

In this work, we presented some facts regarding the relation between the success of a mobile game and the various game development process attributes. The success was given by a novel formula that encodes the general goal of game development i.e. revenue, popularity and reputation. We were able to show that some attributes can impact the success positively or negatively. As a future work, we are planning to consider more datasets and to perform similar kind of analysis to prove the consistency of the found results.

## REFERENCES

- [1] C. Gough, "Video game industry - statistics and facts," <https://www.statista.com/topics/868/video-games/>, May 2020.
- [2] —, "Apple app store: number of available gaming apps as of q1 2020," <https://www.statista.com/statistics/780238/number-of-available-gaming-apps-in-the-apple-app-store-quarter/>, August 2020, online; accessed 06 August 2020.
- [3] Á. Moreira, V. Vicente Filho, and G. L. Ramalho, "Understanding mobile game success: a study of features related to acquisition, retention and monetization," *SBC*, vol. 5, no. 2, 2014.
- [4] K. M. Alomari, T. R. Soomro, and K. Shaalan, "Mobile gaming trends and revenue models," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2016, pp. 671–683.
- [5] G. Lee and T. S. Raghu, "Determinants of mobile apps' success: Evidence from the app store market," *Journal of Management Information Systems*, vol. 31, no. 2, pp. 133–170, 2014.
- [6] G. Lee and T. Raghu, "Product portfolio and mobile apps success: Evidence from app store market." in *AMCIS*, 2011.
- [7] H. Jylhä and J. Hamari, "An icon that everyone wants to click: How perceived aesthetic qualities predict app icon successfulness," *International Journal of Human-Computer Studies*, vol. 130, pp. 73–85, 2019.
- [8] J. Yi, Y. Lee, and S.-H. Kim, "Determinants of growth and decline in mobile game diffusion," *Journal of Business Research*, vol. 99, pp. 363–372, 2019.
- [9] Tristan, "17k mobile strategy games," <https://www.kaggle.com/tristan581/17k-apple-app-store-strategy-games>, May 2020.
- [10] R. Agrawal, "I mielinski t and swami a," *Mining association rules between sets of items in large databases*, vol. 993, 1993.
- [11] D. P. Muino and C. Borgelt, "Frequent item set mining for sequential data: Synchrony in neuronal spike trains," *Intelligent Data Analysis*, vol. 18, no. 6, pp. 997–1012, 2014.
- [12] C. Borgelt and R. Kruse, "Induction of association rules: Apriori implementation," in *Compstat*. Springer, 2002, pp. 395–400.
- [13] C. Borgelt, "Efficient implementations of apriori and eclat," in *FIMI'03: Proceedings of the IEEE ICDM workshop on frequent itemset mining implementations*, 2003.
- [14] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, pp. 1–27, 2011.
- [15] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.