What Are the Characteristics of High-Rated Apps? A Case Study on Free Android Applications

Summary

Aim
The paper tries to find characteristics of successful Android applications in order to help future app developers gain better insights.

Dataset

1,492 high-rated and low-rated free apps mined from the Google Play store are investigated against 28 factors along 8 dimensions

Results
The size of an app, the number of promotional images that the app displays on its web store page, and the target SDK version of an app are found out to be the most influential factors.

Strengths

1. This paper is very **significant for future app developers** as it gives them something to keep in mind and focus on to gain a higher chance of success.
   Also, these **findings will serve as a starting point in relevant research**. For instance, I personally am extending this paper and its findings to identify the characteristics of ~~high-rated~~ the most downloaded applications.
2. The findings of this paper **challenge** some of the **conventional beliefs,** hence giving a new perspective.
   For instance, one would think that having a complicated code base and a high number of dependencies is a recipe for disaster since there is a greater probability of bugs and overall inefficiency. However, this paper finds out that it is the opposite: high-rated apps are the ones that usually have more complex code and depend more on libraries.
3. **The approach for building the model is pretty good.**
   The data is randomly separated into ten folds and ten rounds of experiments are performed. For each round, nine folds are selected to train the random-forest classifier, and then the learned classifier is applied on the remaining fold. This process is repeated ten times and the aggregate of the results is taken to measure the overall performance for the whole data.
   To add on to that the model is built using all factors from the eight dimensions to ensure better performance (and since app rating is impacted by multiple factors).

Weaknesses/ Controversial points

1. There are some **concerns about the overall methodology** in the paper.
   a. With over 2.7 million applications in the Google Play Store, the paper analyzes just 746 high-rated apps and 746 low-rated free apps. Such a **small dataset** raises questions about the gathered insights.

b. The **paper never talks about the rationale behind the 28 factors** chosen to evaluate apps against.
c. **Biased selection of apps based on category**
The paper claims to have chosen apps across several categories. While this is partially true, the different categories aren't represented equally. For example, 10.21% of apps are from the tools category while only 0.71% are weather apps.

2. This paper considers (high) ratings as a **metric for application success**.
However, there are plenty of examples of applications that are really popular and successful but struggle to maintain a good rating.
    a. Tinder - 3.5 with 100M+ downloads
    b. TikTok - 1.6 with 1B+ downloads
    c. Google Classroom - 2.5 with 100M+ downloads

    Hence, user ratings might not be the best choice for determining application success.
3. The paper has a **narrow focus.** It only focuses on free applications, not paid ones. Furthermore, only Android applications are considered. iOS applications or any other mobile store applications are not a part of the dataset.
Perhaps, this might be the reason for the various inconsistencies (with regards to the findings) with related papers.

Suggestions to improve the paper

These suggestions below will allow to overcome some of the above weaknesses.

1. **Increasing the size and quality of the dataset**. This will improve the accuracy and credibility of the paper.
2. A lot of apps have a greater number of ratings because they ask the user for feedback while they use the app. This rating can be positive or negative based on the time or event performed.
For example: prompts for feedback after a successful in-app transaction usually results in positive ratings.
Hence, **bases** like these **need to be addressed and accounted for** in the overall methodology.
3. **Analyze additional factors** relevant to app rating like
    a. App Update frequency
    b. Number of Languages supported by the app
    c. Privacy policy

    This will allow for a more realistic model.