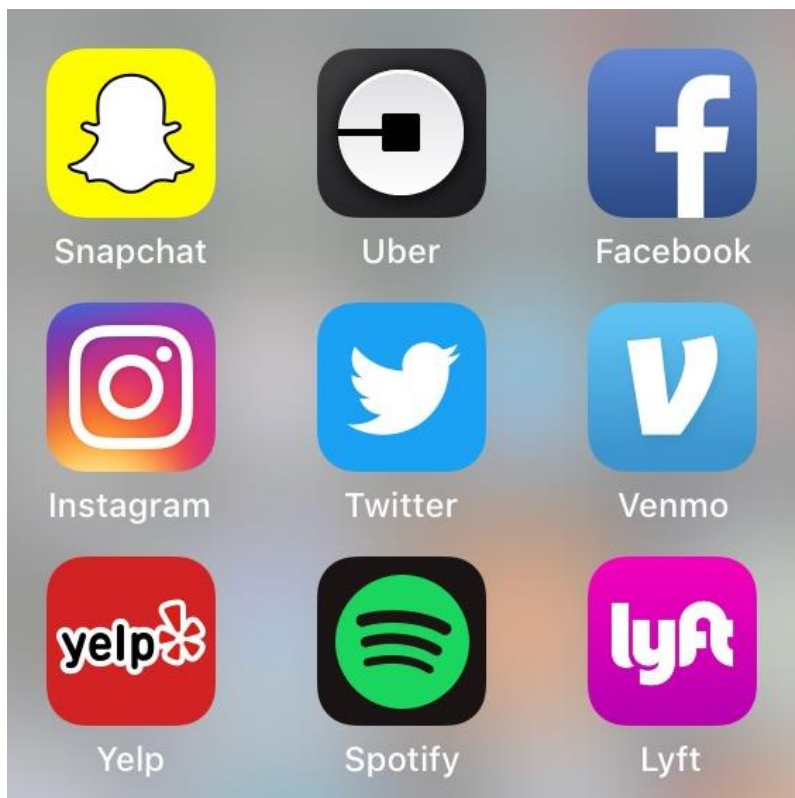


Google Play store Apps Business case

What makes an app in the Google Play Store popular?





Introduction

There are a lot of popular apps in the play store from lots of different categories ranging from games to social.

What is it that makes those apps so popular? Is there some kind of formula or algorithm behind it?

And if there is one what could it be? Is it the same one every time or does it differ per app?

Even the most simple apps are very well received. A perfect example of this is Flappy Bird. The user can only press the screen to make the bird fly and yet it comes on the number 1 popular app spot in the Google Play Store.



The goal of the project is to clarify the following question:

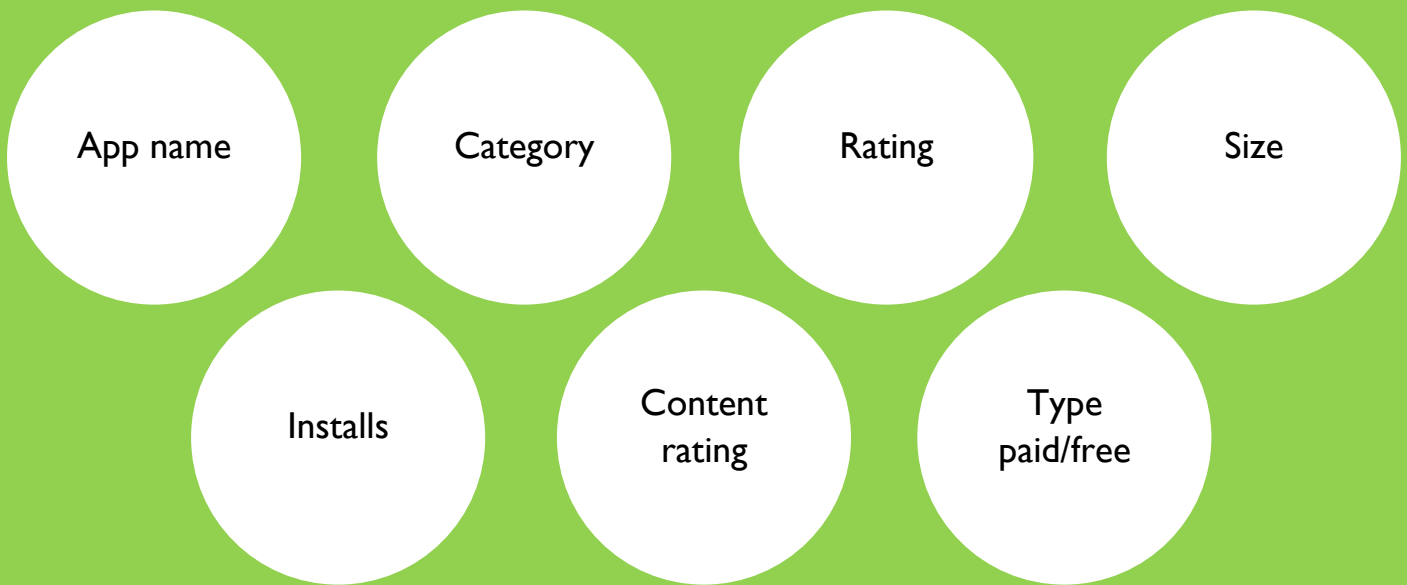
What makes an app in the Google Play Store popular?

The data

For our data we have 2 datasets available from Kaggle.

The first dataset is about the apps on the Google Play Store. The second dataset we have is about the most relevant user reviews on apps.

Apps on the Google Play Store features



There are a lot of interesting features to look at. We can already division in the data between before and after.

Before the app was released:

App name: name of the app.

Category: category of the app (beauty, business, communication).

Size: download size.

Content rating: age group.

Type: paid or free.

After the app was released:

Installs: number of user downloads.

Rating: overall user rating.

Most relevant user reviews on apps features

App name

Translated
review

Sentiment

In this dataset there are not a lot of features. This dataset is more like an extend of the main dataset.

The apps on the Google Play Store dataset consists of 10.8 thousand rows, the most relevant user reviews on apps dataset consists of 64.3 thousand rows.

With the two datasets combined together we can see if users are positive or negative about the app.

Machine learning

To find the solution for the question I will be using machine learning. The machine learning algorithm that I am going to use is a neural network. A neural network is vaguely inspired by biological neural networks. It is used to model complex relationships between inputs and outputs or to find patterns in data.

A neural network uses the concept of deep learning in the form of multiple hidden layers. This approach tries to model the way the human brain processes light and sound into vision and hearing.





Conclusion

For this project I will use the Jupyter Notebook to explore the dataset and create machine learning algorithm. The Jupyter Notebook is a document where you can write plain text and python scripts.

What makes an app in the Google Play Store popular? Is it determined by one property or a combination of properties? Could it be that it is just pure luck?

The answer to those questions are going to be revealed in my project.

