Google Play Apps Rating Analysis

Abstract

This project aimed to identify the apps that are going to be good for Google to promote by analyzing the App ratings using machine learning models to predict which apps will have high ratings.

Question:

1. Which application has the highest rating?
2. What is the most popular category?
3. Which app has the largest number of installs?

Design

This project is one of the T5 Data Science BootCamp requirements. Data provided by Kaggle has been used in this project. Classifying how this feature will help bring more attention to newer apps that have the potential. By using machine learning algorithms would enable understanding most helpful apps for google to promote.

Data

The dataset is provided in `.csv` format. The used data in this project is provided by Kaggle. It contains 10842 apps, each app has 13 features. The most relevant feature to this project is the rating which contains the apps rating.

Tools

- Pandas for data manipulation
- Scikit-learn for modeling
- Matplotlib and Seaborn for plotting