

Linear Regression Application Price Prediction



Raghad Aloraini
Raghad Alawad
18 September 2021

Introduction

In this project, we will be doing app price prediction using linear regression models. The Play Store Apps data has enormous potential to drive app-making business to success. Hence by drawing meaningful insights from data by doing analysis it can really help developers improve their work.

Dataset

The datasets we will be using are taken from the Google Play website (<https://play.google.com/store?gl=SA>). By using this dataset, we are trying to make predictions of the price of the apps, using features such as Rating, Installs, Reviews, Category.

Target variable: Price [numerical]

Features: Ratings & Installs & Reviews[numerical], Category[categorical].

Algorithms

For data cleaning and pre-processing, we will start by deleting the duplicate records and check if there are any null values then drop them. Also, we will not encode the categorical variables Category because we want numerical values only.

Tools

To predict the price, we will be using different tools such as Jupyter notebook, Excel. Also, we will use different libraries with python such as pandas, sklearn, BeautifulSoup, numpy.

Conclusion

To conclude, we expect the regression model will predict the app price to help developers with their work and also give insights to users based on what app is popular.