Mobile Games and User Ratings

Final Capstone Project

Christine Li

Overview

1. Business Problem
2. Methods
3. Model Building and Model Evaluation
4. Conclusions
5. Future Work
6. Q&As

Business Problem

The mobile phone games industry has become increasingly popular over the years with the introduction of smartphones. As technology has advanced, the smartphone has a gaming platform has become a major source of profit for companies.

User ratings are just one factor among many that affect gaming success. Game developers can leverage positive reviews to increase excitement and attract players. Games with higher ratings are more likely to be more visible in App stores and can stand out from competitors. Higher-rated games are also more successful in monetising their user base.

Source: Adobe Stock Images

Our goal is to create a predictive machine learning model that will predict the user rating of a mobile strategy game. Using this tool, game developers can create a mobile strategy game that capitalises on features that players enjoy leading to increased user ratings.

Data

This dataset, '17K Mobile Strategy Games' was downloaded from Kaggle.

This can be accessed with the URL https://www.kaggle.com/datasets/tristan581/17k-apple-app-store-strategy-games?resource=download

The data was collected by the owner of the dataset by mostly using the iTunes API, App Store sitemap, along with some web scraping on 3rd of August 2019.

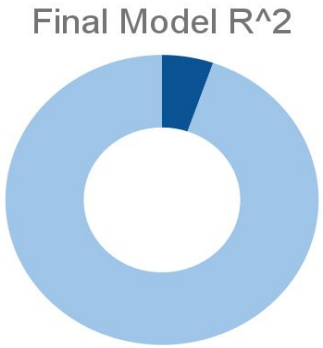
This original dataset contains 17007 records and 18 columns.

- *URL* - The URL for the mobile game
- *ID* - The assigned ID for the game
- *Name* - The name of the game
- *Subtitle* - The secondary text under the name.
- *Icon URL* - The URL for the game icon
- *Average User Rating* - Rounded to nearest 0.5, requires at least 5 ratings.
- *User Rating Count* - Number of ratings internationally, null means it is below 5
- *Price* - Price in USD
- *In-app Purchases* - Prices of available in-app purchases
- *Description* - Mobile Game App description
- *Developer* - App developer
- *Age Rating* - Either 4+, 9+, 12+ or 17+
- *Languages* - Language codes using ISO Language Codes
- *Size* - Size of the app in bytes
- *Primary Genre* - The main genre of the app
- *Genres* - Genres of the app
- *Original Release Date* - When it was released
- *Current Version Release Date* - When it was last updated

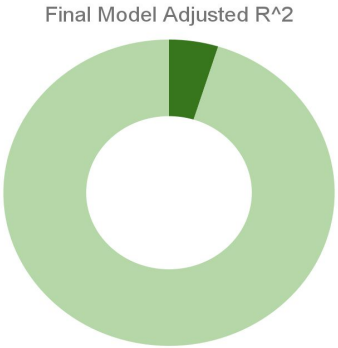
Methods

1. Data collection
2. Data preparation
    a. Deal with missing values and outliers, fix data values and data types
    b. Data exploration to understand the data
    c. Data Preprocessing:
        i. Encoding categorical variables, scaling numerous features
3. Model building
    a. Multiple regression
    b. Split data into training and test sits
    c. Fit the model
    d. Checked met multiple linearity assumptions
4. Model evaluation

R^2: 0.053



Adjusted R^2: 0.048



Prob (F-statistic)< 0
Reject the null hypothesis

Final Model

Dep. Variable:	Average User Rating	**R-squared:**	0.053
Model:	OLS	**Adj. R-squared:**	0.048
Method:	Least Squares	**F-statistic:**	10.30
Date:	Mon, 26 Feb 2024	**Prob (F-statistic):**	2.13e-102
Time:	19:11:03	**Log-Likelihood:**	-8795.3
No. Observations:	12476	**AIC:**	1.773e+04
Df Residuals:	12407	**BIC:**	1.824e+04
Df Model:	68		
Covariance Type:	nonrobust		

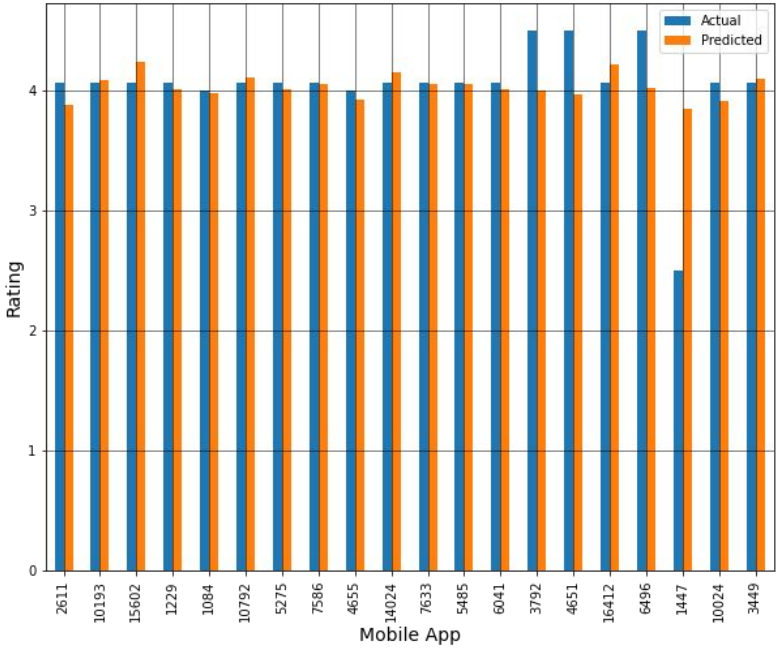
Model Evaluation

R2: 0.053

Train MSE: 0.23

Test MSE: 0.22

```
Accuracy: 0.0576543548999371
```




```
Train Mean Squared Error: 0.2398071778731519
Test Mean Squared Error: 0.22223161203951125
```

Feature importance

Top 5 coefficients:

1. Current Version Year
2. Original Year
3. User Rating Count
4. Secondary Genre Puzzle
5. Secondary Genre Magazines and Newspapers

	Coefficients
Current Version Year	0.059908
Original Year	0.049366
User Rating Count	0.038149
Secondary Genre_ Puzzle	0.009813
Secondary Genre_ Magazines & Newspapers	0.007247
Secondary Genre_ Travel	0.007054
Price	0.004515
Secondary Genre_ Casual	0.004197
Primary Genre_Music	0.003977
Secondary Genre_ Navigation	0.003711
Secondary Genre_ Medical	0.003610
Secondary Genre_ Health & Fitness	0.003529
Secondary Genre_ Photo & Video	0.003233
Secondary Genre_ Music	0.002900
Secondary Genre_ News	0.002450
Languages	0.002197
Primary Genre_Food & Drink	0.002141
Primary Genre_Shopping	0.001807
Secondary Genre_ Finance	0.001668
Size	0.001502

Conclusion

To improve user ratings for mobile games, developers should focus on optimising features such as:

- Continuing to update games regularly
- Continue to gather player ratings to increase the User rating count
- Identify popular genres and incorporate them into the Strategy game, for example Puzzles.

Any questions?

Thank you