

Recommendation System

By Victor Mwatu

Business Understanding

This project aims to create a machine learning model that recommends movies to a user based on the user's ratings of other movies

The main stakeholders of this project are movie streaming service providers or movie stores online

Due to explosion in digital products being offered due to nature of stores being online there is need for easier options to customers

This helps in customer satisfaction and retention therefore revenue



Data Understanding

The dataset: the 'Movielense' dataset from the GroupLens research lab at the University of Minnesota.

This is a well known dataset used in many recommendation system academic papers and machine learning proof of concept.

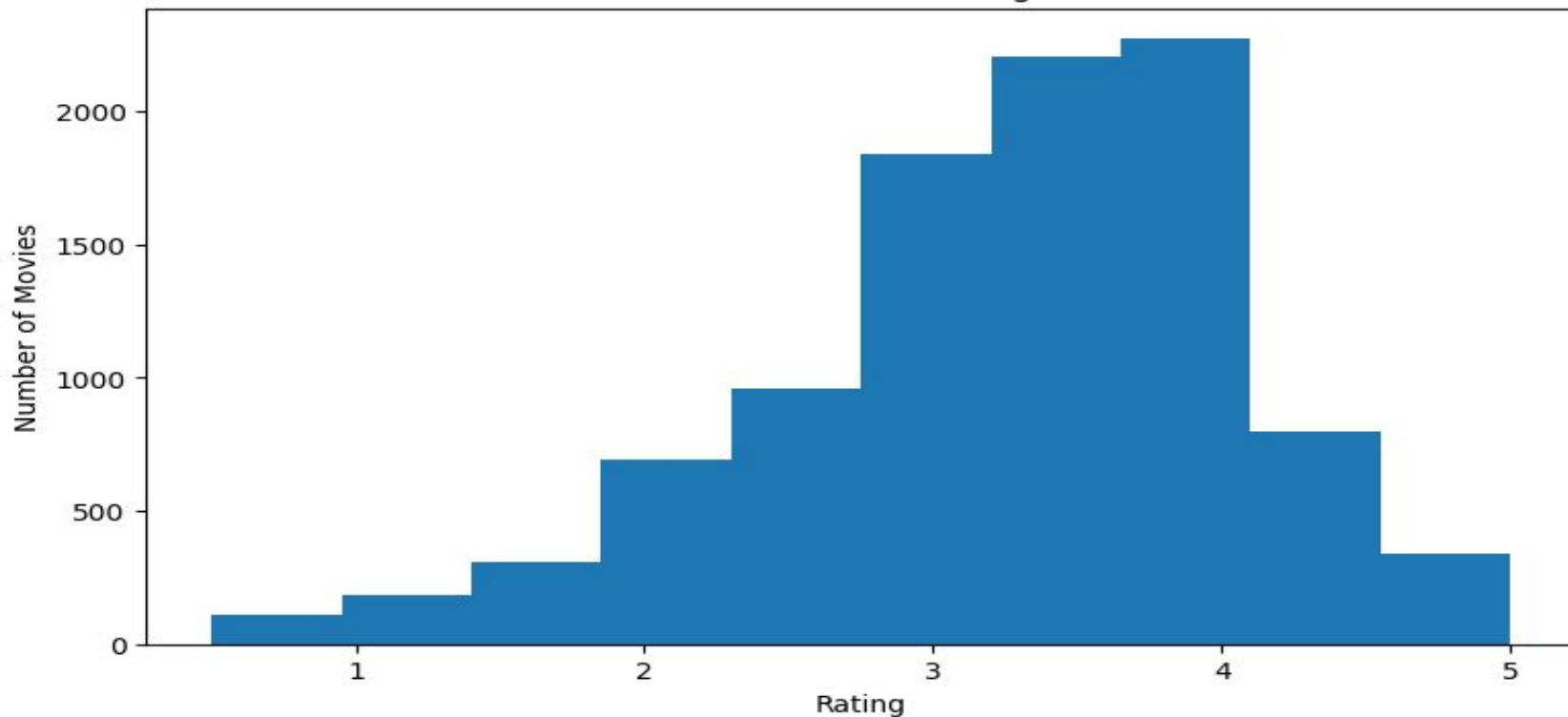
It was clean with no missing values

Some insights from the data as follows:

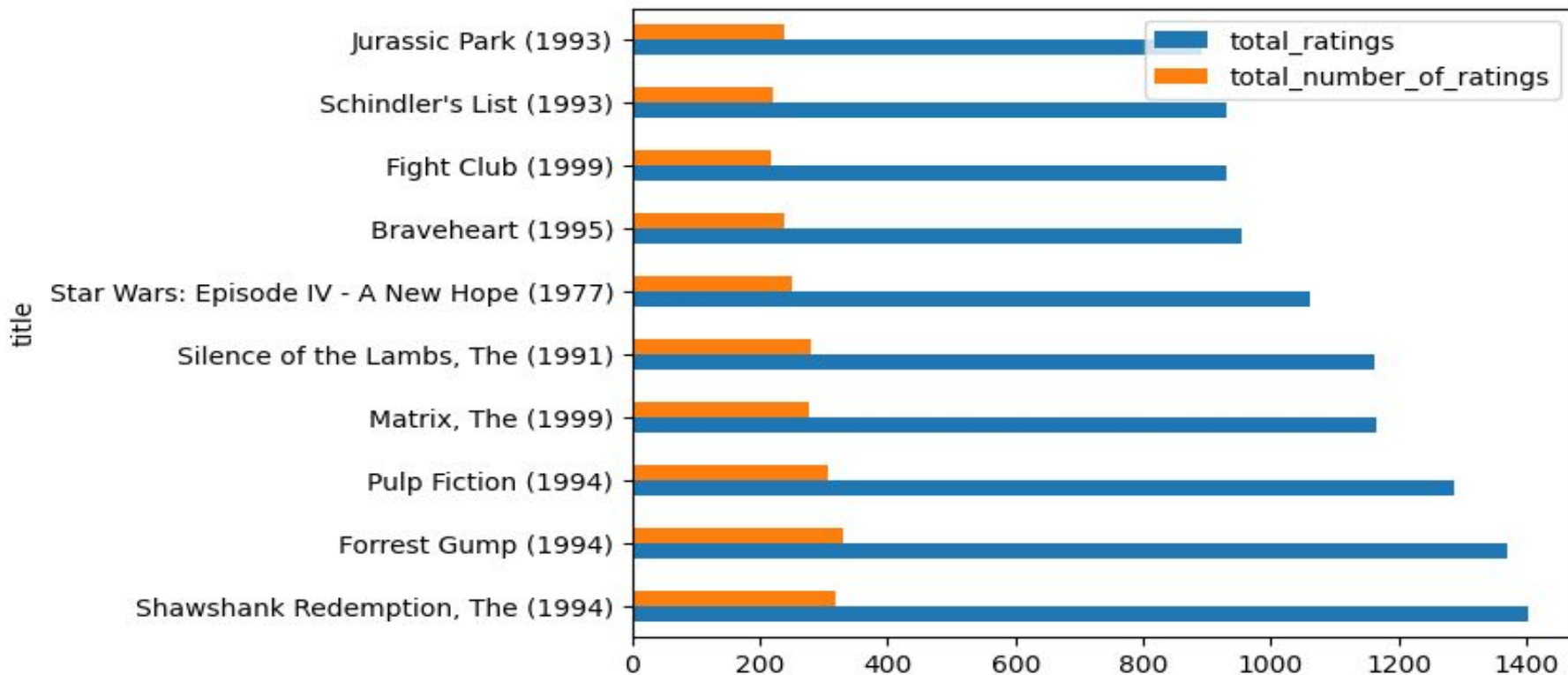


Data Insights: Few high rated movies

Distribution of Ratings



Data Insights: Top rated movies



Modelling ML

Modelling done using Surprise Python library

The best performing basic model turns out to be the simple SVD with an RMSE of 0.877 and the worst is the KNN-basic with RMSE of 0.977

The model was optimised and trained on the whole dataset



Conclusion

Ratings can be accessed from the model in many ways from application

Basic prediction method:

A new user's data is merged with the dataset

The dataset is compiled

Top n ratings given by selecting top n predictions

