HarvardX PH125.9xData Science: Capstone

Simon Gibson

Movielens Project - Machine Learning Submission

Introduction

For the 9th Course in the HarvardX Data Science course we have been asked to build a movie recommendation system using the MovieLens dataset.

The movie recommendation system will predict movie ratings.

The Movielens data package can be found at the MovieLens homepage.

MovieLens is a project run by GroupLens - a research lab run at the University of Minnesota in North America. MovieLens is a non-commercial collection of movie data

Note: code in this package is written on Linux. If running on windows you may need to change the line endings. This code has not been tested on Windows.

Methods

Processes and techniques used

Data cleaning

Data Exploration and Visualisation

Modelling Approach

Results

Conclusion

Limitations

Future work