AUGUST 5, 2021

FINAL PROJECT-EXECUTIVE-SUMMARY
Movies Recommender System

ANWESH PRAHARAJ
STA 6704

Table of Contents

1. Introduction

A successful recommendation system explores the relation between the users and items to get the dependency, which help to produce better suggestions. There is a huge demand of online movies, series in today's era, especially because of covid as families are spending time with each other. There are many competitors in market and user will get the subscription who can provide a better recommendation of movies. Industries are spending lot of wealth on building a best and accurate recommendation system which can attract the targeted audience and increase the marketing and sale. For this project I have used Movie-lens 100K dataset which include item data contains Movie and genres information and user data contains user , movie and rating mapping.

2. Methods

To build a better predictive recommendation system I used collaborative filtering and content-based filtering approach. Collaborative filtering is based on collaboration of different user which is more advance and beneficial while comparing to content-base filtering where all the predictions are based on one user profile. In content-based method we used term frequency and inverse term frequency which is used to find the relative similarity between users and movies. Cosine similarity is used in case of collaborative filtering to find the similarity between user and movies. The value is tends to 1 if there is a similarity or else it will tends to zero. Collaborative filtering we divided into memory based and model based algorithms. Memory based algorithms include item-item filtering and user-item filtering to achieve different case studies. For model-based filtering used Matrix factorization with single value decompose and K mean clustering algorithm. Cluster algorithms help to find out similar users and movies in different scenarios club into one group. This give us an idea related to user like or dislike a movie and how many users are in same situation. In Matrix factorization we decomposed the User Movie matrix into two non empty matrix user and Movie and finally used the dot product to create the final matrix which we used to find the similarity.

3. Evaluation

As content-based filtering is based on genres so we have used KNN classifier to evaluate the model and found a hit rate 0.93 which is a good model. For memory-based algorithms like item-item filtering and user-item filtering used straight forward function for recommendation. Whereas for model-based algorithm used Single vector decompose to calculate RMSE. In our case we got RMSE around 0.94 which again illustrate a good model. For K mean clustering used Elbow graph to find best K value for clustering visualization and prediction.

4. Conclusion

Because of the high usages of applications now a days there is a increase demand in building a good recommendation system. In my project I am trying to accomplish different scenarios using filtering and clustering algorithms. Matrix factorization is a good method used to reduce the dimension as there are lot of users don't have any rating for users so it created sparse matrix. So dimension reduction methodology enhance the model prediction. I feel in future work we can use Deep learning algorithm, Bayesian network and combination of content-base and collaborative method can also increase the recommendation capability in better extent. With our analysis we found that collaborative filtering (Model based) perform better than all other algorithms.