

Movie Recommendation System

(Information Retrieval Project Proposal)

Sai Vamshi Dobbali[†]
School of Computing
University of Utah
Salt Lake City, Utah, USA
u1266122@utah.edu

Abishek Krishnan
School of Computing
University of Utah
Salt Lake City, Utah, USA
u1261980@utah.edu

ABSTRACT

We In this project, we attempt to build a hybrid movie recommendation system. We have divided our task into the following sub-problems.

- a) To determine the similarities among movies.
- b) To determine the similarities among users.
- c) To recommend a new movie to the current user, based on viewing/rating history of similar users.

As a stretch goal, we also intend to,

- d) To recommend new movies based on the current user's search query.

KEYWORDS

Movie, recommendation, information, retrieval

ACKNOWLEDGMENTS

We would like to the School of Computing, University of Utah and Prof. Qingyao Ai for giving us this opportunity as part of the CS 6550 Introduction to information retrieval course.

REFERENCES

- [1] <https://grouplens.org/datasets/movielens/100k/>
- [2] <https://www.kaggle.com/netflix-inc/netflix-prize-data>

STRATEGY

We have divided our project on a higher level into two steps:

1. Find similarity between movies by filtering movies based on attributes of the movie like cast, director, genre etc. So that we can recommend similar movies he viewed/rated so far.
2. Find similarity between the users using collaborative filtering and recommend new movies to the active/current user.

DATASET

We are yet to finalize our data, the most likely datasets we will use are a) MovieLens 100K Dataset [1] and/or b) Netflix Prize Data [2].

EVALUATION MATRIX

We propose to evaluate our model, based on the RMSE between the predicted movie rating and the actual rating and also against the top-N ranking results. Based on the progress of our project, we are open to include other metrics too.

CONTRIBUTION

We are proposing to explore the solutions to the sub-problems mentioned using a mixture of techniques like collaborative filtering (Nearest Neighbor Matching, Dimensionality Reduction), Content based filtering and clustering based search. Based on the progress of our project, we are open to include other information retrieval techniques too.