Movie-based Recommendation System to Connect Users

Sumanth V Rao
Computer Science Dept.
PES University
Bangalore, India.
sumanthvrao@gmail.com

Sumedh Basarkod
Computer Science Dept.
PES University
Bangalore, India.
sumedhpb8@gmail.com

Suraj Aralihalli
Computer Science Dept.
PES University
Bangalore, India.
suraj.ara16@gmail.com

***Abstract*—General Movie Recommendation Systems identify the user preferences in movies, to recommend new movies for each user to watch. In this report, we present a novel idea of finding an ideal partner to watch a new movie with. Considering simple demographic details of individuals that comprise of age, gender, occupation, locality of residence along with user's ratings and preferences in movies, the model looks for finding those ideal pairs with similar personalities who can go for a recommended movie together.**

***Index Terms*—Recommendation System, Collaborative Filtering, Content-based Filtering, Knowledge-based Filtering, Clustering.**

I. Introduction

There has been substantial decline in the figures of people visiting theatres as a result of multiple reasons New York Film Academy Article. The prime reasons are lack of identifying the right movies that matches users interests. The already existing recommendation systems do a reasonable job in recommending similar movies based on one's interests, but fail in other aspects. These mainly include offering the right company for the user to watch a movie with. We address these problems, offering solutions to improve the overall situation. This indicates the need for further improvements to make recommendation systems more effective. Our model achieves this by pairing up the individuals to watch a recommended movie together. In order to do this, the model considers various other factors that include lot more than similar preferences in movies. Each movie has its own story-line that can interest only a subset of people and is not sagacious to gauge the movie on a common criteria. The model aims at alleviating this situation by identifying the right subset of people to whom this movie should be recommended. Our focus hence shifts from the paradigm of suggesting movies to suggesting similar users.

II. Prior Work

A. Assumptions

The entire model of identifying the ideal pairs is contingent on few basic assumptions that revolve around the conjecture that Movies play a significant role in defining the personality of the individuals who watch them. If we classify users based on general traits then teens or indignant individuals are expected to prefer action, adventure over fantasy or drama. Similarly young, bold individuals are likely to prefer romance, mystery genres over sorrow, tragedy. Users can be grouped into clusters based on various parameters like age, occupation that contribute significantly to user's personality and behaviour, which in turn play a significant role in their preferences for different genres.There are numerous research experiments, most of them documented vouch for this argument.

B. Example

Shauna M. Bowes classifies human personalities into multiple overlapping categories based on wide range criteria in his research work [1]. Psychopathy and entertainment preferences: Clarifying the role of abnormal and normal personality in music and movie interests. Each category is further divided into different sub-groups.Gist of these subgroups include Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, Openness, Narcissistic Personality, Leadership-Authority, Entitlement-Exploitativeness, Fearless Dominance, Self-centered Impulsivity, ColdHeartedness etc.It turns out that age and gender were correlated with psychopathic traits and entertainment preferences. Fearless Dominance showed small to medium positive association with entertainment genres like War and Western, Comic and Science-Fiction, Horror and Thriller, and Documentary and Biographica. However, there was a significant correlation between Fearless Dominance and War and Western genre over other genres. As expected Cold heartedness showed moderate negative correlation with

Romance and Romantic Comedies. This substantiates the assumption that individual's preferences in entertainment reflects his/her traits,behaviour patterns and overall personality to a reasonable extent.
Similar User Clustering [2] based on MovieLens Data Set talks about social clustering, involves clustering similar users belonging to social recommender network that is generated from the dataset of MovieLens but does not mention anything on the lines of connecting these users with regard to age, gender or movie preferences

C. Related Work

Movie Recommended systems can be broadly classified into four important groups [3] Demographic Filtering, Collaborative filtering, Content-based filtering, and Hybrid filtering.

1) **Demographic Filtering** This is the naive way of filtering with limited accuracy. It mainly considers the available demographic details like gender,age,locality of residence,and profession of the users.
2) **Collaborative filtering** takes into account the interest of the user, by considering the interests of other similar users. The intuition behind this approach is that if user A and user B can be considered as having the same interests, it is reasonably to assume user A has also the same opinion about a new item only user B has already an opinion of.
   Instead of solely banking on item similarities, user similarities are also considered to make the recommendations more effective. Collaborative filtering can be further categorized into two sub categories: Memory-based collaborative filtering algorithms and model-based collaborative filtering algorithms.
   Collaborative Filtering considers all the available user-item information to make a prediction. Based on the available data ,it determines the most related users, similar to target user.
   The strategy is to first develop a model of user ratings only.To achieve this it uses numerous machine learning techniques that include clustering Model-Based Collaborative Filtering clustering, rule-based and Bayesian network strategies.
3) **Content-based Filtering** The approach used in this filtering technique is to recommend movies to the user on the descriptions of previously evaluated movies. Hence, it recommends movies because they are similar to the movies that the user has showed interest in the recent past.
4) **Hybrid Filtering** This filtering technique relies on the fact that content-based and collaborative filtering characteristics are almost complementary. Combining the two strategies can result in enhanced performance and reliability.

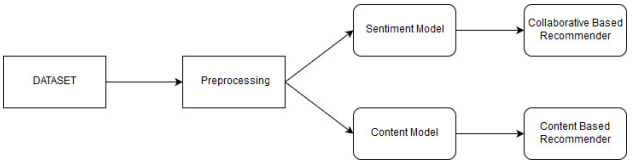

Fig. 1. Basic work flow model

III. Proposed work

A. Data

The dataset under consideration is offered by MovieLens [5]. It contains 100,000 ratings by 943 users on 1682 movies. Each user has rated at least 20 movies, ensuring consistency. The movies have been categorized into 18 genres with each movie falling into multiple categories. The movies also contain the IMDb link to the movie profile giving us further information on its background.

The user demographics play a key role in our case. Roughly 3/10 individuals are females and the average age of individuals is in 20's. Additional information about the users include their occupation and zip code which helps us factor in distance as a metric. The major data includes the ratings given by user for a movie. In addition, information about release date and time stamp give us ample information.

B. Approach

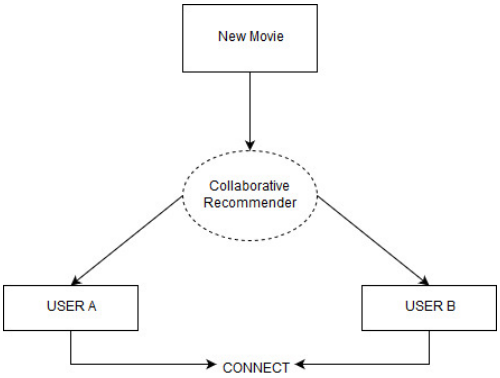

Fig. 2. Connecting users through the recommender

Most of the recommendation systems only look at recommending movies based on user preferences and talk nothing about connecting similar users together in any means. Our model aims at bringing individuals of similar entertainment preferences together to form an ideal pair who can go out to watch a movie together.User demographic details like age, gender, profession and location are known.

In contrast to general recommendation systems, this model aims at considering numerous other insights of the movie.The dataset offers link to the movie description, So additional information related to the movie like the cast-crew, box-office

collection, plot can be used to improve the overall accuracy of the model.

REFERENCES

[1] S.M.Bowes, A.L.Watts, T.H.Costello, B.A.Murphy , Scott.O.Lilienfeld. " Clarifying the role of abnormal and normal personality in music and movie interest ", 27 February 2018.

[2] L. Haesung, K. Joonhee. "Similar User Clustering based on MovieLens Data Set ", Advanced Science and Technology Letters, Vol.51, pp.32-35, CES-CUBE 2014.

[3] Steven Postmus. "Recommender system techniques applied to Netflix movie data", Vrije Universiteit Amsterdam, February 2018.

[4] C.Christakou, A.Stafylopatis. "A hybrid movie recommender system based on neural networks", ISDA'05, 8-10 Sept, 2005.

[5] MovieLens Dataset. "https://grouplens.org/datasets/movielens/100k/".