A Project Report

On

**Personalizes movie recommendations using collaborative filtering**

**-** Unlocking Personalized Viewing Experiences

**PRESENTED BY:**

TUMMA LAHARIKA

YERRA DURGAPRASAD

SEELOJU DIVYA

VIGNANA BHARATHI INSTITUTE OF TECHNOLOGY

HYDERABAD,INDIA.

15/04/2024

# ABSTRACT

Recommendation systems indeed play a crucial role in filtering vast amounts of data to provide users with personalized information for decision-making. At the heart of recommendation systems lies the concept of similarity among objects, particularly items being recommended. Collaborative filtering is a prominent recommendation technique that relies on the ratings or preferences provided by previous users. By analyzing these ratings and identifying similarities between users, the system can effectively recommend items to users based on the preferences of similar users. In essence, collaborative filtering harnesses the collective wisdom of users to make recommendations. By understanding the preferences of users who have similar tastes or behaviors, the system can suggest relevant items to the current user, aiding in their decision-making process. Collaborative filtering, which analyzes user preferences and behaviors to make recommendations, has been supplemented with content-based methods that consider the characteristics of items being recommended. Hybrid approaches combine both collaborative and content-based techniques for improved accuracy. Matrix factorization, user-based, and item-based recommendation are discussed as specific techniques within collaborative filtering. These methods involve analyzing user behavior and item characteristics to make personalized recommendations. Furthermore, content-based filtering, which converts text into vectors for recommendations and detects similarity between vectors, is highlighted as a key aspect of recommendation systems. This content-based approach has revolutionized the way users discover items on various platforms. Overall, recommendation systems play a crucial role in various applications such as OTT platforms, search engines, and music streaming services by predicting or filtering selections based on user preferences. The focus on movie recommendation systems underscores the importance of providing personalized and relevant suggestions to users in today's digital landscape.

# I. INTRODUCTION

Recommendation systems indeed play a crucial role in filtering vast amounts of data to provide users with personalized information for decision-making. At the heart of recommendation systems lies the concept of similarity among objects, particularly items being recommended. Collaborative filtering is a prominent recommendation technique that relies on the ratings or preferences provided by previous users. By analyzing these ratings and identifying similarities between users, the system can effectively recommend items to users based on the preferences of similar users.

In essence, collaborative filtering harnesses the collective wisdom of users to make recommendations. By understanding the preferences of users who have similar tastes or behaviors, the system can suggest relevant items to the current user, aiding in their decision-making process.

Recommendation systems are indispensable components of many online services, contributing significantly to user satisfaction and increased usage of the service. In the realm of movie recommendation systems, major players like Netflix and Amazon have leveraged these systems to engage customers by providing them with highly relevant content.

Netflix, for instance, utilizes recommendation systems to suggest movies or TV shows that users haven't yet watched but are likely to enjoy based on their viewing history and preferences. Amazon, on the other hand, relies on a combination of item ratings, customer buying patterns, and search history to tailor recommendations to individual users. This approach allows the system to recommend content more effectively by targeting users with similar tastes and preferences.

By identifying similarities between users based on their ratings and genre preferences, the system can recommend content that is more likely to resonate with individual users. Overall, this contributes to the advancement of movie recommendation systems by refining the process of identifying similar users and delivering personalized recommendations. By understanding user behavior and preferences, recommendation systems can better serve users and drive engagement with the platform.

**1.1 Problem statement:**

The movie recommendation system will be built using artificial algorithms that analyze user's favorite genres and recommend movies according to their liking. The response will be based on the liking of the user. The User will submit queries depending on their liking of their movies. The System analyses the liking and then recommends the user movies. Providing related content out of relevant and irrelevant collection of items to users of online service providers.

Our aims to recommend movies to users based on content of items rather than other user’s opinions. The main issues with the method for suggesting movies include false user reviews, an exaggerated rating, and irate consumers. As an output, take into account not just a rating choice but also an "explanation" of why the user would enjoy the movie. A recommendation systems objective is to offer users the best possible solution.

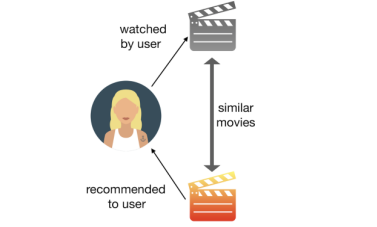
**1.2 Objectives:**

* To increase the impact of movie review endorsement. To make it easier for people to manage their time.
* To provide consumers with a pleasant user experience that enables them to use flexible.

**II. TYPES OF RECOMMENDATION SYSTEM**

There are many ways of recommending movies using Content-based, Collaborative (User-item, User-user), context-based, hybrid methods, and nowadays deep learning is also used to solve this problem. Different terminology used in implementation of movie recommender system is discussed below.

1. **Content-based Filtering:**

This approach for recommending movies does not involve other users. Based on what we like, our algorithm will pick similar items i.e. items having similar content and recommend us. In this approach, the diversity in recommendations will be the least as it only takes into consideration what the user specifically likes.

E.g., A user that says they like Action movies will only be recommended other action movies until they try some other genre autonomously and decide to give it a like. Of course, there are many categories we can calculate the similarity on: as in our case of movies, we can decide to find similarity based on genre, keyword, cast, dire.

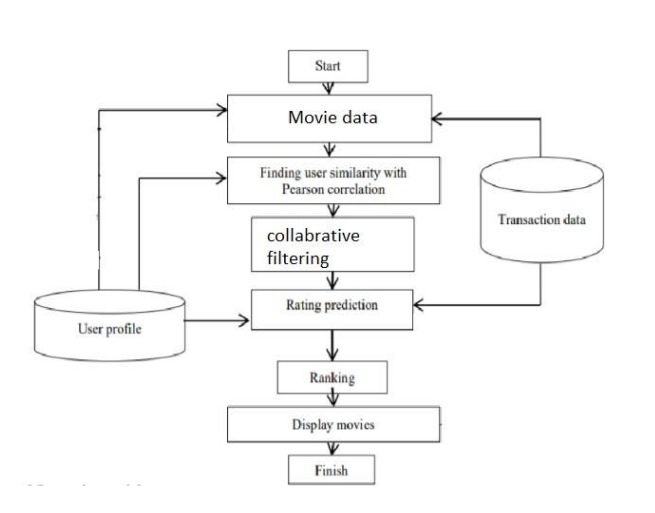
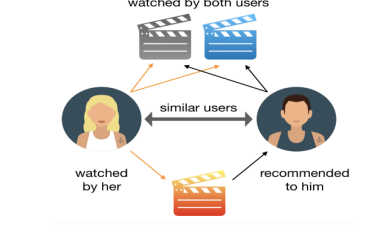


Figure 1: collaborative filtering

1. **Collaborative Filtering Recommendation System:**

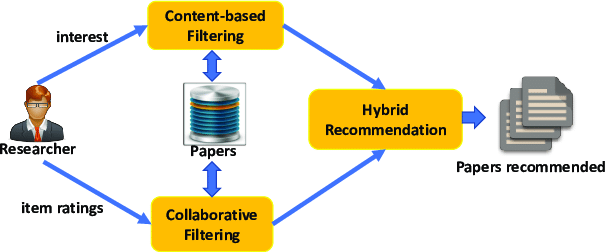
The previous approach didn’t involve other users and in so it had some shortcomings. Such limitations involve the recommendations not being diverse as discussed before. To solve such problems, we use the collaborative filtering technique. This approach is based on the idea that the user rates, and the system will recommend different movies that the user has not watched but the other users similar to our test user have watched and liked. This type of collaborative filtering approach is called the User-to-User Collaborative filtering approach as we find similar users to our user.

To determine whether the two users are similar or not, we consider the movies watched by both of them and how they rated them. Thus, by looking at items in common, we will predict the ratings a user will give to a movie who hasn’t watched it yet, based on its similar user rates

****

**C. Hybrid recommender system (HRS)-approaches:**

A hybrid recommendation system is a special type of recommendation system which can be considered as the combination of the content and collaborative filtering method. Combining collaborative and content-based filtering together may help in overcoming the shortcoming we are facing at using them separately and also can be more effective in some cases. Hybrid recommender system approaches can be implemented in various ways like by using content and collaborative-based methods to generate predictions separately and then combining the prediction or we can just add the capabilities of collaborative-based methods to a content-based approach (and vice versa).



There are several studies that compare the performance of the conventional approaches with the hybrid methods and say that by using the hybrid methods we can generate more accurate recommendations.

**Types of collaborative filtering:**

Collaborative filtering can be categorized into two main types: user-based and item-based.

* **User-Based Collaborative Filtering**:

User-based collaborative filtering recommends items to a user based on the preferences of similar users. It identifies users with similar tastes and suggests items liked by those users that the target user hasn't interacted with yet. This approach relies on calculating user similarity, selecting a group of similar users (neighborhood), and predicting ratings for items not yet rated by the target user based on the ratings of similar users. Finally, it recommends items with the highest predicted ratings to the target user

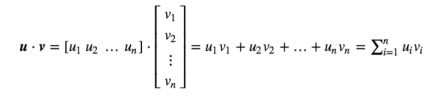
* **Item-Based Collaborative Filtering**:

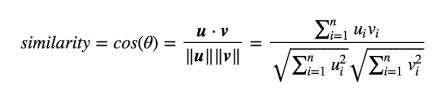
In item-based collaborative filtering, recommendations are made based on the similarity between items. The system identifies items that are similar to the ones the user has liked or interacted with and recommends those similar items. This approach assumes that if a user has liked a particular item, they are likely to also like items that are similar to it.

**III. METHODOLOGY**

In the present system the method used is Item k-NN with Item Recommendation and Item Rating Prediction.

* **COSINE SIMILARITY:**

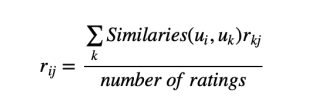
To find similar content for our item, we used the cosine similarity algorithm. The dot product between two vectors is equal to the projection of one of them on the other. Therefore, the dot product of two identical vectors is equal to their squared modules. On the other hand, if the two vectors do not share any directions, the product will be zero. General formula for calculating dot product is given below:

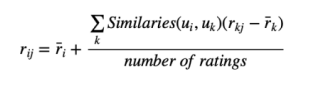
This dot product is important when defining the similarity as it is directly connected to it. The definition of similarity between two vectors u and v is in fact the ratio between their dot products and product of their magnitudes.

Thus, this will be equal to 1 if the two vectors are identical or it will be 0 if the two are orthogonal.

* **K Nearest Neighbors:**

The standard method of Collaborative Filtering is known as Nearest Neighborhood algorithm. We have an n × m matrix of ratings, with user uᵢ, i = 1, ...n and item pⱼ, j=1, …m. Now we want to predict the rating rᵢⱼ if target user i did not watch/rate an item j. The process is to calculate the similarities between target user i and all other users, select the top X similar users, and take the weighted average of ratings from these X users with similarities as weights



However, not all users have the same baseline for giving ratings to movies. Some users may tend to give high scores generally while some are pretty strict with their ratings even though they are satisfied with the items. To avoid such bias, we will subtract each user’s average ratings of all the items when computing weighted average, and add it back for the target user as shown:

* **Matrix Factorization:**

Sparsity is a big issue that needs to be addressed while creating collaborative filtering recommendation systems. Our approach creates matrices where rows are unique users in our environment and the columns represent different movies and the values within are the ratings that different users give to movies. However, it is rather obvious that not all movies will be rated by each user. Thus, this matrix of ours faces the problem of sparsity that needs to be solved. For this purpose, we use Matrix Factorization. In this method, we decompose the original sparse matrix to low-dimensional matrices with latent features. Therefore, matrix factorization gives us how much a user is aligned with a set of latent features, and how much a movie fits into this set of latent features. The advantage of this approach over the previous algorithm is that even though two users haven’t rated same movies, it is still possible to find out the similarity between them if they share similar latent features.

**Implementation Steps:**

1. User Rates movies.
2. A matrix with user, movies and the respective ratings is created.
3. Sparsity of the matrix is dealt with by using matrix factorization.
4. Similar users to our target user are found using similarity algorithms such as Pearson Correlation or Cosine similarity.
5. Ratings for a movie by the target user is predicted by comparing them with other similar users.
6. Latent features can also be compared and used for making recommendations.
7. Such predictions are passed through neural network embeddings to find out the probability of ratings.

**IV. RESULT ANALYSIS**

In order to demonstrate the proposed method, implementation of collaborative filtering item recommendation and collaborative filtering item rating prediction. In this work, User Rating Dataset that contains Rated and Unrated Movies. The rated movies are treated as training set and an Unrated as testing set. In item recommendation, the ranking is given to the different item on the based-on user id and item id. In item rating prediction, the value generated by the system which can be used to predict the rating of different items.

**V. CONCLUSION**

The collaborative filtering has many potential problems but in this proposed the main focus on two main problems i. e., cold start, and data sparsity. The cold start is a situation where a recommender does not have adequate information about a user or an item in order to make relevant predictions. The proposed work can make prediction for the new user who has not rated any movies and new to the system. The data sparsity problem occurs when only a few of the total number of items available in a database are rated by the users. This leads to the sparse user item matrix. The proposed system deals with a large number of sparse data because dataset has huge quantity of unrated movies. The future scope of this work is to recommendation of movies to the different user according user’s choice who rated a movie as well as new user who has not rated any movie in the system.