

# Movies and Books Recommendation System Based on Cosine similarity & K-Nearest Neighbour algorithm

Mohammad Hassaan , Mohammed Usman, Aman Sultan  
Department of Computer Engineering  
School of Engineering and Technology  
D Y Patil , Ambi  
Pune, India  
mdhassaan56@gmail.com, usmanmomin44@gmail.com  
amansultan49@gmail.com

Prof. Rita Kadam  
Department of AI & DS  
School of Engineering and Technology  
D Y Patil , Ambi  
Pune, India  
rita.kadam09@gmail.com

## Abstract:

**A Recommendation System is a filtering program whose primary goal is to predict the “rating” or “preference” of a user towards a domain-specific item. In our case, this domain-specific item are movies and books. Hence the main focus of our recommendation system is to provide a total of five movies and books recommendations to users who searched for a movie or book that they like. This paper focuses on how recommendation system works in real world. The goal of recommender system is to generate meaningful recommendations to a collection of users for items or products that might interest them. The recommendation system is one of the major technologies for implementing personalization services.**

## I. INTRODUCTION

Recommendation system recommend an item to which a user prefers by using automatic filtering method. It deals with the detection and delivery of information that the user is likely to find interesting or useful. In this era of technology the Recommendation system is an application that filtered personalized information and gives the way to understand a user’s taste and to suggest appropriate things to them by considering the patterns among their likes and ratings of various things. In the area of both e-commerce and Recommendation system recommend an item to

which a user prefers by using automatic information filtering method. It deals with the detection and delivery of information no e-commerce recommendation system is broadly studied and used to achieve maximum profit and to fulfill the precision marketing goal. With the increase of information, the access of people to useful information is more difficult, Hence the role of recommendation systems have become inevitable.

## II. LITERATURE SURVEY

[1]Shubam Pawar,Pritesh Patne,Shree Jaswal, Priya Ratanghayra, Simran Dadhich used a cosine similarity algorithm for recommending ten movies to users who searched for movie that they like.

[2] Ashrita Kashyapl introduced Movie REC, a recommender system for movie recommendation, which used Blender and CAD tools.

[3] Hrisav Bhowmick have implemented eight different methods for recommending movies. An example of a genre-based recommendation technique was that movies associated with a particular genre were checked first, then based on the scores, recommended. In genre based recommendation, however, there remains a high chance that the recommended movies may not be liked by the target user since the recommendation is based on only genre, not user profile similarity.

Using the Pearson Correlation Coefficient Based recommended system the similarity between users can be easily determined, but it is a long formula-based method that requires a lot of computation and memory.

[4]Snehitha S, Adithya R, Dr. Sujithra M used a K nearest neighbor algorithm and collaborative filtering technique for recommending five books to users who searched for book they like.

[5] Proposed Amazon.com recommendations: an item-to-item collaborative filtering. This approaches solved recommendation problems with the use of an algorithm different from the traditional collaborative filtering algorithm, cluster models algorithm and search-based algorithm methods. The item-to-item collaborative filtering algorithm applied to Amazon’s online shop computations was able to scale independently of the number of customer’s and items in the product catalog

[6]Y.N.Bhagirathi, P.Kiran used KNN algorithm to recommend different books and the details of books like author name, published year,etc and all.

### III. METHODOLOGY

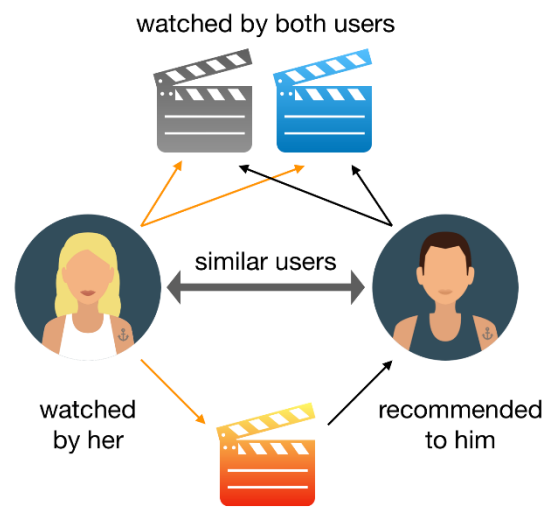
The main purpose of this project is to give a comparison between content based filtering and collaborative filtering techniques, which are used for recommendation system.

The project aims to build a platform that will recommend movies and books to users, provide a poster and name of the searched movies and books. The information provided will surely cut down the time spent in selecting a movie to watch and book to read.

In this project we have used content based filtering technique and collaborative filtering technique for recommending five movies and books on users interest.

#### [1] Content based filtering technique recommendation system:

Content Based recommendation systems recommend an item based upon a description of the item and a profile of the user’s interest. Such systems are used in recommending web pages, TV programs and news articals etc. All content based recommender systems has few things in common like means of description of items, user profiles and techniques to compare profile to items to identify what is the most suitable recommendation for a particular user.



**Fig.1 Content based filtering technique**

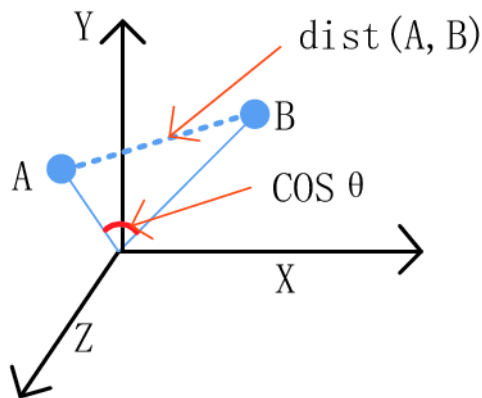
Here for movies recommendation we used cosine similarity algorithm. Cosine similarity is a measure used to determine the similarity between two items. Mathematically it can be determined as the cosine angle between two vectors in a three-dimensional plane. Cosine Similarity is a type of Content-based filtering approach. It is one of the most popular techniques used in recommendation systems. The attributes of a thing are termed as "content". Based on these attributes we are able to classify whether the two things are similar or not. The attributes can be words specified in the database such as genre, cast names, director names, description, and so on. If the attributes match or have a high similarity then the five movies can be classified as similar movies. Many libraries such as scikit-learn, matplotlib has

cosine similarities inbuilt which is of great use. The formula for cosine similarity is

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

**Fig.2 Formula of cosine similarity**

Where **A** and **B** are the vectors we are considering,  $\|\mathbf{A}\|$  and  $\|\mathbf{B}\|$  are their norm (length). The  $A_i$  and  $B_i$  in the formula are the components of each vector. Count vectorizer matrix is used to calculate the occurrence of word in description of movie, after that we have compute the cosine similarity between the different movies.



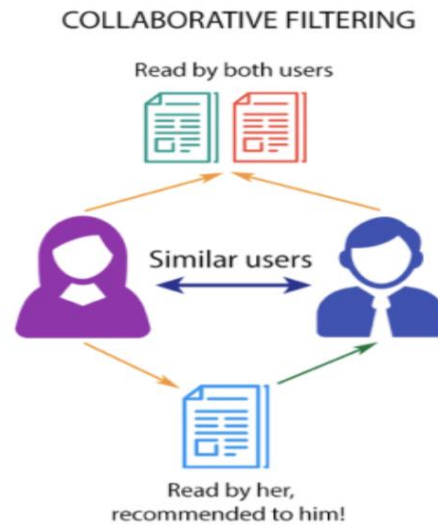
**Fig.3 Graph of Cosine similarity**

**[2] Collaborative filtering based recommendation system:**

Collaborating filtering approaches build the system by considering user’s past behavior (rating is given to those items, previously parched or chosen items) and an additionally similar decision made by different users, then use the system to calculate the item or else rating that the user may perhaps interested. For example, in a movie recommendation application, collaborative filtering system tries to

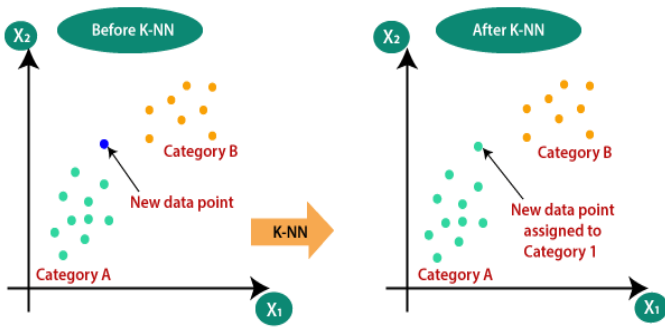
find other like-minded users and then recommends the movies that are most liked by them.

Here for book recommendation we use KNN(K-nearest neighbor) algorithm. KNN is a machine learning algorithm to find clusters of similar users based on common book ratings, and make predictions using the average rating of top-k nearest neighbors . It is a type of collaborative filtering technique . K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.



**Fig.4 Collaborative filtering technique**

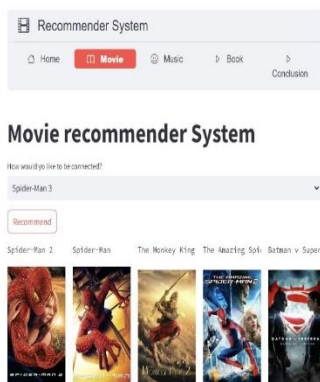
For example, we first present ratings in a matrix, with the matrix having one row for each item (book) and one column for each user. Then the table is converted to a 2D matrix, and fill the missing values with zeros (since distances between rating vectors will be calculated) .Then transformed the values(ratings) of the matrix dataframe into a scipy sparse matrix for more efficient calculations. Before applying the above methods, it is necessary to factorize the dataset in case if we identify the negative values present.



**Fig.5 Graph of working of KNN algorithm**

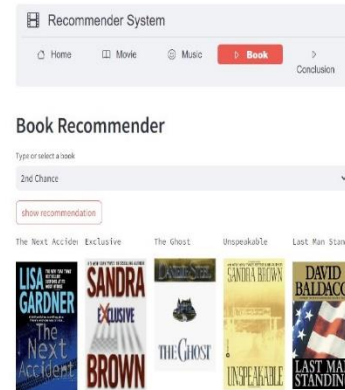
Where it has an  $n \times m$  matrix of ratings, with user  $u_i$ ,  $i = 1, \dots, n$  and item  $p_j$ ,  $j = 1, \dots, m$ . Now we want to predict the rating  $r_{ij}$  if target user  $i$  did not watch/rate an item  $j$ . The process is to calculate the similarities between target user  $i$  and all other users, select the top  $X$  similar users, and take the weighted average of ratings from these  $X$  users with similarities as weights. The next stage is to find other members who are closest to the members who want to find book recommendations and rank them. In this, looking for 5 members who are most similar to members who want to find book recommendations. The results obtained and displayed in tabular form using the Pandas library.

#### IV. RESULTS



**Fig.6 Result of Movie Recommendation System**

In the Fig.6 we can see the five different movies recommended to user on users interest.



**Fig.7 Result of Book Recommendation System**

In the Fig.7 we can see the five different books recommended to user on users interest.

#### V. CONCLUSION

Recommendation systems are software tools and techniques that providing suggestions for items to be of use to a user. In this paper, we try to describe the difference between content based filtering and collaborative filtering techniques used in recommendation systems by using general algorithms of recommendation systems. In this paper we have also describe the working of algorithms in recommendation systems. We are optimistic that in near future research on recommendation systems will witness several new and innovative avenues.

#### VI. REFERENCES

- [1] Shubham Pawar, Pritesh Patne, Priya Ratanghayra, Simran Dadhich, Shree Jaswal have established “A Movies Recommendation System using cosine Similarity”, April- 2022.
- [2] Nayan Verma, Kedaresh Petluri have established “Movie Recommender System”, 2020.
- [3] Prem Kotak, Parth Kotak have established “Movie Recommendation System using Filtering approach”, 2021.

[4] Y.N Bhagirathi, P.Kiran have established “Book Recommendation system using KNN algorithm”, June – 2019.

[5] A. Jayashekar, K. C. Akshay Karan have established Collaborative filtering using Euclidian Distance in Recommendation Engine, 2016

[6]Snehitha S, Adithya R, Dr. Sujitha M have established a “Book Recommendation using K-Nearest Neighbor and collaborative filtering”, Nov – 2022.

[7]Vuong Khuat had established a “Music Recommendation Using Collaborative Filtering”, Dec – 2018.

[8] Sophia Ciocca, “How does Spotify know you so well?”, 2017.