Movie Recommendation System Using Filtering Techniques

(Vaibhav Ghule, Tanmay Ghormade , Sujal Lanjewar, Pallavi Sambhare)
vaibhav.ghule.it@ghrce.raisoni.net, tanmay.ghormade.it@ghrce.raisoni.net,
sujal.lanjewar.it@ghrce.raisoni.net
Information Technology Department
G H Raisoni College of Engineering Nagpur, Maharashtra (India)

Abstract: - In today's world many prestigious applications used recommendation systems. Some of the information filtering algorithms are used to predict preferences as per user requirements. This system saves users time in searching and provides the best user experience. Books, news, music, articles, movies, videos, etc. are the most popular areas where recommendation systems are used. When we take reference from several research papers, one thing is clear Content-Based Filtering Algorithm is used for the recommendation system. In this paper, we work on several algorithms and manipulation of results of this we get a final recommendation list based on ratings like dislike content, related content, etc.

***Keywords*:** -Contest Based Filtering, Movie Recommendation System, Vector Similarity, Text to vector, K NN Algorithm, Cosine Similarity.

Introduction: As we know, till the 21st century lots of information in the form of data are collected on the internet related to any particular topic. This leads to confusion about what one should have to consume or what not to consume. Let's see about YouTube, there is lots of content available related to any topic on YouTube. When we search for any topic, we get several videos. But these videos are ranked as per we had spent more time on which videos and which videos will satisfy user needs.

Suppose you search for something on a website and closed it. But when you visit the next time to same website it will show you suggestions as per the previous search which you might like. In short, this is the work of a recommendation system to suggest the best relevant content for the user. Recommendation system used by many top-most applications like YouTube, Amazon Prime, Flipkart, and Netflix. YouTube used a recommendation system to recommend videos. Flipkart and Amazon used a recommendation system to recommend products. Likewise, amazon prime and Netflix are used for recommending movies. Some of the algorithms are working together in the backend and they tracked user behavior and then the best suitable content to suggest to the user.

Out of all the recommendation systems in this research paper, we will deal with the movie recommendation system. Already several datasets are available on the internet for movie recommendation systems. For example TMDB movie dataset, Netflix has its own dataset. A movie recommendation system is most beneficial for websites like Amazon Prime, and Netflix to increase their revenue as well as user experience. Some filtering techniques are used to filter this dataset for recommendations. These techniques are: -

1. Collaborative Filtering
2. Content-Based Filtering
3. Hybrid Filtering

Existing System: - Memory-based collaborative filtering technique exists. But it has some challenges. To overcome these challenges this model-based collaborative filtering recommendation system has been introduced. Throughout this technique, from available past data systems discover a ratting pattern and provide an accurate recommendation to the user. The existing system is work on rating movies given by users but it could be useless sometime. Because every user is

different, some users like Bollywood movies while some of the users like Hollywood movies, and horror movies. The choice is varying from user to user so sometimes it will not be beneficial to suggest high-rating movies to each user because it might be possible that this movie is not part of that user's interest area.

Propose System: A movie recommendation system can be designed using a collaborative filtering approach, which involves analyzing user behavior to suggest movies based on their preferences. Here's an overview of the system:

Data collection: Collect movie data such as titles, descriptions, genres, actors, directors, ratings, and reviews from various sources such as IMDb, Rotten Tomatoes, and user reviews on social media.

Data pre-processing: Clean and process the data by removing irrelevant or duplicate entries, handling missing values, and standardizing the format of the data.

User profiling: Analyze user behavior such as viewing history, ratings, and reviews to create user profiles. This helps to understand the user's preferences and suggest movies that align with their interests.

Movie profiling: Analyze the attributes of each movie such as genre, director, actors, and rating to create movie profiles. This helps to group similar movies together and recommend movies that are likely to be enjoyed by a user based on their preferences.

Recommendation engine: Use collaborative filtering techniques such as user-based, item-based, or matrix factorization to recommend movies to users. These techniques involve finding similar users or movies and recommending movies that the user has not seen but are likely to enjoy based on their preferences.

User interface: Design an intuitive user interface that allows users to browse recommended movies, view details about the movie, and rate or review the movies they have watched.

Feedback loop: Continuously improve the recommendation engine by collecting feedback from users about the movies they have watched and adjusting the recommendations accordingly.

Deployment: Deploy the system on a web or mobile platform, and monitor its performance regularly to ensure that it is providing accurate and relevant recommendations to users.

By following these steps, a movie recommendation system can be designed that provides personalized recommendations to users based on their preferences, helping them to discover new movies that they are likely to enjoy.
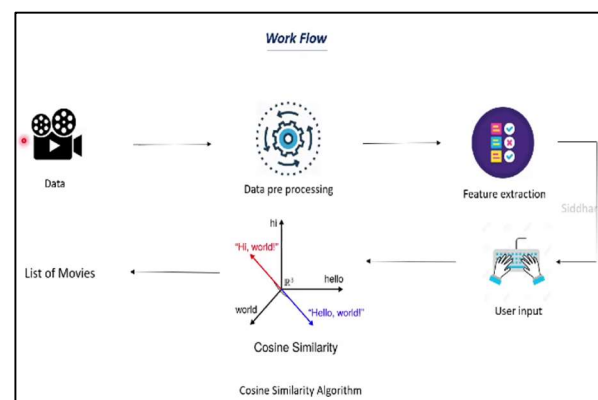
Algorithms Work:

➢ **Collaborative Filtering:** - Collaborative filtering is a popular technique used in movie recommendation systems to provide personalized movie

recommendations to users based on their past movie preferences and ratings. Collaborative filtering is based on the idea that people who have similar tastes and preferences in movies will also have similar ratings for the same movies.

Collaborative filtering works by analyzing the user's ratings and finding other users who have similar tastes and preferences. It then recommends movies that the user has not watched but have been rated highly by other users with similar tastes.

There are two main types of collaborative filtering methods: user-based and item-based. In the user-based collaborative filtering method, the system identifies users with similar ratings and recommends movies that these users have rated highly. In the item-based collaborative filtering method, the system identifies movies that are similar to the movies that the user has already rated highly and recommends these similar movies.

Collaborative filtering is widely used in movie recommendation systems because it provides a personalized and relevant movie recommendation experience to users. However, it does have its limitations, such as the cold-start problem, where new users with no or limited ratings may not receive accurate recommendations, and the sparsity problem, where there may be insufficient data for some movies or users to provide accurate recommendations. To overcome these limitations, hybrid recommendation systems that combine collaborative filtering with

other techniques such as content-based filtering and matrix factorization are often used.

Content-Based Filtering: -

Content-based filtering is another popular technique used in movie recommendation systems to provide personalized movie recommendations to users based on the characteristics of the movies. This technique relies on analyzing the features of the movies that the user has previously watched and rated, such as the genre, actors, directors, and plot.

Content-based filtering works by creating a user profile based on the characteristics of the movies the user has watched and liked. Then, the system recommends other movies that have similar features to the ones the user has liked before.

For example, if a user likes action movies starring Tom Cruise, the system can recommend other action movies starring Tom Cruise or other movies that have similar characteristics, such as other action movies or movies starring similar actors.

Content-based filtering is useful in situations where the user has a limited number of ratings or when the user has a particular preference for a specific type of movie. However, it does have its limitations, such as the inability to recommend new or novel movies that do not have similar features to previously watched movies.

To overcome these limitations, hybrid recommendation systems that combine content-based filtering with other techniques such as collaborative filtering and matrix

factorization are often used. Hybrid systems provide more accurate and diverse recommendations by combining the strengths of different techniques.

Cosine Similarity: -

Cosine similarity is a mathematical technique that is commonly used in movie recommendation systems to determine the similarity between movies or users. In a movie recommendation system, cosine similarity is often used in content-based filtering to calculate the similarity between movies based on their features such as genres, actors, directors, and plot.

Cosine similarity is a measure of the similarity between two vectors in a high-dimensional space. It measures the cosine of the angle between two vectors, where the vectors represent the features of two movies or two users.

Cosine Similarity   Soft Cosine Measure
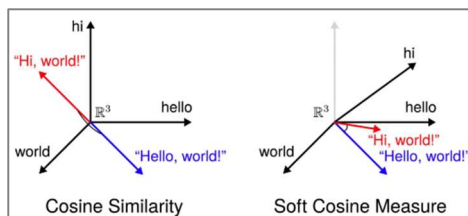
To calculate the cosine similarity between two movies, the system first represents each movie as a vector of its features. Then, it calculates the cosine of the angle between these two vectors. The cosine similarity score ranges from -1 to 1, where a score of 1 indicates that the movies are identical, and a score of -1 indicates that the movies are completely dissimilar.

For example, if two movies have similar features such as the same genre, director, and actors, their cosine similarity score will be high,

indicating that they are similar. On the other hand, if two movies have completely different features, their cosine similarity score will be low, indicating that they are dissimilar.

Cosine similarity is a useful technique in movie recommendation systems because it can help identify movies that have similar features to those that the user has previously watched and rated highly. The system can then recommend these similar movies to the user, increasing the chances that the user will enjoy them.

Overall, cosine similarity plays an important role in content-based filtering by enabling the system to calculate the similarity between movies based on their features and providing accurate and personalized movie recommendations to users.

Future Scope: -

1. The future scope of movie recommendation systems is vast, with many exciting developments on the horizon. Here are some of the future trends and advancements that we can expect to see in movie recommendation systems:

2. Improved personalization: With the advancements in machine learning and artificial intelligence, movie recommendation systems are becoming increasingly personalized. In the future, these systems will use more advanced algorithms and techniques to analyze user data, including social media activity and browsing history, to provide more accurate and personalized recommendations.

3. Integration with virtual assistants: As virtual assistants such as Siri, Alexa, and Google Assistant become more popular, movie recommendation systems will integrate with these assistants to provide personalized movie recommendations based on voice commands.
4. Integration with streaming services: With the rise of streaming services such as Netflix, Amazon Prime, and Hulu, movie recommendation systems will become more integrated with these services to provide seamless and personalized recommendations to users.
5. Use of augmented reality and virtual reality: In the future, movie recommendation systems may integrate with augmented reality and virtual reality technologies to provide users with immersive movie recommendations and experiences.
6. Inclusion of diverse and underrepresented voices: Movie recommendation systems of the future will strive to include and highlight diverse and underrepresented voices in the film industry. These systems will provide recommendations for movies made by and featuring people from diverse backgrounds, helping to promote inclusion and diversity in the film industry.

Overall, the future of movie recommendation systems is exciting, with advancements in technology and data analytics enabling more accurate and personalized recommendations for users.

Conclusion: - In conclusion, a movie recommendation system is an essential application that helps users discover new movies based on their preferences and viewing history. Collaborative filtering and content-based filtering are two popular techniques used in movie recommendation systems to provide personalized recommendations to users. Collaborative filtering is based on the behavior and preferences of other users, while content-based filtering relies on the features of the movies that the user has previously watched and liked. Both techniques have their strengths and limitations, and a hybrid system that combines these techniques can provide more accurate and diverse recommendations.

In addition, cosine similarity is a mathematical technique commonly used in content-based filtering to determine the similarity between movies based on their features. The future scope of movie recommendation systems is vast, with advancements in machine learning and artificial intelligence enabling more personalized and accurate recommendations. Integration with virtual assistants and streaming services, the use of augmented and virtual reality technologies, and inclusion of diverse and underrepresented voices are some of the exciting trends that we can expect to see in movie recommendation systems in the future.

Overall, a movie recommendation system project involves data collection, processing, and analysis to provide accurate and personalized recommendations to users. By implementing different recommendation techniques and taking into account user feedback, a movie recommendation system can continuously improve and provide a better user experience.

References: -

1. Zhang, J.; Wang, Y.; Yuan, Z.; Jin, Q.; "Personalized Real-Time Movie Recommendation System: Practical Prototype and Evaluation", Tsinghua Science And Technology, vol: 25, 2020, pp: 180-191

2. Zhang, R.; Mao, Y.; "Movie Recommendation via Markovian Factorization of Matrix Processes", IEEE, vol: 7, 2019, pp: 13189-13199

3. Pavithra, M.; Sowmiya, S.; Tamilmalar, A.; Raguvaran, S.; "Searching an Optimal Algorithm for Movie Recommendation System", International Research Journal of Engineering and Technology, vol: 6, 2019, pp: 216-221

4. Zhou, T., Kuscsik, Z., Liu, J. G., Medo, M., Wakeling, J. R., & Zhang, Y. C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. Proceedings of the National Academy of Sciences, 107(10), 4511-4515.

5. Movie Recommendation System Using Collaborative filtering, 978-1-5386-65657118/2018 IEEE