Movie Recommendation System

[1]Uma S, [2]Deepika E, [3]Mohana Priya B, [4]Monisha D

*Department Of Information Technology, Panimalar Engineering College

[1]umaokj@gmail.com,[2]deepikaelango03@gmail.com,
[3]rekhamohana2001@gmail.com,[4]monisri31@gmail.com

ABSTRACT Now-a-days people are consuming content in form of movies, series, etc. for entertainment. In this modern era, people always look up to entertainment and in that process, they waste their time in searching for movies. Everyone wants to watch good films that have great content. It takes lot of time to search for a movie they like. Recommendation system comes into play in such situations. It helps to people by recommending movies. This paper develops a Movie Recommendation System to recommend movies based on different parameters. The principal objective of the project is to construct a movie recommendation framework to prescribe pictures to users. There are many algorithms that help to build a recommendation system. Here, the Content-based algorithm has been employed to recommend movies based on the similarity with other films by analysing the content of the movie. To find the similarity, the cosine similarity method has been used. Here, the cosine similarity has been computed by using linear kernel, where the parameters are taken by the result of TF-IDF vectorization. Then the most similar movies are recommended

KEYWORD Recommendation, Movies, Cosine similarity, Films

1.INTRODUCTION

Recommender systems are more popular and increase the production costs for many service providers. Today the world is an over-crowded so that the recommendations are required for recommending products or services. However, recommender systems minimize the transaction costs and improves the quality and decision-making process to users It is applied in various neighbouring areas like information retrieval or human computer interaction (HCI). Movie recommendation system design a big problem since other recommendation systems require fast computation and processing service from service providers and product distributors. To recommend movies, first collects the ratings for users and then recommend the top list of items to the target user. In addition to this, users can check reviews of other users before watching movie. A different recommendation schemes have been presented includes collaborative filtering, content-based recommender system, and hybrid recommender system. However, several issues are raised with users posted reviews. There are 3 typesof recommendation systems

1. Popularity based recommendation engine

2.Content based recommendation engine

3.Collaborative filtering based recommendation engine

Content-Based methods (or cognitive filtering) on the other hand, use information and metadata about the content to find similarities among them, without incorporating user behaviour in any way. Items similar to those ‘accessed ’or ‘searched ’by the user are recommended here. Some approaches

analyze the audio and visual features (video frames, audio clips, movie posters etc.), as in using image and signal processing techniques while some analyze textual features (metadata like plots, subtitles, genre, cast etc.) via Natural Language Processing methods like tf-idf, as in , and word2vec, as in. Content based recommendation engine takes in a movie that a user currently likes as input. Then it analyzes the contents of the movie to find out other movies which have similar content. Then it ranks similar movies according to their similarity scores and recommends the most relevant movies to the user.

2.LITERATURE SURVEY

1. Sang-Min Choi, et. al.--mentioned about the shortcomings of collaborative filtering approach like sparsity problem or the cold-start problem. In order to avoid this issue, the authors have proposed a solution to use category information. The authors have proposed a movie recommendation system which is based on genre correlations. The authors stated that the category information is present for the newly created content. Thus, even if the new content does not have enough ratings or enough views, still it can pop up in the recommendations list with the help of category or genre information. The proposed solution is unbiased over the highly rated most watched content and new content whichis not watched a lot. Hence, even a new movie can be recommended by the recommendation system.

2. Muyeed Ahmed, et. al. --proposed a solution using K-means clustering algorithm. Authors have separated similar users by using clusters. Later, the authors have created a neural network for each cluster for recommendation purpose. The proposed system consists of steps like Data Pre-processing, Principal Component Analysis, Clustering, Data Pre-processing for Neural Network, and Building Neural Network. User rating, user preference, and user consumption ratio have been taken into consideration. After clustering phase, for the purpose of predicting the ratings which the user might give to the unwatched movies, the authors have used neural network. Finally, recommendations are made with the help of predicted high ratings.

3. S. Rajarajeswari, et. al. --discussed about Simple Recommender System, Content-based Recommender System, Collaborative Filtering based Recommender System and finally proposed a solution consisting of Hybrid Recommendation System. The authors have taken into consideration cosine similarity and SVD. Their system gets 30 movie recommendations using cosine similarity. Later, they filter these movies based on SVD and user ratings. The system takes into consideration only the recent movie which the user has watched because the authors have proposed a solution which takes as input only one movie.

4. Jiang Zhang, et. al. --proposed a collaborative filtering approach for movie recommendation and they named their approach as 'Weighted KM-Slope-VU'. The authors divided the users into clusters of similar users with the help of K-means clustering. Later, they selected a virtual opinion leader from each cluster which represents the all the users in that particular cluster. Now, instead of processing complete user-item rating matrix, the authors processed virtual opinion leader-item matrix which is of small size. Later, this smaller matrix is processed by the unique algorithm proposed by the authors. This way, the time taken to get recommendations is reduced.

5. Debashis Das, et. al. --wrote about the different types of recommendation systems and their general information. This was a survey paper on recommendation systems. The authors mentioned about Personalized recommendation systems as well as non-personalized systems. User based collaborative

filtering and item based collaborative filtering was explained with a very good example. The authors have also mentioned about the merits and demerits of different recommendation systems.

6. Md. Akter Hossain, et. al. --proposed NERS which is an acronym for neural engine-based recommender system. The authors have done a successful interaction between 2 datasets carefully. Moreover, the authors stated that the results of their system are better than the existing systems because they have incorporated the usage of general dataset as well as the behaviour-based dataset in their system. The authors have used 3 different estimators in order to evaluate their system against the existing systems.

7. Harper, et. al. --mentioned the details about the movie Lens Dataset in their research paper. This dataset is widely used especially for movie recommendation purpose. There are different versions of dataset available like movie Lens 100K / 1M / 10M / 20M / 25M /1B Dataset. The dataset consists of features like user id, item id / movie id, rating, timestamp, movie title, IMDb URL, release date, etc. along with the movie genre information.

8. V. Subramaniyaswamy, et. al. --have proposed a solution of personalized movie recommendation which uses collaborative filtering technique. Euclidean distance metric has been used in oder to find out the most similar user. The user with least value of Euclidean distance is found. Finally, movie recommendation is based on what that particular user has best rated. The authors have even claimed that the recommendations are varied as per the time so that the system performs better with the changing taste of the user with time.

9. Pavithra, M. et al --designed and implement a movie recommendation system. There are different genres, cultures and languages to choose from in the world of movies. Such a system can suggest a set of movies to users based on their interest, or the popularities of the movies. On an average of one year movie survey 600 movies are released in Hollywood. For streaming movie services like Netflix, recommendation systems are essential for helping users find new movies to enjoy. So far, a decent number of works has been done in this field. But there is always room for renovation.

10. Xi, W. et al. --proposed a novel recommendation algorithm based on Back Propagation (BP) neural network with Attention Mechanism (BPAM). In particular, the BP neural network is utilized tolearn the complex relationship of the target users and their neighbors. Compared with deep neural network, the shallow neural network, i.e., BP neural network, can not only reduce the computational and storage costs, but also prevent the model from over-fitting. In addition, an attention mechanism is designed to capture the global impact on all nearest target users for each user.

2.a. EXISTING SYSTEM

- Recommendation systems are software applications that suggest or recommend movies or product to users.

TYPES

1. CONTENT BASED: It is also known as cognitive filtering. It provides recommendation by comparing representation of content describing an item or a product to representation of the content describing the interest to the user. It is suitable in situation or domains where items are more than users.

2. POPULARTY BASED: It is used to check movies that are in current trend or most popular among the users and it directly recommended it.
3. COLLABORATIVE BASED: It is a family of algorithms where there are multiple ways to calculate rating based on ratings of similar users i.e., previous data collected from other individuals.
4. We are Creating a system in python where user can give name of this favourite movies and based on this input, we are going to recommend certain movies to them.
5. In this we are going to do content based and certain kind of popularity based.

2.b. PROPOSED SYSTEM

With our proposed system we aim at building a system that gives more accurate results. Using Python language, we aim at creating a source code that is also compatible with our GUI. With the help Machine learning library like NumPy and Panda, we are making the system that can do mathematical dimensional array and matrices calculation on its own. In our project we use Python language as the main source code. The database which we are going to use contains Movie information and Ratings and as per that given information the system is going to give a recommendation, our system is going to start giving recommendations to the users which will be the final phase of our system.

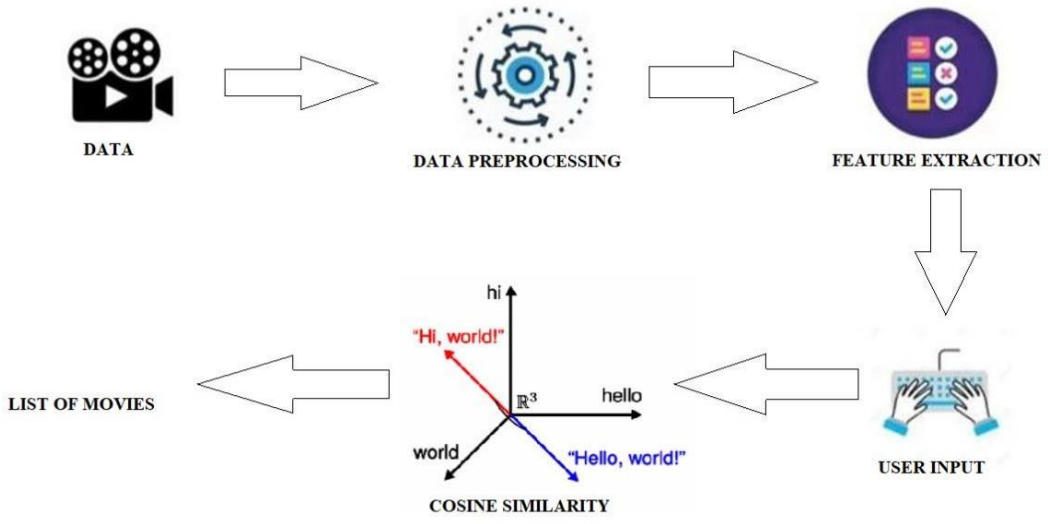

FIGURE 1: ARCHITECTURE DIAGRAM

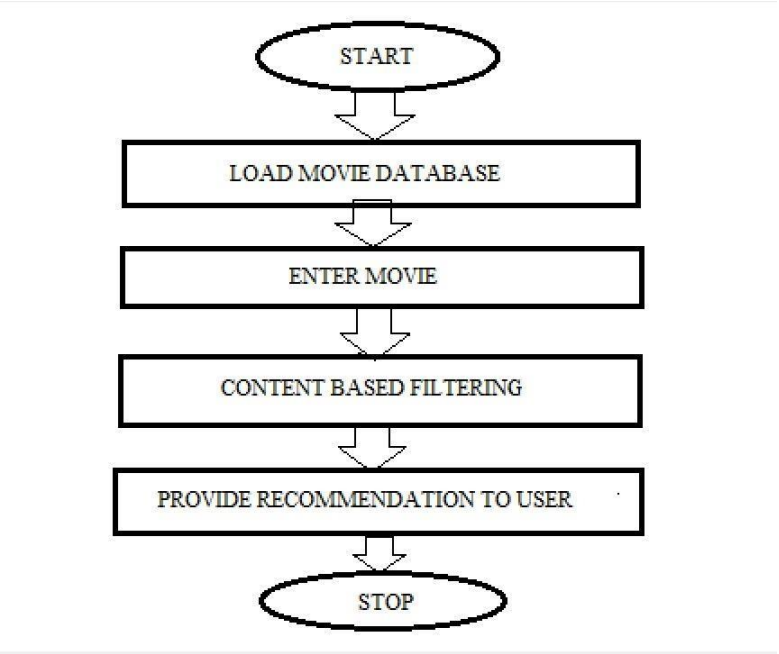

FIGURE 2: FLOW CHART

2.c. MODULES

In this, we plan the outline and execution of the project. There are four modules which we are going to explain:

1. Importing the dependencies.
2. Data collection and Pre-processing.
3. Cosine similarity.
4. Getting the movie name from the user.

1. Importing the dependencies:

The dependencies are nothing but the libraries and functions that we need.We are going to import libraries like numpy,pandas and difflib. Then we import TD-IDF and cosine-similarity.

2. Data collection and Pre-processing:

Data collection:

Collecting data for training the ML model is the basic step in the machine learning pipeline.

Pre-processing:

Data pre-processing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data pre-processing task.

Data pre-processing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

3. Cosine similarity:

Cosine similarity is a metric used to measure how similar two items are. Mathematically it calculates the cosine of the angle between two vectors projected in a multidimensional space.

We need to find similarity of all movies we will do that by cosine similarity it will give some similarity score for all different movies compared to other movies we use this to recommended. Similarity score is numeric value which range from 0 to1.

4. Getting the movie name from user:

We should get the movie name from user. The data filter the movies given by the user and they recommend the similar movies to the user.

3.RESULTS AND DISCUSSION

By using Recommended System, we predicted top 25 movies based on the requirements of the user and the output is shown in Fig.This is to conclude that recommended presented in this paper is very helpful to study the priorities of the customers and recommend other movies similar to their interest.

	index	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity	...	runtime	spoken_languages	status
0	0	237000000	Action Adventure Fantasy Science Fiction	http://www.avatarmovie.com/	19995	culture clash future space war space colony so...	en	Avatar	In the 22nd century, a paraplegic Marine is di...	150.437577	...	162.0	[{"iso_639_1": "en", "name": "English"}, {"iso...	Released
1	1	300000000	Adventure Fantasy Action	http://disney.go.com/disneypictures/pirates/	285	ocean drug abuse exotic island east india trad...	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, ha...	139.082615	...	169.0	[{"iso_639_1": "en", "name": "English"}]	Released
2	2	245000000	Action Adventure Crime	http://www.sonypictures.com/movies/spectre/	206647	spy based on novel secret agent sequel mi6	en	Spectre	A cryptic message from Bond's past sends him o...	107.376788	...	148.0	[{"iso_639_1": "fr", "name": "Fran\u00e7ais"},...	Released
3	3	250000000	Action Crime Drama Thriller	http://www.thedarkknightrises.com/	49026	dc comics crime fighter terrorist secret ident...	en	The Dark Knight Rises	Following the death of District Attorney Harve...	112.312950	...	165.0	[{"iso_639_1": "en", "name": "English"}]	Released

FIGURE 3 MOVIES LIST

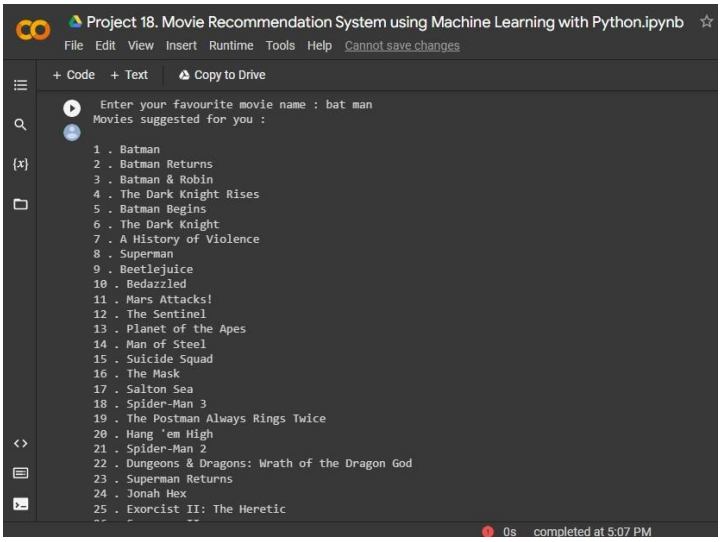

FIGURE 4: SUGGESTED MOVIES

4 .CONCLUSION

In this paper we have designed a technique of content based filtering on a data base of movies. It collects movies from the user known as Data Collection and then pre-processes it. The user enters in the search bar the movie name and gets recommended 25 movies depending on the user's interest. The model's efficiency increases with a better data set and prediction.

REFERENCES

[1] R. Sandeep, S. Sood, and V. Verma, "Twitter sentiment analysis of real-time customer experience feedback for predicting growth of Indian telecom companies," in Proceedings of the 2018 4th International Conference on Computing Sciences (ICCS), pp. 166–174, IEEE, Phagwara, India, August 2018.

[2] Nilashi, M., Ibrahim, O., Bagherifard, K.: 'A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques', Expert Syst. Appl., 2018, 92, pp. 507–520 [3] Tarun Bhatia: Recommendation System: Technology Review, ResearchGate 2018.

[4] Urszula Kuzeleswska: Clustering Algorithms in Hybrid Recommender System on MovieLens Data, SILGR 2014.

[5] Patrikainen, A., Manilla, H.: Subspace clustering of high-dimensional binary data - a probabilistic approach. In: Workshop on Clustering High Dimensional Data and its Applications, SIAM International Conference on Data Mining (2004).

[6] Zan Wang, Xue Yu, Nan Feng, Zhenzua Wang, An improved collaborative movie recommendation system using computational intelligence, Elsevier 2014

[7] T. Chen and Y. H. Chuang, “Fuzzy and nonlinear programming approach for optimizing the performance of ubiquitous hotel recommendation,” Journal of Ambient Intelligence and Humanized Computing, vol. 9, no. 2, pp. 275–284, 2018.

[8] Y. H. Hu, P. J. Lee, K. Chen, J. M. Tarn, and D.-V. Dang, “Hotel recommendation system based on review and context information: a collaborative filtering appro,” in Proceedings of the Pacific Asia Conference on Information Systems PACIS, p. 221, Chiayi City, Taiwan, June-July 2016.

[9] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, Recommender systems survey, Knowledge-Based Systems, 46 (2013) 109-132.

[10] P. Resnick, H.R. Varian, Recommender systems, Communications of the ACM, 40 (1997) 56-58.

[11] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, IEEE Transactions on Knowledge and Data Engineering, 17 (2005) 734-749.

[12] D. Goldberg, D. Nichols, B.M. Oki, D. Terry, using collaborative filtering to weave an information tapestry, Communications of the ACM, 35 (1992) 61-70.

[13] J.B. Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filtering recommender systems, in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.) The Adaptive Web, Springer Berlin Heidelberg2007, pp. 291-324.

[14] M. Pazzani, A framework for collaborative, content-based and demographic filtering, Artificial Intelligence Review, 13 (1999) 393-408.

[15] A. Bellogin, I. Cantador, F. Diez, P. Castells, E. Chavarriaga, An empirical comparison of social, collaborative filtering, and hybrid recommenders, ACM Transactions on Intelligent Systems and Technology (TIST), 4 (2013) 1-29.

[16] J. O'Donovan, B. Smyth, Trust in recommender systems, Proceedings of the 10th International Conference on Intelligent User Interfaces, ACM, San Diego, California, USA, 2005, pp. 167-174.

[17] A.K. Dey, G.D. Abowd, D. Salber, A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications, Human-Computer Interaction, 16 (2001) 97-166.

[18] W. Woerndl, M. Brocco, R. Eigner, Context-aware recommender systems in mobile scenarios, International Journal of Information Technology and Web Engineering (IJITWE), 4 (2009) 67-85.

[19] S. Stabb, H. Werther, F. Ricci, A. Zipf, U. Gretzel, D.R. Fesenmaier, C. Paris, C. Knoblock, Intelligent systems for tourism, IEEE Intelligent Systems, 17 (2002) 53-66.

[20] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic, E. Duval, Context-aware recommender systems for learning: a survey and future challenges, IEEE Transactions on Learning Technologies, 5 (2012) 318-335.