International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.108

Movie Recommender system using Sentiment Analysis

r.A.Shobana Devi
Asst.Professor
Data Science and Business Systems
SRM Institute of Science and Technology
Chennai, India
shobanaa3@srmist.edu.in

Phanindhar.T
Student
Data Science and Business Systems
SRM Institute of Science and Technology
Chennai, India
tp5754@srmist.edu.in

Baby Sarojini.T
Student
Data Science and Business Systems
SRM Institute of Science and Technology
Chennai, India
ta9564@srmist.edu.in

Abstract—Recommendation systems are the most important intelligent systems that play a role in providing information to users in today's world. Content-based filtering and collaborative filtering are two previous approaches in recommendation systems (RS). As a result, these approaches have some limitations, such as the requirement of the user's history as they visit. To mitigate the impact of such dependencies, this research paper proposes a hybrid RS, which combines collaborative filtering, content-based filtering, and sentiment analysis of movies. In this research paper, we created are commander system based on the user's sentiment to recommend a movie to the user based on their viewing history.

Keywords—Recommendation Systems, KNN algorithm, Collaborative filtering, Item-based collaborative filtering, Content based filtering.

I. INTRODUCTION

In today's world, the Internet has become a common source of information overload. Recommendation systems are primarily intended to help users make decisions based on their previouschoices.Thesearecommonlyfoundine-commerceapplicationsandknowledgemanagementsystemssuchasentertainment,online shopping, and tourism.

Movie Recommendation Systems assist us in finding our preferred movies while also shortening the time it takes to find them. The first step is to look at the movies we've seen and visited in the past, and then RS will suggest a movie for us to watch. With the increase in the amount of online data, RS are becoming increasingly useful for making decisions in a variety of day-to-day activities. RS are divided into two types: content-based filtering (CBF) and collaborative filtering (CF)(CF).

During the creation and operation of the movie recommender system as created based on the user's sentiment and comments on the specific movie. Our technology will recommend the best movie to the user based on their prior viewing history and rating. The user's sentiment is recorded as favorable and poor. If the user likes the movie, they may give it a Good smiley, and if they don't like it, they can give it a Sad response, and a movie recommendation is offered to the user based on that.

II. RELATED WORK

There are several strategies that have been examined in relation to the recommender system. Some are based on the amount of weighting of the data, while others are based on user interest. There are several algorithms that have previously been created to lower the user's time and difficulty level.

It necessitates great deal of prior knowledge on the fundamentals of the user's rating of the movie. It mostly use movie datasets for assessment and testing. However, the developed algorithm and system are inefficient, but research is underway to tackle this issue and make the system more flawless and precise.

- Collaborative Filtering System
- Content based System
- Hybrid System

A. *Collaborative Filtering System*

The prediction of recommender systems based on the two formulas narrow and general is where collaborative filtering is most frequently applied. The definition of the term" narrow means predictions" now refers to those that must be created or assessed based on automated prediction and assist users in making choices based on the preferences of other users. Let's use an example where Mohan loves a product X and expresses his opinion while Ram also expresses his approval of the product and provides good feedback. In this case, Mohan will receive further input on the other product based on the opinions of others. The majority of data, including financial data and mineral prospecting, uses CF. These are also divided into the following two groups:

- Memory Based approaches
- Model Based approaches

B. *Content-based Filtering*

The goal of CBF, technically known as Content Based Filtering, is to utilise the dataset's features to propose items that are similar to and close to the user's likes and dislikes based on their prior behaviour and ratings of those specific items.

Content-based filtering also makes recommendations for values or films that have already been seen or deleted. It automatically detects our reviews of a certain item for the dataset we have provided and suggests items that are comparable to that item.

Based on the user's interests and the genres they have previously viewed, our algorithm recommends a movie to them. From the enormous dataset, users may rate aspecific movie that they have seen and propose similar content based on their interests.

DOI: rp-9788770040723.108

C. KNN Algorithm

The recommender system will employ this algorithm, which is one of the most significant ones. K closest neighbouris the full name of this method. The closest neighbour algorithm works as if the majority of the things that are close to the item it belongs to will be placed together or fall into the same cluster, where the distance between the item and its neighbours is the smallest. Consider the following example: T item is a member of the cluster B family since it is near to and identical to Bcluster in terms of distance.

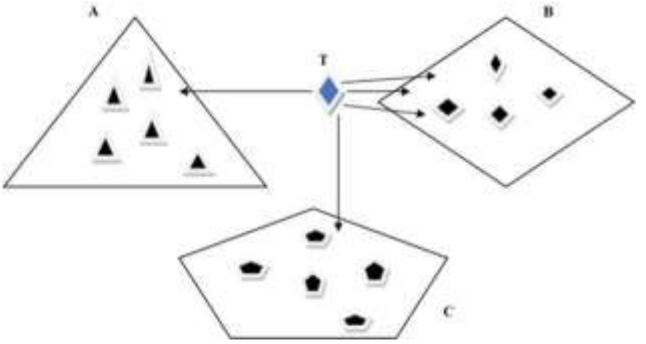


Fig.1.KNNalgorithmExample

D. Collaborating Filtering Algorithm

The primary goal of this algorithm is limited to two tasks: one is project-based, and the other is user-based. However, the user-based algorithm, which is mostly employed for this purpose, is quite successful in meeting the user's need for recommendations.

This algorithm's primary function is to anticipate user needs based on what other users are interested in and then propose related products to the user. Our recommendation engine uses this technique to find movies based on the user's interests.

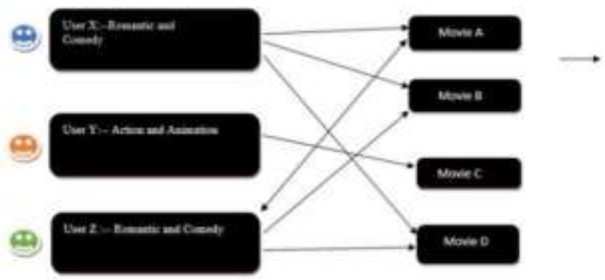


Fig2.User CFAlgorithmExample

Let's look at an example of this algorithm to help you better grasp it.

Assume user X enjoys romantic comedies and tends to watch them more often than other types of movies (A,B,D).

Second, user Z like the same romantic and comedic films, but he just wants to watch movies(B,D).

In this instance, we can see that User X and User Z have parallel interests, yet we can still recommend User Zwatch User X's choice of User A.

The same user Y only enjoys watching the movie C, but user X and Z didn't enjoy it, thus user X and Z won't be recommended it. It is an action and animation movie.

III. Research Methodology

A. KNN Collaborative Filtering Algorithm

We have utilized both the algorithm in the same manner used on the KNN algorithm to determine the neighbour of the item with the smallest distance, which is also known as the collaborative filtering algorithm. The most frequent task accomplished by this algorithm is the creation of a neighbour and recommendation or prediction of the obtained score.

a) Calculating Similarity between Users

The value of an item assessed and projected by two users is calculated to determine how closely the agent and user are related.

For ease of understanding, we have provided an example: To formulate the closeness of X1 and X3,first we must choose our list of movies, which will be labelled as"M1, M2, M4, M5", and then after the parallel scores of these movies. Each user who wants to predict a movie will assignanN-dimvectortoshowtheitemscore.ThefoundscorevectorsforX1andX3are1,3,4,2and2,4,1,5 respectively. The cosine parallel formula is usedtodeterminehowcloselyX1and X3arerelated.

X/M	M1	M2	M3	M4	M5
X1	1	3	3	4	2
X2	3	1	4		
X3	2	4		1	5
X4	2		2		

Fig. 3. users' similarity evaluation

The definition of the correspondence of m or m' is sim (m,m'), and the only formula that is likely to be utilised for the correspondence is cosine similarity.

Cosine parallel is used to determine how similar two users are based on the angle of cosine between them as a vector.

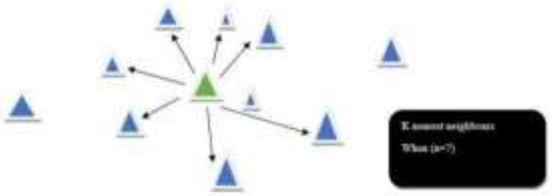
$$sim(x,y') = \cos(\vec{X},\vec{Y}) = \frac{\vec{X} * \vec{Y}}{|\vec{X}| * |\vec{Y}|} = \frac{\sum_{s \in sxy} r_{x,s} r_{y,s}}{\sqrt{\sum_{s \in sxy} (r_{x,s})^2 \sum_{s \in sxy} (r_{y,s})^2}}$$

b) KNN Selection of Nearest Neighbour

Now that the users' similarities have been assessed in the form of sim(u,u'), the KNN algorithm uses this information to determine how many users match U's neighbours, which is denoted by the letters u'. Now we must choose to initialise the K value for the neighbour selection,

DOI: rp-9788770040723.108

which will identify K of the most similar neighbours in the form of a neighbour like value for a user.



c) Predict Score Calculation

Now that we have identified the K nearest neighbours, we must determine the score of the item that is near its neighbour. The primary formula utilised to determine the score prediction is provided as follows:-

$$r_{u,i} = \overline{r_u} + k \sum_{u' \in U} sim(u, u'') * (r_{u',i} - \overline{r_{u'}})$$

$$(k = 1/ \sum |sim(u, u')|)$$

The actions that re listed below will only be utilized forecast the score.

Step 1: Create the user as a 2D score matrix in the RmXN format.

Step 2: Use the Cosine Similarity Formula to help you determine how similar the users are who wish to watch the movie, which will help you create a matrix that reflects what the consumers are seeing**.**

Step 3: In order to discover the N number of the score that indicates the largest amount and will be the same as the K with the neighbours, which is u, we must use the result of step 2 as our starting point.

Step 4: Analyze the value of I for the goal u usingthe predict score algorithm.

In order to forecast movies for the user, KNN collaborative filtering method is used. This project will be used to recommend the movie based on user sentiment based on that as well. Also, it can propose movies, just like Netflix and Amazon Prime Video, which will do so based on the user's prior searches, and it can advise the user's preferences based on their login information to our prototype's server.

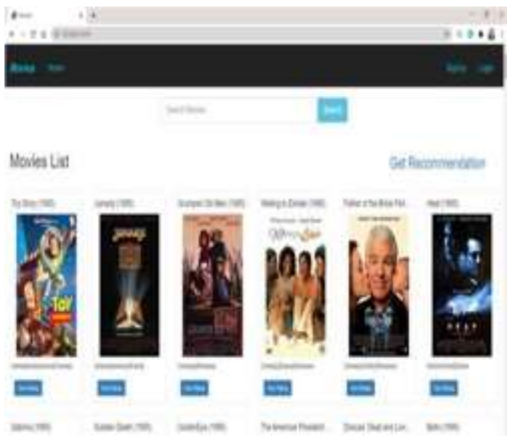
IV. System Designing of the System

A. Architecture Diagram

This architectural diagram demonstrates how our project functions when it is being executed on a server. Based on the user's prior viewing choices, it will suggest a movie to the user.

V. Effects of the Operation on Our System

The user login system that has to gather all the different behavioural traits of the user and save them in the user database will be suggested by our system when it is in operation. Upon a successful login, the system will automatically propose a movie to the user based on their recommendations.



VI. Conclusion

In light of the facts provided, a movie recommender system may greatly benefit us by saving us time and making it easier to find acertain movie that a user requests.This research paper will be created for the movie suggestion based on the user's sentiment analysis, and we have employed several ML algorithms including Collaborative filtering in this article.

KNN algorithm-based algorithm. I have assessed and tested the effectiveness of our system using a sizable dataset that will be used to recommend movies to users, and it performs quite well. This essay will provide suggestions and provide a framework for a movie recommendation system that takes into account user emotion.

DOI: rp-9788770040723.108

REFERENCES

[1] M. Connor,J. Herlocker, "Clustering Items for Collabrative filtering,"InProceedingsofSIGIR2001WorkshoponRecommenderSystems, 2001.

[2] J.Golbeck, Generating redictive Movie Recommendation from Trustin Social Networks,"In:Stolen,k.,Winborough,W.H.,Martinelli, F., Massacci, F. itrust 2006. LCNS, Springer, Heidelberg, vol. 3986,pp. 93-104, 2006.

[3] H. Mohapatra,S. Panda, A.Rath, S, Edalatpanah, andR. Kumar,"Atutorial on Powers hell pipeline and its loop holes," International journal of Emerging Trends in Engineering Research, 2020.

[4] X.Xu,"Fromcloudcomputingtocloudmanufacturing,"Roboticsandcomputer-integratedmanufacturing,vol. 28, no. 1,pp. 75-86, 2012.

[5] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabaharan, N., .&Eswaran, T. (2023). Smart microgrid with the internet of things for adequate energy management and analysis. Computers and Electrical Engineering, 106, 108556

[6] U.S.Manikrao,and.V.Prabhakar,"Dynamicselectionofwebserviceswith recommendationsystem,"International conference on next generation web services practices(NWESP'05), August2005,DOI:10.1109/NWESP.2005.32

[7] K.Choi,D.Yoo,G.Kim, and Y. Suh,"A hybrid online-product recommendation system: combining implicit rating-basedcollaborativefilteringandsequentialpatternanalysis,"Electroniccommerceresearchand applications, vol. 11, no. 4,pp. 309-317,2012.

[8] D.Goldberg, D.Nichols,B. M.Oki, and D. Terry, "Using collaborative filtering to weave an information tapes try," Communications of the ACM,vol. 35, no. 12,pp. 61-70, 1992.

[9] Moshika, A., Thirumaran, M., Natarajan, B., Andal, K., Sambasivam, G., &Manoharan, R. (2021).Vulnerability assessment in heterogeneous web environment using probabilistic arithmetic automata. IEEE Access, 9, 74659-74673..

[10] B. Bhatt, P.P. J. Patel, and P.H. J.Gaudani, "A review paper onmachinelearningbasedrecommendationsystem,"Internationaljournal ofengineering development and research, vol. 2, no. 4, pp.3955-3961, 2014.

[11] P.Campadelli,E.Casiraghi,andC.Ceruti,"Neighborhood selection for dimensionality reduction," International conference on image analysis and processing, Cham: Springer, pp. 183-191,September2015.

[12] B.Bhatt,P. P. J.Patel,and P. H. Gaudani,"A review paper on machine learning based recommendation system,"InternationalJournal of Engineering Development and Research,vol. 2, no. 4,pp. 3955-3961,2014.