_____

Movie Recommender System using Collaborative Filtering

Nupur Kalra
Master of Engineering
Department of CSE
Chandigarh University
Gharuan (Mohali), India
nupurkalra2109@gmail.com

Deepak Yadav
Bachelor of Engineering
Department of CSE
Chandigarh University
Gharuan (Mohali), India
deepakyadav.j23@gmail.com

Gourav Bathla
Assistant Professor
Department of CSE
Chandigarh University
Gharuan (Mohali), India
gourav.cse@cumail.in

_Abstract-_Recommender systems have been a crucial research area in late years. It's a tool that provide recommendation which may be helpful to the user to select item of their interest among thousand other items. In this paper we have given a brief description of collaborative and content based filtering. It contains difference between the two and the flaws they contain. It also reviews the literature of recommender system. The intent of the paper is to study the working of collaborative filtering method using film-trust dataset. The results obtained present a list of recommendation.

Keywords- _Recommender system, issues, collaborative filtering method ,movie recommender system._

_____*********_____

Recommender system uses the opinions of different age group users to help individuals more efficiently to find the item of their interest from a set of choices. There is a wide range of applications for recommendation system. They have become increasingly popular in last decades and are now used in every sector of online services that we approach to. These platforms are which contains movies ,music, books, social media etc. for example facebook ,it start recommending you pages similar to those you have liked recently or if u join a new organization it will start recommending you people in the same place.[1]

Two main approaches widely used for recommender systems. One is content based filtering where we try to recommend things according to his past history of purchase or like dislikes. Other one is collaborative based filtering where we try to group similar users and further use their information to make recommendations.

ther types on the basis of approach are demographic filtering, hybrid recommender system , knowledge based recommender system. In demographic we don't need information like reviews or rating it simply recommend by views of items. They are obtained by machine learning techniques. Hybrid systems are more efficient and overcome the issues faced in other techniques.[2] It's a combination of other approaches. Collaborative and content -based method are mostly used in hybrid recommender system. Knowledge based systems work either on explicit recommendation rules or based on the common characteristics of user needs and products.

Recommender systems became an independent research area in the mid 90's. and have attracted multiple disciplines ,such as mathematics ,physics and computer science. The increasing popularity of social media has greatly enriched people's social life. Real life applications like product recommender in amazon , movie recommender in Netflix and amazon prime, twitter, facebook, linkedinetc[3]

1 .Collaborative filtering based recommender system

Its one of the most popular techniques used to build RS. It works by predicting unexpected behavior and patterns from user's past history or behavior, without any domain knowledge . Methods these days used are categorized as [4]

1.1 Memory based collaborative filtering

This method uses the user matrix or a sample to generate predictions. It is further divided as-:

1. User oriented methods

2. Item oriented method

2 .Content based recommender system

Content based RS work on the principle of information retrieval and processing it to find patterns .They recommend similar items to the user as they have opted in the past ie it uses your past history to further recommend products to the user[5]. Most of the recommendation system works with information in text form such as news ,books and documents. The content in this system usually uses keywords and in formativeness of a keyword is measured by TFIDF weight[6]. It denotes the frequency of occurrence of the keyword in the document where as IDF is inverse document frequency of the keyword.

_____

_____

They have limitations such as-:

1. Limited content analysis

2. Over specialization

3. New user problem

These two systems can work on similar techniques. Mostly in memory based collaborative filtering techniques use rating matrix to store user item database to give recommendations. They are used to refer to those users whose ratings are same as other users or items.

3. Difference between collaborative and content -based methods

 A key advantage of collaborative filtering is here is no need to maintain the profile of the user or item .Lets take an example of an ecommerce organization ,it will create products on the basis of latest trends and emerging needs among their users. they will also create some products which are not in trend but are preferred by some users ie it leverages the behavior of user with similar interest. [7] Where as in content -based , a radio station prefers to play content of music on which they have higher positive feedback. User feedback refines the play list or content . user like or dislike in certain choice of music emphasized by some other attributes.[12]

4. Issues in collaborative filtering

It suffers issues such as sparsity ,scalability and cold start problem.[8]

4.1 Cold start – it appears in the early stages of a system or an item when it does not have much reviews or ratings. A limited amount of information available for a product creates limitations in the expansion and extraction for a system. It appears both in collaborative and content -based method.[9]

4.2Scalability and big data – as the no of users and items are rapidly increasing so the load on the system is also increasing simultaneously which lower the accuracy of recommendations. Large datasets acquires more time to process ,to extract information which compromises the quality ,processing time of the recommender system. Accuracy congestions ,changing datasets are considerate part of it.[10]

4.3 Sparsity – its very common problem observed ,when a user do not prefer to give feedback or like dislike any item. This creates a difficulty in finding nearest neighbor or similar user, according to which the system makes recommendation. It start making random recommendations which may lower the interest of user.

4.4 Privacy- Datasets used in processing contain various attributes of a user in form of information. There is risk of leakage of some confidential information present Or this information can be modified to use in a inappropriate way.[11][12]
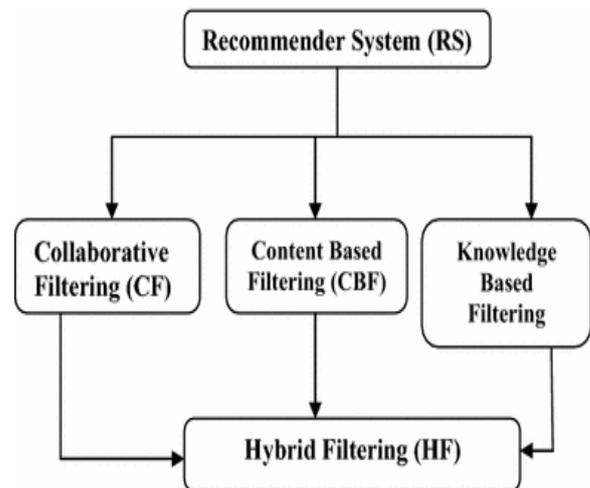

Fig: Types of Recommender System

5. Comparison table of different approaches

Sr. No.	Author name and Title	Approach used	Techniques and algorithm	Advantages	Disadvantages
1	Sonali R Gandhi, Prof Jaydeep Gheewala "A survey on RS with collaborative filtering using big data"[4]	Collaborative filtering approach	Association rule mining	Good scalability and strength	Sparsity and cold start.
2	Md. Tayeb himel , Md. Nazim Hussain " weight based movie RS using K-means algorithm"[8]	Collaborative and content based filtering	k-means algorithm, weighting and matching algorithm	Better results for smaller sets High	scalability

_____

_____

3	A.K.Pandey, Dharamveer Rajpoot "resolving cold start in recommendation Using demographic approach"[9]	Hybrid Model based approach	Decision stump, bayes ,naïvebayes and rule decision table	Improves cold start problem	Black ship and scalability
4	Rahul Katarya, O.P Verma " An effective collaborative movie recommender system with cuckoo search" [10]	Collaborative filtering approach	k-means cuckoo , MAE, RMSE,SD	performance in reliability , accuracy	Initial partition Decrease efficiency a bit.
5	Taeryong Jeon, Jaewo Cho " a movie rating prediction of user propensity analysis based on collaborative"[12]	Collaborative filtering and fuzzy system	RMSE	Fuzzy tools improves all over performance	-
6	Jesus Bobadilla ,Fernando Ortega "improving collaborative filtering RS using genetic algorithm"[15]	Collaborative filtering	Genetic algorithm, MAE	Improves speed of recommendations	cost effective

6. Movie recommender system

Movies are a part of our daily life and movie recommender system help us to stream movies online with help of many filters as per our mood, interest. It helps us to stream through many movies available as per our choice hence decreasing the wastage of time.[1][7].

It's a very reliable approach helps us to find movies. we desire to watch or those which are similar to those issues like big data, sparsity, privacy, cold start .There is a huge scope in this area. As we see movies and users online both are increasing at huge pace so we found Collaborative filtering method used in movie recommender system also suffers scalability and performance issues with poor recommendations we have seen.

6.1 Dataset

We used the film –trust dataset for our experimentation. The data set consist user id ,movie id and ratings given to them. The dataset contains approx. 2100 user who has given around 35000 ratings to the movies they have seen. Dataset is in numeric form. The ratings range from 1-5 on the scale of 5 at the difference of .5. each user is represented by an id no other information is provided. No demographic knowledge or any other kind of data is shared. We split the dataset in ratio of 4:1 that is 80% of data is for training purpose and rest for experimenting. Larger datasets shows issues like scalability, time consumption

6.2 Implementation

6.2.1 Global average- it serve as the simplest technique used as simple base line technique. It serves as predictor for all the missing entries.

6.2.2 User average- every user has its own rating behavior. Some are linnet in their ratings where as some have strict approach giving lower ratings. This bias nature is to be considered. We computed the average rating for movie

given by different users. This method is expected to perform better than the global average method on account of ratings.

6.2.3 Movie average- some movies are rated highly because of genuine reasons and similarly some are

rated low due to there quality. It also performs better than global average[1].

6.2.4 Adjusted average- this method tries to use information about the user i and movie j when making a prediction for entry rij.[1]

6.2.5 Average rating-when a single item is rated by multiple persons so its average rating is calculated and recommended as per the calculated average.

6.3 Method

Movie recommender predicts movie for a particular user lets say user id 123. It will find the movies ssen by user 123 and users who have seen same movies. These users are called similar users, movies seen by them other than those which are seen by user 123 will be extracted out in form of a list. The respective movies will be sorted by using average rating. The movie list will act as a recommendation for user 123. Number of Recommendation can be given to the system for example it can be 10 or 20 movies.
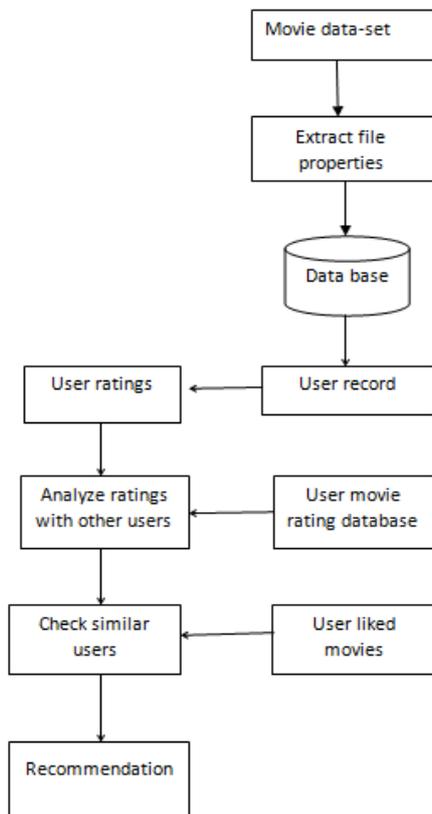
_____

Fig: Flow-chart of movie recommender system

7. Results

Step 1: Input the id of user you want recommendations for

Step 2: Number of recommendation to be given to the user

Step 3: list of recommended movies

8. Conclusion

8. Conclusion- several techniques are proposed for recommender systems are available today based on content based and, collaborative technique. Collaborative filtering technique is the most widely used technique and is able to handle scalability issue. Its provides better recommendation for the users. In this paper it covers collaborative filtering approach on film trust movie data set to provide recommendation of movies for a particular user.

9. References-

[1] Sappadla, P., sadhwani,Y., arora,P.," movie recommender system", search engine architecture, spring 2017, nyu courant.

[2] Herlocker ,J., Konstan, J., Traven,L., Riedl, J., " evaluating collaborative filtering recommender system " ACM transactions on information system , vol 22, no. 1, January 2004.

[3] Bathla, G., Aggarwal,H., Rani, R.," A graph based Model to improve social trust and influence for social recommendations" Springer 2017.

[4] Gandhi ,S., Gheewala, J.," a survey on recommendation system with collaborative filtering using big data", international conference on innovative mechanisms for industry applications (ICIMIA 2017).

_____

_____

[5] Yera ,R.," fuzzy tools in recommender system-a survey", International Journal of Computational Intelligence Systems, Vol. 10 (2017) 776–803.

[6] Jain ,P., Aggarwal, S., " an improved approach for movie recommendation system" (I-SMAC 2017).

[7] Kumar, A ., Sodera ,N., " open problems in recommender systems diversity", international conference on computing, communication and automation(ICCCA2017).

[8] Hamel,T., Hossain ,A.," weight based movie recommendation system using k means algorithm" IEEE 2017.

[9] Pandey,AK., Rajpoot,DS.,"Resolving cold-start problem in recommendation system using demographic approach" International coference on signal processing and communication(ICSPC2016).

[10] Miler,B,N., Grouplens research group, university of Minnesota, Minneapolis, IUT 2003, ACM, Miami ,florida.

[11] Katarya ,R., Verma,O.P.," an effective collaborative movie recommender system with cuckoo search" Full length article, Egyptian informatics journal, science direct,2016.

[12] Tang,J., ,Hu,X., Liu,H.,"social recommendation : A review",full length article.

[13] Bobadilla ,J., Ortega,F., Hernando,A., " Improving collaborative filtering recommender system results and performance using generic algorithm" Elsevier, june 2011.

[14] Balabanovic ,M., Shoham,Y.," content-based, collaborative recommendation" communication of ACM, march 1997,vol 40,no. 3.

[15] Nagarnaik,P., Thomas ,A.," survey on recommendation system" IEEE sponsored 2nd international conference on ECS ,(ICECS 2015).

[16] Jeon, J., Cho, J., Lee, S., " a movie rating prediction system of user propensity analysis based on collaborative filtering and fuzzy system " FUZZ-IEEE 2009, Korea.

[17] Pal, A,.Parhi, P., Aggarwal , M.," an improved content based collaborative filtering algorithm for movie recommendations" 10th conference of contemporary computing (IC3),IEEE 2017.

[18] Peng, X., Liangshan, S., " Improved collaborative filtering algorithm in research and applications of personalised movie recommendation sysem" IEEE, 2013.

[19] Zhou, T., Chen, L., "Movie recommendation system employing the user based CF in cloud computing",IEEE 2017.

[20] Peng, X., Liangshan, S., " Improved collaborative filtering algorithm in research and applications of personalised movie recommendation sysem" IEEE, 2013.

[21] Zhou, T., Chen, L., "Movie recommendation system employing the user based CF in cloud computing",IEEE2017.

[22] Figure1:https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjY3-

TkqsnfAhULQo8KHViaDw4QjB16BAgBEAQ&url=https%3A%2F%2Flink.springer.com%2Farticle%2F10.1007%2Fs10586-017-1616-7&psig=AOvVaw0m3Kf0IpzovDEi2RlwksKc&ust=1546320414151966.

_____