A Review Paper On Collaborative Filtering Based Moive Recommedation System

Nirav Raval, Vijayshri Khedkar

Abstract: The recommendation system plays an essential role in the modern era and used by many prestigious applications. The recommendation system has made the collection of apps, creating a global village, and growth for abundant information. This paper represents the overview of Approaches and techniques generated in the Collaborative Filtering based recommendation system [1]. The recommendation system derived into Collaborative Filtering, Content-based, and hybrid-based approaches. This paper classifies collaborative filtering using various approaches like matrix factorization, user-based recommendation, item-based recommendation. This survey also tells the road map for research in this area.

Index Terms: Recommendation Systems, Collaborative filtering, Model-based, Memory-based machine learning, Model-based methods, user-based collaborative filtering, Item-based collaborative filtering, Matrix Factorization

1 INTRODUCTION

The recommendation system is part of routine life where people rely on knowledge for deciding their interests. The collaborative filtering model takes data from a user's previous behavior (i.e., previously purchased items or chose or numerical ratings provided to the items) as well as similar decisions made by other users. After that, different models are used to forecast items (or ratings for items) that the user might have an interest in. Although there are many approached developed in the past. However, search still goes on due to its often used in many applications, which personalize recommendation and deal with a lack of accuracy. These demands throw some challenges. To solve this, many researchers have used algorithms like Alternating Least Squares, Singular Value decomposition, K-Nearest Neighbor algorithm, and Normal predictor algorithm. Collaborative filtering techniques divided into memory-based and model-based methods. Memory-based methods take action only on a user-item rating matrix and can easily be adjusted to use all the ratings before the filtering procedure; thus, its results updated. On the other hand, a model-based system, like a neural network, generates a model that learns from the information of user-item ratings and recommends new items — following shows the detailed description of all the above approaches. The recommender system still requires improvement to develop a better and accurate method. The recommendation system is a sharp system that provides ideas about the item to users that might interest them. In this paper, different approached

.

2 RELATED WORK

Movie recommendations using several techniques have been extensively studied in the past decades. Examples include a recommendation system using the ALS algorithm, a recommendation based on the weighting technique, item similarity-based collaborative filtering.

• *Nirav Raval is currently pursuing a master's degree program in computer science engineering in SIT, Symbiosis International(Deemed University), Pune, India PH-91 9714223597.E-mail:nirav.raval.mtech2018@sitpune.edu.in*

• *Vijayshri Khedkar is currently an Assistant Professor in the computer science and IT department in SIT, Symbiosis International(Deemed University), Pune, India PH- 91 9850827560 E-mail: vijayshri.khedkar@sitpune.edu.in*

with their techniques are mentioned to compare the limitation of each method in a proper manner to provide appropriate future recommendations

These techniques require prior information about the ratings for the movies which are generated by the user. These techniques majorly use movie lens datasets for evaluation purposes. However, these systems are not a bit accurate, and research is ongoing to improve the real-time performance of these systems.

2.1 Design and Implementation of Collaborative Filtering Approach using KNN

Anshu Sang [2] has represented the recommendation system Using the rating and similarity among the two users; the system recommends an item to the user for the decision making. Then separate the movie data set into an unrated and rated sample set with the help of the KNN model. It can recommend the movies to the unseen users via user registration information, and it can create new and not popular movie recommendations according to the film's history and score. The database in this approach is the MYSQL database. The registration system for a user will snap the user's external and internal behavioral characteristics, and these characteristics are stored in the user database via a login module for the user. The below figure.1. Depicts their effective way of approach for a collaborative filtering approach using KNN.

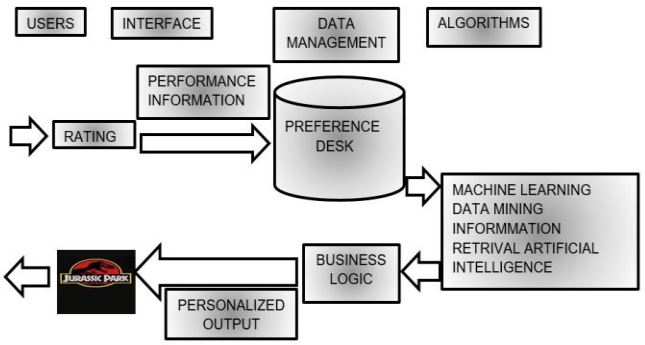

Fig. 1*. The workflow of the recommendation system.*

2.2 Using Alternating Least Squares(ALS)

In recent years, different architectures and models have been developed and used for Movie recommendation via the ALS algorithm. The comparison shows better results with the other state-of-art methods.

2.3 ALS by selecting parameters

By selecting various parameters from the algorithm Mohammed Fadhel Aljunid and D. H. Manjaiah (2019)[3] got better results on the Hadoop platform using KERA's library.

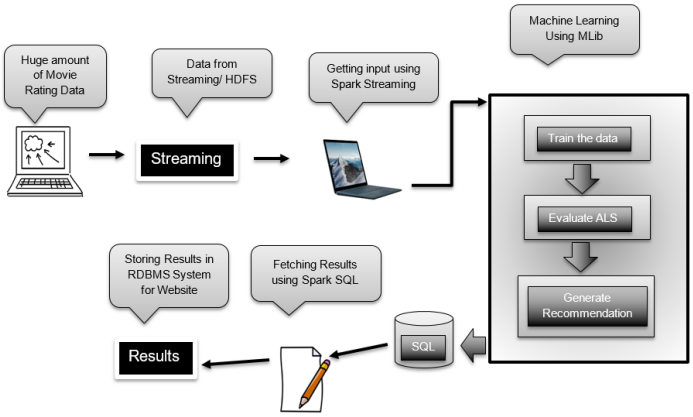

Fig. 2. *Recommendation system using ALS*

However, these systems are not efficient for recommending the movie efficiently to the existing users.

2.4 Comparison with other algorithms.

In [4], Goutham Miryala proposed a comparative study of ALS on other algorithms. However, it is observed that using a more extensive training dataset of 80-20 (Training - Testing) yields a model that has a lower RMSE when compared to the 60-40 (Training - Testing) dataset. The result shows that the higher regularization parameter increases RMSE and vice versa. The ALS algorithm is compared with SVD, KNN, and Normal Predictor, and the results show that ALS is the best algorithm for the recommendation system.

2.5 Weighting Scheme for Collaborative Filtering

Anurag Banerjee in [5] used a weighting technique for the text retrieval system for an item-based collaborative recommender system. Their proposed scheme has been used for effective movie recommendation. The empirical analysis on the benchmark Movie Lens 100K data set has shown improvement over state of the art recommender system algorithms. Also, the performance of the proposed technique needs to be tested on different other applications of the recommender system.

2.6 User-based collaborative filtering

Zhi-Dan Zhao [6] had implemented a user-based CF algorithm on a cloud computing platform, namely Hadoop, to resolve the scalability problem of CF. At first partition, users are essential to get the results accurately. The separation has done via two basic rules. First, a neat arrangement of mapper number to get the better of the starting of mapper and second, partition task uniformly such that all processors finish the job at the scheduled time, can deliver linear speedup.

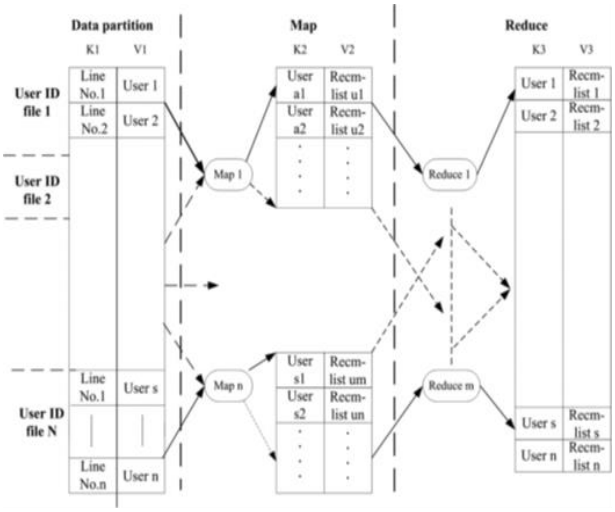

Fig3. *CF's Map Reduce*

Although the main disadvantage of the Map-Reduce framework is that when in the calculation process, whenever a new input file is given, it needs to start a mapper, and this process for some algorithms is very accumulation -absorbing.

2.7 Item-Based collaborative Filtering

Mukesh Kumar Kharita [7] has implemented item based movie recommendation, which uses. For item-based recommendation in the paper, they have used the ratings of those movies that are highly similar to the rating of the movie, which is provided by a proper user from using the item similarity weights in the item similarity weight matrix. Moreover, recommend these movies to the specific user by choosing the K most similar items with higher ratings.

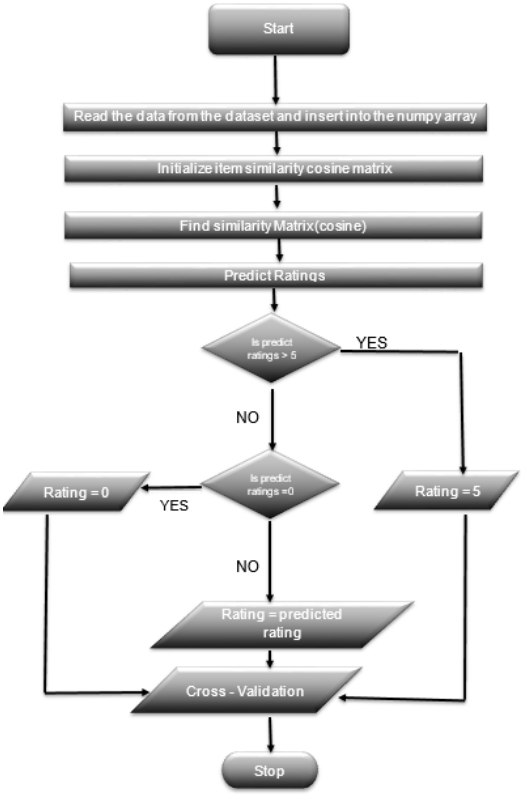

Fig. 4. *Item-based filtering Flowchart*

In the paper, the accuracy of the model is a bit decreased in comparison with contemporary recommendation models. Also, they are using movie rating as the essential element in the whole evaluation process, which can be enhanced in the future by selecting other relevant variables as well.

2.8 collaborative filtering framework using K-means and Cuckoo

Rahul Katarya and Om Prakash Verma [14] came up with the new strategy for the collaborative movie recommendation system. They have used k-means and cuckoo algorithms in order to improve performance. The users have been combined into the clusters, and then the centroid is checked. Users who have lower centroid values are connected more closely. After this stage, the cuckoo algorithms are applied to find the best fitness function based upon the previous and current best solutions. This approach applied to the Movie lens dataset, which contains 100,00 ratings of the users. Furthermore, 943 users have rated around 1682 movies. In the evaluation purpose, it has shown that they have overcome all the state-of-art methods by using standard deviation(SD), Root Mean Square Error(RMSE), Mean Square Error(MSE) and Mean Absolute Error(MAE). The main drawback of this approach is in the initial stage; if the users not clustered well, then it may affect the final evaluation.

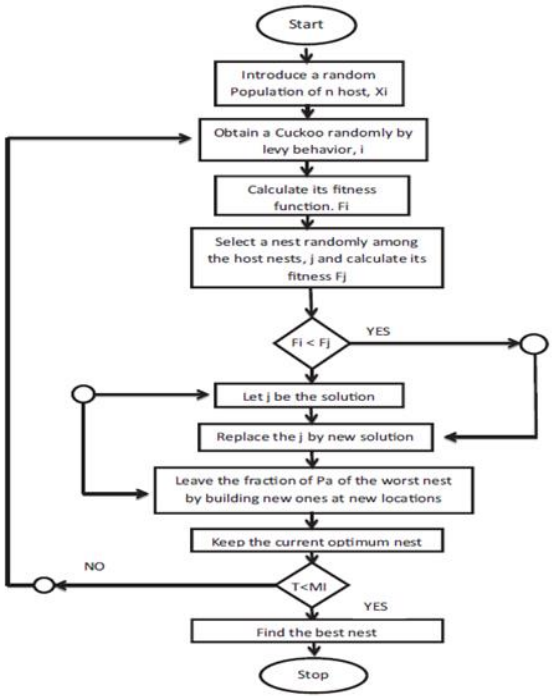

***Fig. 5**. Flowchart depicting the process*

2.9 Recommendation system using k-means and collaborative filtering

This research paper[15] includes the methodology using the k-means algorithm to recommend the movies to the users. Collected movie lens Data then is stored into the clusters. The process divided into three parts User, Movie manager, and system administrator. A comparison can be made by using the Pearson correlation coefficient. Collaborative filtering will calculate the rating for users into the cluster after checking the similarity values in the user-user similarity matrix. K- means algorithm then uses the information via choosing the Euclidean distance and calculate the data for the clustered users. Also, the process will look for a user similarity based on the definition and will create a matrix of data for users in the movie. By using the WEKA software process able to know the number of people in the groups and the centroid dedicated to that group. Fig. 6. Illustrates the working model for the collaborative filtering approach for the recommendation system using the k-means algorithm.

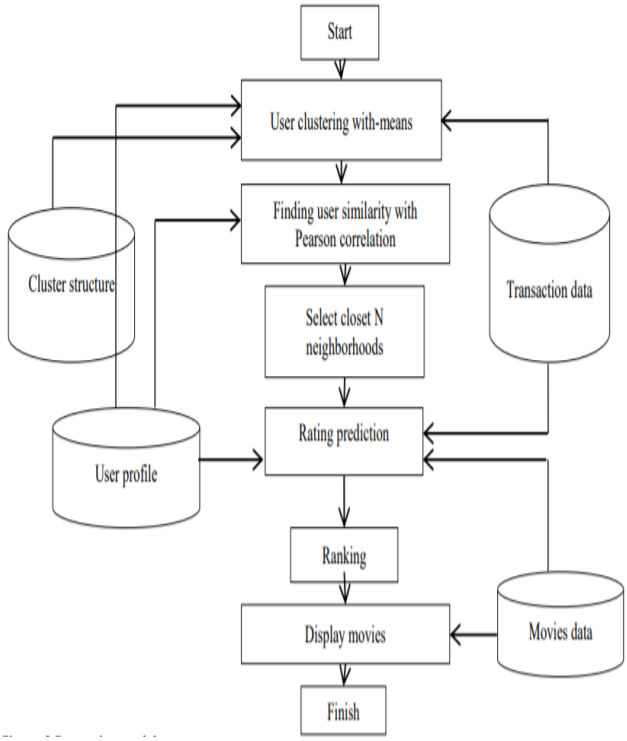

***Fig.6**.Processing model*

2.10 A Recommender System Based on Group Method of Data Handling Neural Network

Meysam Shamshiri[16] proposed a better movie recommendation system that uses a group method for handling the neural networks. In the proposed method, their approach is used to solve the collaborative filtering problem. The network's trust for the users is used to decrease the prediction error of the precise user-oriented collaborative filtering algorithm. The Prediction results of the proposed model are saved in terms of precision and error, and that is compared with many standard algorithms like MLP, Bayesian network. Their main goal was to develop a system that has better accuracy than other models. They have divided the methodology into three parts. Preparation of the data, Pre-processing on data, and GMDH model to give the perfect output for the desired user. The GMDH algorithm, two or more neurons are connected via a polynomial layer in which the subsequent layer is generated. Also, this can be used in modeling mapping inputs to desired outputs. The purpose of this algorithm is to find the unknown coefficients in Voltra function series. For the evaluation purpose, they have used

the Root Mean Square Error (RMSE) and Mean Square Error (MSE) methods, respectively. In the final results, their proposed method outperforms all the state-of-art methods.

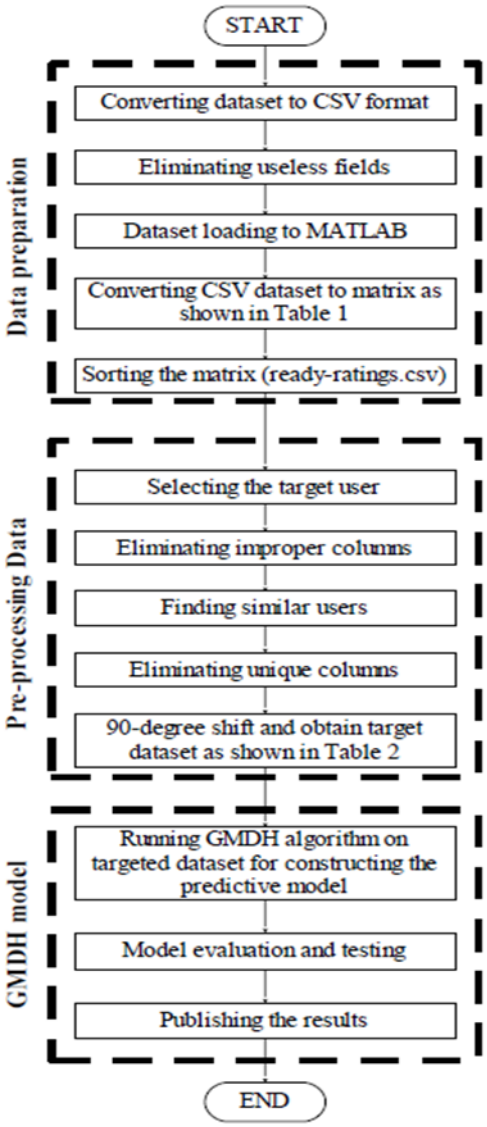

Fig.7. *Proposed method*

2.11 Movie Recommender System using Item-based Collaborative Filtering Technique

In this item-based approach, P.Abhilash [17] implemented its method, which depicts the item-based collaborative filtering approach. They have generated the recommendation based on history. Item-based collaborative filtering technique they identified the User item rating matrix and looked relationships for many items, and then utilized these relationships in terms of computing the recommendations for the user. They have used Netflix user-item database dataset. To evaluate the model Mean Absolute Method used (MAE). The Fig .8 shows the methodology for collaborative filtering.

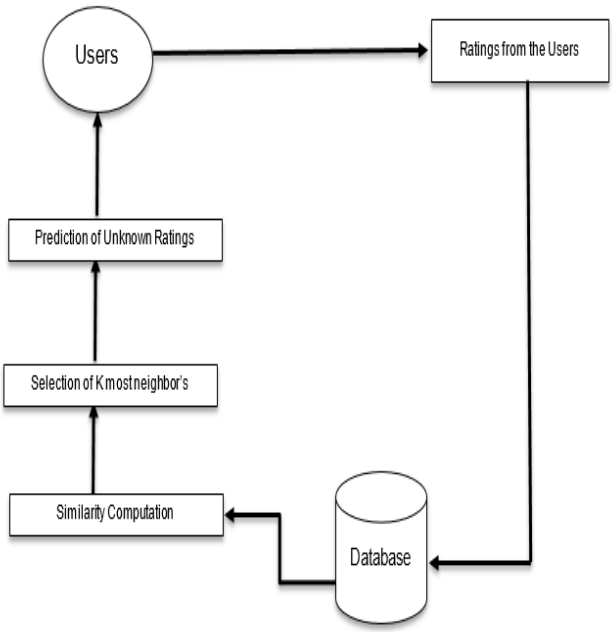

Fig.8.*The architecture of Movie Recommendation*

2.11 Collaborative Filtering for Movie Recommendation using RapidMiner

Arpita Jain [18] has implemented the usual movie recommender system using RapidMiner. The sparsity problem has been neglected by using the rapidminer platform. Also, it permits the persons innovative to convert an accessible introduction to the system and achieve a much-organized arrangement that is working to the datasets. They have used Root Mean Square Method(RMSE) method to evaluate the model for the movie recommendation.

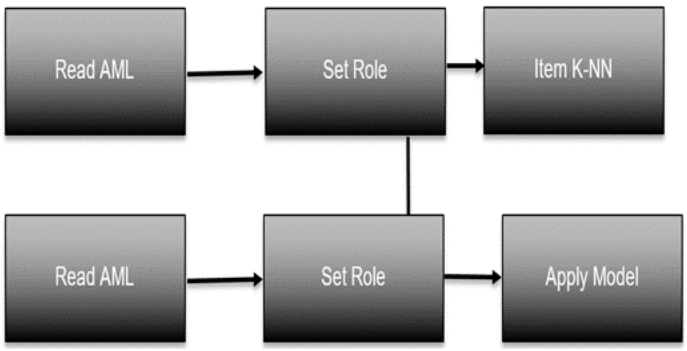

Fig.9*. Workflow for the recommendation*

3. DESIGN METHODOLOGY

Use of Neural network model technique to train a model to learn user-item interaction by addition of user information in user latent factors and item information in item latent factors.
Also, by adding these two random matrices, the results are not accurate, and some loss is there. To remove that, we will use the Gradient descents to fill-up all the gaps and give nearly the same results as per in the given rating metrics. After this, adding the bias terms for two reasons specifically. One,

perhaps, there are certain movies that everybody likes more. Second, probably some users tend to like movies more.

1. Selecting dataset
2. Pre-processing data
3. Split data set (Train - Test)
4. Use Gradient to fill the loss
5. Add Bias for users and movies
6. Train the neural model and test model
7. Experiment the results

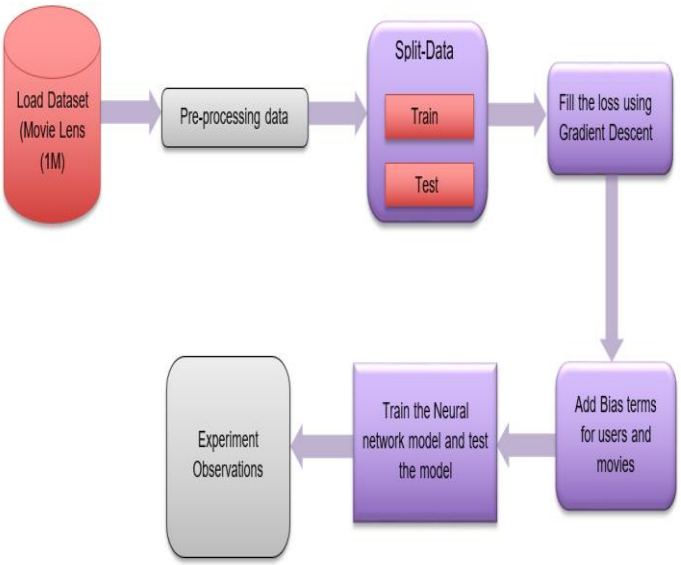

FIG.10. *MOVIE RECOMMENDATION USING PYTORCH*

4. AVAILABLE DATASETS

For the Movie recommendation system, the dataset that is using in the recommendation system majorly is the Movie Lens dataset. It is gathered and handled by Group Lens Organization. Apart from that, most of the experiments are carried on public and standard datasets. Popular datasets Movie Lens, Netflix, and Yelp are the most commonly preferred ones in the operations. The properties of Movie Lens and Netflix datasets are also presented in Table 1.[19]

Name	Domain	Users	Items	Range	Total Ratings
Movie - Lens 100 K	Movie	943	1682	5-star	100,000
Movie - Lens 1 M	Movie	6040	3952	5-star	1,000,209
Movie- Lens 10 M	Movie	71,567	10,681	5-star	10,000,054
Netflix	Movie	408,189	17,770	5-star	100,480,507

Table.1. Most popular Datasets

5. CONCLUSION AND FUTURE SCOPE

This paper includes a summary review of literature studies related to the movie recommendation system based on collaborative filtering. Different approaches, User-based filtering, Item-based filtering, alternating least square methods, KNN method, and for performance measurement of these system Root mean square method (RMSE)[3], Mean Square method(MSE), macro and micro averaged f-measure were used in studies. Each study has its strengths and limitations. In future work, a movie recommendation can improve by using the Pytorch library in which a model would be trained to find the latent (Hidden) factors.

5. REFERENCES

[1] Bobadilla J, Ortega F, Hernando A, Gutiérrez A (2013) Recommender systems survey. Knowl Based Syst 46:109–132. https://doi.org/10.1016/j.knosys.2013.03.012

[2] Cui, Bei-Bei. (2017). Design and Implementation of Movie Recommendation System Based on Knn Collaborative Filtering Algorithm. ITM Web of Conferences. 12. 04008. 10.1051/itmconf/20171204008.

[3] Fadhel Aljunid, Mohammed & D H, Manjaiah. (2018). Movie Recommender System Based on Collaborative Filtering Using Apache Spark. https://doi.org/10.1007/978-981-13-1274-8_22

[4] Miryala, Goutham & Gomes, Rahul & Dayananda, Karanam. (2017). COMPARATIVE ANALYSIS OF MOVIE RECOMMENDATION SYSTEM USING COLLABORATIVE FILTERING IN SPARK ENGINE. Journal of Global Research in Computer Science. 8. 10-14.

[5] Banerjee, Anurag & Basu, Tanmay. (2018). Yet Another Weighting Scheme for Collaborative Filtering Towards Effective Movie Recommendation.

[6] Zhao, Zhi-Dan & Shang, Ming Sheng. (2010). User-Based Collaborative-Filtering Recommendation Algorithms on Hadoop. 3rd International Conference on Knowledge Discovery and Data Mining, WKDD 2010. 478-481. https://doi.org/10.1109/WKDD.2010.54

[7] Kharita, M. K., Kumar, A., & Singh, P. (2018). Item-Based Collaborative Filtering in Movie Recommendation in Real-time. 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC). DOI:10.1109/icsccc.2018.8703362

[8] A. V. Dev and A. Mohan, "Recommendation system for big data applications based on the set similarity of user preferences," 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), Kottayam, 2016, pp. 1-6. DOI: 10.1109/ICNGIS.2016.7854058

[9] Subramaniyaswamy, V., Logesh, R., Chandrashekhar, M., Challa, A. and Vijayakumar, V. (2017) 'A personalized movie recommendation system based on collaborative filtering,' Int. J. High-Performance Computing and Networking, Vol. 10, Nos. 1/2, pp.54–63.

[10] Thakkar, Priyank & Varma (Thakkar), Krunal & Ukani, Vijay & Mankad, Sapan & Tanwar, Sudeep. (2019). Combining User-Based and Item-Based Collaborative Filtering Using Machine Learning: Proceedings of ICTIS 2018, Volume 2. 10.1007/978-981-13-1747-7_17.

[11] Wu, Ching-Seh & Garg, Deepti & Bhandary, Unnathi. (2018). Movie Recommendation System Using Collaborative Filtering. 11-15. 10.1109/ICSESS.2018.8663822.
[12] Verma, J. P., Patel, B., & Patel, A. (2015). Big data analysis: Recommendation system with Hadoop framework. In 2015 IEEE International Conference on Computational Intelligence & Communication Technology (CICT). IEEE.
[13] Zeng, X., et al. (2016). Parallelization of the latent group model for group recommendation algorithm. In IEEE International Conference on Data Science in Cyberspace (DSC). IEEE.
[14] Katarya, R., & Verma, O. P. (2017). An effective collaborative movie recommender system with a cuckoo search. Egyptian Informatics Journal, 18(2), 105–112. DOI:10.1016/j.eij.2016.10.002
[15] Phorasim, P., & Yu, L. (2017). Movies recommendation system
[16] Using collaborative filtering and k-means. DOI:10.19101/IJACR.2017.729004 M Shamshiri, GO Sing, YJ Kumar, International Journal of Computer Information Systems and Industrial Management Applications(2019).
[17] Sri, M. N., Abhilash, P., Avinash, K., Rakesh, S., & Prakash, C. S. (2018). Movie Recommender System using Item-based Collaborative Filtering Technique.
[18] Jain, A., & Vishwakarma, S. K. (2017). Collaborative Filtering for Movie Recommendation using RapidMiner. International Journal of Computer Applications, 169(6), 0975-8887
[19]]GroupLens, Movielens Data, 2019 , http://grouplens.org/datasets/movielens/.