

COLLABORATIVE FILTERING OF MOVIE RECOMMENDATION SYSTEM

N. DURGAHARSHITHA, N. KAVYA, Y. GAYATHRI, N. RAMAKRISHNA, Student,
Department of CSE, NRI INSTITUTE OF TECHNOLOGY, Vijayawada, A.P., India.

Mr.B.VENUGOPAL, Associate Professor, Department of CSE, NRI INSTITUTE OF
TECHNOLOGY, Vijayawada, A.P., India.

ABSTRACT: A recommendation system is a type of information filtering system which attempts to predict the preferences of a user, and make suggests based on these preferences. These have become increasingly popular over the last few years and are now utilized in most online platforms that we use. The content of such platforms varies from movies, music, books and videos, to friends and stories on social media platforms, to products on e-commerce websites. There are different types of recommender systems using association rules. In this project, we are using Collaborative Filtering. It is a Machine learning approach that makes use of the information provided by users, analyzes them and then recommends the movies that is best suited to the user at that time, it is based on cosine similarity using k-nearest neighbours. The recommended movie list is sorted according to the movies that user watching and ratings given to these movies by previous users.

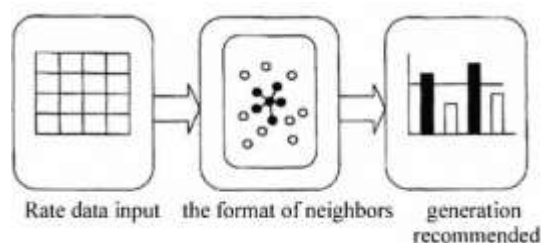
INTRODUCTION

A recommendation system is a type of information filtering system that attempts to

predict the preferences of a user and make suggestions based on these preferences. There are a wide variety of applications for recommendation systems. There are a wide variety of applications for recommendation systems. These have become increasingly popular over the last few years and are now utilized in most online platforms that we use. The content of such platforms varies from movies, music, books and videos, to friends and stories on social media platforms, to products on e-commerce websites. There are different types of recommender systems using association rules. In this project, we are using Collaborative Filtering. It is a Machine learning approach that makes use of the information provided by users, analyzes them and then recommends the movies that is best suited to the user at that time. The recommended movie list is sorted according to the movies that user watching and ratings given to these movies by previous users. In this project we have two data sets. They are :

Movielens dataset : This dataset contains movieId, title, genre.

ratings dataset : All ratings are contained in the file ratings.csv. Each line of this file after the header row represents one rating of one movie by one user, and has the following format: userId,movieId, rating, timestamp.



LITERATURE SURVEY

Zhang, R. et. al. [1] presented a new model family termed Markovian factorization of matrix process (MFMP). On one hand, MFMP models, such as timeSVD++, are capable of capturing the temporal dynamics in the dataset, and on the other hand, they also have clean probabilistic formulations, allowing them to adapt to a wide spectrum of collaborative filtering problems. Two simple example models in this family are introduced for the prediction of movie ratings using time-stamped rating data. The experimental study using MovieLens dataset demonstrates that the two models, although simple and primitive, already have comparable or even better performance than timeSVD++ and a standard tensor factorization model.

Zhang, J.; et al. [2] proposed a high-efficient recommendation algorithm, which exploits users' profile attributes to partition them into several clusters. For each cluster, a virtual

opinion leader is conceived to represent the whole cluster, such that the dimension of the original user item matrix can be significantly reduced, then a Weighted Slope One-VU method is designed and applied to the virtual opinion leader-item matrix to obtain the recommendation results. Compared to traditional clustering based CF recommendation schemes, our method can significantly reduce the time complexity, while achieving comparable recommendation performance. Furthermore, we have constructed a real personalized web-based movie recommendation system, MovieWatch, opened it to the public, collected user feedback on recommendations, and evaluated the feasibility and accuracy of our system based on this real- world data.

Haruna, K. et. al. [3] derived an alternative method based on the concept of co-occurrence from multi-soft sets to handle conflict situations and first used an illustrative example of a movie selection problem to demonstrate the proposed approach and provide an extensive elaboration using a publicly available dataset and provided a new measure based on support, strength, certainty and coverage of soft set on movie selection problem and findings have revealed that the proposed approach achieved less computational time when compared with the rough set-based approach of up to 8.05%.

EXISTING SYSTEM

ASSOCIATION RULES:

Association rule mining is the task of identifying patterns in basket data – transactions that possibly consist of multiple items. The aim of this thesis is to better understand the applications of association rule mining for recommender systems.

CONTENT BASED RECOMMENDER SYSTEM:

A content-based recommender learns a profile of the new user's interests based on the features present, in objects the user has rated. In a content-based recommender system algorithms used are such that it recommends users similar items that the user has liked in the past or is examining currently.

UTILITY BASED RECOMMENDER SYSTEM:

Make suggestions based on a computation of the utility of each object for the user. In utility based system, every industry will have a different technique for arriving at a user specific utility function and applying it to the objects under consideration.

KNOWLEDGE BASED RECOMMENDER SYSTEM:

Knowledge based recommendation works on functional knowledge: They have knowledge about how a particular item meets a particular user need, and therefore reason about the

relationship between a need and a possible recommendation.

DEMOGRAPHIC BASED RECOMMENDER SYSTEM:

This system aims to categorize the users based on attributes and make recommendations based on demographic classes. The benefit of a demographic approach is that it does not require a history of user ratings like that in collaborative and content based recommender system.

DISADVANTAGES:

Limited content analysis: If the content doesn't contain enough information to discriminate the items precisely, the recommendation itself risks being imprecise.

Over-specialization: Content-based filtering provides a limited degree of novelty since it has to match up the features of a user's profile with available items

PROPOSED SYSTEM

COLLABORATIVE FILTERING APPROACH:

It's the most sort after and most mature technologies that is available in the market. Collaborative recommender systems aggregate ratings or recommendations of objects, recognize commonalities between the users on the basis of their ratings, and generate new recommendations based on inter-user

comparisons. Collaborative filtering is based on the assumption that people who agreed in the past will agree in the future and that they will like similar kind of objects as they liked in the past.

Collaborative filtering approach builds a model from a user's past behaviours (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users.

The feedback about movies falls into one of two categories:

Explicit: Users specify how much they liked a particular movie by providing a numerical rating.

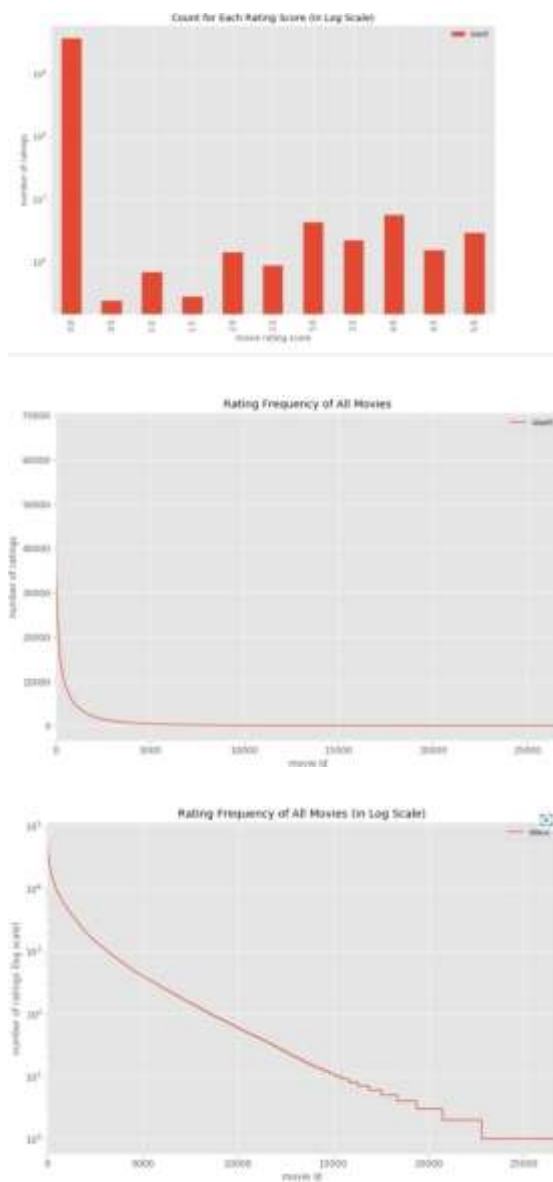
Implicit: If a user watches a movie, the system infers that the user is interested.

ADVANTAGES:

No domain knowledge necessary: We don't need domain knowledge because the embeddings are automatically learned.

Serendipity: The model can help users discover new interests. In isolation, the ML system may not know the user is interested in a given item, but the model might still recommend it because similar users are interested in that item.

SAMPLE RESULTS

[illegible]

CONCLUSION

In this Collaborative filtering is the most successful and popular algorithm in the recommender system's field. It helps customers to make a better decision by recommending interesting items. Even though this algorithm is the best, it suffers from poor accuracy and high running time.

To solve these problems, we proposed a recommendation approach based on user clustering by using the cosine similarity to calculate two users to cluster dataset. This method combines clustering and neighbours vote to generate predictions.

FUTURE SCOPE FOR FURTHER

DEVELOPMENT

Cosine similarity calculation do not work well when we don't have enough rating for movie or when user's rating for some movie is exceptionally either high or low. As an improvement on this project some other methods such as adjusted cosine similarity can be used to compute similarity.

In the future there may be techniques, fuzzy c-means in the group stages of the first system to provide a more effective segmentation.

REFERENCES

- 1.Zhan J, Hsieh CL, Wang IC, Hsu TS, Liao CJ, Wang DW. Privacy-preserving collaborative recommender systems.IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2010; 40(4):472-6.
- 2.Gong S. A collaborative filtering recommendation algorithm based on user clustering and item clustering. Journal of Software. 2010; 5(7):745-52.
- 3.Manvi SS, Nalini N, Bhajantri LB. Recommender system in ubiquitous commerce. In international conference on electronics computer technology 2011 (pp. 434-8).IEEE.
- 4.Pu P, Chen L, Hu R. A user-centric evaluation framework for recommender systems.In proceedings of the fifth ACM conference on recommender systems 2011 (pp. 157-64).ACM.
- 5.Hu R, Pu P. Acceptance issues of personality-based recommender systems. In proceedings of the third ACM conference on recommender systems 2009 (pp. 221-4).ACM.
- 6.Pathak B, Garfinkel R, Gopal RD, Venkatesan R, Yin F. Empirical analysis of

the impact of recommender systems on sales.
Journal of Management Information Systems.
2010; 27(2):159-88.

7.Witten IH, Frank E, Hall MA. Data mining:
practical machine learning tools and
techniques. Morgan Kaufmann Publishers,
Elsevier; 2011.