CS410 Text Information Systems Technology Review – kghate2

Introduction

Due to people changing their roles from consumers to consumers and producers on the internet, there has been an explosion of information and hence, due to high volume of data, there is difficulty in providing the user on the most appropriate search result. Hence, recommender systems are used in various domains on the internet. User likes and dislikes along with preferences can be expressed via reviews or binary attributes like 'like' or 'dislike'. Recommendation consists of rating the item and recommending the item and the most popular is the collaborative filtering technique. One of the other popular is content based filtering [2][3][4]. Others are demographic based [5], knowledge based[6], and community based[6].

Overview

In collaborative filtering, a matrix is used to rank the user preferences accordingly. They also try to address the user item interaction as well. CF approaches are mainly classified into three types i.e., Memory based, Model based and hybrid approach. Memory based has further categories. Category based contains user based and item-based filtering where in user based similar users are grouped together and recommended similar items and item based recommends items similar to the users current item. Hence the system finds similarity and recommends closest neighbor. In model based, there are vectors assigned to items and users and a matrix is created accordingly. Hybrid collaborative filtering combines the above listed methods to output better results.

Challenges

User Cold Start Problem – A new user finds difficulty in being as he is new to the system and has very less to no data feed in it and hence similar users are difficult to find[1].

Item Cold Start Problem – When an item is new to a system, less people have engaged with it and hence it is difficult to link the item to any attribute to rank it[1].

Sparsity – One of the major problems is the number of users that have been exposed to the item[1].

Scalability – The item's credibility decreased as the number of users increases as there is a high chance of different opinions among users[1].

Gray Sheep and Shared Account Problem – Users sharing accounts and some having unusual item tastes can bring discrepancy in the recommendation algorithm[1].

Conclusion

The use of recommender systems to deal with the information explosion is quite beneficial. The task of making recommendations can be accomplished using a variety of approaches. Collaborative filtering is one of the most popular and effective methods among them.

We have also talked about several methods and difficulties associated with CF; while many researchers have offered solutions to these issues, there is still much work to be done before the outcomes are adequate.

References

[1] R. Sharma, D. Gopalani and Y. Meena, "Collaborative filtering-based recommender system: Approaches and research challenges," *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT)*, 2017, pp. 1-6, doi: 10.1109/CIACT.2017.7977363.

[2] K. Lang, "Newsweeder: Learning to filter netnews," in 12th international conference on machine learning, 1995, pp. 331-339.

[3] R. Armstrong, D. Freitag, T. Joachims, and T. Mitchell, "Webwatcher: A learning apprentice for the world wide web," in AAAI Spring symposium on Information gathering from Heterogeneous, distributed environments, 1995, pp. 6-12.

[4] M. Pazzani and D. Billsus, "Learning and revising user profiles: The identification of interesting web sites," Machine learning, vol. 27, pp. 313-331, 1997.

[5] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Recommender Systems Handbook: Springer US, 2011.

[6] R. Burke, Knowledge-Based Recommender Systems: Marcel Dekker, 2000.