Zixiang Li

Professor ChengXiang Zhai

CS 410 Text Information Systems

05 Nov 2021

Collaborative filtering: applications and main challenges

INTRODUCTION

The recommendation system is widely used in various applications such as advertisements, social networking sites, online retails. These recommender systems personalize the application service by predicting what users could like in the future based on the living pattern they have collected from users. recommendation based on collaborative filtering is one of the widely used and effective recommendation strategies in recommendation systems. It began to arise in the 1990s and promoted the development of recommender systems

BODY

Collaborative filtering generally finds out a small part of a large number of users that are similar to your taste. In collaborative filtering, these users become neighbors and then organize them into a sorted directory based on other things they like as recommendations to you. In other words, if many other users who have a similar preference as the current user choose a certain product, the system will predict the current user might also choose the product with a high probability.

Collaborative filtering is divided into three types: user-based collaborative filtering, item-based collaborative filtering, and model-based collaborative filtering. User-based and item-based collaborative filtering are both belong to the memory-based approach.

1) The basic idea of user-based collaborative filtering is to first find a set of users with similar interests to the target user, and then find items in this set that the user likes and has not heard of and recommend to the target user.

2) The basic idea of item-based collaborative filtering is a recommendation strategy based on the evaluation of all users on the recommended objects. If most users have similar ratings for some recommended objects, then the current users might have similar ratings for these recommended objects. Therefore, the products that the user has not evaluated in the similar recommendation objects will be recommended to the user.

3) The basic idea of model-based collaborative filtering is to train a recommendation model based on the preference information of sample users, and then make recommendations based on real-time user preference information. The difference between it and the above two types of collaborative recommendation is that it first applies statistics and machine learning methods to the existing data to obtain a model, and then makes predictions. Commonly used methods include machine learning methods, statistical models, Bayesian models, and linear regression models.

There have many advantages of recommendation based on collaborative filtering.

1) Can be used on complex unstructured objects.

2) Be able to discover new interests and hobbies of users and bring surprises to users.

3) User-centric automatic recommendation, as the number of users increases, the user experience will get better and better.

However, Collaborative filtering is not perfect. It faces many challenges in the real-world scenario.

Synonymy is one of the challenges that need to deal with. Synonyms refer to items that have different names but are very similar. Most recommender systems have difficulty distinguishing the differences between these items, such as baby clothes and baby fabrics. Collaborative filtering usually cannot establish a match between two terms, nor can it calculate the similarity between the two. Automatic term expansion, thesaurus construction, singular value decomposition (SVD), especially latent semantic indexing, can solve the problem of synonymy, but the disadvantage is that some added terms may have different meanings than expected, resulting in a rapid decline in recommendation performance.

Another issue is the cold start problem, that is, in the absence of a large amount of user data, users may be dissatisfied with the recommended results obtained.

Data sparsity can be caused when users rate only a small part of the items available in the database. The larger the data size, the sparser data in general.

Scalability is another challenge that collaborative filter is facing. The calculation usually grows linearly with the number of users and items. When the amount of data set is limited, the recommendation technique is effective and feasible, but when the amount of data set increases, the amount of recommendation generated is not good. In this case, the methods used to solve the scalability problem and accelerate the generation of recommendations will be based on dimensionality reduction techniques, such as singular value decomposition (SVD).

There is a problem of sparseness, that is, as users grow in large numbers, the evaluation differences will become larger and more and more recommended objects, resulting in many recommended objects that have not been evaluated by users, some users cannot obtain recommendation results, and some recommended objects cannot Is recommended.

CONCLUSION

To sum up everything that has been stated so far, the recommender system has found its usefulness in several fields. Among several filtering approaches, collaborative filtering is most popular in building recommender systems, and it has a decent amount of strength. However, there are four main issues suffered by the collaborative filter recommender system which are cold star, Data sparsity, Scalability, Synonymy, and sparseness issues.

CITATION

Singh, Pradeep & Dutta Pramanik, Pijush & Choudhury, Prasenjit. (2020). Collaborative Filtering in Recommender Systems: Technicalities, Challenges, Applications, and Research Trends. 10.1201/9781003007210-8.

Al-bashiri, Hael & Abdulgabber, Mansoor & Romli, Awanis & Hujainah, Fadhl. (2017). Collaborative Filtering Recommender System: Overview and Challenges. Advanced Science Letters. 23. 9045-9049. 10.1166/asl.2017.10020.

Wikimedia Foundation. (2021, October 29). Collaborative filtering. Wikipedia. Retrieved November 8, 2021, from https://en.wikipedia.org/wiki/Collaborative_filtering.