Collaborative filtering: applications and main challenges

Yuanpei Ma

CS 410

November 6, 2022

Collaborative filtering is a method of making recommendations based on the feedback of other users. This approach is used by many recommendation systems, such as those used by Amazon and Netflix, to provide personalized recommendations to users. The basic idea of collaborative filtering is that if two people have similar preferences, then they are likely to like the same things. For example, if two people both like the same movies, then they are likely to have similar preferences for other movies. In the technical review, I would like to discuss how collaborative filtering is applied in a recommendation system and what is the main challenge of collaborative filtering.

In general, collaborative filtering technique can be spliced into two categories: memory-based CF and model-based CF. Under memory-based category, there are two sub-categories: user-based CF and item-based CF.

User-based collaborative filtering is one of the most popular methods in CF. It computes the similarity between users and predicts the rating of a target user by using the ratings of similar users. It is simple and easy to implement. It is a method that is based on the user's rating data, and therefore, it is not affected by the sparsity of the data.

However, it faces the following challenges: First, it is difficult to accurately estimate the similarity between users in a high-dimensional space. Second, the similarity between users is static, and therefore, it is difficult to estimate the similarity between users when a user's preference changes.

Item-based CF is another memory-based CF method. It computes and predicts the rating of a target item by using the ratings of similar items. Item-based CF is

easy to implement and has the same advantages as user-based CF. However, it faces the following challenges: First, it cannot effectively handle sparse and large-scale data. Second, it cannot learn latent factors of users and items. Third, it cannot effectively handle new users and new items.

To solve this problem, model-based CF is involved. Model-based CF is popular because it can deal with sparsity better. Deep neural network can be also introduced to model-based CF to accelerate computation. Since a recommendation system is just a two-way interaction between item features and user preferences, a dual neural network can model the two-way interaction between item and user. The Neural collaborative filtering can replace tradition matrix factorization part in the model-based CF.

Auto-encoder based collaborative filtering is another successful application. It takes user partial vectors or item partial vectors as input to reconstruct the output layer. It is the first auto encoder based collaborative recommendation model in 2019. However, it faces the following challenges: First, it fails to deal with non-integer ratings. Second, the decomposition of partial observed vector increases the sparseness of input data, which will lower the accuracy.

Reference:

Sarwar, B. M., Karypis, G., Konstan, J. A., Riedl, J., & others. (2001). Item-based collaborative filtering recommendation algorithms.

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: a survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 5