Collaborative Filtering Recommendation Algorithm Based on User Preferences

Yongjie Yan, Hui Xie*

School of Mathematics and Computer Science Jiangxi Science and Technology Normal University Nanchang 330038, China

*Corresponding author e-mail: huixie@aliyun.com

Abstract. The personalized recommendation technology provide people an effective method to solve the problem of information overload. Collaborative filtering is one of the key algorithm of recommendation technology. In this paper, the User-Movie (UM) model based on multilayer perceptron algorithm is a kind of item-based collaborative filtering recommendation algorithm. Therefore, the paper puts forward a new method to integrate item similarity in the UM model. Using Jaccard to find the similar items of current item, respectively. Experiments on the well-known datasets Epinions and Movielens show that the algorithm weighted by Jaccard in the case of sparse datasets and less neighbor achieves great improvement of prediction accuracy.

1. Introduction

The rapid development of Internet has brought the information overload problem, finding the content that people want from the vast amounts of information is becoming more and more difficult. In the face of the Internet era of information explosion, the personalized recommendation technology has become a key technology to solve the problem of today's information overload.

Recommendation technology analyzes the interest model from the user's history information. Then, based on the user's interest model, the recommendation system can choose the appropriate items to recommend to the user. This initiative behavior that recommendation system selects items from huge amounts of data for the user, can greatly improve the users’ experience, so the study of recommendation technology has a certain practical significance.

Online service providers and retailers are facing the challenges predicting user preferences and recommending the right, personalized contents that meet their tastes [1, 2]. So recommendation systems becomes an important issue in both research and industry areas. Recommendation systems compares the items and users’ information and find correlations among them, and provide a personalized and ranked items or content to users.

Collaborative filtering recommendation technology is one of the most successful technologies in the recommendation system. It began to research and promote the prosperity of the whole recommendation system research in the 1990s. A large number of papers and research belong to this category.

The basic idea of collaborative filtering is to find other users c_i similar to the current user c (such as similar interests and tastes), calculate the utility value u(c_ij,s) of the object s for the user, and use the utility value to sort or weight all s to find the most suitable object s'. Its basic idea is very easy to...
understand. In daily life, we often use the recommendation of good friends to do one Collaborative filtering is the application of this idea to the recommendation system, that is, to recommend a certain content to the target user based on the evaluation of other users.

In recent years, multiple areas of learning have become a new research direction [3-8], there have been many scholars who began to focus on learning methods to migrate applications to collaborative filtering algorithm. The recommendation system based on collaborative filtering can be said to recommend from the user's point of view, and it is automatic, that is to say, the recommendation obtained by the user is obtained implicitly from the user's purchase or browsing behavior, without the user's initiative to find the recommendation information suitable for their own interests, such as filling in some investigation forms, etc. another advantage is that there is no special need for the recommended object It can deal with unstructured complex objects, such as music, movies, etc. at the same time, it needs a lot of historical data of users' access behavior to study the relationship between users. It has intersection with social network research, rich research basis and broad prospects. Generally speaking, this kind of recommendation algorithm can be divided into two types [9]: heuristic-based or memory-based or memory based methods and model-based methods.

2. Recommendation System Based On Deep Learning

2.1. Classic Collaborative Filtering

A classic Collaborative Filtering (CF) is described in detail in [10]. In the classic CF model, the similarity between two users is calculated using the Pearson correlation over the ratings of their common items. The formula for the Pearson correlation is the following, as stated in [10]:

$$userSim(u, n) = \frac{\sum_{i \in CR_u,u_n} (r_u - \bar{r_u})(r_n - \bar{r_n})}{\sqrt{\sum_{i \in CR_u,u_n} (r_u - \bar{r_u})^2 \sum_{i \in CR_u,u_n} (r_n - \bar{r_n})^2}}$$  

(1)

In the formula, $r$ stands for rating, $u$ denotes the center user and $n$ a neighbor. $CR_{u,n}$ denotes the set of co-rated items between $u$ and $n$. After performing this calculation, we select the top ten most similar users. Next, we rank the articles of these users to recommend to the center user, using the formula of predicted rating for user $u$ with average adjusts described in [10]:

$$pred(u, i) = \bar{r_u} + \frac{\sum_{n \in neighbor(u)} userSim(u, n) \cdot (r_n - \bar{r_n})}{\sum_{n \in neighbor(u)} userSim(u, n)}$$  

(2)

In general, the CF method makes the product of a previously successful search available to other users. A user with low personal experience/knowledge can benefit from the prior experience of others. However, there are a few possible ways the collaborative filtering results can be suggested to the user (e.g., including the user’s own search results vs just the recommended results, no user’s own results), and there are different types of users, with knowledge at different levels and in different subject domains. Which way or combination of ways is effective for what type of user and for which subject domain(s) has not been investigated.

2.2. Example of collaborative filtering

Consider the same user who has rated the movies in the previous section and the following users with their ratings for those movies:
Table 1. Rating matrix as an example

| Movies rated by user1 | Ratings of user2 | Ratings of user3 | Ratings of user4 | Ratings of user5 | Ratings of user6 |
|----------------------|------------------|------------------|------------------|------------------|------------------|
| Movie1               | ?                | 2                | ?                | 3                | 4                |
| Movie2               | ?                | ?                | 3                | 5                | 4                |
| Movie3               | ?                | 1                | 4                | 5                | ?                |
| Movie4               | ?                | ?                | ?                | 1                | ?                |
| Movie5               | 5                | 4                | ?                | 3                | ?                |

The numbers in this matrix is the number of stars that other users, common users, gave to the movies that user1 has rated. A ? means that a common user has not rated a particular movie. The matrix with vectors of the movies that user1 has rated in the training set will become:

Table 2. Rating matrix of common users

| Movies rated by user1 | Ratings of common users | Ratings given by user1 |
|----------------------|-------------------------|------------------------|
|                      | Ratings of user3 | Ratings of user4 | Ratings of user5 | Ratings of user6 |
| Movie1               | 2                | ?                | 3                | 4                | 4                |
| Movie2               | ?                | 3                | 5                | 4                | 4                |
| Movie3               | 1                | 4                | 5                | ?                | 2                |

Notice that the ratings of user2 are not considered in this matrix since user2 has not rated one of the movies that user1 has rated in the training set and the neural networks only trains with the ratings of the training set.

2.3. Deep Learning

Deep learning belongs to a sub field of machine learning. It learns data representation and abstraction at different levels from data, and can be used to solve supervised and unsupervised learning tasks [11]. In this section, we mainly introduce some deep learning technologies that are closely related to the recommendation system field.

Multilayer perceptron (MLP) is a feedforward neural network, which is different from the traditional perceptron with only input layer and output layer. It adds one or more hidden layers between input layer and output layer. Therefore, the multi-layer perceptron can flexibly apply the activation function to make it not only limited to the binary classification problem.

Here is a simple MLP with only one hidden layer as an example in figure 1.
Figure 1. MLP with only one hidden layer

We use 3 dimension vector $X = [x_1, x_2, x_3]$ to represent the input layer and 3 dimension vector $H = [h_1, h_2, h_3]$ to represent the hidden layer. Then, we get

$$H = f(W^T X) = f\left(\sum_{i=1}^{3} W_i^i X_i + b^i\right).$$

Among them, the weight matrix of the first layer is the excitation function and the bias term (bias) of the first layer. Here, the excitation function $f$ is usually nonlinear, and sigmoid function, tanh function, or ReLU function are generally selected. After substitution, the expressions are as follows:

$$f(W^T X) = \text{sigmoid}(W^T X) = \frac{1}{1 + \exp(-W^T X)}$$

$$f(W^T X) = \text{tanh}(W^T X) = \frac{\exp(w^T x) - \exp(-w^T x)}{\exp(w^T x) + \exp(-w^T x)}$$

$$f(W^T X) = \text{ReLU}(W^T X) = \max(0, W^T X)$$

From the hidden layer to the output layer, it can be regarded as a multi class logistic regression, and softmax function is generally selected as the excitation function. The output can be expressed as

$$\text{softmax}\left(\sum_{i=1}^{3} W_i^2 X_i + b^2\right),$$

where $w^2$ is the $3 \times 3$ weight matrix of the second layer, $f$ is the excitation function, and $b^2$ is the bias of the second layer.

To sum up, this three-layer MLP is summed up by the formula:

$$g(X) = G(b^2 + W^2(f(b^1 + W^1 X)))$$

Among them, the function is the softmax function, which maps the output of multiple neurons to the $(0,1)$ interval and is generally used for multi classification. The softmax function is defined as follows:

$$G(X)_j = \frac{e^{x_j}}{\sum_{k=1}^{K} e^{x_k}} \quad \text{for} \quad j = 1, ..., k$$

Here $K$ is the dimension of the input variable $X$. 
3. Neural Networks For Collaborative Filtering Recommendation

In the movie scoring prediction problem, each training sample contains three types of information: user ID, movie ID and scoring data. We need to design a model starting from user ID and movie ID. The correlation between them is found in the training data, and the score data is predicted. In this paper, we use multi-layer perceptron to design the network structure as shown in the figure 2 below to model the relationship between users’ movies.

![Figure 2. User-Movie (UM) model based on multilayer perceptron](image)

Figure 2 shows the first neural network model proposed in this paper. We mark it as um model, where the symbol u represents user and the symbol M represents movie. In the UM model, we accept the user ID and movie ID as input. First, we use the unique heat representation to quantize their numbers. For example, if there are 100 users with the user ID of 1, the user ID will be vectorized into a vector with the length of 100. The value of the first element of the vector is 1, and the value of other elements is 0.

Generally speaking, there are a large number of users and movies in a real recommendation platform, so the unique heat vector has a large dimension, so it is not suitable to be a node in the neural network directly.

So we use embedding technology to map the unique heat vector with higher dimension to the embedded vector with lower dimension. The embedding technology uses the embedding matrix, the number of rows of the embedding matrix corresponds to the number of users, and the number of columns corresponds to the dimension of low-dimensional eigenvector. The embedding technology multiplies the unique heat representation vector and the embedding matrix to get the embedding vector, which is equivalent to taking out the row vector corresponding to the user in the embedding matrix.

\[
Loss = \sum (r - \hat{r})^2
\]  

Here, r represents the real score data, \( \hat{r} \) represents the prediction score data of multi-layer perceptron, and um model needs to minimize the gap between real score and prediction score on all training samples.
4. Experimental Result

To verify the superiority of the improved collaborative recommendation algorithm, we make the following experimental design that compared the traditional collaborative filtering recommendation algorithm.

4.1. Dataset

In the experiment, we used the data of movielens recommendation system. Movielens is a Web-based research recommendation system, which was released in the autumn of 1977. Every week, hundreds of users visit movielens to rate movies and receive recommendations. The movielens dataset contains more than 100000 ratings, more than 940 users, and 1680 movies. In this data set, the user's score is between 1-5, "5" means "very like", "1" means "not like", and the data sparsity is 93.7%. 10000 scores were randomly selected from the movielens data set, which were randomly divided into 70% and 30%, 30% as part of the test set, and the rest 70% as part of the training set.

4.2. Metrics

The accuracy of recommendation system is the most basic index. The ultimate goal of the improved collaborative recommendation algorithm is to improve the accuracy of the results of this paper, so we mainly consider the accuracy of the algorithm. In order to evaluate the recommendation accuracy of the improved recommendation algorithm, root mean square error (RMSE) and mean absolute error (MAE) are used to measure the effect of the recommendation system. Mae and RMSE are measures of the deviation between the recommended value and the real value specified by the user. RMSE and Mae values can be obtained by calculating the score deviation between the actual score and the predicted score between users. The lower the RMSE and Mae values are, the higher the accuracy of the proposed algorithm is. Formally,

\[
MAE = \frac{\sum_{j=1}^{N} |p_{ij} - r_{ij}|}{N} \tag{9}
\]

\[
RMSE = \sqrt{\frac{\sum_{j=1}^{N} (p_{ij} - r_{ij})^2 / N}{N}} \tag{10}
\]

Here, \( p_{ij} \) is the predicted score to target user on the project \( i \), \( N \) is the count of projects predicted and \( r_{ij} \) is the real score to target user on the project \( i \).

4.3. Result Analysis

Slope One algorithm is a project-based collaborative filtering recommendation algorithm proposed by Lemire et al. in 2005 [12]. In practice, this algorithm has the advantages of simple calculation, easy implementation and maintenance, providing recommendation services for new users, efficient query corresponding, reasonable accuracy, etc. It generates prediction by calculating the deviation between users' projects. TrustWalker, a recommendation algorithm based on random walk model, is proposed in literature [13].

In the MovieLens data set, each algorithm is used to calculate the predicted value, and then MAE (Figure 3) and RMSE (Figure 4) are calculated and compared. Among them, the number of neighbor users selected for the current user is 5, 10, 15 50.
As can be seen from figure 3, the MAE of UM model recommendation results is lower than the other two algorithms.

In figure 4, the RMSE of UM model recommendation results is lower than the other two algorithms.

Acknowledgment
This work was sponsored by the National Natural Science Foundation of China (71561013), the Fund of Humanities and Social Sciences in Universities of Jiangxi Province (JC17221, JD18083).

References
[1] Xu Hai-ling, Wu Xiao, Li Xiao-dong, et al. Comparison study of Internet recommendation system [J]. Journal of Software, 2009, 20 (2): 350-362.
[2] Wang Peng, Wang Jing-Jing, Yu Neng-hai. A kernel and user-based collaborative filtering recommendation algorithm [J]. Journal of Computer Research and Development, 2013, 50 (7): 1444s-1451.
[3] Li Bin. Cross-domain Collaborative Filtering: A Brief Survey [C]. Proceedings of the 23rd International Conference on Tools with Artificial Intelligence. [S. l. ]:IEEE Press,2011:1085-1086.
[4] Ning X,Karypis G. Multi-task Learning for Recommendation System [C]. Proceedings of the 2nd Asian Conference on Machine Learning. Tokyo,Japan: [s. n. ],2010:269-284.
[5] Tsai C F, Hung C. Cluster ensembles in collaborative filtering recommendation [J]. Applied Soft Computing, 2012, 12 (4): 1417-1425
[6] Zhang Juan, Hu Xuegang, Zhang Yuhong, et al. An efficient ensemble method for classifying skewed data streams [C]. Proceedings of the 7th International Conference on Intelligent Computing: Bio-inspired Computing and Applications. 2011: 144-151.

[7] Li Bin, Yang Qiang, Xue Xiangyang. Transfer Learning for Collaborative Filtering via a Rating-matrix Generative Model [C]. Proceedings of the 26th Annual International Conference on Machine Learning. Quebec, Canada: [s. n.], 2009: 617-624.

[8] Singh A P, Gordon G J. Relational Learning Via Collective Matrix Factorization [C]. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. [S. l.]: ACM Press, 2008: 650-658.

[9] G. Adomavicius and A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE transactions on knowledge and data engineering, pp. 734-749, 2005.

[10] Schafer, J., Frankowski, D., Herlocker, J. and Sen, S. Collaborative Filtering Recommendation Systems. The Adaptive Web. 2007, 291-324.

[11] Yuyun Gong and Qi Zhang. Hashtag recommendation using attention-based convolutional neural network. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016, pages 2782-2788, 2016.

[12] Lemire D, Maclachlan A. Slope One Predictors for Online Rating-Based Collaborative Filtering [C]. In: Proceedings of the Fifth SIAM International Conference on Data Mining, 2005, 471-480.

[13] Jamali M, Ester M. TrustWalker: a random walk model for combining trust-based and item-based recommendation. In: Pro of KDD’09, New York: ACM Press, 2009, 397-405