Improved Metric Factorization Recommendation Algorithm Based on Social Networks and Implicit Feedback

Bilin Wang1, Jiaxin Han1, Ying Cuan1

1 School of Computer, Xi’an Shiyou University, Xi’an, Shanxi, 710065, China

* Corresponding author’s e-mail: wblhcgd@gmail.com

Abstract: The Metric Factorization algorithm solves the problem of the suboptimal solution caused by the inner product of the traditional matrix factorization algorithm. Although the basic metric factorization model has achieved good results in rating prediction and item ranking tasks, the algorithm ignores the role of implicit feedback and user social information. Considering the social relationship and implicit feedback information between users, this paper improves the basic metric factorization algorithm, and proposes an improved metric factorization recommendation algorithm based on social networks and implicit feedback. We do rating prediction tasks on the Filmtrust and Last.FM datasets, experimental results show that the improved algorithm can further improve the accuracy of prediction.

1. Introduction

How to help users quickly find the information they are interested in in massive data has become the primary task of the current recommendation system. The latent factor model: SVD [1] is widely researched and applied, because of its good performance. However, the SVD model and the algorithm improved by its evolution all use the dot product, which does not satisfy the inequality property and leads to suboptimal solutions. Factorized Metric Learning [3] (FML) overcomes the problem. It assumes that users and items can be placed in low dimensional space and replaces the dot product of vectors with Euclidean distance.

However, the FML ignores the influence of user's implicit feedback and social information on recommendation quality.

To solve the above problem, this paper integrates the user's social network information and implicit feedback information and proposes the MFreg + + algorithm. By comparing the experimental results on the two datasets Filmtrust and Last.FM, we show that the proposed model obtains higher recommendation accuracy in the rating prediction tasks of recommendation.

2. Related Work

2.1. Matrix Factorization model

The SVD is a model derived from the factorization of the user-item rating matrix. It is widely used because of its simple and strong scalability. To solve the problems of data sparseness and efficiency faced in actual scenarios, the Funk – SVD [4] model was proposed. The Funk – SVD algorithm converts the scoring matrix to obtain the user factor matrix and the item factor matrix. Both of them are used to fill the missing part of the rating matrix.

Based on Funk-SVD, considering that different users have different rating preferences. Koren [5] added user and item biases to eliminate the impact between items and between users and proposed the
Basic – SVD model. It obtains a better recommendation result. However, the dot product adopted in Basic – SVD model does not satisfy the inequality property, which may lead to sub optimal solutions.

2.2. Metric Factorization model
To solve the problem caused the dot product, Zhang[3] proposed the Metric Factorization model (FML) which performs better than the most advanced model based on SVD on both rating prediction and item ranking tasks. It assumes that distance and similarity are opposite properties. Based on this assumption, the distance is obtained from score data as follows:

\[ D(u, i) = \text{Rating}_{\max} - R_{ui} \]  \hspace{1cm} (1)

Next, suppose the positions of user and item in a metric vector space are denoted by \( p_u \in R^K \), \( q_i \in R^K \), the distance between user and item is defined by squared Euclidean Distance.

\[ D(u, i) = \|p_u - q_i\|^2 \]  \hspace{1cm} (2)

The formula (2) replaces the form of the dot product in the matrix factorization model. Similar to Basic – SVD, considering the global user and item bias terms, the final prediction distance is:

\[ \hat{r}_{ui} = \mu + b_u + b_i + D(u, i) \]  \hspace{1cm} (3)

The metric decomposition model has achieved good results in the recommendation task [3]. However, it ignores the role of implicit feedback and user social information which are important features in user modelling. This paper will consider the influence of the user's social relationship [8,9] and implicit feedback information [2] obtained from the user's historical behaviour to improve the performance of recommendation.

3. Improvement of Metric Factorization model

3.1. Add implicit feedback information
In this part, we modify the user \( P \) matrix into which is the metric will be decomposed. We do not only consider the rating information but also take into account the user's behavior of rating the Item. The user's rating behavior is regarded as a user's preference for the rating item. Therefore, the user vector \( p_u \) is modified as follows:

\[ \hat{p}_u = p_u + (1/\sqrt{|M(u)|}) \cdot \sum_{z \in M(u)} x_z \]  \hspace{1cm} (4)

Where the \( p_u \) is the user's position vector learned from explicit ratings, while the latter term represents the user's implicit preference information. Among them, \( M(u) \) represents the set of items that the user has evaluated, and \( x_z \) represents the user’s hidden interest in the evaluated item \( z \). Finally, we use \( \hat{p}_u \) to replace the user vector \( p_u \) in equation (2). The regularization term \( |M(u)|^{-1/2} \) is to eliminate the influence of the number of user rating behaviours.

3.2. Add user’s social regulation and item regulation
Studies have found [10] that using social information for recommendation provides users with additional preference information. The social regularization term proposed by Hao Ma [11] gives us a good solution. In the model of the previous section, we consider adding user’s social regularization and item regularization.

3.2.1. User’s social regulation
User’s historical rating information and social information are used to calculate the confidence between users and to accurately locate the user's position in low dimensional space. The user’s social regulation is defined as follows:

\[ \text{Reg}_{user} = \frac{\delta}{2} \sum_{w \in K^+} \text{Sim}(u, w) \|\hat{p}_u - \hat{p}_w\|^2 \]  \hspace{1cm} (5)

Where \( \text{Sim}(u, w) \) is the similarity between user \( u \) and user \( w \). \( K^+ \) represents the set of users who trust user \( u \). The smaller the \( \text{Sim}(u, w) \), the farther the distance between the two users. Here we use Pearson correlation coefficient [12] to measure \( \text{Sim}(u, w) \):
\[
\text{Sim}(u, w) = \frac{\sum_{i \in \text{Sim}(u,w)} r_{iu} - r_{iw}}{\sqrt{\sum_{i \in \text{Sim}(u,w)} (r_{iu} - \bar{r}_u)^2} \sqrt{\sum_{i \in \text{Sim}(u,w)} (r_{iw} - \bar{r}_w)^2}}
\]

(6)

In addition to considering users with a direct social relationship, we also consider the indirect relationship between users. If the user \(u\) has an indirect friend \(y\), we also need to properly reduce the distance between the two users. \(K^-\) represents the set of users who trust user \(w\). Finally, the social regularization items we added are:

\[
\text{Reg}_u = \frac{\delta}{2} \sum \text{Sim}(u, w) \| \hat{p}_u - \bar{p}_w \|^2 + \frac{\delta}{2} \sum \text{Sim}(u, y) \| \hat{p}_u - \bar{p}_y \|^2
\]

(7)

3.2.2. item regulation

When the user's social relationship is relatively sparse, it will limit the expression of the relationship between users. At this point we need to consider the relationship between items. Here we do not consider calculating the similarity between items according to the attributes of the items, but describing based on the existing rating data given by all users. Similarly, the items regulation is as follows:

\[
\text{Reg}_{item} = \frac{\delta}{2} \sum_i \sum_{j \in I} A(U(i,j)) \| q_i - q_j \|^2
\]

The set \(I\) represents the similarity item top – \(k\) similar item set to the item \(i\). Among them, the similarity \(A(U(i,j))\) we adopt the modified cosine similarity of the item mean centralization:

\[
A(U(i,j)) = \frac{\sum_{i \in I} (R_{ij} - \bar{R}_i) (R_{ij} - \bar{R}_j)}{\sqrt{\sum_{i \in I} (R_{ij} - \bar{R}_i)^2} \sqrt{\sum_{i \in I} (R_{ij} - \bar{R}_j)^2}}
\]

(9)

\(\pi_{ij}\) represents a collection of users who have commented on items \(i, j\) at the same time. Besides, we consider that each user's score should be given a different trust value, which is determined by each user's influence in the social network. We define the influence degree \(B\) to be determined by the average similarity of users:

\[
B = \frac{\sum_{u \in \text{Sim}(u,v)} \| R_{uv} \|}{|K^{+}| + 1}
\]

(10)

\(K\) represents the trusted user set of users \(u\), \(A(U(i,j))\) is defined in as follows:

\[
A(U(i,j)) = \frac{\sum_{u \in \text{Sim}(u,v)} \| R_{uv} \|}{\sqrt{\sum_{i \in I} (R_{ij} - \bar{R}_i)^2} \sqrt{\sum_{i \in I} (R_{ij} - \bar{R}_j)^2}}
\]

(11)

To sum up, in the end we get the loss function:

\[
\min \sum_{u \in \text{Sim}(u,v)} 0.5 \cdot (r_{uv} - \hat{r}_{uv})^2 + 0.5 \cdot \text{Reg}
\]

(12)

where: \(\text{Reg} = \gamma_1 (\sum_i b_i^2 + \sum_i f_i^2) + \gamma_2 (\sum_i p_i^2 + \sum_i q_i^2 + \sum_{e \in \text{Sim}(u,v)} x_e^2) + \text{Reg}_u + \text{Reg}_{item}\)

Then we use stochastic gradient descent to solve the loss function:

\[
b_u \leftarrow b_u + \theta (e_{u,l} - y_t b_u)
\]

(13)

\[
b_i \leftarrow b_i + \theta (e_{i,l} - y_t b_i)
\]

(14)

\[
p_u \leftarrow p_u + \theta \left[ e (p_u - q_i + |M(u)|^{-1/2} \sum_{x \in \text{Sim}(u,v)} x_e) + \gamma_2 q_i - \beta \sum_{u \in \text{Sim}(u,v)} A(U(i,j)) (q_i - q_j) \right]
\]

(15)

\[
p_i \leftarrow p_i + \theta \left[ e (p_i - q_i + |M(u)|^{-1/2} \sum_{x \in \text{Sim}(u,v)} x_e) - \gamma_2 q_i - \beta \sum_{u \in \text{Sim}(u,v)} A(U(i,j)) (q_i - q_j) \right]
\]

(16)

\[
\forall e \in \text{Sim}(u,v) x_e \leftarrow x_e + \theta \left[ e \cdot |M(u)|^{-1/2} (p_u - q_i + |M(u)|^{-1/2} \cdot \sum_{x \in \text{Sim}(u,v)} x_e) - \gamma_2 x_e - \delta \sum_{x \in \text{Sim}(u,v)} x_e \right]
\]

(17)

4. Experiments

4.1. Data Set

The Filmtrust and Last.FM datasets are used to evaluate the performance of MFReg++ model. Filmtrust is a small dataset collected from the entire Filmtrust website in June 2011. Last.FM (www.last.fm) collects user-artist listening information. The more the user listens to the artist, the higher the user’s rating of the artist, we pre-process the listening count to \{1,2,3,4,5\} according to certain
criteria. The profile of pre-processed datasets is shown in Table 1.

| Dataset      | # User | # Item | # Ratings | #density | # Relations |
|--------------|--------|--------|-----------|----------|-------------|
| Filmtrust    | 1508   | 2071   | 35479     | 0.0114   | 1853        |
| Last.FM      | 1892   | 17632  | 92834     | 0.0028   | 25434       |

4.2. Evaluation index
This article is an improvement in the rating prediction task. The adopted evaluation index is the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_i^n (R_{ui} - R'_{ui})^2}{N}}$$

(18)

where $N$ is the number of the test set, $R_{ui}$ is the actual rating of item $i$ by user $u$, and $R'_{ui}$ represents the predicted rating of item $i$ by user $u$.

4.3. Model comparison analysis

| Algorithms | Filmtrust | Last.FM |
|------------|-----------|---------|
| Dim        | 5         | 10      | 5       | 10       |
| PMF        | 0.949     | 0.968   | 0.897   | 0.890    |
| SoReg      | 0.878     | 0.875   | 0.886   | 0.881    |
| FML        | 0.838     | 0.835   | 0.877   | 0.870    |
| MFReg++    | 0.818     | 0.815   | 0.860   | 0.855    |

Table 2 is the experimental results on these two datasets with a dimension of 5/10, using 5-fold-cross-validation. The experimental parameters are shown in Table 3:

| Algorithms | Filmtrust | Last.FM |
|------------|-----------|---------|
| PMF        | $\lambda = 0.01$ | $\lambda = 0.01$ |
| SoReg      | $\lambda = \beta = 0.1$ | $\lambda = 0.5$, $\beta = 0.1$ |
| FML        | $\gamma_1 = \gamma_2 = 0.001$ | $\gamma_1 = \gamma_2 = 0.001$ |
| MFReg++    | $\gamma_1 = \gamma_2 = 0.001$ | $\gamma_1 = \gamma_2 = 0.001$ |

4.4. Model parameter exploration
The parameter $\alpha$ in Figure 1 represents the influence of social information on $MFReg^{++}$ model. When the value of $\alpha$ is 0.1, the model obtains the best results. The result shows that the fusion of social information can improve the quality of rating prediction.

The results in Figure 2 show that, in addition to considering the user's social information, the relationship between items affects the quality of the model recommendation. It proves that the regularization of items can better allow the model to measure the positional relationship between items in the low-dimensional space.

5. Conclusion

Through the comparative analysis of the experiments in this paper, the model $MFReg^{++}$ proposed in this paper has been improved in the rating prediction task, proving the effectiveness of this model.

For future work, based on $MFReg^{++}$, we will consider adding the behavior time of the user and the attribute information of the user and the item, and do further exploration to improve the recommendation quality of the model.

Acknowledgments

This work was financially supported by the Xi’an shiyou University Graduate Innovation and Practice Ability Development Project (YCS20113060).

References

[1] Srebro, N., Rennie, J. D. M., & Jaakkola, T. (2004). Maximum-Margin Matrix Factorization. Advances in Neural Information Processing Systems.

[2] Koren, Y. (2008). Factorization meets the neighborhood: A multifaceted collaborative filtering model. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008. ACM.

[3] Zhang, S., Yao, L., Tay, Y., Xu, X., Zhang, X., & Zhu, L. (2018). Metric factorization: recommendation beyond matrix factorization.

[4] Gregory Piatetsky-Shapiro. (2007). Interview with simon funk. acm sigkdd explorations newsletter, 9(1), 38-40.

[5] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30-37.

[6] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). [ACM press the 26thAinternational conference - Perth, Australia (2017.04.03-2017.04.07)] proceedings of the 26th international conference on world wide web - www \"17 - neural collaborative filtering. 173-182.

[7] Shi, Y., Larson, M., & Hanjalic, A. (2014). Collaborative filtering beyond the user-item matrix. Acm Computing Surveys, 47(1), 1-45.
[8] Hsieh, C. K., Yang, L., Cui, Y., Lin, T. Y., Belongie, S., & Estrin, D. (2017). [ACM Press the 26th International Conference - Perth, Australia (2017.04.03-2017.04.07)] Proceedings of the 26th International Conference on World Wide Web - WWW '17 - Collaborative Metric Learning. International Conference (pp.193-201).

[9] Jamali M, Ester M (2010) A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of the fourth ACM conference on Recommender systems. RecSys ‘10, pp.135–142.

[10] Faliagka, E., Tsakalidis, A. K., & Vaikousi, D. (2011). Teenagers' Use of Social Network Websites and Privacy Concerns: A Survey. Panhellenic Conference on Informatics. IEEE.

[11] Cheng-Kang Hsieh, MAH,ZHOUD,LIUC,et al. (2011):Recommender systems with social regularization[C]/Proceedings of the 4th ACM International Conference on Web Search and Data Mining.New York: ACM,287 — 296.

[12] Pearson's correlation coefficient. (1996). Pearson's correlation coefficient. new zealand medical journal, 109(1015), 38.