Rec-GNN: Research on Social Recommendation based on Graph Neural Networks

Gaofei Si[1], Shuwei Xu[2*], Zhaoke Li[2] and Jingyun Zhang[2]

[1]Henan University School of Software, Kaifeng, Henan, China
[2]Henan University School of Software, Kaifeng, Henan, China

gaofei@henu.edu.cn
*Corresponding author:xsw@henu.edu.cn;
ke@henu.edu.cn; zjingy599@163.com

Abstract. Solve the problem that the accuracy of scoring prediction is not high due to insufficient learning of the features of two graph data (user social graph and user item graph) in the social recommendation system (GraphRec), and improve the accuracy of the social recommendation system. In this paper, the user associated with item and the user evaluation of items are improved, and the commodity category attribute is added; And use users as a bridge to connect user-user social graph and user-item graph; Analyze the heterogeneous advantages between nodes, and make full use of the two graph data contained in the social recommendation system to generate the feature vectors that are sufficiently differentiated and representative for users and goods. Compared with the GraphRec basic model without commodity category attribute, the experimental results show that the improved (Rec-GNN) based social recommendation model has a higher accuracy rate in scoring prediction. In this method, RMSE indicators have been improved by 2.5% and 1.5%, respectively, on Ciao and Epinions' two real data sets, and MAE indicators have been improved by 1.3% and 1.3%, respectively. The experimental performance of the model after adding a deeper potential relationship between users and commodities was not tested. This paper analyzes the complex interaction between nodes and adds rich feature information, such as user social information, user ratings on items, and item category attributes, to obtain a more differentiated feature vector representation of users and goods, which has higher accuracy in the social recommendation system score prediction.

Keywords: Graph Neural network; Recommendation system; Social recommendation; Neural network;

1 Introduction

The recommendation system is one of the most widely used information service systems on the Internet today. It can predict the degree of interest in commodities by analyzing users' historical behavior [3-5]. However, most existing recommendation systems have a cold start problem [2]: new users and new products are unfamiliar to the recommendation system without the accumulation of past information, and essential data

H. Wu et al. (Eds.): CSIEDE 2022, ACSR 103, pp. 478–485, 2022.
https://doi.org/10.2991/978-94-6463-108-1_55

needs to be collected through certain user behaviors and interactions. To solve this problem, the existing method is to combine the social information between users [6,7] as auxiliary information for user feature vector learning. The experimental results of Lorenzo. et al. show that the user's social network dramatically affects the user's behavior, such as WeChat's circle of friends, Weibo's attention, the trusted connection of fans, and Taobao. The above research shows that the integration of social relations and users' interaction with related items significantly improves recommendation performance.

The graph neural network social recommendation model uses graph neural network (GNN) [8] to naturally integrate node information, spatial structure information and other advantages, and aggregates information from user-user social graph and user-item graph to process social recommendation tasks [9, 10].

As shown in Figure 1, it includes a user-item graph(left) between user and associated items and a user graph(right) representing the relationship between users. As a bridge, users participate in two drawing data simultaneously. The user-item graph represents the item associated with the user, where the numbers and words on the red line represent the user's ratings of different projects and the category attributes of projects; On the right is the use social graph, and the red line indicates the user's social relationship.

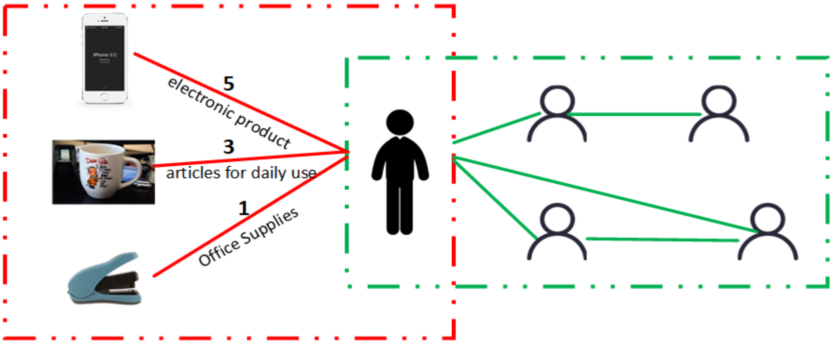

Fig. 1. Two graph data involved in social recommendation

2 Definition and Notations

The social recommendation system mainly includes two types of entities: a user set $U=\{u_1,u_2,u_3,...,u_m\}$ and an item set $T=\{t_1,t_2,t_3,...,t_n\}$, where m and n are the number of users and items used, respectively. Each item t_j has its own category C. If $\mathrm{r}_{i,j}$ not equal to 0, which means that the user u_i has a comment on the item t_j, where $\mathrm{r}_{i,j}$ The value represents the degree of preference. The larger the value r, the more likely the user will buy the product. The value range of r is $\{1,2,3,4,5\}$.

3 Improved model

The model framework of Rec-GNN is shown in Figure 2. Most of the existing recommender systems only consider the user's comments on the item, but do not consider the category information of the item. The category of the item is an important attribute of the item. Each item has a corresponding category, and the items of the same category have similar attributes and functions. This feature is significant for the recommender system. In this paper, the category of the item is added to the existing model. The model mainly includes three parts: user model, item model, and final scoring prediction model.

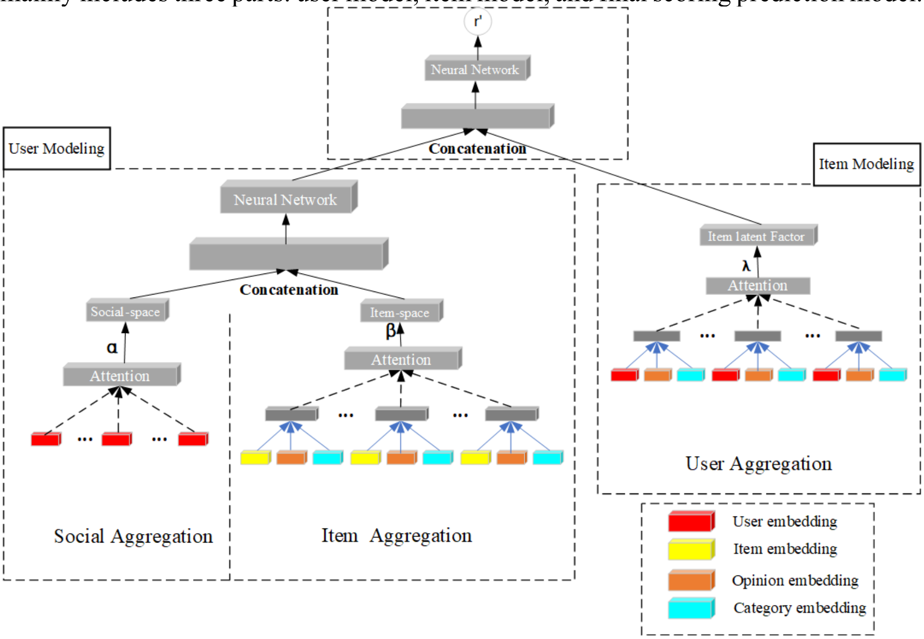

Fig. 2. Improved overall model architecture

3.1 User Model

User modeling is mainly used to learn user u_i feature vectors, expressed as $h_i \in \mathbb{R}^d$. It is necessary to analyze how to combine user social graph and user-item graph better. The Rec-GNN model uses social aggregation and item aggregation to solve this problem, as shown in the left half of Figure 2. Social aggregation learns the eigenvector of user u_i in user social graph, expressed as h_i^S; Item aggregation learns the feature vector of the user u_i in the user-item graph, which is described as h_i^I. Then, the user's feature vector h_i s are obtained by column splicing. Social aggregation and item aggregation are introduced below.

Social aggregation.

The purchase preference of users is affected by their friends or the people around them, and social-related information needs to be added to learn the user feature vector representation further. However, social factors around us cannot be considered equally, because there are strong and weak connections between users. For example, if a user has many social relationships directly connected with him, his influence on the suggestions or preferences of surrounding users will be relatively small; If another user only has a connection to the current user, it will have a great impact on the current user. The Rec-GNN model uses an attention mechanism to solve the problem of social information weight. The specific aggregation formula is as follows:

$$h_i^S = \sigma\left(W \cdot AGG\left(\left\{h_o, \forall o \in N\left(i\right)\right\}\right) + b\right) \tag{1}$$

AGG represents the aggregation function of user trust neighbors and h_o represents the feature vector representation of trusted neighbors. AGG aggregation function is generally expressed as mean aggregation function, it takes the mean value of elements in the vector $\left\{\mathrm{h}_o, \forall o \in N\left(i\right)\right\}$, and the specific formula is as follows:

$$h_i^S = \sigma\left(W \cdot \left\{\sum_{o \in N(i)} \alpha_i h_o\right\} + b\right) \tag{2}$$

Where $\alpha_i = \frac{1}{N\left(i\right)}$, This indicates that all trusted neighbors have the same impact on user u_i. According to the previous description, the influence of trusted neighbors on users varies in strength, so two-layer fully connected neural networks are used to implement an attention mechanism to aggregate the information of these trusted neighbors. It mainly uses the trust neighbor feature vector h_o and user u_i feature vector p_i to calculate the connection strength between nodes. The specific formula for calculating the attention weight is as follows:

$$\alpha_{io}^* = W_2^T \cdot \sigma\left(W_1 \cdot \left[h_o \oplus p_i\right] + b_1\right) + b_2 \tag{3}$$

$$\alpha_{io} = sigmod\left(\alpha_{io}^*\right) \tag{4}$$

α_{io}^* after normalization of the softmax function, α_{io} can be used as the connection strength coefficient between the user u_i and the trusted neighbor.

The eigenvector of the end user after aggregating neighbors in the social space:

$$h_i^S = \sigma\left(W \cdot \left\{\sum_{o \in N(i)} \alpha_{io} h_o\right\} + b\right) \tag{5}$$

Item aggregation.

The other part of user modeling is item aggregation. By using the information in the user-item graph: items directly associated with users, users' evaluation of items, and item category matrix $C \in \mathbf{R}^{m \times n}$. Each item t_j has its own category. The rec-GNN model uses the above information to model users in the item space, and obtains the eigenvector representation h_i^I of users in the item space.

Item aggregation is mainly to learn the user representation of item space according to the interaction between users and items, the user's evaluation of items, and the category of items. It can obtain rich information about the items directly connected to it, so as to better learn the eigenvector representation h_i^I of the user in the item space. The aggregation formula of user in the item space is as follows:

$$h_i^I = \sigma\left(W \cdot \left\{\sum_{a \in S(i)} \beta_{ia} x_{ia}\right\} + b\right) \tag{6}$$

$$x_{ia} = L_v\left(\left[q_a \oplus e_r \oplus e_c\right]\right) \tag{7}$$

x_{ia} is a perceptual representation of interaction between item opinions and categories. It integrates the embedded representation of items, user opinions on items and item category information into the user feature vector representation. β_{io} It is the connection strength coefficient between user u_i and interactive items.

User feature vector representation.

This paper combines the eigenvector representation of users in social space and the eigenvector representation of users in item space through a fully connected neural network to get the final user representation. The user representation h_i is defined as:

$$h_i = \sigma\left(W \cdot \left(h_i^S \oplus h_i^I\right) + b\right) \tag{8}$$

3.2 Item Model

The item model is shown in the right half of Figure 2, which aims to learn the item's feature vector by aggregating users who interact with item t_j. The eigenvector representation of the item is further learned through the interactive user and user evaluation information of the item in the user-item graph and the category of the item.

User aggregation.

It is basically the same as the user feature vector obtained above, without too much explanation.

$$Z_j = \sigma\left(W \cdot \left\{\sum_{t \in Q(j)} \lambda_{jt} K_{jt}\right\} + b\right) \tag{9}$$

3.3 Final Scoring Prediction Model

The model Rec-GNN proposed in this paper is trained on the training set and applied to the scoring prediction recommendation task. By splicing the user's eigenvector representation and the item's eigenvector representation in columns, the item's score is finally obtained through a two-layer fully connected neural network scoring prediction:

$$r = W^T \cdot \sigma\left(W_2 \cdot \sigma\left(W_1 \cdot \left[h_i \oplus z_j\right] + b_1\right) + b_2\right) \tag{10}$$

3.4 Loss function

In order to calculate the model parameters according to the loss value, this paper defines a loss function, where $|\varepsilon_r|$ the number of items evaluated by all users in the data set $r_{i,j}$ is the real evaluation score of u_i for t_j, and the loss function is expressed as:

$$Loss = \frac{1}{2|\varepsilon_r|} \sum_{i,j \in \varepsilon_r} \left(r - r_{ij}\right)^2 \tag{11}$$

4 Experiment

4.1 Datasets

In the experiment, Ciao and Epinions are public and representative data sets for social recommendation research. In the experiment, users' self-connected edges were deleted, and users without social relations were deleted, because they did not belong to the scope of social recommendation. There are 7317 users, 104975 items, and 283320 scores in the Ciao dataset. Epinions has 18069 users, 261246 items, and 762938 scores.

Table 1. datasets

Dataset	Ciao	Epinions
Users	7,317	18,069
Items	104,975	261,246
Num of ratings	283,319	762,938
Social links	111,781	355,530

4.2 Analysis of experimental results

Rec-GNN, GCMC+SN [11], and GraphRec [1] have better-recommended performance than SoRec, SoReg, and SocialMF. Through comparison, we can see that the GNN method has a stronger ability and better effect in aggregating neighbor information. The Rec- GNN proposed in this paper shows promising results on both datasets and is superior to the existing recommendation methods.

Table 2 shows the RMSE and MAE indicators of each recommended method in Ciao and Epinions datasets:

Table 2. Compare the performance of different recommendation systems on two datasets

Datasets	Ciao		Epinions	
	RMSE	MAE	RMSE	MAE
PMF	1.1967	0.952	1.2739	1.0211
SoRec	1.0738	0.8489	1.1563	0.9086
SoReg	1.0947	0.8987	1.1936	0.9412
SocialMF	1.0592	0.8353	1.1410	0.8965
GCMC+SN	1.0221	0.7697	1.1004	0.8602
GraphRec	1.0093	0.7540	1.0878	0.8441
Rec-GNN	**0.9848**	**0.7409**	**1.0730**	**0.8314**

4.3 Conclusion

In recent years, research on social recommendation has made significant progress and has been widely used in various fields. However, the existing social recommendation systems only consider the social relationship between users and the user and the user's rating of the item, ignoring the category of the item, which is an important characteristic of the item itself. In this paper, we use graph neural network to add the category of items to user modeling and item modeling, and through a reasonable way, we can aggregate and learn the interaction opinions of user social graph and user item graph and the category of items. In addition, we analyzed the heterogeneity of the connection, and the connection with different strengths gave them different weight coefficients. So as to better learn the feature vectors of users and items. In this paper, the experimental results on two real-world datasets, Ciao and Epinions, show that Rec- GNN can exceed the most advanced baseline. For future research, we can pay more attention to the interaction between items. There is a social relationship between users, and the corresponding items should be related to each other, not independent. This is worth exploring in the future.

Acknowledgments

This study was funded by Engineering Ethics, a quality course of Henan Province Postgraduate Quality Improvement Project in 2022 (Project Approval No.: YJS2022KC29).

References

1. Fan W, Ma Y, Li Q, et al. Graph neural networks for social recommendation[C]//The world wide web conference. 2019: 417-426.
2. Zhu Y, Xie R, Zhuang F, et al. Learning to warm up cold item embeddings for cold-start recommendation with meta scaling and shifting networks[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 1167-1176.
3. Zhiwei Liu, Xiaohan Li, Ziwei Fan, Stephen Guo, Kannan Achan, and Philip S Yu. 2020. Basket Recommendation with Multi-Intent Translation Graph Neural Network. arXiv preprint arXiv:2010.11419 (2020).
4. Zhiwei Liu, Mengting Wan, Stephen Guo, Kannan Achan, and Philip S Yu. 2020.Basconv: aggregating heterogeneous interactions for basket recommendation with graph convolutional neural network. In Proceedings of the 2020 SIAM International Conference on Data Mining. SIAM, 64–72.
5. Zhiwei Liu, Lei Zheng, Jiawei Zhang, Jiayu Han, and S Yu Philip. 2019. JSCN: Joint spectral convolutional network for cross domain recommendation. In 2019 IEEE International Conference on Big Data (Big Data). IEEE, 850–859.
6. Jiliang Tang, Huiji Gao, Xia Hu, and Huan Liu. 2013. Exploiting homophily effect for trust prediction. In Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013. ACM, 53–62.
7. Jiliang Tang, Xia Hu, Huiji Gao, and Huan Liu. 2013. Exploiting local and global social context for recommendation. In IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013.IJCAI/AAAI, 2712–2718.
8. Zhiwei Liu, Yingtong Dou, Philip S Yu, Yutong Deng, and Hao Peng. 2020. Alleviating the Inconsistency Problem of Applying Graph Neural Network to Fraud Detection. In Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020. ACM, 1569–1572.
9. Zhiwei Guo and Heng Wang. 2021. A Deep Graph Neural Network-based Mechanism for Social Recommendations. IEEE Transactions on Industrial Informatics 17, 4 (2021), 2776–2783.
10. Nan Mu, Daren Zha, Yuanye He, and Zhihao Tang. 2019. Graph Attention Networks for Neural Social Recommendation. In 31st IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2019, Portland, OR, USA, November 4-6,2019. IEEE, 1320–1327.
11. Rianne van den Berg, Thomas N Kipf, and Max Welling. 2017. Graph convolutional matrix completion. arXiv preprint arXiv:1706.02263 (2017).