Deep Modeling of Social Relations for Recommendation

Wenqi Fan, Qing Li, Min Cheng
Department of Computer Science, City Univeristy of Hong Kong
83 Tat Chee Avenue, Kowloon, Hong Kong
wenqifan1990@hotmail.com, itqli@cityu.edu.hk, mc.cheng@my.cityu.edu.hk

Abstract

Social-based recommender systems have been recently proposed by incorporating social relations of users to alleviate sparsity issue of user-to-item rating data and to improve recommendation performance. Many of these social-based recommender systems linearly combine the multiplication of social features between users. However, these methods lack the ability to capture complex and intrinsic non-linear features from social relations. In this paper, we present a deep neural network based model to learn non-linear features of each user from social relations, and to integrate into probabilistic matrix factorization for rating prediction problem. Experiments demonstrate the advantages of the proposed method over state-of-the-art social-based recommender systems.

Introduction

Exploiting social relations for recommender systems can help to mitigate data sparsity issues and potentially improve recommendation performance (Guo, Zhang, and Yorke-Smith 2015). Such social-based recommender systems are developed based on the phenomenon that users usually acquire and disseminate information through their nearest neighbouring users.

The exploitation of social relations for recommender systems attracts increasing interest in recent years. SoRec (Ma et al. 2008) is a co-factorization method, which shares a common latent user-feature matrix factorized by ratings and social relations. TrustSVD (Guo, Zhang, and Yorke-Smith 2015) is a trust-based matrix factorization and uses a weighted regularization to regularize the user- and item-specific latent feature vectors. TrustPMF (Yang et al. 2017) models mutual influences between users, and maps users into two low-dimensional spaces: truster space and trustee space, by factorizing social trust networks. However, most of the methods mentioned above combine the multiplication of social features linearly, which may cause ineffectiveness on learning the non-linear features from social relations. Specially, SoRec and TrustPMF treat the interests of users close to the linear combination of their neighbours' interests. Such a method is hard to model non-linear social features.

In this paper, we propose a **Deep** neural network model on **So**cial **R**elations for recommendation, namely **DeepSoR**, to learn non-linear social features from users' social relations, rather than the linear. To the best of our knowledge, DeepSoR is among the first Deep Neural Network(DNN) based models to bridge the gap between social relations and recommender systems. Experiments demonstrate the advantages of DeepSoR over state-of-the-art social-based recommender systems.

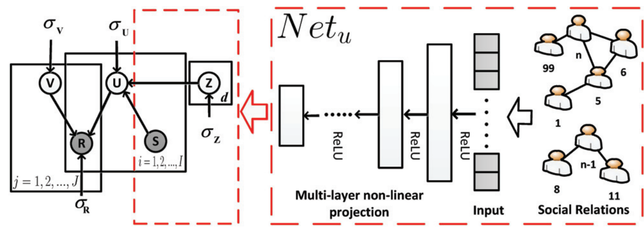

Figure 1: Graphical model of DeepSoR. Deep neural network on social relations part in right (dashed-red).

The Proposed Model

Probabilistic Model of DeepSoR

An overall view of the probabilistic model for DeepSoR is shown in Figure 1. Suppose that we have I users, J items, and the observed rating matrix $R \in \mathbb{R}^{I \times J}$. The goal of DeepSoR is to find user and item latent models ($U \in \mathbb{R}^{l \times I}$ and $V \in \mathbb{R}^{l \times J}$, where l is the dimension size.) whose product ($U^T V$) reconstructs the rating matrix R. The conditional distribution over observed ratings can be defined as: $p(R|U, V, \sigma_R^2) = \prod_i^I \prod_j^J N(r_{ij}|u_i^T v_j, \sigma_R^2)^{C_{ij}}$, where $N(x|\mu, \sigma_R^2)$ is the probability density function of the Gaussian normal distribution with mean μ and variance σ_R^2, and C_{ij} is an indicator function equal to 1 if user i rated item j and equal to 0 otherwise. Also, zero-mean Gaussian distribution are assumed for item V with variance σ_V^2.

By replacing the user latent models of the probabilistic matrix factorization(PMF) (Salakhutdinov and Mnih 2007) with DNN (Net_u)(see Figure 1), user latent feature vector u in our proposed model is approximated by social latent vector generated from DNN(Net_u) as follows: $u_i = Net_u(Z, s_i) + \epsilon_i$, where s_i is social information of user i and will be described later. We also model the internal weights Z (weights and bias) in Net_u and noise ϵ as zero-mean Gaussian distribution with

Table 1: Comparisons of different recommender systems

Dataset	Metrics	Algorithms						
		PMF	SoRec	SoReg	TrustSVD	TrustPMF	**DeepSoR-Random**	**DeepSoR**
Ciao	MAE	0.8622	0.7797	0.8457	0.8109	0.7783	**0.7786**	**0.7664**
	RMSE	1.1466	1.0413	1.1185	1.0891	1.0543	**1.0281**	**1.0160**
Epinions	MAE	1.0489	0.8717	0.8851	0.8707	0.8871	**0.8632**	**0.8518**
	RMSE	1.3834	1.1462	1.1609	1.1225	1.1495	**1.1109**	**1.1003**

variance σ_Z^2 and σ_U^2 respectively. Accordingly, we define a conditional distribution of U given Z and S to model the user latent models as follows:

$$p(U|Z,S,\sigma_U^2) = \prod_i^I N(u_i|Net_u(Z,s_i),\sigma_U^2\mathbf{I}) \quad (1)$$

Deep Neural Network on Social Relations

In this section, we will investigate how to introduce the social relations into the Net_u and learn the parameters.

We adopt the latest feature representations technique in social network — *node2vec* (Grover and Leskovec 2016) which is proposed to learn a mapping of users to a low-dimensional feature vector from their social relations. A user embedding feature $\mathbf{e}_i \in \mathbb{R}^q$ is a q-dimensional vector generated by the *node2vec* from social relations of user i. Moreover, a user's preference is similar to or influenced by her nearest neighbouring users in online social networks (Ma et al. 2011). Therefore, the k Nearest Neighbors (kNN) technique is exploited to compose user's neighbors. The k most similar neighbours are returned for each user using the Euclidean distance based on user's embedding feature. Then, we take the vector average of the embeddings associated with the k most similar neighbours for user i as follows: $s_i = \sum_{a=1}^{k} \frac{\mathbf{e}_a}{k}$. Finally, the averaged embedding vector s_i as input is fed into a non-linear multi-layer networks(hidden layers) architecture. Specifically, each non-linear layer computes as follows:

$$h^{(2)} = f(Z^{(1)}s_i + b^{(1)}), h^{(p+1)} = f(Z^{(p)}h^{(L)} + b^{(p)}) \quad (2)$$

where p is the layer number and f is the non-linear activation function (ReLU, Rectified Linear Units). $b^{(p)}$ and $Z^{(p)}$ is the bias and weight for p-th hidden layer, respectively.

To optimize the variables, we devise an EM-style algorithm for obtaining the maximum a posteriori(MAP) estimation, as follows: $\max_{U,V,Z}[p(U,V,Z|R,S,\sigma_R^2,\sigma_U^2,\sigma_V^2,\sigma_Z^2)] = \max_{U,V,Z}[p(R|U,V,\sigma_R^2)p(V|\sigma_V^2)p(U|Z,S,\sigma_U^2)p(Z|\sigma_Z^2)]$.

For the variables u_i and v_j, the coordinate ascent method which iteratively optimizes a latent variable while fixing the remaining variables is used:

$$u_i \leftarrow (VC_iV^T + \lambda_U\mathbf{I}_k)^{-1}(VC_iR_i + \lambda_U Net_u(Z,s_i)) \quad (3)$$

$$v_j \leftarrow (UC_jU^T + \lambda_V\mathbf{I}_k)^{-1}UC_jR_j \quad (4)$$

where $\lambda_U = \sigma_R^2/\sigma_U^2$, $\lambda_V = \sigma_R^2/\sigma_V^2$, $\lambda_Z = \sigma_R^2/\sigma_Z^2$. Given U and V, we can learn Z for each layer using the backpropagation learning algorithm.

Experiments

Two representative datasets are employed in our experiments: Ciao and Epinions[1]. Two metrics are adopted to evaluate predictive accuracy, namely MAE and RMSE (Yang et al. 2017). We compare our approach with several state-of-the-art recommender systems, including PMF, SoRec, SoReg (Ma et al. 2011), TrustSVD, TrustPMF, and DeepSoR-Random (S is randomly initialized).

In Table 1, the improvements against the five respective competitors (non-deep models) show that the way of learning social features through DNN is better than other social-based recommender systems. Moreover, DeepSoR-Random performs slightly weaker than DeepSoR. It implies that the embedding features is effective for modeling the social information of users from their social relations.

Conclusion and Future Work

In this paper, we proposed a general framework in combining DNN on learning social features of users with probabilistic matrix factorization model for recommender systems. In the future, we will investigate more real-world datasets to further evaluate the effectiveness of our model.

Acknowledgments

This work has been supported by a research grant from City University of Hong Kong (project no. 9610367). The authors also thank their group members for the constructive comments and suggestions.

References

Grover, A., and Leskovec, J. 2016. node2vec: Scalable feature learning for networks. In *KDD*, 855–864. ACM.

Guo, G.; Zhang, J.; and Yorke-Smith, N. 2015. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In *AAAI*, 123–129.

Ma, H.; Yang, H.; Lyu, M. R.; and King, I. 2008. Sorec: social recommendation using probabilistic matrix factorization. In *CIKM*, 931–940. ACM.

Ma, H.; Zhou, D.; Liu, C.; Lyu, M. R.; and King, I. 2011. Recommender systems with social regularization. In *WSDM*, 287–296. ACM.

Salakhutdinov, R., and Mnih, A. 2007. Probabilistic matrix factorization. In *NIPS*, volume 1, 2–1.

Yang, B.; Lei, Y.; Liu, J.; and Li, W. 2017. Social collaborative filtering by trust. *IEEE Trans.PAMI* 39(8):1633–1647.

[1]http://www.cse.msu.edu/∼tangjili/index.html