Machine Learning Concepts

Iman

Contents

[1] The sections colored with red indicate the problems for which a rough idea has been developed.

1 SVM

Support Vector Machine: Consider an input space $\mathbb{X}$ that is a subset of $\mathbb{R}^N$ with $N \geq 1$ and the output or target space $\mathbf{y}$

$$\min_{\mathbf{U},\mathbf{V}} \frac{1}{\Omega_{\mathbf{R}}} \sum_{(i,j)\in\Omega_{\mathbf{R}}} \left(\mathbf{R}_{i,j} - \mathbf{U}_{i:*}^{\top}\mathbf{V}_{*,j}\right)^2 + \lambda_{\mathbf{U}}\|\mathbf{U}\|_{\mathrm{F}}^2 + \lambda_{\mathbf{V}}\|\mathbf{V}\|_{\mathrm{F}}^2 + \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \ell(\mathbf{U}_{t_i,*}, \mathbf{U}_{t_j,*}, \mathbf{U}_{t_k,*})$$

1.1 Separable Case

This problem investigates the network completion problem initiated in [7]. In network completion problem, it is assumed that only a part of the network (e.g., a complete subgraph of the social graph) is observed and we would like to infer the unobserved part of the network. In this problem, we assume that besides the observed subgraph, side information about the nodes such as the pairwise similarity between them is also provided. In contrast to the original network completion problem where the standard methods such as matrix completion is inapplicable due the non-uniform sampling of observed links, the goal here is to show that by effectively exploiting the side information, it is possible to accurately predict the unobserved links.

1.2 Structured Sign Prediction in Social Networks

This problem is related to Problem 2.7. The main disadvantage of exiting link prediction algorithms (e.g., see [9] for a recent survey) is that they mostly reduce the problem to a binary classification over edges. As a result, the structure of the network is somehow ignored in prediction. On the other hand, there are many theories in social network analysis that try to model the structural properties of networks such as balance and status theories– to name a few.

Therefore, an interesting problem is to find efficient algorithms that are capable of exploiting the structural properties of social network in link prediction. One solution to this problem is to directly define the objective of prediction in terms of the structural properties of network instead of individual edges. In [2] an attempt has been made towards this problem, but the proposed algorithm is not efficient and practically useless.

1.2.1 Social Constrained Transductive Classification

Here the goal is to extend the transductive SVM (S^3VM) algorithm [1, 6] to sign prediction problem with node features. The existing methods for link prediction with node features (and similarly edge features or latent features) try to reduce the problem to a binary classification problem which ignores the structure of social network [9].

One idea to exploit the structural properties of social network, and in particular balance theory, is to reduce the problem to transductive classification problem and impose constraint on the unknown labels based on the structural properties of the networks. More specifically, the goal of learning is to transact labels from known links to unknown links by simultaneously maximizing the margin of classification and consistency of links extracted from the social graph over all triads.

1.3 Negative Link Prediction with One Bit Matrix Completion

Although distrust relations/negative links in social networks have been shown to be a rich source of information and in some cases as valuable as trust relations, a major impediment in their effective use is that most social networks do not enable users to specify them explicitly (e.g., Facebook and Twitter do not enable users to explicitly specify negative links) [2]. Therefore, it is natural to explore whether one can predict negative links automatically from the commonly available social network data and other sources of information about users (for more details please check [11]).

There is a close connection between this problem and PU learning for matrix completion [5], which makes it an interesting problem to be investigated from both theoretical and empirical point of view.

2 Social Recommender System

2.1 PushTrust

Let assume there are n users $\mathcal{U} = \{u_1, u_2, \cdots, u_n\}$ and a set of items $\mathcal{I} = \{i_1, i_2, \cdots, i_m\}$. The input is the rating matrix $\mathbf{R} \in \mathbb{R}^{n\times m}$ and the social (trust/distrust) network $\mathbf{S} \in \{-1, +1\}^{n\times n}$ between users.

The overarching goal of social recommender systems is to factorize rating matrix $\mathbf{R} \in \mathbb{R}^{n\times m}$ into the latent features $\mathbf{R} \approx \mathbf{U}\mathbf{V}^\top$ using the social network between users. The only work that is able to simultaneously exploits both trust and distrust relations in factorization is [3] which proposed a ranking idea over latent features but suffers from following main issues:

1. It optimizes over all the triplets which makes it computationally unattractive due to $O(n^3)$ number constraints in the optimization.

2. For each triplet (i, j, k) it ignores the relations between jth and kth nodes.

3. The absence of a link must be considered as a *unknown* relation meaning that it can be potentially trust or distrust link. Therefore, in learning the latent features this fact must be considered.

The focus of this problem would be on resolving these issues and in particular the computational cost.

2.2 Recommendation Using Multiple Sources

The contemporary recommender systems only try to utilize the rating information available in the user-item matrix $\mathbf{R} \in \mathbb{R}^{n\times m}$. But in recent application other sources of information have been emerged that can potentially benefit the recommendation such as:

1. The feature vector for the users

2. The feature vector of items

3. The features of rating, location of rating, time, and etc.

4. The social network among users

5. The reviews (text data) attached to ratings

6. $\cdots$

The main equation here is how to effectively integrate multiple sources of information to boos the accuracy of existing methods.

2.3 Matrix Completion with Side information for Cold-start Recommendation

Unfortunately, the existing methods are not capable of handing cold-start users or items. In particular, the matrix completion or factorization methods are not applicable as an entire row or column is unavailable for cold-start users (items) in the rating matrix.

The question is whether or not if it is possible to exploit other sources of information (similarity between users/items) to apply factorization or matrix completion methods.

2.4 Implicit Distrust for Better Recommendation

In many social networks such as Facebook users are not allowed to make distrust relations. But, the absence of a link between two users, depending on the context of users, might have two different meaning: they know each other but distrusted, or they do not know each other.

The question is that how we could extract predict distrust relations and utilize them in the recommendation and see if it helpful at all. To achieve this goal, an interesting intermediate question is as follows.

2.5 Implicit Subjective Trust Extraction and Utilization

In most of social networks, in particular Epinions, the trust/distrust relations between users is subjective. This means a user might trust another users for a specific category of items while this is not reflected in the relation.

2.6 From Reviews to Ratings

The problem here is to predict a user's numeric rating in a product review from the text of the review. Please see [10, 8, 4]

2.7 Social Regularized Matrix Factorization

There are different theories such as social balance theory and status theory that aim at modeling the social behaviour of users in social networks. In some applications such as sign prediction, these theories also used as a metric to assess the performance of prediction algorithms.

The goal of this project would be a deep understanding of these theories and devising novel algorithms that directly takes these theories into the optimization. In particular, let $\mathcal{T} = \{(i,j,k) \in [n]\times[n]\times[n]\}$ be the set of triangles in the social graphs. The goal of Social Matrix Factorization (SMF) is solving the following optimization problem:

$$\min_{\mathbf{U},\mathbf{V}} \frac{1}{\Omega_{\mathbf{R}}} \sum_{(i,j)\in\Omega_{\mathbf{R}}} \left(\mathbf{R}_{i,j} - \mathbf{U}_{i:*}^{\top}\mathbf{V}_{*,j}\right)^2 + \lambda_{\mathbf{U}}\|\mathbf{U}\|_{\mathrm{F}}^2 + \lambda_{\mathbf{V}}\|\mathbf{V}\|_{\mathrm{F}}^2 + \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \ell(\mathbf{U}_{t_i,*}, \mathbf{U}_{t_j,*}, \mathbf{U}_{t_k,*})$$

The only challenge is above optimization problem is to find an appropriate mapping from the space of latent features to sign of edges to penalize the violation of social theories accordingly.

References

[1] Kristin Bennett, Ayhan Demiriz, et al. Semi-supervised support vector machines. *Advances in Neural Information processing systems*, pages 368–374, 1999.

[2] Kai-Yang Chiang, Cho-Jui Hsieh, Nagarajan Natarajan, Inderjit S Dhillon, and Ambuj Tewari. Prediction and clustering in signed networks: a local to global perspective. *The Journal of Machine Learning Research*, 15(1):1177–1213, 2014.

[3] Rana Forsati, Mehrdad Mahdavi, Mehrnoush Shamsfard, and Mohamed Sarwat. Matrix factorization with explicit trust and distrust side information for improved social recommendation. *ACM Transactions on Information Systems (TOIS)*, 32(4):17, 2014.

[4] Zhen Hai, Kuiyu Chang, Jung-Jae Kim, and Christopher C Yang. Identifying features in opinion mining via intrinsic and extrinsic domain relevance. *IEEE Trans. Knowl. Data Eng.*, 26(3):623–634, 2014.

[5] Cho-Jui Hsieh, Nagarajan Natarajan, and Inderjit S Dhillon. Pu learning for matrix completion. *arXiv preprint arXiv:1411.6081*, 2014.

[6] Thorsten Joachims. *Text categorization with support vector machines: Learning with many relevant features*. Springer, 1998.

[7] Myunghwan Kim and Jure Leskovec. The network completion problem: Inferring missing nodes and edges in networks. In *SDM*, pages 47–58. SIAM, 2011.

[8] Fangtao Li, Nathan Liu, Hongwei Jin, Kai Zhao, Qiang Yang, and Xiaoyan Zhu. Incorporating reviewer and product information for review rating prediction. In *IJCAI*, volume 11, pages 1820–1825, 2011.

[9] Aditya Krishna Menon and Charles Elkan. Link prediction via matrix factorization. In *Machine Learning and Knowledge Discovery in Databases*, pages 437–452. Springer, 2011.

[10] Lizhen Qu, Georgiana Ifrim, and Gerhard Weikum. The bag-of-opinions method for review rating prediction from sparse text patterns. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 913–921. Association for Computational Linguistics, 2010.

[11] Jiliang Tang, Shiyu Chang, Charu Aggarwal, and Huan Liu. Negative link prediction in social media. *arXiv preprint arXiv:1412.2723*, 2014.