Explainable Matrix Factorization for Collaborative Filtering

Behnoush Abdollahi
Knowledge Discovery & Web Mining Lab
Dept. of Computer Engineering & Computer Science, University of Louisville
Louisville, KY 40222, USA

Olfa Nasraoui
Knowledge Discovery & Web Mining Lab
Dept. of Computer Engineering & Computer Science, University of Louisville
Louisville, KY 40222, USA

ABSTRACT

Explanations have been shown to increase the user's trust in recommendations in addition to providing other benefits such as scrutability, which is the ability to verify the validity of recommendations. Most explanation methods are designed for classical neighborhood-based Collaborative Filtering (CF) or rule-based methods. For the state of the art Matrix Factorization (MF) recommender systems, recent explanation methods, require an additional data source, such as item content data, in addition to rating data. In this paper, we address the case where no such additional data is available and propose a new Explainable Matrix Factorization (EMF) technique that computes an accurate top-n recommendation list of items that are explainable. We also introduce new explanation quality metrics, that we call Mean Explainability Precision (MEP) and Mean Explainability Recall (MER).

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences; H.1.0 [**Information Systems**]: Models and Principles—General

Keywords

Explanations, Matrix Factorization (MF), Recommender Systems, Collaborative Filtering (CF)

1. INTRODUCTION

MF is a family of latent factor models that have been used with success in CF recommender systems [4]. They approximate the rating, r_{ij} given by user i on item j using a factorization of the ratings: $r_{ij} \simeq p_i q_j$. To do this, MF learns $p_i \in \mathbb{R}^f$ and $q_j \in \mathbb{R}^f$, which are the lower-rank representations of user i and item j in a joint latent space of dimensionality, f, that is much lower than the typically large number of users or items. To solve for p_i and q_j, different approaches such as stochastic gradient descent can be used to minimize the error between r_{ij} and $p_i q_j$ [4]. One way to communicate user-based neighborhood style explanations in MF-based models is to show to the active user, the preferences of the users who are most similar to her (in terms of their previous ratings) and who have also highly rated or liked the recommended item [1]. Similarity between users can be computed based on their latent factor space representation. The drawback of this method is that the way the neighborhood style explanation is generated may not convey the reasoning behind the recommendation. Therefore, there could be cases where the explanation cannot be generated. Recall that the MF optimization problem is non-convex [4] and does not guarantee that the most similar users to an active user are necessarily those neighbors in latent space, who have liked a recommended item.

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
DOI: http://dx.doi.org/10.1145/2872518.2889405 .

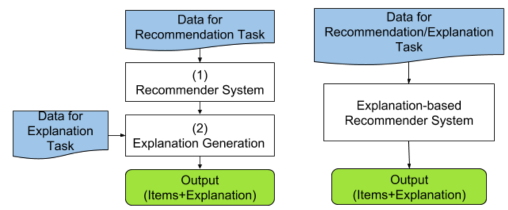

Figure 1: (a) Typical explanation system, (b) proposed explainable recommendation system.

2. EXPLAINABLE-MF (EMF)

Figure 1(a) shows the typical flow of most recommender systems that provide explanations [5], while Figure 1(b) shows the flow of the proposed integrated explanation and recommendation approach. In this approach, we learn a recommendation model that tries to optimize the items' explainability at the same time as the accuracy of the recommendations. In order to to do this, we first need to quantify *explainability.*

One way to formulate explainability is based on the rating distribution within the active user's neighborhood. If many neighbors have rated the recommended item, then this can provide a basis upon which to explain the recommendations, using neighborhood style explanation mechanisms. We thus propose a novel MF-based CF that leverages an explainability bipartite graph with arcs from the set of users onto the set of items. The graph adjacency is captured in an edge weight matrix, W, between user-item pairs, defined as follows:

$$W_{ij} = \begin{cases} \frac{|N'(i)|}{|N_k(i)|} & if \ \frac{|N'(i)|}{|N_k(i)|} > \theta \\ 0 & otherwise \end{cases} \tag{1}$$

where $N_k(i)$ is the set of k nearest neighbors of user i, $N'(i)$ is the set of user i's neighbors who have rated item j, and θ denotes an optional threshold above which we accept item j to be explainable for user i. Neighbors are calculated using the cosine similarity. The idea here is that if item j is explainable for user i, then their representations in the latent space, q_j and p_i, should be close to each other. With this rationale, the new objective function, to be minimized over the set of known ratings, can be formulated as:

$$J = \sum_{i,j \in R} (r_{ij} - p_i q_j^T)^2 + \frac{\beta}{2}(||p_i||^2 + ||q_j||^2) + \frac{\lambda}{2}(p_i - q_j)^2 W_{ij} \quad (2)$$

R is the set of user-item pairs for which the ratings are available, $\frac{1}{2}(||p_i||^2+||q_j||^2)$ is an $L2$ regularization term weighted by the coefficient β, and λ is an explainability regularization coefficient that controls the smoothness of the new representation and tradeoff between explainability and accuracy. To minimize the objective function, we use stochastic gradient descent and derive the following updates for p_i and q_j:

$$\begin{aligned} p_i &\leftarrow p_i + \alpha(2(r_{ij} - p_i q_j^T)q_j - \beta p_i - \lambda(p_i - q_j)W_{ij}) \\ q_j &\leftarrow q_j + \alpha(2(r_{ij} - p_i q_j^T)p_i - \beta q_j + \lambda(p_i - q_j)W_{ij}) \end{aligned} \quad (3)$$

3. EXPERIMENTAL RESULTS

We tested our approach on the MovieLens benchmark data[1] which has 100,000 ratings from 943 users on 1682 items, on a scale of 1 to 5. 10% of the ratings are selected randomly for the test set. Without loss of generality, we chose $\theta = 0$, which means that if at least one of the neighbors of user i have rated item j, then $W_{ij} > 0$.

We compare our results with five baseline methods: Non-Negative Matrix Factorization (NMF), Probabilistic Matrix Factorization (PMF), classical user-based and item-based top-n techniques, and non-personalized top-n most popular items. To assess the accuracy of EMF in terms of rating prediction, we used the Root Mean Squared Error (RMSE) and Normalized Discounted Cumulative Gain (nDCG@10) [3] metrics. Note that RMSE can be obtained for methods that predict ratings but not for top-n algorithms. Each experiment is run 30 times and the average results with varying number of latent factors, f, when $k = 10$, $\alpha = 0.001$, $\beta = 0.01$, and $\lambda = 0.1$ are reported in Figure 2, top row.

We measure explainability using the MEP and MER metrics. *Explainability Precision* (EP) is defined as the ratio of number of explainable recommended items to the number of recommended items for each user; Mean EP (MEP) is the average value of explainability precision over all users. Similarly, *Explainability Recall* (ER) is the ratio of number of explainable recommended items to the number of explainable items for each user; Mean ER (MER) is the mean explainability recall calculated over all users. Figure 2 shows MEP and MER results, for varying number of neighbors, k, when $f = 30$, $\alpha = 0.001$, $\beta = 0.01$, and $\lambda = 0.1$. The results in Figure 2, bottom row, indicate that EMF results in significantly better MEP and MER metrics compared to other baselines.

To study the effect of the explainability regularization coefficient, we varied λ, while fixing all the other parameters ($\alpha = 0.001$, $\beta = 0.01$, $k = 10$, and $f = 30$). Table 1 shows all metrics based on 5-fold cross validation. Increased regularization improves the explainability metrics (MER and MEP) while RMSE and nDCG@10 remain almost unchanged.

4. CONCLUSIONS

Our scope, in this work, was limited to CF recommendations where no additional source of data is used in recommendations or in explanations, and where explanations for recommended items can be generated from the ratings given to these items, by the active user's neighbors only. Thus explainability can be directly formulated based on the rating distribution within the active user's neighborhood.

[1]http://www.grouplens.org/node/12

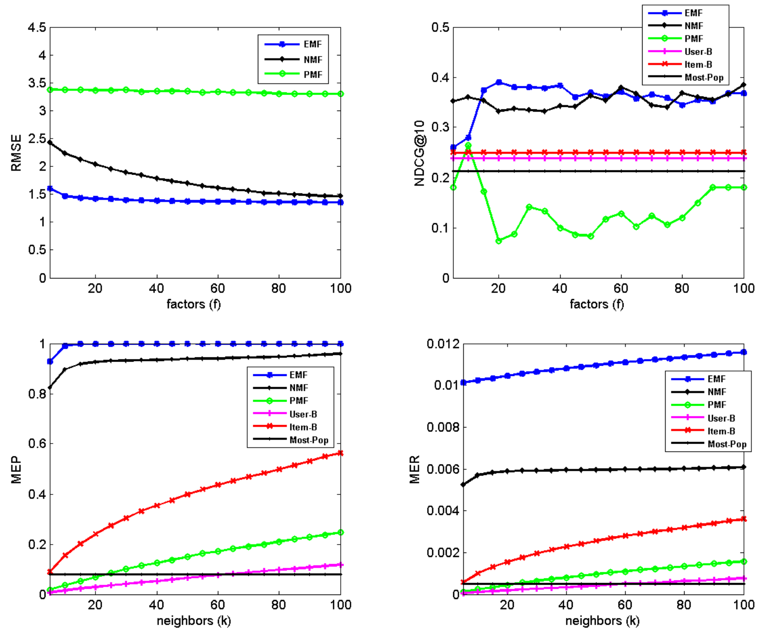

Figure 2: Top-left RMSE & top-right nDCG@10 *vs.* f. Bottom-left MEP & bottom-right MER *vs.* k.

Table 1: Performance of EMF *vs.* λ.

λ	Metrics			
	RMSE	nDCG@10	MER	MEP
0	1.3772	0.3578	0.0525	0.9932
0.01	1.3231	0.3608	0.0091	0.9939
0.05	1.3283	0.3652	0.0102	0.9964
0.1	1.3484	0.3503	0.0105	1
0.5	1.3256	0.3601	0.0128	1
1	1.3992	0.3741	0.0133	1
Avg.	1.3421	0.3587	0.0158	0.9956

We focused our research on CF methods which have been shown to have better serendipity than, and to outperform, Content Based (CB) methods [2]. We have incorporated user-based neighbor style explanations based only on the rating data and without using any additional external data. This is one main distinction of our approach compared to existing explanation approaches in the literature, which are, for this reason, not comparable on a fair basis.

5. ACKNOWLEDGMENTS

This research was partially supported by KSEF Award KSEF-3113-RDE-017.

6. REFERENCES

[1] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 241–250. ACM, 2000.

[2] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53, 2004.

[3] K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)*, 2002.

[4] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, (8):30–37, 2009.

[5] N. Tintarev and J. Masthoff. A survey of explanations in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, pages 801–810. IEEE, 2007.