Towards Optimal Active Learning for Matrix Factorization in Recommender Systems

Rasoul Karimi, Christoph Freudenthaler, Alexandros Nanopoulos, Lars Schmidt-Thieme
University of Hildeshiem
Samelsonplatz 1, University of Hildesheim, D-31141 Hildesheim, Germany
karimi,freudenthaler,nanopoulos,schmidt-thieme@uni-hildesheim.ismll.de

Abstract

This is the extended abstract of the paper which has already been published in the proceeding of IEEE ICTAI 2011 Conference (http://www.cse.fau.edu/ictai2011/). It applies active learning technique to new user problem in recommender systems.

1 Introduction

Recommender systems help web users to address information overload in a large space of possible options [1]. In many applications, such as in e-commerce, users have too many choices and too little time to explore them all. Moreover, the exploding availability of information makes this problem even tougher.

Collaborative filtering is the popular technique for recommender systems. Nevertheless, recent research (especially as has been demonstrated during the Netflix challenge[1]) indicates that Matrix Factorization (MF) is a superior prediction model compared to other approaches [2].

Evidently, the performance of collaborative filtering depends on the amount of information that users provide regarding items, most often in the form of ratings. However, a well identified problem is that users are reluctant to provide information for a large amount of items [3; 4]. This fact impacts negatively the quality of generated recommendations. A simple and effective way to overcome this problem, is by posing queries to new users in order that they express their preferences about selected items, e.g., by rating them. Nevertheless, the selection of items must take into consideration that users are not willing to answer a lot of such queries. To address this problem, *active learning* methods have been proposed to acquire those ratings from users, that will help most in determining their interests [4; 3].

2 Proposed Method

In this paper, we propose a novel method for applying active learning in recommender systems. Due to the rapidly increasing interest in MF as a powerful prediction model in recommender systems, the proposed method introduces an active learning approach designed to take into account the characteristics of MF in order to improve its accuracy. The proposed method is inspired from optimal active learning for regression problem. Assuming the distribution of the test data is known, it is possible to find the optimal active learning algorithm for specific regression models [5]. As MF is actually a regression problem, it makes sens to use the same approach for active learning in MF. Given the test items are known, we develop a method which approximates the optimal active learning for MF. It capitalizes on the updating mechanism of MF and allows us to formulate a new criterion for the selection of the queried items, in terms of reducing the expected prediction error. A detailed experimental evaluation is performed, whose results demonstrate the superiority of the proposed method. Our results provide insight into the effectiveness of the proposed criterion for selecting the queried items, as it compares favorably to methods that use MF but are based on simplistic criteria.

3 Experimental Result

In this section, we examine experimentally the performance of the proposed method.

3.1 Experimental set up

The main challenge in applying active learning for recommender systems is that users are not willing to answer many queries in order to rate the queried items. For this reason, we report the performance of all examined methods in terms of prediction error (MAE) versus the number of queried items, which is simply denoted as the *number of queries*. Non-myopic active learning [6], and random selection are used as the baseline.

We use the MovieLens(100K)[2] dataset in our experiments. MovieLens contains 943 users and 1682 items. The dataset was randomly split into training and test sets. The training dataset consists of 343 users (the same number used in [4]) and the rest of users are in the test dataset. Each test user is considered as a new user. The latent features of the new user are initially trained with three random ratings. 20 rated items of each test user are separated to compute the error. The test items are not new item and already appeared in the training data. The remaining items are in the pool dataset, i.e the dataset that is used to select a query. For simplicity, we assume that the new user will always be able to rate the queried item. In our experiment, 10 queries are asked from each new user. Therefore, the pool dataset should contain at east 10 items which exist in the training data. Considering 10 queries and 20 test items, each test user has given ratings to at least 30 items.

3.2 Results

Figure 1 illustrates the comparison between the proposed method, non-myopic active learning [6], and random selection in terms of MAE as a function of the number of

[1]www.netflixprize.com

[2]www.grouplens.org/system/files/ml-data0.zip

queried items. The non-myopic method works well in the first queries but it finally converges to the random selection. This convergence also happens for active learning in AM [3]. Generally, this evidence holds for active learning methods aiming to improve the new user parameters using some heuristics. In the optimization theory, usually the heuristics provide a good performance only if the difference between current solution and optimal solution is high. At first, as the new user has provided a few ratings, the new user parameters are inaccurate and are far away from the optimal parameters. But as more ratings are provided by the new user, the accuracy of the estimated parameters also increases and the heuristic-based methods do not gain much improvement. However, the proposed method in this paper has a different approach. It aims to directly optimize the test error. That is why its performance continues and does not converge to the random selection. Therefore, if the new user is ready to provide more ratings, the proposed method can efficiently use them to improve the accuracy.

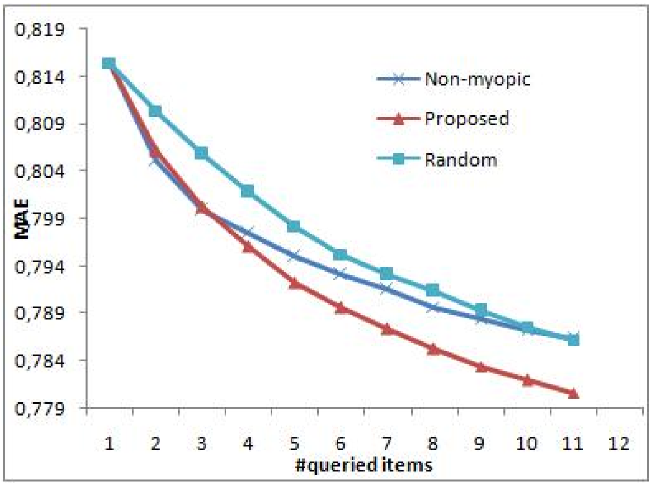

Figure 1: MAE results of the proposed active learning, non-myopic and random. Smaller MAE means better accuracy

References

[1] R. Burke, “Hybrid recommender systems,” *User Modeling and User Adapted Interaction*, vol. 12, no. 4, pp. 331–370, 2002.

[2] Y. Koren, R. Bell, and C. Volinsky, “Matrix factorization techniques for recommender systems,” *Computer*, vol. 42, pp. 30–37, 2009.

[3] R. Jin and L. Si, “A bayesian approach toward active learning for collaborative filtering,” in *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, 2004.

[4] A. S. Harpale and Y. Yang, “Personalized active learning for collaborative filtering,” in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2008, pp. 91–98.

[5] D. A. Cohn, G. Z., and M. Jordan, “Active learning with statistical models,” in *Advances in Neural Information Processing Systems(NIPS)*, 1995.

[6] R. Karimi, C. Freudenthaler, A. Nanopoulos, and L. Schmidt-Thiemee, “Non-myopic active learning for recommender systems based on matrix factorization,” in *IEEE Information Reuse and Integration (IRI)*. IEEE, 2011.