Information Sciences xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Regression-based three-way recommendation

Q1 Heng-Ru Zhang[a], Fan Min[a,*], Bing Shi[b]

[a] School of Computer Science, Southwest Petroleum University, Chengdu 610500, China
Q2 [b] College of Computer Science, Sichuan University, Chengdu 610065, China

A R T I C L E I N F O

Article history:
Received 19 August 2015
Revised 6 January 2016
Accepted 6 March 2016
Available online xxx

Keywords:
Granular computing
Numerical prediction
Promotion cost
Regression
Three-way decision

A B S T R A C T

Recommender systems employ recommendation algorithms to predict users' preferences to items. These preferences are often represented as numerical ratings. However, existing recommender systems seldom suggest the appropriate behavior together with the numerical prediction, nor do they consider various types of costs in the recommendation process. In this paper, we propose a regression-based three-way recommender system that aims to minimize the average cost by adjusting the thresholds for different behaviors. This is undertaken using a step-by-step approach, starting with simple problems and progressing to more complex ones. First, we employ memory-based regression approaches for binary recommendation to minimize the loss. Next, we consider misclassification costs and adjust the approaches to minimize the average cost. Finally, we introduce coupon distribution action with promotion cost, and propose two optimal threshold-determination approaches based on the three-way decision model. From the viewpoint of granular computing, a three-way decision is a good tradeoff between the numerical rating and binary recommendation. Experimental results on the well-known MovieLens data set show that threshold settings are critical to the performance of the recommender, and that our approaches can compute unique optimal thresholds.

1. Introduction

Recommender systems have been studied extensively to manage items, such as movies [8,22,24] and music [1,49,70]. One of the most successful technologies for recommender systems is memory-based collaborative filtering (CF) [16], which uses a database of user preferences to predict additional topics or products that may appeal to a new user. These preferences are typically expressed as numerical ratings. Many CF approaches have been designed to minimize mean absolute error (MAE) [55]. However, as indicated in [16], minimizing MAE can produce a so-called "magic" barrier, where natural variability prevents obtaining good accuracy. In practice, the aim of recommender systems is to present to the user a reasonable suggestion rather than a numerical prediction.

Granular computing is a general computational theory for using granules such as classes, clusters, subsets, groups, and intervals to build an efficient computational model for complex applications [58]. Rough set is a leading special case of granular computing approach [30]. Three-way decision [31,32,64,68] is an extension of decision theoretical rough sets [57,62,69] for dealing with situations in which three different decisions can be made, namely, accept, reject, and wait-and-see. Within the trisecting-and-acting framework [67], three-way decision is described as two separated tasks of trisecting

Q3 * Corresponding author. Tel.: +86 135 4068 5200.
E-mail address: minfanphd@163.com (F. Min).

http://dx.doi.org/10.1016/j.ins.2016.03.019
0020-0255/© 2016 Published by Elsevier Inc.

and acting. With respect to trisecting [67], a universal set is divided into three regions as regions I, II, and III, respectively. With respect to acting [67], there are strategies I, II, and III, respectively. Recently, there is a trend to applying three-way decision to different applications, such as email spam filtering [77], risk decision making [27], face recognition [26], concept lattices [42] and recommender system [2,74].

In this paper, we propose a regression-based three-way recommender system, the aim of which is to minimize the average cost by adjusting the thresholds for different behaviors. We are essentially dealing with three problems, where the last problem is more general than the first. The first problem is regression-based binary recommendation. The regression subtask is fulfilled using the slope one [25] or k-nearest neighbors (kNN) [44] algorithm to predict the ratings. To convert the numerical prediction into a binary recommendation, a threshold is needed. An item with an above-threshold rating is recommended, while one with a below-threshold rating is not. We design a threshold learning approach to determine the threshold r_t^* minimizing classification loss.

The second problem involves misclassification costs [14,23] corresponding to incorrect recommendation behavior, including recommending items to users who dislike them, and non-recommending items to users who like them. In existing works, misclassification cost is the most widely considered cost since classification is one of the main tasks in data mining (see, e.g., [12,20,78]). Because misclassification costs are considered, the work essentially involves cost-sensitive learning [14,34–36,79]. A cost-sensitive learning approach is designed to determine the optimal threshold r_t^c according to the misclassification costs. Naturally, the objective is to minimize the average misclassification cost.

The last, but crucial problem introduces the coupon distribution action, including promotion cost, to enrich recommender behavior. Promotion cost derives from consultation with the user about the actual decision. We propose optimal threshold determination approaches based on the three-way decision model. This kind of decision often begins with a cost matrix including misclassification and delay costs. In our scenario, we consider promotion cost instead of delay cost. Consequently, we have three actions, namely, recommend, non-recommend, and promote. Determining the threshold pair (r_l^*, r_h^*) involves three steps. First, two parameters, α^* and β^*, are computed according to the cost matrix. Second, the probability PR that the user likes an item is predicted using the slope one or kNN algorithm. Third, the threshold pair is determined based on α^*, β^*, and PR. If the prediction for an item is greater than r_h^*, the item is recommended to the user, while a prediction less than r_l^* results in the item not being recommended. Otherwise, we consider user tendency, which incurs promotion cost.

In our scenario, numerical prediction is exceedingly fine for the recommendation, while binary recommendation is rather coarse, with three-way decision a good tradeoff between these. From the viewpoint of granular computing [29,53,56,59,72], three-way decision has good granularity.

Experimental results, obtained using the well-known MovieLens data set (http://www.movielens.org/), show that: 1) the loss of regression-based binary recommendation (where the minimum loss of the slope one algorithm is obviously lower than that of the kNN one) is a convex function with respect to threshold r_t, and has a unique minimum; 2) the misclassification cost settings directly influence the optimal setting of the recommendation threshold r_t^c, where the average cost considering unequal misclassification costs is obviously lower than that considering equal misclassification costs; and 3) the optimal threshold (r_l^*, r_h^*)-pair determined by three-way decision is optimal not only on the training set, but also on the testing set. With the introduction of promotion cost, the three-way approach often achieves a significantly lower average cost compared with the two-way approach.

The rest of the paper is organized as follows. Section 2 presents some preliminary knowledge including the rating system and memory-based recommendation. Sections 3–5 discuss regression-based binary recommendation, misclassification cost minimizing recommendation, and three-way-decision-based recommendation, respectively. Section 6 presents the experimental results on the MovieLens data set for the three models. Finally, our conclusions are given in Section 7.

2. Related works

Collaborative filtering recommender systems usually use the rating system as input, and recommender accuracy as a kind of evaluation metric. Our recommendation behavior considers both misclassification and promotion costs. Through cost-sensitive learning, we build proper classifiers to find the minimum average cost.

2.1. Rating system

First, we revisit the rating system proposed in [75]. Let $U = \{u_0, u_1, \ldots, u_{n-1}\}$ be the set of users of a recommender system and $V = \{t_0, t_1, \ldots, t_{m-1}\}$ be the set of all possible items that can be recommended to users. Then, the rating function is given by

$$R : U \times V \to V_k, \tag{1}$$

where V_k is the rating domain used by the users to evaluate items, and r_w and r_g are the lowest and highest ratings, respectively. For convenience, we represent the rating system with an $n \times m$ rating matrix $R = (r_{i,j})_{n\times m}$, where $r_{i,j} = R(u_i, t_j)$, $0 \le i \le n-1$, and $0 \le j \le m-1$.

Example 1. An example rating system is depicted in Table 1, where $V_k = \{1, 2, 3, 4, 5\}$. In Table 1, some elements are zero, indicating that the users do not watch the corresponding movies.

Please cite this article as: H.-R. Zhang et al., Regression-based three-way recommendation, Information Sciences (2016), http://dx.doi.org/10.1016/j.ins.2016.03.019