Non-Compensatory Psychological Models for Recommender Systems

ID: 335

Abstract

The study of consumer psychology reveals two categories of procedures used by consumers to make consumption related choices: compensatory rules and non-compensatory rules. Existing models assume the consumers follow the compensatory rules, which are to make decisions based on a weighted or summated score over different aspects. Our main contribution in this paper is to improve performance of recommender systems by adopting non-compensatory rules to make ranking decisions. We present non-compensatory versions for three most commonly adopted ranking models in this area, i.e. BPR, BTL and SVD++. We show that, the non-compensatory versions all outperform their original models on a wide range of real data sets.

Introduction

Ever since the dawn of psychological science, many attempts have been made to explain human choice behavior. It is well regarded that there are two categories of procedures to make consumption related choices: compensatory rules and non-compensatory rules (**?**). Consumers who adopt compensatory rules evaluate every product over multiple relevant aspects and compute a weighted or summated score for each product. Then they will select the product with the highest score. The shortcomings of a product are balanced out by its attractive features. On the contrary, non-compensatory rules do not allow a good performance on one aspect of a product to compensate for poor performances on other aspects. Non-compensatory rules include lexicographic, conjunction and disjunction rules. Under lexicographic rules, products are compared on the most important aspect. Under conjunction and disjunction rules, the consumer imposes requirements for minimally acceptable values on each aspect separately. The conjunction and disjunction rules are often used in conjunction with lexicographic rules.

Related Work

Compensate Recommendations Models

Most successful recommendation models are built upon latent factor models, i.e. they assume a universe of K factors in which the user preferences and item features are represented. Hereafter, unless stated otherwise, we use lower-case letters for indices, upper-case letters for universal constants, lower-case bold-face letters for vectors and upper-case bold-face letters for matrices. Specifically, $\mathbf{X} \in \mathcal{R}^{M \times N}$ denotes the observation matrix, $\mathbf{p}, \mathbf{q} \in \mathcal{R}^K$ denotes the item features, $\mathbf{u} \in \mathcal{R}^K$ denotes the user features.

Point-wise Prediction Model

MF In matrix factorization, the predicted rating can be approximated by an inner product of $\hat{\mathbf{X}_{u,p}} = \sum_k \mathbf{p}_k \mathbf{u}_k$

Factorization with Neighborhood As in (**?**), in this paper, we adopt the asymmetric matrix factorization (AMF) (**?**), each item i is associated with two factor vectors q_i and p_i. The representation of user u is through the sum $\sum_{j \in R(u)} p_j / \sqrt{|R(u)|}$, where $R(u)$ is the set of positive items.

$$\hat{r_{ui}} = u + b_u + b_i + q_i(\sum_{s \in R(u)} p_s / \sqrt{|R(u)|}) \qquad (1)$$

Local Low-Rank Matrix Approximation. The intuition behind the LLORMA is that the entire rating matrix M is not low-rank but a submatrix M_s, restricted to certain types of similar users and items. (**?**) minimizes a pair-wise loss function, which is constructed using observed entries. In the experiments (**?**), the log-loss function is shown to achieve best results, thus it is adopted in this paper.

$$L = \sum_{u \in U} \frac{1}{s_u} \sum_{k=1}^{s_u} \Delta M_{u,i,j} \log(1 + \exp -g(u,i,j)), \qquad (2)$$

where s_u is the number of ordered items rated by the user u. The loss function approximates the observed rating difference with predicted rating difference $g(u,i,j)$. LCR is inspired by LLORMA, which assumes that the rating matrix itself is not low-rank. Instead, M is locally low-rank where locality is defined by a neighborhood with respect to the given kernel K. Therefore, the prediction difference can be approximated by a set of low-rank factorizations around q anchor points: $g(u,i,j) = \sum_{t=1}^{q} \sum_k u_{t,k} [\frac{K((u_t,i_t),(u,i))}{\sum_{s=1} qK((u_s,i_s),(u,i))} i_{t,k} - \frac{K((u_t,j_t),(u,j))}{\sum_{s=1}^{q} K((u_s,j_s),(u,j))} j_{t,k}]$.[1]

Copyright © 2019, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

[1] This equation is not the same as Equ.15 in the paper

Ranking Models

point-wise loss function
pair-wise loss functions
Bradley-Terry Model
Thurstone Model

Table 1: default

Modeling

Non-Compensatory Ranking Models for Recommendations

Experiments