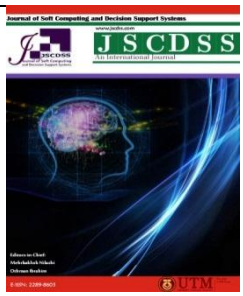
An International Journal
JSCDSS
http://www.jscdss.com
Vol.3 No.5 October 2016: 24-30
Article history:
Accepted 10 July 2016
Published online 10 July 2016

Journal of Soft Computing and Decision Support Systems



E-ISSN: 2289-8603

A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques

Mehrbakhsh Nilashi [a,b,*], Mohammad Dalvi-Esfahani [a], Morteza Zamani Roudbaraki [b], T. Ramayah [c], Othman Ibrahim [a]
[a] Faculty of Computing, Universiti Teknologi Malaysia, Malaysia
[b] Department of Computer Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran
[c] Department for Operations Management Section at the School of Management, USM, Malaysia

*** Corresponding author email address**: nilashidotnet@hotmail.com

Abstract

Traditional Collaborative Filtering (CF) recommender systems recommend the items to users based on their single ratings which are used to match similar users. In multi-criteria CF recommender systems, however, multi-criteria ratings are used instead of single ratings which can significantly improve the accuracy of traditional CF algorithms. This research proposes a new recommendation method using Classification and Regression Tree (CART) and Expectation Maximization (EM) for accuracy improvement of multi-criteria recommender systems. We also apply Principal Component Analysis (PCA) for dimensionality reduction and to address multi-collinearity induced from the interdependencies among criteria in multi-criteria CF datasets. Experimental results on Yahoo! Movies and TripAdvisor datasets demonstrated that the proposed method significantly improves recommendation accuracy of multi-criteria CF.

Keywords: Multi-criteria recommender systems, Accuracy, CART, Collaborative Filtering

1. Introduction

During the last decade the amount of information available online increased exponentially and information overload problem has become one of the major challenges faced by Information Retrieval (IR) and Information Filtering (IF) systems. Recommender systems are one solution to the information overload problem. They support the online customer in his/her decision making and buying process (Bagherifard et al., 2013; Nilashi et al., 2013; Nilashi et al., 2014a; Nilashi et al., 2014b; Vahid et al., 2016). Recommender systems based on Collaborative Filtering (CF) are particularly popular and used by large online retailers (Nilashi et al., 2015a; Nilashi et al., 2015b; Farokhi et al., 2016).

The ratings provided by users for items are the key input to CF recommender systems. They present information regarding the quality of the item along with the preference of the user who shared the rating. Principally, the large numbers of recommender systems are developed for single-valued ratings. For example they consider single rating for each user and item that indicates how much the item is liked by the user in system. However, another idea is multiple components (criteria) that sometimes recommender systems are design by them. For example, the Zagat system considers rates of restaurants on four components for its customer's servings as food, decor, services and cost. Correspondingly, in a Yahoo! movies recommender system, four criteria such as Acting, Visual Effects, Direction, and Story could be rated by users. The objective of a CF algorithm is to recommend new products or to estimate the utility of a certain product for a specific user depending on the customer's past likings and the views of other like-minded users. There are two tasks that a CF can perform, leading to two unique kinds of result. The first task is the rating prediction process which is predicting the rating that a given unseen product will have for the target user. Recommendation task as a second task in CF provides top-N recommendation contains of unseen relevant items for the target user.

CF algorithms can be divided into two categories: memory-based algorithms and model based algorithms (Nilashi et al., 2013; Adomavicius and Tuzhilin, 2005; Deshpande and Karypis, 2004). Memory-based (also called neighborhood-based) algorithms are also known as lazy recommendation algorithms, because they defer the actual computational effort of predicting a user's interest in an item to the moment a user requests a set of recommendations. The training phase of a memory-based algorithm consists of simply storing all the user ratings into memory. The idea of memory-based methods is that the rating predictions for a user directly depend on his/her similar users' ratings on similar items. Since the entire rating database is kept in memory, new ratings can