



# Learning from explanations in recommender systems



Sergio Cleger<sup>a,1</sup>, J.M. Fernández-Luna<sup>b</sup>, Juan F. Huete<sup>b,\*</sup>

<sup>a</sup>Departamento de Informática, Facultad de Informática y Matemática, Universidad de Holguín, Cuba

<sup>b</sup>Departamento de Ciencias de la Computación e Inteligencia Artificial, E.T.S.I. Informática y de Telecomunicación, CITIC-UGR Universidad de Granada, 18071 Granada, Spain

## ARTICLE INFO

### Article history:

Received 27 June 2013

Received in revised form 11 June 2014

Accepted 24 July 2014

Available online 1 August 2014

### Keywords:

Recommender system

Recommendation explanation

Learning

## ABSTRACT

Although recommender systems are extremely useful when it comes to recommending new products to users, it is important that these applications explain their recommendations so that users can consider and trust them. There is also another reason: by analyzing why the system recommends a particular item or proposes a certain rating, the user might also consider the quality of the recommendation and, if appropriate, change its predicted value. On this basis, this paper presents a new technique to improve recommendations based on a series of explanations that should be given when various already-known items are recommended. To summarize, the aim of our proposal is to learn a regression model from the information presented in these explanations and, where appropriate, use this model to change the recommendation for a target item. In order to test this approach, we experimented with the MovieLens data set. A number of lessons can be learned: firstly, it is possible to learn from a set of explanations, although this is highly user-dependent; and secondly, we can use an automatic procedure to analyze the role of the different features presented in an explanation. We consider these results to be interesting and to validate our novel approach.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

Ever since they were first conceived in the last decade of the 20th century [10,32,33], recommender systems (RS) have been used in many different areas. There are currently a large number of RSs on Internet sites such as Amazon, Netflix or Last.fm. These systems are designed to learn from a user's behavior to discover their preferences and either help them find what they are specifically looking for or what they might find useful in a vast amount of information.

Despite the popularity of such systems, there are various reasons why users might feel uneasy about using them and/or relying on their recommendations. These systems are usually seen as black boxes in which there is no other choice than to trust their recommendations [14]. Much research has been conducted to focus on this problem and among the proposed solutions we focus our attention on the use of explanation facilities which have also been considered as a critical function for recommender systems [8,9,20,36]. These facilities provide users with an explanation of the rationale behind such recommendations, enabling them to better understand the workings of the system. Tintarev and Masthoff [41] list a number of reasons why recommender systems should provide such an explanation and these include:

\* Corresponding author.

E-mail addresses: [sergio@facinf.uho.edu.cu](mailto:sergio@facinf.uho.edu.cu) (S. Cleger), [jmfluna@decsai.ugr.es](mailto:jmfluna@decsai.ugr.es) (J.M. Fernández-Luna), [jhg@decsai.ugr.es](mailto:jhg@decsai.ugr.es) (J.F. Huete).

<sup>1</sup> Visiting Academic at the Universidade do Estado do Amazonas, Brazil.

1. Transparency (illustrates how a recommendation is computed).
2. Trust (increases user confidence).
3. Effectiveness (helps the user make a good decision).
4. Scrutability (allows users to tell the system it is wrong).

An algorithm's ability to explain both recommendation results and the explanation itself strongly depends on the recommender model used, more especially so in the case of memory-based approaches [36] which are based on the similarities extracted from known data and which can serve as the basis for explaining recommendation results. In this paper, we shall assume a nearest-neighborhood-based RS in order to predict the rating for a target item. In this recommendation paradigm, there are two good alternatives for explaining recommendations [15]: (i) the use of a bar chart histogram which explains that "the system suggests 3 stars because it has been rated by other similar users as ..."; and (ii), based on the system's previous performance, i.e. "The system correctly recommended for you 80% of the time". This second alternative is valuable for increasing user trust but does not fulfill any of the other criteria.

Focusing on the first more informative alternative, this partly explains *how* the predicted rating is computed (transparency), for instance by combining the ratings into a weighted average prediction. In this case, we can obtain several explanations for the same recommended rating, let us say 3. If we assume that there are five neighbors (who each contribute equally to the prediction), this rating can either be obtained when the suggested ratings are {1, 2, 3, 4, 5} or when the ratings are {3, 3, 3, 3, 3}. Users might find both explanations useful but while the user is unsure of the quality of the prediction in the first, the explanation reinforces the recommendation in the second (this is related to trust). In the same way, we can consider a situation where the system predicts the value 4 because the neighbors rated the target item as {1, 5, 5, 5, 5}. The information given in the explanation can then be used to change the prediction by modifying the rating to 5 (fact related to scrutability).

In our previous research [4], we found that having observed an explanation, users often modified their proposed rating (around 35% of the time). Moreover, better predictions about the quality of the proposed ratings were obtained once a large number of explanations had been observed. Certain learned processes must exist in the user's mind to cause them to act on the prediction. In this paper, we talk about such an action in terms of trusting, mistrusting or even possibly changing the predicted rating. This situation underlies our research proposal, which can be formulated as the following question: *Would it be possible to learn from a set of explanations?* If the answer is "yes", we must explore *how the learned information can be used to improve recommendations*. With the aim of learning from explanations, we propose the use of a machine-learning [13] algorithm that is capable of inducing general rules about the user's actions from a set of observed instances.

In order to obtain the set of training instances, a first approach might be to gather the required information from user experience of the system, by either explicitly or implicitly analyzing their actions over the observed explanations. This task involves various problems: (a) user burden: we need to know what action the user takes after observing the explanation for each recommended item even when they have no interest in it; (b) dependence on the recommendation strategy: when considering a neighborhood-based RS, it could be that the set of neighbors used to obtain the predictions (and their associated explanation) varies with the recommended items; and (c) the cold-start problem: we need enough feedback to start the learning process.

In order to solve this problem, we propose that it be tackled from a different perspective: let us consider a given (unobserved) target item,  $I_t$ , where  $N$  is the set of neighbors selected to compute the pair prediction/explanation. We will use *this* neighborhood to obtain a prediction for all the items rated by the user, let us say  $\mathcal{R} = \{I_1, \dots, I_k\}$ , and their associated explanations. For these items, since we know the real and predicted ratings, we can automatically presume what the user's actions should be, such as, for instance, to trust or increase the predicted value. Using this approach, we shall obtain a corpus of tuples (predictions, explanations and actions) that can be used as training data to induce the particular action for the target item. We shall say that this approach is based on previous explanations.

We would like to stress that the main focus of our paper is to determine whether the general idea of learning from explanations may work. As a result, we conclude that explanations represent a valuable source of information that might be exploited since this is (as far as we know) the first attempt to explore their impact on a recommending environment. By means of this approach, we should explore the costs and benefits of scrutability for improving recommendations, something which is considered to be an open problem in the field [21,23,42]. Although our approach focuses on a neighborhood-based RS, it could also be applied to other recommending strategies, such as those using matrix factor [22] or additional informational elements such as the quality of the items [40] or social-based knowledge [35].

This paper is organized as follows: the next section discusses the reasons for our approach; Section 3 presents related work on recommender systems; Section 4 describes our approach; Section 5.1 outlines empirical experimentation with MovieLens data sets; and finally Section 6 presents our concluding remarks.

## 2. Explaining by using predictions for already known items

As we mentioned previously, our research stems from the use of predictions on a set of observed items which serve as the basis for explanation facilities [4]. More specifically, this research considered a collaborative RS to predict how an active user,  $a$ , might rate a target item,  $t$ ,  $\hat{r}(a, t)$ . Following a user-based approach, the predictions are obtained using a weighted combination of the suggestion given by the nearest  $N$  neighbors, i.e.

Download English Version:

<https://daneshyari.com/en/article/393378>

Download Persian Version:

<https://daneshyari.com/article/393378>

[Daneshyari.com](https://daneshyari.com)