A Highly Automated Recommender System Based on a Possibilistic Interpretation of a Sentiment Analysis

Abdelhak Imoussaten[1,⋆], Benjamin Duthil[2], François Trousset[1], and Jacky Montmain[1]

[1] LGI2P, Ecole des Mines d'Alès,
Site EERIE Parc Scientifique G. Besse, 30035 Nîmes Cedex, France
[2] L3i, Université de La Rochelle,
Avenue M. Cré peau, 17042 La Rochelle cedex 01, France
{abdelhak.imoussaten,francois.trousset,jacky.montmain}@mines-ales.f,
benjamin.duthil@univ-lr.fr

Abstract. This paper proposes an original recommender system (RS)[1] based upon an automatic extraction of trends from opinions and a multicriteria multi actors assessment model. Our RS tries to optimize the use of the available information on the web to reduce as much as possible the complex and tedious steps for multicriteria assessing and for identifying users' preference models. It may be applied as soon as i) overall assessments of competing entities are provided by trade magazines and ii) web users' critics in natural languages and related to some characteristics of the assessed entities are available. Recommendation is then based on the capacity of the RS to associate a web user with a trade magazine that conveys the same values as the user and thus represents a reliable personalized source of information. Possibility theory is used to take account subjectivity of critics. Finally a case study concerning movie recommendations is presented.

Keywords: Possibility theory, Intervals merging, Multicriteria aggregation, Recommender system, Opinion-mining.

1 Introduction

In recent years, many companies and web sites have set up systems to analyze the preferences of their users in order to better meet their expectations. To date, recommendation systems are present in many areas such as tourism /leisure, advertising, e-commerce, movies, etc. Due to the exponential growth of the quantity of data available on the Internet in recent years, searching and finding products, services and relevant contents become a difficult task for the user often drowned out by the mass of information. This explains the growing interest in recommendation systems (RS) both by users as by commercial sites.

⋆ Corresponding author.

[1] This work is an extension of [1].

A. Laurent et al. (Eds.): IPMU 2014, Part I, CCIS 442, pp. 536–545, 2014.

The task of recommendation has been identified as a way to help users to find information, or elements that are likely of interest. Roughly speaking, we consider a set of users and a set of items (products or services) that may be recommended to each user. In addition, a multicriteria recommendation improves the quality of a RS because it makes explicit the characteristics for which an item was proposed to the user [2] [3]. A RS that takes advantage of evaluation related to multicriteria preference elicitation provides users with more relevant detailed recommendations [2]. However, the implementation of such a model requires a knowledge base where the items are evaluated *w.r.t* a set of criteria. This constraint imposed by the model is very heavy for the user.

In [4], an unsupervised multicriteria opinion mining method is proposed. It allows users to free themselves from constraining partial evaluations *w.r.t* each criterion: users simply submit their critics in natural language to express their opinions, and system analyzes them automatically. It first dissects the critics according to the evaluation criteria (thematic segmentation) before calculating the polarity or opinions of each of the extracts resulting from the segmentation step (opinion-mining/sentiment analysis).

Combining this method with an interactive multicriteria decision support system makes possible to have a highly automated system for recommendation purposes. However, the automated assignment only provides imprecise scores related to items that are modeled by intervals in an adequate multicriteria analysis process. Our possibility theory based approach then manages multiple imprecise assessments derived from sentiment analysis on each evaluation criterion (intervals fusion) and then aggregate them on all criteria. Finally, we try to match a user and an adequate specialized magazine in the domain of concern (movies in our application) that will provide the most suitable personalized recommendation to the user.

Section 2 summarizes opinion mining approach to extract Internet user's critics and compute opinion scores on a set of criteria. Section 3 explains how to merge these imprecise opinion scores for each criterion and introduces the notion of matching of a distribution with data available on a criterion. Section 4 describes how to deduce the multicriteria model used by a specialized magazine to assess items. Section 5 shows how these approaches can be combined to address the multicriteria recommendation problem in the case study of movie recommendations.

2 Opinion-Mining Process

On the basis of statistical methods, the Opinion-mining approach of [4] allows us to build a lexicon of opinion descriptors for a given thematic. This lexicon is used to automatically extract the polarity of text segments that are related to the criterion. Two stages are distinguished in this multicriteria evaluation: - firstly, the segments of text related to each of the evaluation criteria are extracted with the *Synopsis* approach described in [5]. The text is first segmented into criteria. Then, for each criterion, the polarities of the segments that have been identified