Decision Tree Induction Techniques for E-Commerce Recommendation Systems

Jong Woo Kim
Chungnam National University, Department of Statistics
220 Kung-dong, Yusung-gu
Taejon, 305-764 KOREA
jwkim@stat.cnu.ac.kr

SeokHoon Lee
Chungnam National University, Department of Statistics
220 Kung-dong, Yusung-gu
Taejon, 305-764 KOREA
shlee@stat.cnu.ac.kr

Ha Bin Lee
Paichai University, Department of Electronic Commerce
436-9 Doma-dong, Seo-gu, Taejon, 302-735 Korea
hanbin@mail.paichai.ac.kr

1. Recommendation Techniques

Personalized recommendation techniques suggest products or services to a customer of Internet storefronts based on demographics of the customer or based on an analysis of the past purchasing behavior of the customer. Current available recommendation tools are mainly based on collaborative filtering or rule-base approaches. There are several commercial back-end products for personalization services including BroadVision's One-to-One, Net Perceptions's E-Commerce, and Andromedia's Likeminds[1].

In the rule-based approach, recommendation rules from marketing experts are a core component to provide personalized advertisement. It is however difficult to extract recommendation rules from marketing experts, and also difficult to validate and to maintain the constructed knowledge base. In this study, we propose a marketing rule extraction technique for personalized recommendation on Internet storefronts using decision tree induction techniques[2].

2. Recommendation Rule Extraction and Personalization

In order to extract useful recommendation rules, we first define a hierarchy tree of product categories to extract recommendation rules at various levels of product categories. Learning data set

includes customer purchase data and customer profile (demographic information and initial product category preference). The process of recommendation rule extraction consists of four steps, (1) target variable generation, (2) data partitioning, (3) decision tree construction, (4) recommendation rule selection. The process of rule induction begins with purchase history data, which is used to construct target variables based on historical levels of purchases in different product categories. Decision trees are then constructed for each target variable using data mining techniques, and a filter is applied to preserve only decision trees above a threshold validation value. Branches in the filtered decision trees, i.e. alternative recommendation rules are filtered again based on their performance on learning data set. The filtered rules are then encoded in an expert systems, and the rules then are applied to new customers, based on their category preference and demographic data, to select advertisements for display.

3. Experiment

An experiment has been performed to evaluate the effectiveness of the proposed approach with preference scoring, random selection, and collaborative filtering. Using an Internet survey, data from 330 respondents was gathered. The survey was designed to elicit product choice and advertising evaluation judgments for two product types: MP3 music files, and Sporting and Leisure Goods. The results of the experiment shows that in the case of sporting goods and leisure equipment, decision tree induction techniques give best results among the four techniques. But, in the case of MP3 music files, the preference scoring techniques give best results among the four techniques.

REFERENCE

[1] Schafer, J. Ben, Joseph Konstan, and J. Reidl, "Electronic Commerce Recommendation Applications," *Journal of Data Mining and Knowledge Discovery*, January, 2001.
[2] Quinlan, J.R., "Induction of Decision Trees," in Reading in Machine Learning (edited by Chavlik, Jude W. an dDietterich, Tomas G.), Morgan Kaufmann Publishers, Inc, CA, 1990.

RESUME

Jong Woo Kim, Assistant Professor at Department of Statistics, Chungnam National University
SeokHoon Lee, Professor at Department of Statistics, Chungnam National Universtiy
Ha Bin Lee, Full-time Instructor at Department of Electronic Commerce, Paichai University