

Interactive Marketing Data Analysis Using Inductive Learning, Genetic Algorithms, and Pairwise Comparison

Takao TERANO ⁽¹⁾, and Yoko ISHINO ⁽²⁾

(1) Graduate School of Systems Management, University of Tsukuba, Tokyo
3-29-1 Otsuka, Bunkyo-ku, Tokyo 112, Japan,
terano@gssm.otsuka.tsukuba.ac.jp

(2) Interdisciplinary Course on Advanced Science and Technology,
The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153, Japan, ishino@ai.rcast.u-tokyo.ac.jp

Abstract: In this paper, to analyze questionnaire data on consumer goods for marketing decision making, we use *inductive learning*, *genetic algorithms* and the *Analytic Hierarchy Process*. The basic idea of the method is to integrate *inductive learning* to acquire decision trees or sets of decision rules, *genetic algorithms* to get the effective features and *the AHP* to relatively evaluate each acquired knowledge in order to develop simple, easy-to-understand, and accurate knowledge from noisy data. The unique characteristics of the method is that the offspring (decision trees) are evaluated by human-in-a-loop phase with *simulated breeding* and *pairwise comparison* methods. The proposed method has been qualitatively and quantitatively validated by a case study on consumer product questionnaire data of 2,400 entries with 16 attributes.

Marketing decision makers must develop promotion strategies from noisy questionnaire data. Unlike popular learning-from-example methods, in such tasks, they must interpret the characteristics of the data without clear features of the data nor pre-determined evaluation criteria. AI-based inductive learning techniques for concept learning seem promising, however, they involve combinatorial feature selection problems. Feature selection is the problem in machine learning and statistical researches to choose a small subset of features that is necessary and sufficient to describe the target concept.

To overcome the problem, we have developed a novel method SIBILE, which stands for *Simulated Breeding and Inductive LEarning* method, to acquire efficient decision rules from questionnaire data using both *simulated breeding* and inductive learning techniques. *Simulated Breeding* is one of the GA-based techniques which subjectively or interactively evaluate the qualities of offspring generated by genetic operations with human-in-a-loop manner.

In this paper, we employ the *pairwise comparison* technique used in the AHP, which enables a marketing decision maker or the user to easily evaluate decision trees or sets of decision rules in order to to develop offspring with 'effective' features. By the word 'effective', we mean the resulting decision trees or sets of rules are simple, accurate, and easy-to-understand.

The experimental results have suggested that the proposed integrated method is superior to the conventional methods (e.g., both statistical ones and the older version of SIBILE) from the points of the effectiveness of the knowledge and the easiness of the usage. The pre-requisites of the proposed method are quite simple and the algorithm is easy to implement. Future directions of the work include to generalize SIBILE for a portable tool applicable to the other decision making problems and to specify SIBILE for efficient decision knowledge acquisition by improving the inductive learning techniques.