Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 56–64
Expert Systems with Applications
www.elsevier.com/locate/eswa

Rough Set Theory in analyzing the attributes of combination values for the insurance market

Jhieh-Yu Shyng [a,*], Fang-Kuo Wang [b,c], Gwo-Hshiung Tzeng [d], Kun-Shan Wu [e]

[a] *Department of Management Information Science, Lan-Yang Institute of Technology, No. 79, Fu-Shin Rd., To-Chen, I-Lan 621, Taiwan*
[b] *Department of Risk Management and Insurance, Ming-Chuan University, No. 250, Sec. 5, Chong-San N. Rd., Taipei 111, Taiwan*
[c] *Graduate Institute of Management Sciences, Tamkang University, 151 Ying-chuan Road, Tamsui, Taipei 25137, Taiwan*
[d] *Department of Business Administration, Kainan University, No. 1 Kainan Rd., Luchu, Taoyuan County 338, Taiwan*
[e] *Department of Business Administration, Tamkang University, 151 Ying-chuan Road, Tamsui, Taipei 25137, Taiwan*

Abstract

Based on Rough Set Theory, this research addresses the effect of attributes/features on the combination values of decisions that insurance companies make to satisfy customers' needs. Attributes impact on combination values by yielding sets with fewer objects (such as one or two objects), which increases both the lower and upper approximations. It also increases the decision rules, and degrades the precision of decisions. Our approach redefines the value set of attributes through expert knowledge by reducing the independent data set and reclassifying it. This approach is based on an empirical study. The results demonstrate that the redefined combination values of attributes can contribute to the precision of decisions in insurance marketing. Following an empirical analysis, we use a hit test that incorporates 50 validated sample data into the decision rule so that the hit rate reaches 100%. The results of the empirical study indicate that the generated decision rules can cover all new data. Consequently, we believe that the effects of attributes on combination values can be fully applied in research into insurance marketing.

Keywords: Rough Set Theory; Combination values; Insurance marketing; Decision rule; Expert knowledge

1. Introduction

An understanding of customer consumption trends, and purchasing intentions and behavior is very important in sales and marketing. It is well known that customer satisfaction with products or services is an important key to achieving successful business operations and sustaining competition. Customer satisfaction is a critical issue in keeping customers, so measuring the impact of buying behavior on customer satisfaction is critical to understanding customer needs. Customers' repurchase intentions and the retention of customers is driven by customer satisfaction (Kim, Ferrin, & Rao, 2003). The fulfillment of customer needs is related to satisfying customer expectations, which achieves customer satisfaction.

The insurance industry needs face-to-face contact with customers in order to provide services that satisfy customers' needs so that they continue buying insurance, and/or to repurchase products. This is the main source of revenue for the industry. Due to its specialized nature, the insurance industry needs up-to-date information in order to modify products or services to attract potential customers. The best source of data is a market survey, the results of which may provide information about customers, such as their needs, purchase intentions, and service requirements.

We have designed a questionnaire about insurance products, purchase intentions, the budget for the premium, and participants' basic data that may serve as the basis for understanding their needs. The consumers' purchasing decisions and processes are analyzed, and proper marketing strategies and management operations are proposed. The results may

* Corresponding author. Tel.: +886 3 9777997.
E-mail address: shyng@mail.fit.edu.tw (J.-Y. Shyng).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2005.11.002

be fully applied by managers to make decisions about strategies and processes related to consumer purchasing.

Most papers deal with insurance audits, purchase intentions, purchase channel studies, methodologies for investigating customer purchasing intentions, and customer satisfaction (Hennig-Thurau & Klee, 1997). Many research papers have quantified the problem in order to simplify the parameters, such as social parameters, and use statistical tools to analyze data. This approach, however, is only good for crisp types of data sets and certain data values. If the value of data is continuous or uncertain we must apply fuzzy theory (Zadeh, 1965).

Rough Set Theory is used in this study to analyze the content and features of data. The theory, which was developed by Pawlak (1982), is a rule-based decision-making technique that can handle crisp datasets and fuzzy datasets without the need for a pre-assumption membership function, which fuzzy theory requires. It can also deal with uncertain, vague, and imperceptible data. Until now, analysis of the attributes of combination values using Rough Set Theory has only been addressed by a few papers. In this study, a questionnaire with single-choice and multi-choice answers is used to apply Rough Set Theory to investigate the relationship between a single value and a combination of values of the attributes. Based on expert knowledge, the value class of the questions with multi-choice answers is reclassified in order to simplify the value complexity, which is useful in the decision-making procedure.

After the study, we applied a hit test to check the feasibility of the decision rules. As the hit rate reaches 100%, it is clear that new data fits the decision classes. The results of this research demonstrate customers' insurance needs. The results are as follows: the purchase expectation is endowment; the age of target customers is 25–35 years; and the most purchased product is a mixture of products.

The remainder of this paper is organized as follows. Section 2 describes the methodology of Rough Set Theory. In Section 3 a real case of insurance marketing is presented to show the process of the effects of attributes on combination values. Finally, in Section 4, we present our conclusions.

2. Concepts of Rough Set Theory and the algorithms for decision-making

In this section we briefly introduce Rough Set Theory and its use in analyzing the attributes of combination values for making insurance marketing decisions. In Section 2.1 the history of Rough Set Theory is described, and in Section 2.2 the algorithms of the theory for decision-making are presented.

2.1. History of development

Rough Set Theory can deal with inexact, uncertain, and vague datasets (Walczak & Massart, 1999). Both Fuzzy Set Theory and Rough Set Theory are used with the indiscernibility relation and perceptible knowledge. The major difference between them is that Rough Set Theory does not need a membership function; thus, it can avoid pre-assumption and one-sided information analysis. A detailed discussion of Rough Set Theory can be found in Walczak and Massart (1999). Rough Set Theory was developed by Pawlak (1982, 1984, 2004). It has been applied to the management of a number of the issues, including: medical diagnosis, engineering reliability, expert systems, empirical study of materials data (Jackson, Leclair, Ohmer, Ziarko, & Al-kamhwi, 1996), machine diagnosis (Zhai, Khoo, & Fok, 2002), business failure prediction (Beynon & Peel, 2001; Dimitras, Slowinski, Susmaga, & Zopounidis, 1999), activity-based travel modeling (Witlox & Tindemans, 2004), travel demand analysis (Goh & Law, 2003), solving linear programs (Azibi & Vanderpooten, 2002), data mining (Li & Wang, 2004; Hu, Chen, & Tzeng, 2003; Chan, 1998), and α-RST (Quafafou, 2000). Another paper discusses the preference-order of attribute criteria needed to extend the original Rough Set Theory, such as sorting, choice and ranking problem (Greco, Matarazzo, & Slowinski, 2001). The Rough Set method is useful for exploring data patterns because of its ability to search through a multi-dimensional data space and determine the relative importance of each attribute with respect to its output.

Rough Set Theory applies the indiscernibility relation and data pattern comparison based on the concept of an information system with indiscernible data, where the data is uncertain or inconsistent. The data is grouped into classes called *elementary sets*. Feature/attribute selection is crucial in data processing that consists of relevant (or maybe irrelevant) object patterns, but it may be redundant in data pattern recognition. More detailed information regarding attributes can be found in the works of Swiniarski and Skowron (2003), Polkowski (2004), and Inuiguchi (2004). The objects in a class may have a relationship with the corresponding features/attributes, and expert knowledge is used to process attribute extraction. Each elementary set is independent of the others. We can extract knowledge from each elementary set used in the real world.

2.2. Rough set algorithm for decision-making

Rough Set Theory is a mathematical approach to managing vague and uncertain data or problems related to information systems, indiscernibility relations and classification, attribute dependence and approximation accuracy, reduct and core attribute sets, and decision rules. The remainder of this section discusses the above areas in detail.

2.2.1. Information systems

Given a questionnaire model QM (an information system),

$$\mathrm{QM} = (U, A, V, \rho)$$

$$U = \{x_1, x_2, x_3, \ldots, x_n\}$$

$$A = \{\text{features/attributes}\} = \{a_1, a_2, a_3\}$$