Available online at www.sciencedirect.com
ScienceDirect
Omega 35 (2007) 365–383
omega
The International Journal of Management Science
www.elsevier.com/locate/omega

Rough set-based approach to feature selection in customer relationship management

Tzu-Liang (Bill) Tseng[a], Chun-Che Huang[b,*]

[a]Department of Mechanical and Industrial Engineering, The University of Texas at El Paso, El Paso, TX 79968, USA
[b]Department of Information Management, National Chi Nan University, Nan-Tau 545, Taiwan

Received 4 February 2004; accepted 19 July 2005
Available online 03 October 2005

Abstract

In this paper, application of the rough set theory (RST) to feature selection in customer relationship management (CRM) is introduced. Compared to other methods, the RST approach has the advantage of combining both qualitative and quantitative information in the decision analysis, which is extremely important for CRM. To derive the decision rules from historical data for identifying features that contribute to CRM, both the mathematical formulation and the heuristic algorithm are developed in this paper. The proposed algorithm is comprised of both equal and unequal weight cases of the feature content with the limitation of the mathematical models. This algorithm is able to derive the rules and identify the most significant features simultaneously, which is unique and useful in solving CRM problems. A case study of a video game system purchase is validated by historical data, and the results showed the practical viability of the RST approach for predicting customer purchasing behavior. This paper forms the basis for solving many other similar problems that occur in the service industry.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Rough set theory; Feature selection; Customer relationship management; Mathematical programming; Heuristic algorithm

1. Introduction

Customer relationship management (CRM) is more necessary today because of the increasing rate of change in the consumer market. The rapid changes in the requirements of customers are distinct from each other. CRM is the main means by which businesses can face these challenges, and it is able to help them grasp the varied demand of customers and then earn competitive advantage [1,2]. CRM can be defined as a dynamic process of managing a customer–company relationship such that customers elect to continue mutually beneficial commercial exchanges and are dissuaded from participating in exchanges that are unprofitable to the company [3]. CRM is an enterprisewide business strategy, designed to optimize revenue and customer satisfaction by organizing the institution around customer segments [4,5], and it is accomplished through a process and technology that can translate customer information into customer knowledge [6].

Most enterprises are product oriented and blindly use the 'push' strategies, rather than using 'pull' strategies with customer orientation for selling products [7]. To improve customer feedback rate, loyalty, Web sales, fame, and satisfaction, one-to-one marketing is seen as the most effective approach for CRM. To succeed, companies must be proactive and anticipate what a customer desires [8]. However, with the great number of customers, how do we identify their interests? The answer to this question is to build personalized service [7] and it can be practiced through understanding of customer preference. Through customers' purchasing history, the product relevance, such as brand, material, size, color, appearance, price, quality, etc., can be

* Corresponding author.
E-mail address: chuang@im.ncnu.edu.tw (C.-C. Huang).

0305-0483/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2005.07.006

studied to understand customers' preference toward particular product features [9]. For example, which are the customer-oriented features in the video game market (e.g. with respect to shapes, favor of culture style, in terms of functions, level of age, understanding of video game, demand degree of network, and the level of comfort) that are critical and can be used to segment consumers? Obviously, feature selection is a core and effective tool for exploring the critical customer features.

In CRM, information related to the customer-preferred features is collected through research, interviews, meetings, questionnaires, sampling, and other techniques. These type of data are often discretized and are frequently in "qualitative" format (e.g. salary level, preference level, etc.). Analysis of these qualitative data to extract useful information to boost promotion sales is critical in CRM. Numerous approaches have been applied to feature selection, e.g. genetic of algorithms [10,11,4], artificial neural network (ANN) [12], Tabu Theory [13], branch-and-bound algorithm [14], Fuzzy C-means Algorithm [15] and SOM [16]. However, these approaches are not used for processing qualitative information. They are not suitable for feature selection of CRM because the aforementioned methodologies are population-based approaches which may require several statistical assumptions and they have limitations in handling qualitative data in CRM. An individual object model-based approach that acts a very good tool for analyzing data is preferred. One of the promising approaches to deal with qualitative information and provide an individual object model-based approach is the rough set approach [17].

The rough set theory (RST) is of fundamental importance in artificial intelligence (AI) and cognitive sciences, especially in the areas of machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems, decision support systems, inductive reasoning, and pattern recognition [18]. The rough set approach is suitable for processing qualitative information that is difficult to analyze by standard statistical techniques [19]. It integrates learning-from-example techniques, extracts rules from a data set of interest, and finds data regularities [20]. Furthermore, RST also complements the fuzzy set theory [21] and the usefulness of RST has been demonstrated in a variety of applications [22–24]. Consequently, RST [22] is able to facilitate CRM in feature selection.

Numerous rough set-based feature selection methodologies can be found in the literature. For example, Bredensteiner and Bennett [25] proposed an approach based on a linear program with an equilibrium constraints (LPEC) formulation. The proposed approaches include both heuristic and mathematical formulations. The heuristic approach (e.g., the filter approach) provides feasible solutions but it has some disadvantages, and the performance of induction is not considered. The mathematical approach (e.g., the wrapper approach) guarantees finding optimal solutions, but it is not easy to use because of complexities of time and space [26]. Furthermore, numerous approaches have applied the rough set in other fields, rather than CRM. For example, Bazan [27] analyzed the dynamic and non-dynamic rough set methods to extract laws (decision rules) in decision support systems. Swiniarski et al. [28] use rough sets and hidden layer expansion to select features for rupture prediction in a highly automated production system. Swiniarski and Nguyen [29] developed a rough set expert system to perform classification based on 2D spectral features. Lee and Vachtsevanos [30] applied the rough set to identify defects on a backlight (a rear window of a vehicle with a defrost circuit). Shang and Shen [31] presented an approach that incorporated a rough set-assisted feature reduction method and a neural network-based classifier for image classification. Li et al. [32] described the application of the rough sets method to feature selection and reduction in texture images recognition. Swiniarski and Skowron [33] presented applications of rough set methods for feature selection in pattern recognition. Shen and Jensen [34] proposed a feature selection technique that employs a hybrid variant of rough sets, fuzzy-rough sets, to avoid this information loss. Hu and Cercone [35] presented a method to learn maximal generalized decision rules from databases by integrating discretization, generalization, and rough set feature selection. Basically, there very little effort is needed to apply a rough set-based approach in CRM, where the qualitative data constitutes the primary information in this domain.

In CRM, the feature selection approach attempts to eliminate as many features as possible in the problem domain, and still obtain useful and meaningful outcomes with acceptable accuracy. Having a minimal number of features often leads to establishment of simple models that can be more easily interpreted. This paper mainly focuses on eliciting a minimum number of features from n-dimensional feature space to derive inductive rules. This problem approach can be formulated as a mathematical programming problem with an objective function that will attempt to minimize the average distance among the reducts. Here the "reduct" is defined as the minimum data content including input and output features necessary to represent an object. Note that the data content comprises the designated features and their corresponding values. A more detailed definition regarding the "reduct" can be found in [22]. After the reducts have been derived through the proposed approach, the preferred reducts which contain strong support from different objects and have been examined by domain experts are determined to be decision rules related to CRM.

In this paper, a rough set-based methodology which is able to support rule induction more effectively is proposed. The methodology includes both mathematical and heuristic aspects. Moreover, it is also able to handle conditions, like the weight of each feature, and objects are assigned. Through the weight analysis algorithm, the features are highly correlated to customers' characteristics that are identified. The methodology is able to achieve the following

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات