Available online at www.sciencedirect.com
ScienceDirect
Advanced Engineering Informatics 21 (2007) 293–301
ADVANCED ENGINEERING INFORMATICS
www.elsevier.com/locate/aei

Applying knowledge engineering techniques to customer analysis in the service industry

Sung Ho Ha *

School of Business Administration, Kyungpook National University, Sangyeok-dong, Buk-gu, Daegu 702-701, Republic of Korea

Received 31 October 2006; accepted 22 December 2006

Abstract

Using the customer relationship management perspective to investigate customer behavior, this study differentiates between customers through customer segmentation, tracks customer shifts from segment to segment over time, discovers customer segment knowledge to build an individual transition path and a dominant transition path, and then predicts customer segment behavior patterns. By using real-world data, this study evaluates the accuracy of predictive models. The concluding remarks discuss future research in this area.

Keywords: Knowledge engineering; Data mining; Customer relationship management; Customer analysis; Segmentation

1. Introduction

A successful company today does well in keeping and managing its customers through providing a number of attractive, personalized services that satisfy customer needs. This is due to the premise that it is less expensive to 'cross-sell' an incremental product or service to existing customers, and that attracting new customers is expensive [1].

Therefore, a company needs to understand its existing customers and their needs better than ever before. Satisfying customer needs and building strong relationships with customers entail good customer relationship management (CRM). The goal of CRM is to forge closer and deeper relationships with customers and to maximize the lifetime value of a customer to an organization [2].

From this perspective, it is important to understand customer behavior by analyzing customer information to differentiate between customers, to identify the most valuable customers over time, and to increase customer loyalty by providing customized products and services [3]. Moreover, it is also important to predict the customer purchasing behavior.

In today's environment, most companies contact and serve customers or customer groups by utilizing a range of commercially viable channels. To understand their customers with a unified view, companies try to integrate an abundance of data collected via multiple channels. These include Web browsing, purchasing behavior, complaints, and demographics. Furthermore, companies divide customers into numerous groups with similar preferences and examine distinct characteristics of each group in order to determine the most profitable segments.

Experience, however, shows that business is ceaselessly changing and that customers continue to evolve over time. Customer segments and related knowledge discovered from multiple data sources change over time as the customer base changes [4]. This means that knowledge and predictions about customers are valid during a particular period. However, much research on customer, until now, has assumed that customer segments and their members are stable. In addition, most existing prediction methods are fundamentally based on numerical and historical data

* Tel.: +82 53 950 5440; fax: +82 53 950 6247.
E-mail address: hsh@mail.knu.ac.kr

1474-0346/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aei.2006.12.001

patterns using a simple regression or neural network technique. In a real-world situation, therefore, because of sudden fluctuations or peaks caused by internal and external events such as promotions, new product launchings, and customer support policies, the assumption of the status quo is not appropriate.

To resolve these problems and to focus on customers, this study keeps track of customer shifts among segments in order to monitor changes in the segments over time, and then investigates segment knowledge in order to predict behavior patterns of customer segments. By observing segment shift behavior of each customer, this builds a transition path for each customer. Aggregating each customer's transition path reveals the dominant transition paths that the majority of customers follow. It is possible to predict the next path that each customer is likely to shift by examining the dominant paths. This is helpful in responding appropriately to customers and in exercising customer-centric strategies.

2. Literature review

Customer segmentation assumes that customers exhibit heterogeneity in their preferences and buying behavior [5]. Focusing on customer segments, with relatively homogeneous requirements, can be a basis of satisfying these diverse customers more effectively [6]. In this context, customer segmentation is defined as either a process of aggregating individual customers into groups of likely behavior, or an analysis method for the identification and allocation of resources among identified segments [7].

In general, the most widely accepted and applicable process to segment customers can be summarized by the following four steps: (1) Choose a basis for segmentation and select appropriate variables, e.g., demographic or behavioral; (2) Use multivariate analysis to group together or split-up customers; (3) Evaluate and validate the output; and (4) Analyze results in economic terms.

According to both academic and practitioner's literature, segmentation design schemes greatly depend on factors such as the measures used for segmentation, the number of resulting segments, the view about changes over time, the segmentation techniques used, and the number of customers selected.

2.1. Segmentation measures

Segmentation variables consist of either one or a combination of the following: Demographic, geographic, psychographic, or behavioral purchasing patterns [8]. Behavioral segmentation, including RFM (Recency, Frequency, and Monetary) or FRAT (Frequency, Recency, Amount, and Type) schemes, provides more knowledge of each customer's actual spending preferences and more accurate behavior predictions than other segmentations [9–11]. This is because behavioral measures provide information on how customers think and shop.

2.2. Number of resulting segments

Bonoma and Shapiro [12] minimized the combined direct and opportunity costs of segmentation, as a criterion for determining the optimal number of segments. Ladany and Arbel [13] allowed the derivation of equal-sized segments. Dibb and Stern [14] stated that judgmental decisions are the basis of determining the number of segments, and the interpretability and applicability determine the management utility of a particular solution.

2.3. View about changes over time

In a customer segmentation design, the most common assumption is that a market is relatively stable, segments are unchanging, and the people who belong to them are unchanging over time. If the market is unstable, however, this assumption should be relaxed. One way to predict instability is through an occasion-based design, assuming that people vary in their needs across product purchasing occasions [15]. Another way is to consider time-segmented customers by studying the same customers at different points in time [16].

2.4. Segmentation techniques

In general, methods for customer segmentation are divided into two distinct areas. One area deals with conventional statistical techniques, including the k-means algorithm, discriminant analysis, and logistic regression. The other considers machine learning techniques, such as neural networks. The potential of neural networks as an alternative methodology has been explored [17]. Fish et al. [18] and West et al. [19] suggested that neural networks are more accurate in terms of classification than statistical techniques. A self-organizing map (SOM) does well in segmentation, compared with statistical methods [20]. A strategy has been proposed for acquiring customer requirement patterns by using a laddering technique and ART2 neural network [21].

2.5. Number of customers

Customer segmentation can incorporate all customers or it can be limited to a sample [15]. If a sample is used as the basis for segmentation, management should predict how other customers fall into each group. Management then has to draw conclusions about the universe of all customers via inferential statistics [9,22].

When transaction patterns identify several types of customer segments, it is important to predict changes in the segments, as well as to derive static characteristics of the segments. This is because changes in segments are closely related to the increase or decrease in profitability obtained from segments.

West et al. [19] utilized neural networks to predict consumer choice and decision processes. Ratner [23]