Expert Systems with Applications 38 (2011) 1716–1722

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

To mine association rules of customer values via a data mining procedure with improved model: An empirical case study

Wen-Yu Chiang *

Department of Industrial and Business Management, Aletheia University, Taiwan

ARTICLE INFO

Keywords:
Association rules
Customer values
RFMDR model
Supervised Apriori algorithm

ABSTRACT

This paper proposes a new procedure and an improved model to mine association rules of customer values. The market of online shopping industry in Taiwan is the research area. Research method adopts Ward's method to partition online shopping market into three markets. Customer values are refined from an improved RFMDR model (based on RFM/RFMD model). Supervised Apriori algorithm is employed with customer values to create association rules. These effective rules are suggested to apply on a customized marketing function of a CRM system for enhancing their customer values to be higher grades.

1. Introduction

The objective of this paper is to propose a new procedure for mining association rules of customer value. That is, to transfer and refine customer value to be type of rules. Through mined association rules, businesses can easily to perform customized marketing project, as well as be leaded in customer relationship management for enhancing customer values.

The Jupiter Research indicated that 71% of Internet users will be engaged in shopping activities in 2010, online retailers will spend about 1440 billion US dollars on cost of trading. However, conclusion of this report was: online consumers need more discount and free shipping and reverse logistics fees (Jupiter Media, 2006).

However, in accordance with estimates of MIC/Nelson Media Research online shopping market was approximately 8 billion US dollars in 2008, compare with 2007, its growth rate was 32.3%. The share of B2C market was approximately 4 billion US dollars; C2C market was approximately 3.3 billion US dollars. In addition, the expected market will reach to 10 billion US dollars in the end of 2009. In 2008, two categories of clothing (includes shoes, bags, and adornments) and cosmetics (includes cosmetics and facial/skin care products) for online shopping market were growth rapidly and whose Compound Annual Growth Rates were 88% and 49% (http://www.cyberone.tw, 2009). Thus, experienced consumers of online shopping are focused to be the research objects.

The trends of Taiwan online shopping market are: trading (paying and receiving) in convenient chain stores, group purchasing, extending trial period, and free shipping and reverse shipping. Those factors of trend caused profit decreasing and trade cost increasing (Chiang, 2009). Nevertheless, how to enhance customer value will be a principal goal on electronic marketing strategy of Internet-Shopping websites.

The paper applies Data Mining technologies to process data analysis. Data Mining is defined as using auto or semi-auto way to analyze meta-data for something meaningful models or rules. Usually, the methods of processing are classification, estimation, prediction, association/classification rules, clustering, and visualization (Linoff & Berry, 2002). Recently, businesses apply Data Mining technologies on marketing planning. Their objective is to gain customer loyalty and contribution by customer value discovering. Businesses apply different marketing plans on different customer clusters for extending customers' life cycle. Businesses can get higher growth rate for high profit/expense products, or, to make them to become long-term customers by implement the discovered rules (Berry & Linoff, 2004) Fig. 1.1.

Fig. 1.2 shows that 53% of online shoppers are female in the past 12 months, 68% of online shoppers are 15–34 years old, there are 85% of online shoppers are graduated from high school or college. Thus, objects of this research are those consumers who have been studied in high schools or colleges.

RFM model can process clustering by customer value effectively. Businesses plan to extend customers life cycle (Linoff & Berry, 2002) by implementing marketing projects in various ways for the purpose of enhancing shopping rates, increasing high profit/price products sales, or retaining customers to become long-term customers.

This research improves RFMD model (Chiang, 2009), where *D* various is defined as discount-price products, thus customer value can be identified clearly by RFMD model. However, this paper improves RFMD model to be more visible to recognize customer value. According to the Jupiter Research, environment of online shopping has been maturing gradually, requirements of online

* Tel.: +886 989820339; mobile: +886 6 570 3100.
E-mail address: ianc888@gmail.com

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.07.097

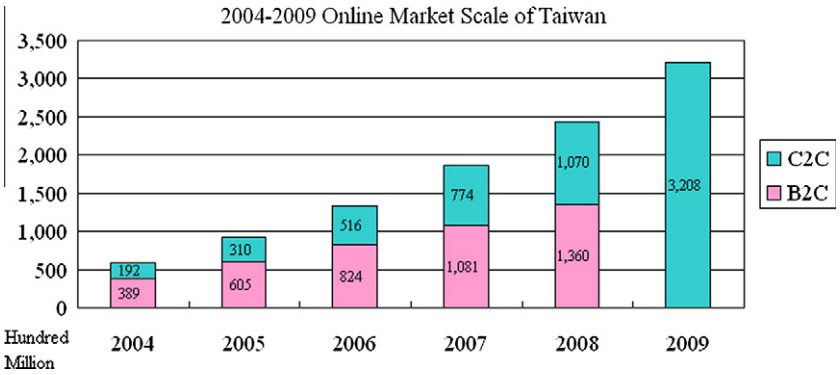

Fig. 1.1. 2004–2009 online market scale of Taiwan.

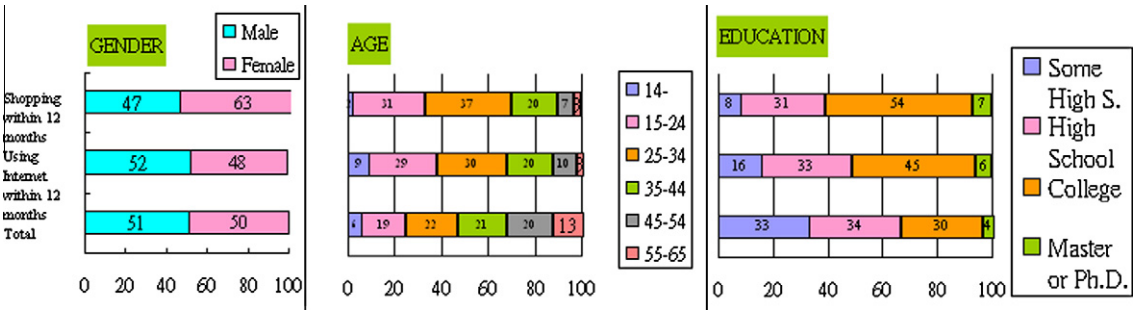

Fig. 1.2. Chart of online shopping populations in Taiwan.

shoppers are free shipping, free reverse shipping, and more discount (Jupiter Media, 2006). In accordance with Berry and Linoff (2004) stated that profit increasing is the basis of customer value. Therefore, as above-mentioned, this paper proposes a negative variable of customer value: *R* variable (Return Cost, an observed variable), thus an improved model of RFMB is formed to be Recency, Frequency, Monetary, Discount, Return Cost (RFMDR) model for testing profit earning rate (positive customer value) and return cost (negative customer value). With RFMDR model, the real customer value can be discovered.

2. Literatures review

2.1. RFM model and customer value

RFM was defined by Hughes (1994), R (Recency) was defined as "Last purchasing time"; F (Frequency) was defined as "Purchasing frequency in a specific period"; M (Monetary) was defined as "Average amount of purchasing in a specific period".

A research for RFM model was applied on a particular industry, Lo, Wu, Chang, and Cheng (2008) adopted RFM model to analyze members of an outdoor sports store. The result has found that higher expense customers are male and from 26 to 35, Gemini, Virgo, and Leo. Moreover, Shih and Liu (2003) applied Analytic Hierarchy Process (AHP) to determinate the weights of RFM model and sort the customer life value. This research has found eight customer clusters, and two kinds of golden customers.

Chao and Yang (2003) applied back-propagation neural network on transactions of medical equipment industry based on RFM model. The result has shown that total expense index of 76.4% customers' value is >0.7.

Processing RFM analysis of consumer for various industries, Miglautsch (2002) has stated some points of view for 1-1-1: (a) About 50% and up of customers, only shopping one time in many of stores. (b) Only few of ordering amount from 50% of advertisement addresses list. (c) There are more than 50% of customer have not shopping in a ten years old company. The researcher has found one time shopper reached 80% from a RFM list of "Business to Business Catalogue Company".

Su, Lin, and Lee (2006) predicted that customers' trading base on RFM model of comprehensive industries. The result has shown the high value customer was only about 18.59% and it approved the rule of 20/80 as well.

RFM model is identified for discovering customer value of direct sales (Roberts, 1992). Lin and Tang (2006) applied RFM model to analyze customers of music product. They classified same value of customers into a cluster. This study defined each variable of RFM model as two levels: High (H) and Low (L) level, thus eight clusters was established. Besides, the researcher applied Apriori algorithm (Agrawal, Imielinski, & Swami, 1993) to create association rules, according to the rules to recommend music for other customers in a same cluster. The recommendation rate was 0.78.

In the research of classifying and segmentation with RFM field, Cheng and Chen (2009) employed RFM attributes and K-means algorithm into Rough set theory to mind accuracy classification rules which can be used for driving an CRM in an enterprise. Huang, Chang, and Wu (2009) applied K-means method, Fuzzy C-means clustering method and bagged clustering algorithm to the analysis of customer value for an outfitter in Taipei, Taiwan. Their study concludes that bagged clustering algorithm outperforms the other two methods.

In addition, RFM model can be processed in market segmentation for finding some valuable customers (Goodman, 1992). The RFMD (*D* for Discount, introduced by Chiang (2009)) model was applied on market segmentation by Chiang (2009) for mining useful customer value, there were three Internet-Shopping markets were found: Cost, Risk, and Convenience. Five and four association rules were established for urban and suburb university students.