Scientia Iranica E (2012) 19 (3), 849–855

Sharif University of Technology

Scientia Iranica

Transactions E: Industrial Engineering

www.sciencedirect.com

Customer lifetime value prediction by a Markov chain based data mining model: Application to an auto repair and maintenance company in Taiwan

C.-J. Cheng [a,1], S.W. Chiu [b,2], C.-B. Cheng [c,*], J.-Y. Wu [c,3]

[a] *Department of Information Engineering, Kun Shan University, 949 Dawan Rd., Yongkang Dist., Tainan City, Taiwan*

[b] *Department of Business Administration, Chaoyang University of Technology, 168, Jifeng E. Rd., Wufeng District, Taichung, 41349, Taiwan*

[c] *Department of Information Management, Tamkang University, 151 Yingzhuan Rd., Tamsui Dist., New Taipei City, Taiwan*

Received 17 March 2011; revised 5 October 2011; accepted 21 November 2011

KEYWORDS

Customer lifetime value;
Auto repair and maintenance industry;
Data mining;
Markov chain;
Decision tree;
Neural networks.

Abstract The present study attempts to establish a framework for computing customer lifetime values for a company in the auto repair and maintenance industry. The customer lifetime value defined in this study consists of the current and future values of a customer, which involve an estimation of lifetime length, future purchasing behavior and the profit associated with each behavior of the customer. The proposed framework contains three groups of techniques to obtain these estimates from historical customer transactions. The first group includes a logistic regression model and a decision tree model to estimate the churn probability of a customer and to, further, predict the lifetime length of the customer. The second group comprises a regression analysis to identify the critical variables that affect a customer's purchasing behavior, and a Markov chain to model the transition probabilities of behavior change. Finally, the third group contains two neural networks to predict the profits contributed by a customer under various purchasing behaviors. The proposed framework is demonstrated with the historical customer transactions of an auto repair and maintenance company in Taiwan.

1. Introduction

The automobile industry in Taiwan used to be protected under a high tariff imposed on imported cars by the government, before 2002. After joining the World Trade Organization (WTO) in year 2002, and opening the market to foreign automobiles, Taiwan's auto industry has encountered the impact of declining sales. In 2005, 445,000 vehicles produced by local makers were sold, but the number dropped to 234,000 in 2009 [1]. Together with economic decline and high fuel prices during past years, the entire auto industry of Taiwan is pessimistic about the future sales of new cars.

In Taiwan, car dealers also operate auto repair and maintenance factories (referred to as the original brand maintenance factory), providing major services to cars of their own brand. As the demand of new cars is weak, repair and maintenance services have become the major source of profit for car dealers. However, these factories still need to face competition from more than 8600 garages that provide the same services with lower prices. In general, when a car is within its warranty period, the owner will return to the dealer's factory for maintenance, but when the warranty is expired, 42% of the owners would choose different garages [2]. Such high customer defection has a great impact on the profitability of companies in this industry. On the other hand, a study by Reichheld and Sasser [3] pointed out that, in the auto service industry, reducing customer defection by 5% would boost profits by 30%. Thus, it is important for an auto repair and maintenance company to take action to retain customers.

* Corresponding author. Tel.: +886 2 26215656x3332; fax: +886 2 26209737.
E-mail addresses: cjcheng@mail.ksu.edu.tw (C.-J. Cheng), swang@cyut.edu.tw (S.W. Chiu), cbcheng@mail.tku.edu.tw (C.-B. Cheng), Jun8954145@gmail.com (J.-Y. Wu).

1 Tel.: +886 6 2727175x65101; fax: +886 6 2050207.

2 Tel.: +886 4 23323000x4379; fax: +886 4 23742331.

3 Tel.: +886 2 26215656x2645; fax: +886 2 26209737.

Peer review under responsibility of Sharif University of Technology.

Production and hosting by Elsevier

1026-3098 © 2012 Sharif University of Technology. Production and hosting by Elsevier B.V. All rights reserved.
doi:10.1016/j.scient.2011.11.045

The ability to identify profitable customers and build long-term loyalty with them is a key factor in today's highly competitive business environment. To achieve this goal, companies have adopted the concept of Customer Relationship Management (CRM) as a business strategy to integrate their sales, marketing and services across multiple business units and customer contact points. Under the concept of CRM, customers are not equal and, thus, it is unreasonable for the company to provide the same incentive offers to all customers. Instead, companies can select only those customers who meet certain profitability criteria based on their individual needs or purchasing behaviors [4]. Precise evaluation of customer profitability is a crucial element for the success of CRM [5]. However, evaluation of customer profitability is not an easy task, because it involves prediction of future contributions by the customer. The major factors in consideration of this prediction include: for how many years in the future will the customer stay with the company, and how much will the customer contribute to the company each year? The time length that a customer will stay with a company is referred to as the lifetime of this customer, and the profitability attributed to a customer is called the lifetime value (LTV) of the customer.

The use of customer lifetime values in marketing, for segmenting customers or formulating strategies, has been found in literature. Hwang et al. [6] and Kim et al. [7] proposed a customer lifetime value computation model considering the past profit contribution, potential benefit, and defection probability of a customer. They also covered a framework for analyzing customer value, and segmenting customers based on their values, and applied their approach to formulating marketing strategies for a wireless communication service company. Shih and Sohn [8] used three clustering methods (k-means, self-organizing map, and fuzzy k-means) to segment stock trading customers to determine differentiated commission rates based on customer potential values. Shih and Liu [9] proposed using a collaborative filtering technique via customer lifetime value and customer demand to develop a product recommendation system. Chan et al. [10] also used a customer lifetime value to predict the company long-term return on investment.

In particular, Chan [11] also adopted the concept of the customer lifetime value in segmenting customers of an automobile retailer. He formulated a genetic algorithm that combined customer segmentation with campaign decisions where the fitness function of the algorithm was defined as a customer's lifetime value. Computation of the lifetime value was based on the potential profit that is obtained from the customer, by adopting a certain campaign strategy in which the adoption of a strategy was considered as a probability distribution. However, how such a probability distribution could be obtained was not discussed, neither did the estimation of profit generated from a campaign strategy. The missing link between the customer lifetime value computation and the historical customer transaction data in Chan's approach reduces the practical usefulness of the approach. Thus, the present study attempts to develop a clear framework of customer lifetime value computation for the auto repair and maintenance industry, which can be used as a road map for the industry to compute customer lifetime values.

The length of the lifetime of a customer depends on his loyalty to the company. This study predicts the lifetime of a customer by a loyalty measure, which is estimated based on customer demographic data and historical transaction records. Data mining techniques are employed to obtain the loyalty measure of a customer. To estimate the contribution of a customer during his lifetime with the company, the purchase frequency and profit generated from each visit are further predicted. A Markov chain is used to model the possible purchasing frequency of a customer in this study, while neural network approaches are used to estimate the profit generated from customer purchases. Our approach is formulated as a step-by-step framework that synthesizes various data mining techniques, including regression analysis, decision tree and neural networks, to provide a roadmap for a company to compute the lifetime values of its customers. The proposed customer lifetime value computation framework is applied to an auto repair and maintenance company in Taiwan.

2. LTV prediction model

The basic concept of customer lifetime value computation is based on the Net Present Value (NPV) received from a customer over his lifetime of transactions with the company (e.g., [12–14]). The common formulation of the NPV-based model is as follows:

$$LTV = \sum_{t=0}^{n} \pi(t) \cdot \frac{1}{(1+r)^t}, \tag{1}$$

where $\pi(t)$ is the profit contributed by a customer at time t, r is the interest rate, and n is the total period of projected life of a customer staying with the company.

Hwang et al. [6] expanded the NPV-based model to comprise, not only the projected value of a customer, but also his past profit contribution (called current value). The present study adopts the same concept, and computes the lifetime value of the kth customer by the following equation:

$$LTV_k = CV_k + FV_k, \tag{2}$$

where CV_k and FV_k denote the current and future values of the kth customer, respectively. Computation of the current value is simple. It is the present value of past profits contributed by a customer. Let the number 0 denote the current time point, and -1, -2, etc. denote one period, two periods, etc. before the current time. Considering that the bank interest rate is usually changed every year, the present study suggests computing the current value of the kth customer as:

$$CV_k = \sum_{t=-m}^{-1} \pi_k(t) \prod_{d=t}^{-1} (1 + r(d)), \tag{3}$$

where $\pi_k(t)$ is the profit contributed by customer k at period t, $r(d)$ is the bank interest rate at period d, and m is the number of periods before the current time point with which we compute a customer's current value.

The prediction of future value is more difficult and complicated. The computation of the future value of a customer involves two dimensions, namely:

1. The profit that the customer will contribute in a certain period in the future.
2. The length of time that the customer will stay with the company.

Traditionally, the future profit contributed by a customer is a projection of the customer's past contribution. However, the purchasing behavior of a customer may change in the future and, thus, make such a projection inaccurate, as it is purely based on the customer's historical behavior. Considering the uncertainty of customers' future behavior, Colombo and Jiang [15] developed a stochastic RFM model to rank customers