



# Customer profitability forecasting using Big Data analytics: A case study of the insurance industry



Kuangnan Fang<sup>a</sup>, Yefei Jiang<sup>b</sup>, Malin Song<sup>c,\*</sup>

<sup>a</sup> School of Economics, Xiamen University, Xiamen, China

<sup>b</sup> Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, Taipei, Taiwan

<sup>c</sup> Research Center of Statistics for Management, Anhui University of Finance & Economics, Bengbu, China

## ARTICLE INFO

### Article history:

Available online 15 September 2016

### Keywords:

Random forest regression  
Customer profitability  
Big data  
Insurance industry

## ABSTRACT

This paper proposed a new customer profitability method for the insurance industry by adding liability reserve. Considering the historical purchasing behavior and the foreseeable future cash flow, the proposed method can measure the real insurance customer contribution effectively. In addition, this paper firstly applies random forecast regression, a method for Big Data analytics, to forecast insurance customer profitability. Comparing with other models, we find that random forest outperforms traditional forecasting methods, such as linear regression, decision tree, SVM and generalized boosted model. Empirical study finds that customer's region, age, insurance status, sex and customer source are the most important factors to predict insurance customer profitability.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Since the general recognition that keeping customers is more profitable than developing new customers, relationship marketing attracts more and more attention from the managers and the scholars in marketing. However, not every relationship between customers and companies is worth maintaining. Instead of treating all customers equally, it is more effective to develop customer-specific strategies. Customer profitability is one of the most important segmentation factors for distinguishing valuable and non-valuable customers. Customer profitability refers to the net dollar contribution made by a customer to an organization (Mulhern, 1999). There are several terms in the literature refer to customer profitability including lifetime value, customer lifetime value, customer value, customer lifetime value, customer relationship value, and customer equity (Boyce, 2000; Mulhern, 1999; Pfeifer, Haskins, & Conroy, 2005). Analysis of the profitability was once concentrated at the level of firm, product or brand, the customer level profitability has not come into practice until the availability of large-scale customer databases containing the customers' historical purchase behavior.

With the advent of customer profitability, the traditional marketing practice has been reconsidered by treating the customer

as one kind of asset analogous to other economic units. Thus the marketing decisions are the same as investment decisions in which expected revenues are evaluated. The well-known 80/20 rule indicates that for most companies, profits mainly sourced from a small set of customers. Accordingly, knowledge of customer profitability can enhance the decision-making in marketing and provide a metric for the allocation of marketing sources to customers and market segments. It is possible for companies to invest in the customers that are valuable but minimize the investments in non-valuable customers. Although the concept of customer profitability has been discussed a lot in the literature, little attention has been paid to how to measure it. Particularly, different companies in different industries may have different definition of customer profitability. Thus there is no one universally applicable measure of customer profitability for all companies in all industry. Specific measure should be proposed in line with the characteristics of the target organization and industry.

As a customer-oriented industry with customers being the driving power and the development foundation, the research on insurance customers is of great importance (Peng et al., 2007). Like many competitive industries, the customers in the insurance industry are free to choose their insurers. We need to know which customers are leaving and why, whether it is deserve the efforts to retain them or not. Fortunately, the availability of the effective data warehouse that records details of every financial transaction and claim enables the decision makers to gain the valuable information at the customer level. With information regarding the customer

\* Corresponding author.

E-mail address: [songmartin@163.com](mailto:songmartin@163.com) (M. Song).

profitability, customer behavior as well as customer preference, managers are able to make long-term customer-related decisions in order to yield a customer mix that will generate the greatest returns. Therefore, how to subdivide the customer reasonably is the crucial problem.

For the insurance industry, the premiums of customers are the main profit sources and the settlement of claims are the major expenditures, so the premium profit and the claim risk are the two important aspects for insurance companies for evaluating customers. By taking premium profit and claim risk into consideration, customers can be divided into four different types: high profit and high risk, high profit and low risk, low profit and high risk as well as low profit and low risk. For insurance companies, the best customer is the high profit and low risk type. However, dividing customers with only one aspect of profit or risk will inevitably lead to some bias. In order to overall consider the premium profit and claim risk, according to the particularity of insurance industry, this paper brings in the concept of liability reserve funds to construct an insurance customer profitability (ICP) aimed at insurance industry based on the traditional customer profitability (CP). The traditional CP reflects the premiums and the claim events incurred, however the proposed ICP in this paper not only use the information of CP but it also further considers the future claim events.

Since the technology development for storing mass data, the analysis of profitability at the customer level is possible. There has been a growing interest in customer-oriented accounting and client profitability (Anandarajan & Christopher, 1987; Guilding & McManus, 2002; Robin & Kaplan Robert, 1999; Ryals, 2002). In the marketing literature, the analysis of customer profitability was proposed in the literature of direct marketing while cost accounting literature has focused on the customer profitability analysis on measuring customer revenues and customer costs (Jacobs, Johnston, & Kotchetova, 2001). These studies of customer profitability may be divided into three groups: (1) The concept of customer profitability. (2) The measure of customer profitability. (3) The application of customer profitability.

Marketing strategies are increasingly based on customer profitability information, so it is important to get this information as accurate as possible (Ryals, 2002). Mulhern (1999) proposed the profitability models and customer history contribution model based on direct-marketing pattern, which include the lifelong profitability computation model. For the lifelong profitability computation model requires the accurate prediction of the future customers' behavior. Boardman and Vining (1996) and Zeithaml, Lemon, and Rust (2001) conducted the customer segmentation and the customer oriented contribution analysis. Niraj, Gupta, and Narasimhan (2001) proposed the profitability model based on supply chain which inserted the intermediate vendors into the supply chain and took the transportation costs into consideration. From the perspective of customer lifetime value (CLV), Ryals (2002) compared several customer profitability evaluation methods, and found that the economic value of risk is a good measure of consumer price. Fornell, Mithas, and Morgeson (2009) analyzed the customer value by taking the company stock price and other effects into consideration. Petersen et al. (2009) compared and assessed numerous methods for measuring customer profitability.

According to Ryals (2002), there are two categories of approaches to calculate customer profitability: top-down method and bottom-up method. Top-down method starts from the profit of all customers, while the bottom-up approach calculates the profit contribution of a single customer first. For bottom-up approach, there are two methods to measure a single customer's profit contribution: the first method is RFA which evaluates the customer profit contribution according to the history data such as the customer trading frequency in the retail industry, the defect of RFA is that it does not consider the cost and it is not able to

forecast the future contribution; the second method is the Customer lifetime value discount theory (CLTV) which predicts and discounts the customers' future cash flow, however, the premise of the theory is the precise prediction of the customers' future purchase pattern and benefit-cost cash flow.

The application of customer profitability has a dramatic increase of literatures in the past few years. Noone and Griffin (1999) illustrated how customer profitability was implemented in a hotel environment. Anderson and Mittal (2000), Bowman and Narayandas (2004), Yeung and Ennew (2000), and Helgesen (2006) investigated the relationship between satisfaction, loyalty and customer profitability at the individual consumer level. Reinartz, Thomas, and Kumar (2005) noted that measuring, managing and maximizing customer profitability requires both the revenues and costs of marketing, sales and customer interactions to be considered in resource allocation, and they investigated the problem of balancing the resource allocation between the customer acquisition and customer retention which optimizes customer profitability. Larivière and Van den Poel (2005) used the random forest to predict the customer retention and customer profitability. Hallowell (1996), Noone and Griffin (1999) and Van Raaij, Vernooij, and van Triest (2003) had conducted the empirical research about the profit contribution. Epetimehin and Ekundayo (2013) used descriptive analysis and regression analysis to investigate the impact of pricing of risk on the profitability of Nigeria insurance market. Ogbonna and Ogwo (2013) adopted a survey research methodology to examine the market orientation strategies of insurance firms for their desired profitability and other performance potential, and the paper used Spearman's Rank correlation coefficient, multiple regression and partial correlation analyses. Karamizadeh and Zolfaghari (2016) identified affecting factors in the profit and loss of the third party insurance of Iran insurance company auto, using two ruled-based and three clustering algorithm.

The calculation of historical profitability has been studied a lot, but it is rare to predict future profitability, and the task is challenge. Considering the characteristics of the insurance industry, this paper proposes a new customer profitability method for the insurance industry with considering liability reserve. Then, we firstly apply random forecast regression to forecast insurance customer profitability which measured by our proposed method. RF regression outperforms than other methods such as linear regression, decision tree, support vector machine and generalized boosting model. The empirical study finds that customer's region, age, insurance status, sex and customer source are the most important factors to predict insurance customer profitability.

The remainder of the article is organized as follows: Section 2 discusses the theory of the customer profitability and reviews the related literature; Section 3 constructs the insurance customer profitability (ICP) model; the insurance customer profitability forecast model based on random forest is introduced in Section 4. Section 5 is the data illustration and preprocessing; the empirical results are provided in Section 6; at last the conclusion and discussion of this paper are provided.

## 2. Insurance customer profitability

### 2.1. Mulhern's customer profitability

Mulhern (1999) from Northwestern University proposed the customer lifelong profitability computation model:

$$CP_i = \left[ \sum_{t=1}^T \frac{\sum_{j=1}^{J_t} (p_{ijt} - c_{ijt}) - \sum_{k=1}^{K_t} mc_{ikt}}{(1+r)^t} \right] \quad (1)$$

Download English Version:

<https://daneshyari.com/en/article/5127915>

Download Persian Version:

<https://daneshyari.com/article/5127915>

[Daneshyari.com](https://daneshyari.com)