Expert Systems with Applications 38 (2011) 5005–5012

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Facilitating cross-selling in a mobile telecom market to develop customer classification model based on hybrid data mining techniques

Hyunchul Ahn [a], Jae Joon Ahn [b], Kyong Joo Oh [b,*], Dong Ha Kim [b]

[a] School of Management Information Systems, Kookmin University, 861-1, Jeongneung-Dong, Seongbuk-Gu, Seoul 136-702, South Korea
[b] Department of Information and Industrial Engineering, Yonsei University, 134, Shinchon-Dong, Seodaemun-Gu, Seoul 120-749, South Korea

ARTICLE INFO

Keywords:
Customer classification
Genetic algorithms
Logistic regression
Artificial neural network
Decision tree
Mobile telecom market

ABSTRACT

As the competition between mobile telecom operators becomes severe, it becomes critical for operators to diversify their business areas. Especially, the mobile operators are turning from traditional voice communication to mobile *value-added services* (VAS), which are new services to generate more *average revenue per user* (ARPU). That is, cross-selling is critical for mobile telecom operators to expand their revenues and profits. In this study, we propose a customer classification model, which may be used for facilitating cross-selling in a mobile telecom market. Our model uses the cumulated data on the existing customers including their demographic data and the patterns for using old products or services to find new products and services with high sales potential. The various data mining techniques are applied to our proposed model in two steps. In the first step, several classification techniques such as logistic regression, artificial neural networks, and decision trees are applied independently to predict the purchase of new products, and each model produces the results of their prediction as a form of probabilities. In the second step, our model compromises all these probabilities by using genetic algorithm (GA), and makes the final decision for a target customer whether he or she would purchase a new product. To validate the usefulness of our model, we applied it to a real-world mobile telecom company's case in Korea. As a result, we found that our model produced high-quality information for cross-selling, and that GA in the second step contributed to significantly improve the performance.

1. Introduction

Along with the rapid growth of mobile telecom market in recent years, the mobile telecom markets in the world are becoming saturated. As traditional voice communication services have become widespread, mobile telecom operators have experienced difficulties in attracting more customers, which has led to fierce competition amongst the mobile operators. To overcome the current crisis, the mobile operators are focusing more on the *value-added services* (VAS)[1] market such as communication, entertainment, transaction, and information services, instead of the ordinary voice communications market since the VAS market returns higher *average revenue per user* (ARPU). In other words, the mobile operators are trying to cross-sell their new mobile telecom service (Kuo & Chen, 2006).

However, it is never easy for the operators to make their subscribers to use mobile VAS. There are two types of the subscribers who do not buy or use mobile VAS. The first type of subscribers are the people who have no interests in using mobile VAS or are not able to use it. They simply think that mobile VAS do not offer any benefits. The second type is the subscribers who may have interests and may find benefits in using mobile VAS, but are not aware of the availability of the services. These second type subscribers offer great business opportunities for mobile service operators. Thus, it becomes critical for the operators to find the appropriate prospects for using mobile VAS.

Recently, most mobile telecom operators utilize CRM (customer relationship management) systems in order to cumulate various types of data from their subscribers, such as demographic information and the patterns for using voice communication, in their DB (database) or DW (data warehouse). These data can be useful to find the appropriate prospects for using mobile VAS since they may be used as the cues for understanding customers' life styles and value system. As a result, the effective application of the cumulated data on the existing customers and their usage pattern for old

* Corresponding author. Tel.: +82 2 2123 5720.
E-mail addresses: hcahn@kookmin.ac.kr (H. Ahn), redsamzang@yonsei.ac.kr (J.J. Ahn), johanoh@yonsei.ac.kr (K.J. Oh), donghakim@yonsei.ac.kr (D.H. Kim).

[1] Mobile VAS combine the characteristics of mobile telephones and the Internet (Lee, 2003). Shih (2003) defined mobile VAS as, "The addition of digital services (excluding voice services) to the basic mobile phone network, such as games, image download, ringing tones download, messages, coupons and electronic transactions". Mobile VAS can bring the five values to consumers: "time-critical needs and arrangements", "spontaneous needs and decisions", "entertainment needs", "efficiency needs and ambitions" and "mobility-related needs" (Anckar & D'Incau, 2002).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.09.150

products or services may serve as a core competence of companies including mobile telecom operators.

To meet these kinds of needs, we propose an innovative customer classification model, which may be used to facilitate cross-selling. This model can predict whether a customer would purchase new products or services by using the cumulated data on the customer. Our model consists of two steps, which applies several data mining techniques. The first step builds multiple classification models simultaneously by using data mining techniques for classification such as logistic regression, artificial neural networks, and decision trees. Each of these models produces the probability of a customer to purchase or use new products. In the second step, all of these probabilities are compromised by genetic algorithms (GA), and final scores for customers are calculated, which can be used as a guide to find prospects for cross-selling new products. Furthermore, the thresholds that are used to interpret the final score to the final decision are also optimized by GA in this step.

Though the customer classification model proposed in this paper can be applied to various kinds of industries or markets, we apply it to an existing mobile telecom company's case in Korea as cross-selling is crucial for competence and survival in the mobile telecom industry.

We organize this article as follows. In Section 2, we briefly review the theoretical background, and propose our research model – a new customer classification model – in the next section. Section 4 presents the research design and experiments. The empirical results are presented and discussed in Section 5. The final section suggests the contributions and limitations of this study.

2. Background

We first review the concept of customer classification. And then, we review the prior studies on data mining techniques for classification, and the methods for the combination of multiple classification methods. Finally, genetic algorithms are briefly introduced as a tool for global optimization.

2.1. Customer classification

CRM involves creating, maintaining and enhancing strong relationships with customers and other stakeholders (Kotler & Amstrong, 1999). The main focus of CRM is to establish a long term and profitable relationship with customers. The core of CRM is not to satisfy all the customers, but to completely satisfy the customers who have proven to be beneficial to the company. Thus, in order to implement CRM, it is essential to understand who our beneficial customers are first. In addition, using this kind of knowledge, the service provider should be able to determine whether their new clients would be beneficial or not (Ahn, Kim, & Han, 2006). Thus it is essential to find the 'right' customers by utilizing and mining the previous database and to use such data to successfully conduct a campaign to suit the customer's preference. This process to find the appropriate customers using their data is called customer classification.

Customer classification is an important issue in marketing. It is believed that the more understanding the corporation has about its customer behavior patterns, the greater the chance that more effective marketing strategies can be developed. Thus, there have been several studies on building effective customer classification models (see Ahn, Kim, & Han, 2007; Chiu, 2002). However, most of these studies are based on a single data mining technique. There have been few attempts to apply several techniques simultaneously and combine their outcomes for customer classification.

2.2. Data mining techniques for classification

In general, data mining techniques are used for association, estimation, classification and segmentation. Among them, logistic regression (LR), decision trees (DT) and artificial neural networks (ANN) are the typical data mining techniques for classification.

LR is a model used for prediction of the probability of occurrence of an event by fitting data to a logistic curve. It is a kind of generalized linear models, which makes use of several predictor variables that may be either numerical or categorical. LR is used extensively in the medical and social sciences as well as marketing applications such as prediction of a customer's propensity to purchase a product or cease a subscription, bankruptcy prediction, and so on (Flagg, Giroux, & Wiggins, 1991; Laitinen & Laitinen, 2000; Lau, 1987; Suh, Noh, & Suh, 1999).

DT is a predictive model that maps target values from observations. It is a flow-chart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and leaf nodes represent classes or class distributions (Han & Kamber, 2001). DT models have an advantage in that they can be directly recognized into sets of If-Then rules, which are easy to apply in practical areas. As a result, DT models have been in used in many application areas ranging from market segmentation, fraud prediction, churn prediction, credit embezzlement, and even stock trading (Langley & Simon, 1995; Lejeune, 2001; Wu, Lin, & Lin, 2006).

ANN is a mathematical model or computational model, which mimics biological neural networks. It captures the salient fundamental features of the inputs and recognizes the pattern in the data. Compared to LR or DT, ANN models can reflect unknown or complex nonlinear relationships in a certain data, so the prediction accuracy is generally high. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Due to its excellent performance, ANN model has been applied to various data mining tasks such as exchange rate estimation, stock price estimation, market segmentation, churn prediction and box office success predictions (Davies, Moutinho, & Curry, 1996; Enke & Thawornwong, 2005; Sharda & Delen, 2006).

2.3. Combination of multiple classifiers

It has been a controversial issue to integrate several data mining techniques. Especially, combining traditional quantitative methods such as statistical methods and numerical programming, and AI (artificial intelligence) methods has been a popular research topic. Combination of multiple data mining methods can be done in many ways. *Three-architecture model* is one of the most popularly used schemas for the combination of several methods. This model proposes three types of integration: (1) loosely coupled or distributed, (2) tightly coupled, and (3) embedded or full integration (Medsker & Turban, 1994). Fig. 1 shows four types of combination, which are the extension of three-architecture model (Jo, 1999).

In type (a), method A is the preprocessor of method B, whereas method B is embedded in method A in type (b). In type (b), method A becomes the main part of the whole model. In type (c), main problem is decomposed into sub-problems, and then method A and B are applied to the different sub-problems. Method A and B solve the problem simultaneously, and the final solution is generated by adjusting the results from method A and B in type (d).

There have been many prior studies tried to combine several classifiers. Liang, Chandler, and Han (1990) combined two statistical methods – discriminant analysis and factor analysis – and two inductive learning methods including ID3 and CRIS. This study constructed four different combination models, and analyzed the performance of each model. Tsujino and Nishida (1995) proposed the