Cluster analysis in industrial market segmentation through artificial neural network

R.J. Kuo[a,*], L.M. Ho[b], C.M. Hu[c]

[a]*Department of Industrial Engineering, National Taipei University of Technology, 1 Section 3, Chung-Hsiao East Road, Taipei 106, Taiwan, ROC*
[b]*Department of Leisure, Recreation and Tourism Management, Shu-Te University, Kaohsiung County 824, Taiwan, ROC*
[c]*Division of System Information, Foxconn Industrial PCE Production Group, Hsing-Chu, Taiwan, ROC*

Abstract

Market segmentation has commonly applied cluster analysis. This study intends to make the comparison of conventional two-stage method with proposed two-stage method through the simulated data. The proposed two-stage method is the combination of self-organizing feature maps and K-means method. The simulation results show that the proposed scheme is better than the conventional two-stage method based on the rate of misclassification. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Market segmentation; Cluster analysis; Monte Carlo simulation; Self-organizing feature maps; K-means

1. Introduction

Wendal Smith presented the concept of market segmentation in 1956. It has become one of the fundamental concepts of marketing (Smith, 1956). By following the market segmentation strategy, a firm could increase the expected profits. Though a number of market segmentation methods have been presented to solve this problem for many decades, the advancement of market segmentation research requires narrowing the gap between the academically oriented research on segmentation and real-world application of segmentation research (Wind, 1978).

Thus, this study is intended to propose a new two-stage scheme by integrating ANN and multivariate analysis, though a two-stage method was suggested by Punj and Steward (1983). At that time, the two-stage method consists of a hierarchical method, like Ward's minimum variance method, and followed by a non-hierarchical method, such as K-means method. In the current study, a modified two-stage method, which first uses the self-organizing feature maps to determine

---

the number of clusters and the starting point and then employs the K-means method to find the final solution, is proposed. The numerical simulation data are applied to validate the feasibility of the proposed method. The simulation results show that the proposed two-stage method is more accurate than the conventional two-stage method (Ward's minimum variance method followed by K-means method) based on the rate of misclassification.

The rest of this paper is organized as follows. Section 2 presents the general idea of market segmentation and applications of ANNs in marketing segmentation, while the proposed two-stage method is explained in Section 3. Section 4 shows the simulation algorithm and results. Finally, the concluding remarks are made in Section 5.

2. Artificial neural networks in market segmentation

An artificial neural network (ANN) is a system, which has been derived through models of neurophysiology. In general, it consists of a collection of simple non-linear computing elements whose inputs and outputs are tied together to form the network (Hertz, Krogh, & Palmer, 1991). A number of studies have successfully demonstrated the learning capability of ANNs and their applications in the area of engineering. Though most investigations are interested in the performance of ANNs as compared with the conventional statistical methods, these networks have been employed recently to solve the management problems. The results are very promising.

Proctor (1992) presented that ANN is an alternative expert system for the solution of marketing decision problems. He showed two examples, which are the sales forecasting and new product evaluation. Besides, Venugopal and Baets (1994) showed the possible applications of ANNs in marketing management. They employed three examples, retail sales forecasting, direct marketing and target marketing, to demonstrate the capability of ANNs. For more works on this, refer to Dhar and Roger (1996) and Goonalilake and Treleaven (1995).

For the area of market segmentation, a network with six inputs and three outputs is proposed by Venugopal and Baets (1994). The inputs are six attributes of market including demographic information, socio-economic information, geographic location, purchase behavior information, consumption behavior information, and attitude to product, while the three outputs represent three segments. Either the adaptive resonance theory models (ART-1 and ART-2) or self-organizing feature maps can be used for clustering. Fish, Barnes, and Aiken (1995) also suggested that ANNs can be utilized for market segmentation. In addition, Bigus (1996) suggested that ANNs can be employed as a tool for data mining and presented a network with three different dimensions of data, population (sex, age, and marriage), economic information (salary and family income), and geographic information (states, cities, and level of civilization). (Balakrishnan, Cooper, Jacob, and Lewis (1994)) compared self-organizing feature maps with the K-means method. The results reveal that the K-means method has a higher rate of classification through the Monte Carlo algorithm. Two years later, Balakrishnan, Cooper, Jacob, and Lewis (1996) employed the frequency-sensitive competitive learning algorithm (FSCL) and the K-means method for clustering the simulated data and real-world problem data. Also, the combination of these two methods was presented. Neither the simulated nor real-world problem data can determine which method alone is better. However, the combination of the two methods seems to provide better managerial explanation for the brand choice data.