

Customer Segmentation based on Neural Network with Clustering Technique

XIANG-BIN YAN, YI-JUN LI

Department Of Management Science & Engineering

Harbin Institute of Technology

Mailbox 1223, Harbin Institute of Technology, 150001 Harbin, China

CHINA

Abstract: - A novel approach for customer segmentation based on neural network is proposed in this paper. Customer purchase behavior is considered in the neural network by clustering technique. Compared with existing research, more information is included in the segmentation process. The experimental results indicate that our method is effective in customer segmentation.

Key-Words: - Neural network, clustering, customer segmentation, enterprise, feature, evaluation metric

1 Introduction

The limited range of high profit customer and fierce market competition has evoked unprecedented demand of understanding customer. In order to meet customers' demands and improve their satisfaction, enterprises should have abilities of customer segmentation. Different customers have different characteristics such as preference, values, profit, etc.. Customer segmentation theory proposes that groups of customers with similar needs and purchasing behaviors are likely to demonstrate a more homogeneous response to marketing programs. With proper market segmentation, enterprises can arrange the right products, services and resources to each target customer cluster and build a close relationship with them. Market segmentation has consequently been regarded as one of the most critical elements in achieving successful modern marketing and customer relationship management [1].

BP neural network is widely used in customer segmentation for its simplicity and well classification performance [2]. However, most current neural network based customer segmentation approaches are based on customer's static attributes, such as demographics, and have not included the customer behavior data in the segmentation process. In fact, the behaviors that customer exhibited are very important for customer segmentation. The reason that it was not being considered is largely because it is difficult for neural network to incorporate customer behaviors data in the process of segmentation.

This paper develops a new customer segmentation technique base on neural network with clustering technique. We first cluster customers based on their purchase behaviors, and

after that, combined with static attributes. The clustering result is regarded as one of input attributes of neural network to classify customers. Compared with existing research [1]-[3], more useful information is included in the classification process. Theoretically, it should get a more satisfactory result. The remainder of this paper is organized as follows. Section 2 gives a general approach of our customer segmentation method. In Section 3, customer purchase behavior clustering method is provided. A BP neural network for customer segmentation is given in Section 4. Section 5 provides an experiment to demonstrate the benefit of proposed method. A summary and conclusion are presented in Section 6.

2 General Approach of Customer Segmentation

Segmentation accuracy largely depends on the choice of features and the segmentation method. For customer segmentation problem, the features mainly come from the data accumulated in the enterprise database, which involves customer's demographic data, transaction data, and some other mutual contact data between enterprise and their customers. According to the characteristic of the data, we divide them into two categories.

(1) Static data which refer to the data that does not often change, and reflect some basic states of customer. Such as customer's name, address, age, sex, income, etc.. Generally, the static attributes are relatively stable and easy to analyze.

(2) Dynamic data which refer to the data about customer will be changing with time going on, such as customer transaction data, mutual contact data.

Those data constitute large proportion of customer data and reflect customer characteristic from many different angles, such as customer's consuming habit, changing behavior, and so on.

Many researches have devoted themselves to using general variables to partition customers because the variables are intuitive and easy to operate, such as Chou used customer demographics to identify prospective customers [3]. The dynamic attributes are neglected more or less, which often lead to an unsatisfactory classification results [4]. To avoid this limitation, our general approach is shown in Fig. 1.

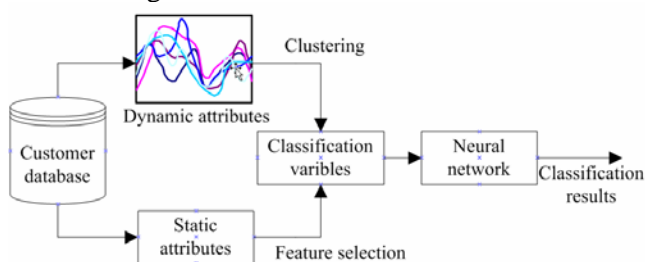


Fig. 1. A general approach of customer segmentation

Customers are grouped by their purchase history, and one attribute is used to substitute clustering result, which will be included in the classification process. We use feature selection method to get valuable attributes from static attributes. All the selected static attributes combined with clustering result are treated as the input vector of the neural network to construct segmentation model. The method can be used for different types of customer segmentation problems, such as credit based, value based segmentation. There should be noted that for different segmentation objective, the features selected for segmentation may be different.

3 Customers Clustering Based on Purchase Behaviors

Customer purchase history usually exhibited as transaction data. In order to cluster customer by purchase history we select some dynamic attributes from transaction data as the clustering features. In this paper, we take following attributes.

(1) Trend attributes trend attribute is one of the most widely used attributes in the transaction data, which describes an upward, downward or balance trend of the customer's transaction amount. It is used to reflect long-term trend of customer. We use least square method to get the trend, which is shown in the formula (1).

$$T = \frac{\sum_{t=1}^n ty(t)}{\sum_{t=1}^n t^2} \tag{1}$$

Where t is the time coordinate, and the $y(t)$ is the customer's monetary expense at time t .

(2) Monetary monetary is used to describe a customer consumption amount. We use mean value of customer's purchase amount. It reflects customer's contributions to the overall turnover of enterprise. The computation method is s institutive and easy to get by formula (2).

$$\mu = \frac{\sum_{t=1}^n y(t)}{n} \tag{2}$$

(3)Purchase frequency it is used to reflect the customer's purchase habit. Different types of customer may have different transaction frequency and have different contribution to the enterprise.

(4) Standard variance it is used to reflect the fluctuation of customer purchase sequence. Its computation method is shown in formula (3).

$$\sigma = \sqrt{\frac{\sum_{t=1}^n (y(t) - \mu)^2}{n}} \tag{3}$$

In fact, there are some complex methods to extract features for clustering. The reason we select those features is that the feature extraction process for them will not be influenced by the length of customer's transaction records. The reason we are not put those features into the neural network directly is that we can improve the clustering quality by including more features in near future.

4 BP Neural Network for Customer Segmentation

Our segmentation model is a hybrid of the clustering and ANN procedures. Clustering process identifies customer's purchase history with one descriptor. ANN is used to segment customer based on the objective. We focus here on the structure of neural network and the criteria used to evaluate the model.

4.1 Structure of the Neural Network Model

Neural network describes a family of analytical models that is based on the physiological properties of animal nervous systems. As with animal nervous systems, neural networks are composed of interconnected 'nodes' that are capable of processing and transmitting information. How the interconnections are modeled determines how

information is transmitted in the network [5]. Since BP neural network is the most widely used architecture, we adopt BP neural network in this research[6]. All layers use sigmoid function as the transfer function. Since it is difficult for the network to output extreme values (i.e., near 0 and 1), in order to overcome this difficulty, we studied the effect of scaling output in the range of x to 1-x linearly so that the effective output range was between 0 and 1. We select x=0.2, so during the training stage, if the neural network predict event will occur, it should output the value greater than 0.8, otherwise the output should be smaller than 0.2.

4.2 Feature Selection Process

Feature selection is defined as the process of choosing a subset of the original predictive variables by eliminating features that are either redundant or possess little predictive information. If we extract as much information as possible from a given data set while using the smallest number of features, not only can we save a great amount of computing time and cost, but also we can build a model that generalizes better to customer segmentation[7]. On the other hand, reducing the dimensionality of the input space tends to reduce overfitting in the segmentation model, and make it easy to be understood by the decision makers. Correlations analysis is used for feature selection in our approach, which will determine the length of feature vector of neural network.

4.3 Evaluation Metric

Table 1. Example of customer static attributes

CID	Sex	Birth	Age	City	Edu	Marital	Income	Title	H_loan	Title	...
321596	1	1978-10-8	2	5	2	2	1	3	Y	2	...
321597	1	1963-8-21	3	3	3	1	2	1	N	1	...
321598	2	1971-7-6	2	1	2	2	3	2	N	3	...
...

For each customer there is one transaction sequence depict customer's behaviors, which in fact constitute a time series. At different time point there are two attribute, one is the transaction code denote transition category, such as customer consumption through credit card, and so on. Another attribute is monetary amount of the transaction category. There are altogether 144 transaction codes. After analysis, 56 transaction codes are selected as valuable transaction category to identify different customer's behaviors.

5.2 Experimental results

We undertake feature selection process after remove redundancy attributes, such as birth date and age refer

The input of neural network is an ordered features vector. All features in the vector are substituted by a unique numeric value distribute evenly between 0 and 1. The prediction precision is defined in formula (4), which expressing the percentages of input which are correctly assigned to the desired class. In the formula, $\delta_{ij}=1$ if and only if $i=j$, e_i is the known class label of input i , \bar{e}_i is the predicted label of i that output by the neural network and N is the number of data that predicted by the neural network.

$$precision = \frac{1}{N} \sum_{i=1}^N \delta_{\bar{e}_i e_i} \tag{4}$$

5 Experiments

The proposed method is applied to classify customers based on their credit. To validate our approach, we contrast the performance of our classification method with a traditional classification method which not includes the clustering result.

5.1. Data Description

The data of customer whose retaining time is greater than 1 year are taken from a commercial bank of China. In our analysis, we use 2000 customer data for the training, 1000 customer data for the validation, and 1000 customer data for the test. Each customer's record contains 34 static variables, include information on both demographic characteristics such as sex, age, etc., and various service opened in the bank. Examples of customer static attributes are shown in Table 1.

to the same thing, we only reserve one attribute. In this experiment, our objective is to segment customer based on their credit. All the customer is partitioned into two categories with credit value of high or low, the high credit is treat as 1, and the low credit is treated as 0. Based on the segmentation objective, 9 static attributes were selected in the experiment. Adding clustering result, we have 10 attributes all together. Those attribute constitute the input vector of the neural network.

Table 2. Comparison of experimental result get by neural network with clustering technique and without clustering technique

Segmentation precision	Training	Validation	Test
------------------------	----------	------------	------

Neural network with clustering	0.823	0.835	0.794
Neural network without clustering	0.757	0.733	0.716

Management Science, Vol. 51, No.2, 2005, pp.264-276.

Experimental result of comparison between BP neural network with clustering technique and BP neural network without clustering technique is shown in table 2. It is easy to see that, compared with the static attributes based segmentation, our methods have improved the segmentation accuracy to a large extent.

6 Conclusion

In this paper, we presented a novel approach for customer segmentation based on neural network. We used clustering algorithm to identify the behavior characteristics of customer. When the clustering result is included in the classification process, a satisfactory result is obtained. It also provides evidence that there exist strong dependencies between customer cluster and their behaviors. The experimental results indicated that our method is effective in the customer classification.

References:

- [1] TSAI C Y, CHIU C C. A purchase-based market segmentation methodology. *Expert Systems with Applications*, Vol. 27, No. 5, 2004, pp. 265-276.
- [2] Natter, M.. Conditional market segmentation by neural networks: a Monte-Carlo study. *Journal of Retailing and Consumer Services*, Vol.6, No.4, 1999, pp.237-248.
- [3] Chou, P. B., Grossman, E., Gunopulos, D., Kamesam, P.. Identifying prospective customers. *Proceedings of the sixth international conference on knowledge discovery and data mining*. Boston, MA, 2000, pp. 447-456.
- [4] Beane, T. P., & Ennis, D. M.. Market segmentation: a review. *European Journal of Marketing*, Vol. 21, No.5, 1987, pp.20-42.
- [5] Derrick S. Boone, Michelle Roehm. Retail segmentation using artificial neural networks. *International Journal of Research in Marketing*, Vol. 19, No.3, 2002, pp.287-301.
- [6] J. Holland. *Adaptation in natural and artificial system*. Ann Arbor: The University of Michigan Press, 1975.
- [7] YongSeog Kim, Gary J. Russell, Filippo Menczer. Customer Targeting: A Neural Network Approach Guided by Genetic Algorithms.