

A Hybrid Customer Churn Model Using Discriminant Analysis, Logistic Regression Analysis, and Artificial Neural Networks in Health Club

Chen Chi-Wen, National College of Physical Education and Sports
Kao Chin-Hsung, National College of Physical Education and Sports

**Session 14: Marketing
Presentation (25-minute)**

Friday, June 1, 2007

9:30 AM - 9:55 AM

As the step of information technology (IT) is making a leap forward constantly, health clubs have been taking advantage of IT when it comes to managing data related to the member's affairs, including establishment of the member data, management of the member files, and the making of the financial statement. These affairs, however, are the first step to build a database. But how to understand the members' needs and characteristics from the database in order to increase the health club's profit, lower the churn rate and do the customer relationship management (CRM) well is the goal in the future.

Traditionally, the way of using market research or questionnaire to realize the consumer behavior or customer satisfaction may lead to sample error or analytical bias because of bad sampling, authenticity of the responses, subject's reaction to the survey questions and so forth, and then influences the accuracy of the result. Moreover, a wealth of research into data mining technologies applied to enterprise database analysis has proven to yield a more valuable result than do traditional methods. In addition, the customer churn problem is the general classification problem in social science research. To solve this problem, many classification tools have been developed, including traditional statistical methods, non-parameter statistical methods and artificial intelligent methods. Typically, the applications of traditional statistics methods are discriminant analysis and logistic regression analysis. Also, there are many techniques developed successfully in non-parameter statistical methods and artificial intelligent methods. The artificial neural networks (ANNs) is frequently discussed and used among them recently.

These methods mentioned above have their strength and weakness, respectively. Therefore, excepting for using these methods to analyze the customer churn problem, this study also tries to incorporate these methods to build a hybrid customer churn model according their merits and shortcomings. Consequently, the purposes of this study are not only to build and evaluate three classification models which were constructed by discriminant analysis, logistic regression analysis, and ANNs, but to incorporate discriminant analysis with ANNs, or logistic regression analysis with ANNs. The reason of incorporation is that ANNs can not choice the important variable while discriminant analysis and logistic regression analysis can. A Taiwanese health club (following with another name of H club because of the business secret) provided a database for this study, and the original data were 1,287 records. After eliminating the incomplete data and the unreasonable data, there were 1,152 records totally. A random sample of 922 records (80% of 1,152 records) was extracted for training model, and the sample of remaining 230 records (20% of 1,152 records) was for testing model. The results were as followed:

1. The whole correct classification rate of discriminant analysis was 73.04%.
2. The whole correct classification rate of logistic regression analysis was 68.70%.
3. The whole correct classification rate of ANNs was 84.78%.
4. Considering the important variables which were chosen by discriminant analysis as input variables of ANNs model, the whole correct classification rate of this hybrid model was 86.09%.
5. Considering the important variables which were chosen by logistic regression analysis as input variables of ANNs model, the whole correct classification rate of this hybrid model was 90.43%.
6. By all accounts, the characteristics of churn customer through the hybrid customer churn model using logistic regression analysis and ANNs were the duration time of membership of one-year-below, the monthly fee of 2,500 New Taiwanese dollars, the method of payment of cash, and the Taipei city resident.