Tamkang University

Available online at jims.ms.tku.edu.tw/list.asp

International Journal of Information and Management Sciences
20 (2009), 39-53

International Journal of
Information
and
Management
Sciences
jims.ms.tku.edu.tw

Enhancing Consumer Behavior Analysis by Data Mining Techniques

Nan-Chen Hsieh
Department of Information Management
National Taipei College of Nursing
R.O.C.

Kuo-Chung Chu
Department of Information Management
National Taipei College of Nursing
R.O.C.

Abstract

Analyzing consumer behavior is a costly implementation of sophisticated information technology, which requires detailed planning and business knowledge for successful adoption. The current trend on consumer behavior analysis has been recognized on the business problem rather than on the information technology. This study presents a two-stage framework of consumer behavior analysis, and the key feature is a cascade involving self-organizing map (SOM) neural network to divide customers into homogeneous groups of customers and a decision-tree simplified method to identify relevant knowledge. Identifying consumers by this approach is helpful characteristic of customers and facilitates marketing strategy development.

Keywords: Data Mining, Neural Network, Decision Tree, Consumer Behavior Analysis, Credit Card.

1. Introduction

Many enterprises have gathered significant numbers of large databases. The database marketing technique uses modern data analysis methods to acquire new customers and apply to develop new business strategies and opportunities. Unlike most data summaries are usually a summary of the data, data mining involves the automated analysis of data to produce useful knowledge in a highly summarized form. Data mining thus is very useful in market segmentation, customer profiling, risk analysis, and other applications. Data mining can also produce rules and models that are useful in replicating or generalizing decisions that can be applied to determine marketing strategies.

Economic theory has established that there are a large number of customers with a small income and a small number of customers with a large income. However, instead of targeting all prospects equally or providing the same incentive offers to all customers,

Received November 2006; Revised March 2008; Accepted May 2008.

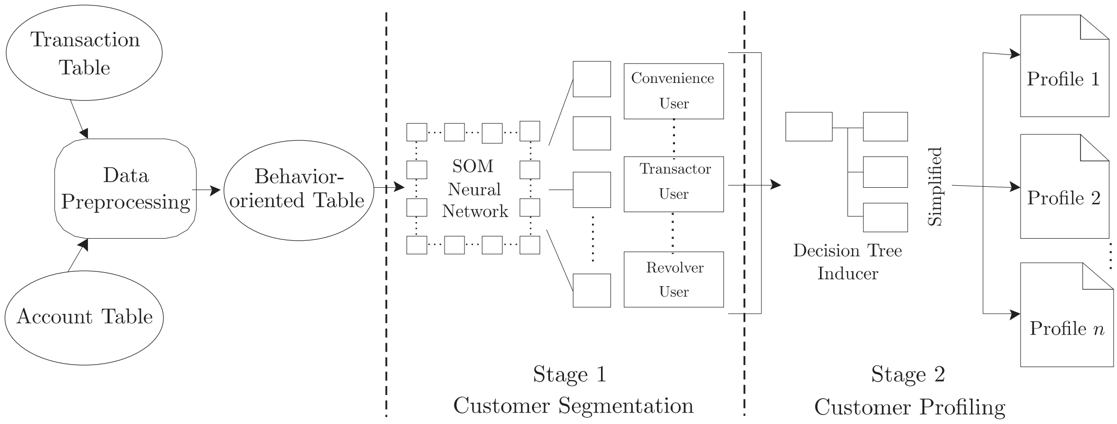

Figure 1. The two-stage framework of consumer behavior analysis.

enterprises can select only those customers who meet certain profitability criteria based on their individual needs or consumer behaviors [13]. Therefore, assuming that consumer behavior follows a similar pattern seems reasonable. For a bank, most existing data mining approaches were discovered rules [1, 27] and predicted personal bankruptcy [12, 10, 30] in a bank database. Hadden et al. [17] review literature for the development of a customer churn management platform. As shown in figure 1, a two-stage framework of consumer behavior analysis was established to predict profitable customer based on demographic characteristics and previous consumer behavior [22].

The key feature of the two-stage framework of consumer behavior analysis is a cascade involving self-organizing map (SOM) neural network [21] and a decision tree inducer [26]. In the first stage of the approach presented here, SOM was used to divide customers into homogeneous groups of customers based on consumer behavior and customer value. Customers then divided into the following profitable groups of customers: revolver user, transactor user, and convenience user [4].

Once SOM was used to identify the profitable groups of customers, the statistical summarized data and the decision tree inducer were used to characterize the groups of customers by creating customer profiles, and a tree simplified mechanism was used for finding relevant classification rules. The customer profile then utilized to describe a representative case in the profitable group, and served as a tool for the establishment of the bank marketing strategies. After analyzing the bank databases, this study demonstrates that consumer behavior analysis is an important trend for bank database marketing.

For a better understanding of our solutions, this study is organized as follows. Section 2 introduces the concept of data preprocessing of bank databases. Based on consumer behavior, Section 3 proposes a conceptual framework of analyzing three profitable groups of customers. Section 4 presents the neural networks to the customer segmentation. Subsequently, Section 5 profiles profitable groups of customers according to their feature attributes as determined using a decision tree inducer. Section 6 presents the method of creation more relevant classification rules. Finally, conclusions are made in Section 7.

2. Experimental Dataset

For this study, bank databases were provided by a major Taiwanese credit card issuer. Data preprocessing was required to ensure data field consistency in consumer behavior model building. Obviously, not all the data are related to the chosen purposes, so knowledge extraction from the bank databases included the following three sub-actions. The first sub-action was intended to organize the raw data. Two data tables were obtained: a table containing effective credit card account information of 150,000 customers until June 2005, and another table storing over 20 million individual transaction records for these accounts from January 2003 to June 2005. Then, two data tables were joined using a customer identifier to create a single behavior-oriented data table. The second sub-action was the extraction of only that data considered useful for the analysis. Unnecessary data fields and records containing incomplete or missing data were removed from the data tables. The third sub-action was the application of simple statistics to calculate an aggregate of new behavioral predicators.

The calculating of the aggregate variables was used to emphasize the customer consumer behavior hidden in the 12 months observation period. In this case, the values derived from the database such as maximum, minimum and average of a set of variables (e.g. number of purchases, repayment cycle days, credit line, yearly amount of consumption, and so on) for the monthly activity over the past 12 months were considered. As mentioned, the desired outcome is to be able to predict which customer belongs to which profitable group. The ranges of values of numerical predictor are split into intervals so that each interval contains as many customers as possible that have a significant homogeneous behavior. Multiple predictors can be grouped together to obtain the same effect. To derive the most profitable customers, it was chosen to identify similar consumer behavior with respect to customer values found in the real world.

3. Analyzing the Consumer Behavior

An important observation on the current state of the art of customer segmentation is the use of past transaction data. The results produced are based on the assumption that the consumer behavior follows patterns similar to past pattern and repeats in the future. The decisions to be made include which target groups of customers will be encouraged to spend more, what credit line to assign, whether to promote new products to particular groups of customers, and, if the repayment ability turns bad, how to manage debt recovery. However, attempts to make good consumer behavior analysis may be limited by the poor data relevance and quality, or the volume of data needing to be processed.

Database marketing is an approach to generate integrated and accessible customer information to help the marketers' better target their markets efforts to existing customers [28]. To date, many database marketing tools were developed, including the RFM (recency, frequency, monetary), the behavior segment of the existing customers and the customer value of a customer [15]. Based on these tools, customer segmentation is a method of achieving more targeted communication with customers and is a pioneering step towards classifying individual customers. The process of customer segmentation describes the characteristics of the customer groups within the data, and putting the customers into segments according to their affinity or similar characteristics.

The original tables could not be used directly to predict consumer behavior, so extra variables need to be constructed for predication. More precisely, consumer behavior analysis tried to group customers that represent shared behavioral patterns. This is carried out by assigning behavior scores to each customer and grouping customers into classes of similar score value using an SOM neural network. The behavior score is given by a mathematical function of the form:

$$behavior\ score = f_{SOM}\ (predicator_1,\ predicator_2, \ldots).$$

As mentioned, banks have three types of profitable customers: revolver users, transactor users and convenience users. Revolver users always carry a credit card balance, rolling over part of the bill to the next month, instead of paying off the balance in full each month. Revolver users are highly profitable customers because every month they pay considerable interest on their outstanding balance. Meanwhile, transactor users pay in full on or before the due date of the interest-free credit period and do not incur any interest payments or finance charges. Transactor users do not contribute significant revenue through interest on their credit balances, but the discount on each transaction they make still provides an important source of bank revenue. Finally, convenience users are customers who periodically charge large bills, such as for vacation or large purchases to their credit card, and then pay these bills off over several months. Convenience users thus contribute significant amounts of interest on their credit balance.

Figure 2 presents the conceptual framework used to answer the questions posed in this study. This figure shows two components, customer segmentation and customer profiling, which serve as major issues studied here. In this framework, consumer behavior and customer value are assumed to be variables affecting customer segmentation. On the other hand, account and transaction tables are assumed to be input tables to customer segmentation. Generally, credit card issuers make money from annual fees, interest on credit balance, and the discount collected from merchants on each transaction. Consequently, the customer value (CV) is first calculated by the summation of interest on credit balance and the discount on each transaction, then multiplied by yearly number of credit card purchases, and finally normalized by the average repayment cycle day for each customer.

Next, variables such as customer attributes and credit card usage are assumed to influence customer profiling. Finally, clusters and the associated profiles are assumed to be outputs, as well as influencing of credit card marketing strategies. In Figure 2, consumer behavior is highly related to customer segmentation, but is an implicit variable which cannot be retrieved directly from the data table. We need to develop a method for modeling the customer consumer behavior.

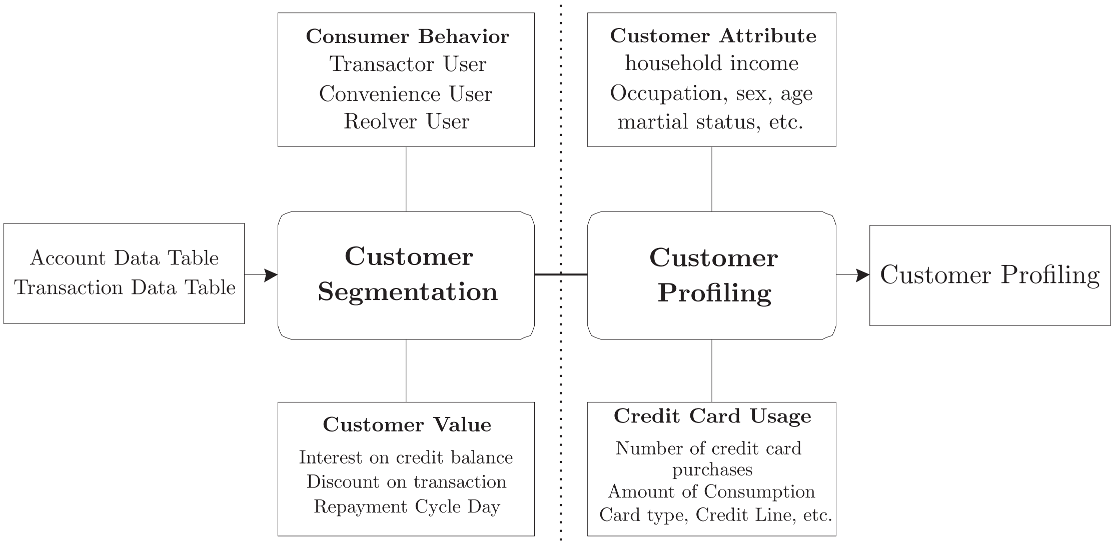

Figure 2. A conceptual framework of this study.

As shown in the following equation, this study employs "Pay-off Ability" (PA) as a behavior score variable to model consumer behavior,

$$\text{Pay-off Ability} = \frac{\text{no. of months without delayed pay off}}{\text{no. of months of holding the card}}.$$

The default observation range is assumed to be 12 months, and PA is computed as the "no. of months without delayed pay off" divided by the "no. of months of holding the card". For example, a customer carries a credit card balance for three months, and then the degree of PA is computed as 0.25. For each customer, if PA is approaching to one, then the consumer behavior of that customer is considered a transactor user. Meanwhile, if PA is between zero and one then the consumer behavior of that customer is considered a convenience user. Finally, if the value of PA is approaching to zero then the consumer behavior of that customer is considered a revolver user.

4. Neural Networks to the Customer Segmentation

For customer segmentation, many studies have presented that neural networks perform significantly better than statistical techniques such as linear discriminate analysis (LDA), multiple discriminate analysis (MDA), logistic regression analysis (LRA) and so on [23, 25, 30, 8]. The application of neural networks to segmentation analysis is a promising research area and is a challenge for a variety of marketers [29]. Baesens et al. [2] employed Bayesian neural networks to repeat purchase behavior modeling in direct marketing. Davies et al. [11] analyzed how different bank customer groups represent different expectations of the automatic teller machines (ATMs) service. Rather than profiling segments based on demographic or geographic characteristics, Dasgupta, Dispensa, and Ghose [10] characterized potential customer segments in terms of lifestyle variables. Balakrishnan et al. [3] accomplished a six-segment classification study using coffee brand switching probabilities derived from the scanner data at a sub-household level. Setiono, Thong and Yap [27] utilized a rule-extraction neural network to aim at companies for the promotion of new information technology. Fish, Barnes and Aikenl [16] proposed a new methodology for industrial market segmentation by neural networks. Lee et al. [24] explore the performance of credit scoring by integrating the back propagation neural networks with traditional discriminate analysis. Kim and Sohn [20] used neural networks to manage customer loans. Chan [7] used SOM and BPN two neural networks to perform data clustering and price prediction after information is crawled and stored into a database.

Typically, the clustering algorithm affects the performance of the clustering result. However, the quality of input samples leads to misclassification, and the misclassification problem is due to unrepresentative samples that drastically reduce the utility of the clustering result. A way of identifying the unrepresentative samples is to look for various consumer behaviors of customers. Handling various consumer behavior clusters in detail and investigating their properties will be very helpful for understanding customers. A way of achieving this goal is to employ clustering techniques to preprocess the training samples by consumer behavior variables. Finally, we stand to gain if we can keep the feature values of input samples more homogeneous, would allow efficient classifiers to be build. The input samples filtered by clustering are used to successively train the decision tree classifiers.

The proposed data mining framework is a hybrid approach using clustering techniques to preprocess input samples into homogeneous clusters, and decision tree techniques to build customer profiles. The clustering process depends on sample's feature values and optionally makes use of the original class label to generate clusters. The cluster labels are thus added as new class labels for each sample. This is a so-called class-wise classification process [19], that is, the clustering process generated clusters of samples belonging to class and eliminates the unrepresentative samples from each class.

In this study, the SOM [21] is built with data from existing customers, which include variables from account and transaction data tables. All of the existing customer's data are used to build the consumer behavior model in order to predicate potential consumer behavior. The utilization of SOM for customer segmentation was used PA and CV as predicated variables to classify each customer into several homogenous clusters. As shown in Figure 3, we arranged three profitable groups of customers as 16 clusters in a two-dimensional square grid of 4×4 neurons. The consumer behavior, number of cases, ratio of number of customers relative to the overall customers, average PA and CV values were shown for each neuron. The mass cases were distributed over neurons 9$\sim$16, the number of customers was 105,252 and consumer behavior was revolver user. Neurons 3, 4, 7, and 8 indicated convenience user totaling 21,202 customers. Moreover, neurons 1, 2, 5 and 6 indicated transactor user totaling 31,945 customers. Figure 4 represents the statistical summarized data of transactor user as an example. We only represented four most differentiating variables to each cluster in group of transactor user by Chi-square test.

Revolver User 5, 303 3.2% $PA: 0$ $CV: 0.8$ $C13$	*Revolver User* 12, 623 8% $PA: 0$ $CV: 0.5$ $C14$	*Revolver User* 1 0% $PA: 0$ $CV: 1.5$ $C15$	*Revolver User* 63, 096 40% $PA: 0.03$ $CV: 2255$ $C16$
Revolver User 13, 598 8.6% $PA: 0.04$ $CV: 947$ $C9$	*Revolver User* 3, 020 1.9% $PA: 0$ $CV: 8$ $C10$	*Revolver User* 7, 611 4.8% $PA: 0$ $CV: 4$ $C11$	0 $C12$
Transactor User 4, 314 2.7% $PA: 0.95$ $CV: 378$ $C5$	*Transactor User* 597 0.37% $PA: 1$ $CV: 4$ $C6$	*Convenience User* 1, 508 0.95% $PA: 0.5$ $CV: 6.8$ $C7$	*Convenience User* 2, 007 1.26% $PA: 0.5$ $CV: 4$ $C8$
Transactor User 22, 458 14.2% $PA: 0.954$ $CV: 780$ $C1$	*Transactor User* 4, 576 2.8% $PA: 0.983$ $CV: 99$ $C2$	*Convenience User* 6, 006 3.7% $PA: 0.47$ $CV: 200$ $C3$	*Convenience User* 11, 681 7.3% $PA: 0.47$ $CV: 1320$ $C4$

Figure 3. Neurons arranged in a 4×4 map, each neuron defined a cluster.

On the basis that no meaningful conclusions can be drawn from small numbers of customers, no future analysis was performed on clusters fewer than 1,000 cases (i.e. cluster 6, 12 and 15). The next major step is to choose the target groups of customers, so as to choose the target customers for direct marketing or encourage consumption [23]. The consumer behavior can be used to indicate the risk of customers, the risk degrees among three profitable groups of customers are "transactor user" $\leq$ "convenience user" $\leq$ "revolver user". Moreover, the clusters of customer values tend to larger of each profitable group are selected as target ones, all customers who belong to these clusters become candidates for conducting suitable marketing strategies for a bank, which attract the most attention.

5. Create Customer Profiles by Decision Tree

Once the clusters and the associated statistical summarized data are made, the decision tree inducer is assisted to create and verify the customer profiles. The table obtained after data preprocessing originally contained 86 attributes, 32 character attributes and 54 continuous attributes. These attributes may contain irrelevant or un-important attributes which must be removed for the purpose of creating more accuracy customer profiles. This study uses neural network sensitivity analysis to reserve the relative importance attributes, pay-off ability choose as predicated variable. As recommended by Hornik et al. [18], one hidden layer network is sufficient to model a complex system with any desired accuracy, and the employed neural network model has just one hidden layer.

Figure 4. Statistical summarized data of transactor user.

Table 1 lists the distribution of the relative importance for each input variable using the neural network. The sensitivity analysis of the neural network and the order of most significant input variables indicate those variables that are worth looking at in more detail. Finally, factors with a relative importance of 0.00997 and above were used in successive customer profiling. In table 1, No_of_Purchases, and Repayment_Cycle and Creditline, were the three most differentiating variables. On the other hand, Age_Segments, CardAge_Month and Sex were the least differentiating variables.

Table 1. Sensitivity analysis the relative importance of input variables.

Neural Network Model		
Input Layer (no. of neurons)	32	
First Hidden Layer (no. of neurons)	20	
Output Layer (no. of neurons)	1	
Predicted Accuracy	96%	
Relative Importance to PA&CV		
Variable Name	**Relative Importance**	**Comments**
No_of_purchases	0.40658	Number of purchases / month
Repayment_Cycle	0.30048	Repayment cycle days: 1∼22 days
CreditLine	0.19980	Credit line
Total_Consumption	0.15002	Yearly amount of consumption
Blockcode	0.12431	Card usage limit or not
Occupation	0.03722	Encoded field
CardType	0.02546	Encoded field
Marital_Status	0.01779	0: single, 1: married, 2: divorced 3: separation
Customer_value	0.01680	Customer value
Age_Segments	0.01499	1:<25, 2:25∼30, 3:30∼35, 4:35∼40, 5:40∼45, 6:45∼50, 7:50∼60, 8:>60
CardAge_Month	0.01350	Number of months for which the card has been held
Sex	0.00997	1: male, 0: female
:	:	

Customer profiling provides a basis for enterprises to offer customers better services and retain good customers. Classification is a customer profiling technique for creating model to predict class membership and has been applied to numerous applications. A number of different classification techniques have been proposed [5, 6], and the induction of decision trees is a well-known approach. Decision tree learning function is a technique for approximating discrete-valued target functions that performs tests and attempts to identify the best sequence for predicating the target classification variables. The commonly implemented decision trees, including Chi-squared automatic interaction detection (CHAID), classification and regression trees (CART), GID3 and C4.5 [6, 14, 26]. The decision tree inducer can be assisted marketers to fulfill proper applications.

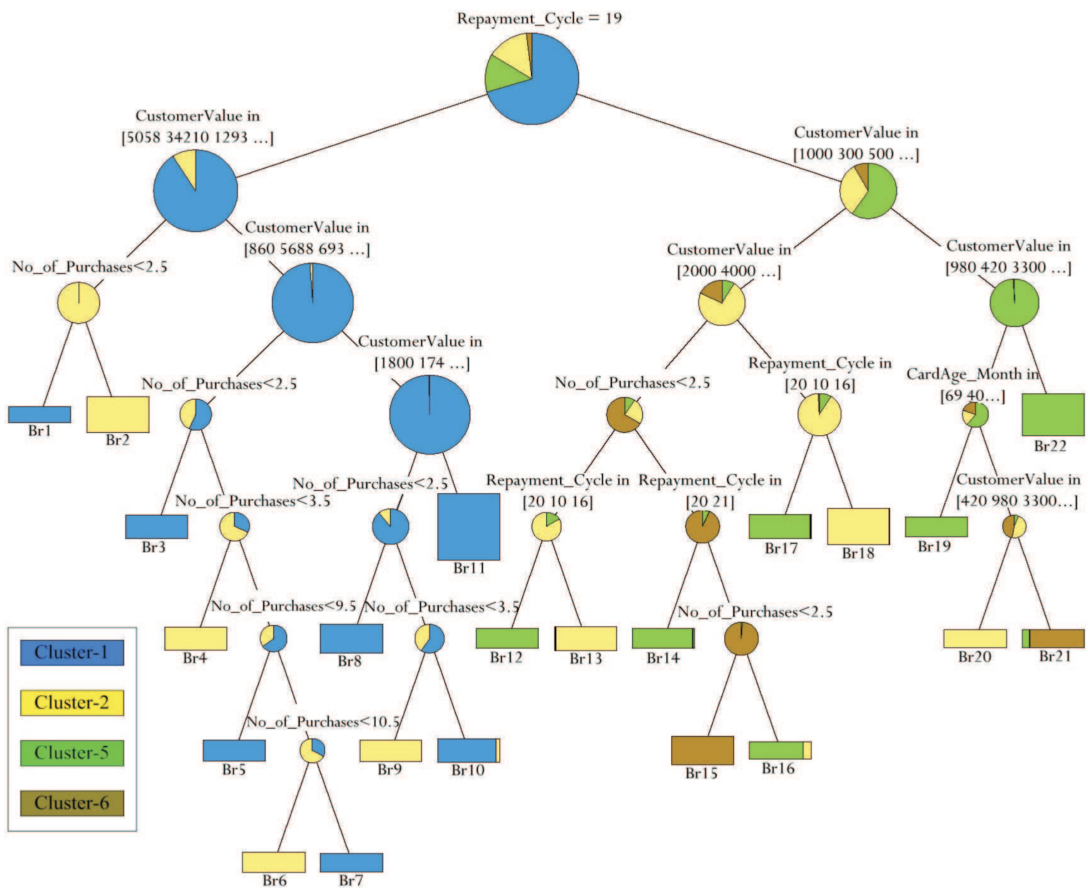

Figure 5. A decision tree for transactor user (Cluster-1, 2, 5, 6).

For each profitable group of customers, the inputs to the decision tree inducer consisted of those customers that were classified by the variable Cluster-ID. The leaf classifications provide preliminary classifying of the customers into distinct cluster patterns (e.g. transactor user; cluster-1, cluster-2, cluster-5 and cluster-6). The 70% of samples were used to generate a decision tree. The resulting decision tree was then tested against the remaining 30% samples. Parameters were set up to identify rules that had at least 10 individuals to the total number of customers in each decision tree branch, and the classification error rate was 0.1377%, with a pruning severity of 95% imposed on the decision tree.

In Figure 5, the decision tree allows further examination of the attributes to the importance in customer behavior classification. Traversing the decision tree from the root to the target leaf node provides the classification rule that contributed to explain the properties of clusters. After briefly reviewing the 16 clusters using cluster profiles, the customers with values tend to $CV\uparrow$ can be targeted with greater accuracy. However, the risk arising from the different profitable groups of customers in practical applications should be considered. After the decision tree is grown, it is further processed to generate classification rules, where each rule represents a customer profile that was dominant or most strongly associated with the customers matching that cluster. From this, marketers can create more accurate campaigns towards each target group of customers for strategies of adjust customers' annual membership fees and interest rate, cross-selling, and encouraging consumption.

6. Removal of Irrelevant Values within a Decision Tree

The methods of decision trees and classification rules have been significantly applied to numerous data mining systems. However, the complex data set with high dimensional attributes, which may influence the product of a simple tree in its product opaque structure to be incomprehensible or inefficient. The removal of irrelevant values within the decision tree leads to simplified classification rules. It also avoids unnecessary checking and decreases the difficulty of the application in the domain knowledge. Fayyad [14] has proposed GID3 and GID3* algorithm to solve the over-specialization problem. In these two algorithms, before attribute selection, irrelevant values were removed at the attribute phantomization step. In this study, we used a decision tree simplified method, which depends only on the semantics of the decision tree, and can be applied to any decision tree algorithms.

In order to figure out irrelevant values, a decision tree can be represented by a set of virtual branches. Let $T = \bigcup_{\forall B_r} B_r$ be a decision tree, $A = \{A_1, \ldots, A_n\}$ be a set of attributes, $C = \{C_1, \ldots, C_S\}$ be a set of classes, and (A_i, a_{ij}) denotes an attribute-value pair, where $a_{ij} \in dom(A_i)$ and $dom(A_i)$ is the attribute domain of attribute A_i. Then, a branch B_r of a decision tree can be represented as the form $(B_r[A_1], \ldots, B_r[A_n], C_i)$ or $(B_r[A_1, \ldots, A_n], C_i)$, where $1 \leq i \leq s$. Moreover, according to [26], this branch B_r can be converted to a classification rule: $B_r[A_1] \wedge \cdots \wedge B_r[A_n] \rightarrow C_i$. We defined virtual branches for a decision tree as the following:

Definition 1. Let $B_r = (B_r[A_1, \ldots, A_n], C_i)$ be a virtual branch in a decision tree, and the virtual branches of B_r to attributes $A_1, \ldots, A_n$ are:

$$\left\{(B_r[A_1, \ldots, A_n], a_{j1}, \ldots, a_{jn}, C_i) | (a_{j1}, \ldots, a_{jn}) \in dom(A_1) \times \cdots \times dom(A_n)\right\}.$$

In a decision tree, not all attributes A_j need to be nodes of branches in the decision tree. That is, the corresponding branch values $B_r[A_j]$ may be absent from the branches. These branch values can be seen as irrelevant in these branches, and the irrelevant value of the rule, $B_r[A_j]$, means this branch value can be removed or replaced by any values from the same attribute domain without affecting the correctness of the classification rule.

According to the semantics of the irrelevant value, if the branch value, $B_r[A_j]$, is irrelevant of a rule, means this value can be removed or replaced by any values from the same domain value without affecting its correctness of the classification rule. There are two cases in the removal of irrelevant values. First, let B_r be a branch in a decision tree, where $B_r = (B_r[A_1, \ldots, A_j, \ldots, A_k], C_{i1})$. When $A_{j-m}, \ldots, A_j$ are the same attributes and the criterion of $B_r[A_{j-m}, \ldots, A_{j-1}]$ contained in $B_r[A_j]$, then $B_r[A_{j-m}, \ldots, A_{j-1}]$ are irrelevant values and can be removed from the branch. Second, let B_r and $B_{r'}$ be two distinct branches in a decision tree, where $B_r = (B_r[A_{k1}, \ldots, A_{k2}], C_{i1})$. If B_r is in conflict with $B_{r'}$ to attributes $A_{k1}, \ldots, A_{k2}$ if and only if $(B_r[A_{k1}, \ldots, A_{k2}], C_{i2})$ is a part of the virtual branch $B_{r'}$, and $C_{i1} \neq C_{i2}$. The following definition presents how to identify irrelevant values of a virtual branch.

Definition 2. Let $T = \bigcup_{\forall B_r} B_r$ be a decision tree and $B_r = (B_r[A_1, \ldots, A_j, \ldots, A_k], C_{i1})$ be a branch in the decision tree:

(1) When $A_{j-m}, \ldots, A_j$ are the same attributes and the criterion of $B_r[A_{j-m}, \ldots, A_{j-1}]$ contained in $B_r[A_j]$, then $B_r[A_{j-m}, \ldots, A_{j-1}]$ are irrelevant values and can be removed from the branch.
(2) When B_r is not conflict with any other branches to attributes $A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_k$, then the branch value $B_r[A_j]$ is an irrelevant value of the branch; otherwise, $B_r[A_j]$ is not an irrelevant value of the branch.

Let us consider the left sub-decision tree in Figure 5 After applying the first rule in Definition 2, the virtual branches of $(B_{r1}, \ldots, B_{r11})$ with respect to attributes Repayment_Cycle, CustomerValue, and No_of_Purchases in the decision tree are as Table 2.

By the second rule in Definition 2, since B_{r1}, B_{r3} and B_{r8} are not in conflict with other branches to attributes Repayment_Cycle and No_of_Purchases. Therefore B_{r1} (CustomerValue in [5058 34210 1293$\cdots$]), B_{r3} (CustomerValue in [860 5688 693$\cdots$]) and B_{r8} (CustomerValue in [1800 174$\cdots$]) are irrelevant values in B_{r1}, B_{r3}, B_{r8} respectively, and can be removed from the decision tree. The process of removing irrelevant values not only solves the problem of over-specialized branches but also carries a side-effect to help the problem of missing branches [9, 14]. Table 3 shows the classification rules without irrelevant values. The irrelevant values of the right sub-decision tree can be removed accordingly.

In practical application, a combinatorial explosion was made in the number of comparisons to all the branches to identify all irrelevant values of a virtual branch. Consequently, we can remove irrelevant values from the corresponding classification rules and applied the final classification rules according to the sequence of rule's identifier. Once the classification rules are made and agreed upon, the statistical summarized data of each cluster and classification rules can be useful tool to show how profitable a customer is.

7. Conclusion

This study proposes a two-stage framework of consumer behavior analysis for analyzing bank databases. For discriminating purposes, we grouped customers with shared

Table 2. The virtual branches of left sub-decision tree.

B_{r1} = (Repayment_Cycle = 19, CustomerValue in [5058 34210 1293$\cdots$], No_of_Purchases < 2.5, Cluster-1)
B_{r2} = (Repayment_Cycle = 19, CustomerValue in [5058 34210 1293$\cdots$], No_of_Purchases > 2.5, Cluster-2)
B_{r3} = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 2.5, Cluster-1)
B_{r4} = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 3.5], Cluster-2)
B_{r5} = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 9.5, Cluster-1)
B_{r6} = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 10.5, Cluster-2)
B_{r7} = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases > 10.5, Cluster-1)
B_{r8} = (Repayment_Cycle = 19, CustomerValue in [1800 174$\cdots$], No_of_Purchases < 2.5, Cluster-1)
B_{r9} = (Repayment_Cyclev19, CustomerValue in [1800 174$\cdots$], No_of_Purchases < 3.5, Cluster-2)
B_{r10} = (Repayment_Cycle = 19, CustomerValue in [1800 174$\cdots$], No_of_Purchases > 3.5, Cluster-1)
B_{r11} = (Repayment_Cycle = 19, CustomerValue not in [1800 174$\cdots$], x, Cluster-1 | $x \in dom$(No_of_Purchases) $\cdots\cdots$

Table 3. The classification rules without irrelevant values.

Rule_1 = (Repayment_Cycle = 19, No_of_Purchases < 2.5, Cluster-1)
Rule_2 = (Repayment_Cycle = 19, CustomerValue in [5058 34210 1293$\cdots$], No_of_Purchases > 2.5, Cluster-2)
Rule_3 = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 3.5, Cluster-2)
Rule_4 = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 9.5, Cluster-1)
Rule_5 = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases < 10.5, Cluster-2)
Rule_6 = (Repayment_Cycle = 19, CustomerValue in [860 5688 693$\cdots$], No_of_Purchases > 10.5, Cluster-1)
Rule_7 = (Repayment_Cycle = 19, CustomerValue in [1800 174$\cdots$], No_of_Purchases < 3.5, Cluster-2)
Rule_8 = (Repayment_Cycle = 19, CustomerValue in [1800 174$\cdots$], No_of_Purchases > 3.5, Cluster-1)
Rule_9 = (Repayment_Cycle = 19, CustomerValue not in [1800 174$\cdots$], Cluster-1) $\cdots\cdots$

consumer behavior and customer value. After briefly reviewing the customer profiles using the decision tree inducer, the customers with a higher customer values might be the target customer groups of precedence. Then marketers can made more accuracy decisions on each target group of customers for specific marketing strategies. This study provides good experience of analyzing credit card database. Beyond simply understanding customer value, the bank also attempts to establish a best customer relationship, increase customer loyalty and revenue.

Additionally, for bankruptcy application, the present available scoring models apply single classification rule to every customer monotonously. This kind of classification might be ineffective because a single classification rule cannot catch the fine nuance of various customers. The proposed two-stage framework also can be applied to predicate personal bankruptcy among bank customers to the bank databases. By the proposed class-wise classification process, the scoring models can predicate the credibility of customers by setting separate classifiers for each cluster at various consumer behaviors. This enables the lenders to predict the borrowers' credibility more accuracy.

References

[1] Au, W. H. and Chan, K. C. C., *Mining fuzzy association rules in a bank-account database*, IEEE Transactions on Fuzzy Systems, Vol. 11, 2003.

[2] Baesens, B., Viaene, S., Poel, D., Vanthienen, J. and Dedene, G., *Bayesian neural network for repeat purchase modelling in direct marketing*, European Journal of Operational Research, Vol. 138, pp.191-211, 2002.

[3] Balakrishnan, P. V. S., Cooper, M. C., Jacob, V. S. and Lewis, P. A., *Comparative performance of the FSCL neural net and K-means algorithm for market segmentation*, European Journal of Operational Research, Vol. 93, pp.346-357, 1996.

[4] Berry, M. J. A. and Linoff, G. S., *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 2nd Edition*, John Wiley & Sons, 2004.

[5] Cai, Y., Gercone, N. and Han, J., *An attribute-oriented approach for learning classification rules from relational databases*, Proceedings of Conference on Data Engineering, pp.281-288, 1990.

[6] Cendrowska, J., *PRISM: an algorithm for inducing modular rules*, International Journal of Man-Machine Studies, Vol. 27, pp.349-370, 1988.

[7] Chan, C.-C. H., *Intelligent spider for information retrieval to support mining-based price prediction for online auctioning*, Expert Systems with Applications, Vol. 34, pp.347-356, 2008.

[8] Chan, C.-C. H., *Online auction customer segmentation using a neural network model*, International Journal of Applied Science and Engineering, Vol. 3, pp.101-109, 2005.

[9] Chiang, D. A., Chen, W., Wang, Y. F. and Hwang, L. J., *Rules generation from the decision tree*, Journal of Information Science and Engineering, Vol. 17, pp.325-339, 2001.

[10] Dasgupta, C. G., Dispensa, G. S. and Ghose, S., *Comparing the predictive performance of a neural network model with some traditional market response models*, International Journal of Forecasting, Vol. 10, pp.235-244, 1994.

[11] Davies, F., Moutinho, L. and Curry, B., *ATM user attitudes: a neural network analysis*, Marketing Intelligence & Planning, Vol. 14, pp.26-32, 1996.

[12] Donato, J. C., Schryver, G. C., Hinkel, R. L., Schmoyer, J., Leuze, M. R. and Grandy, N. W., *Mining multi-dimensional data for decision support*, Future Generation Computer Systems, Vol. 15, pp.433-441, 1999.

[13] Dyche, J. and Dych, J., *The CRM Handbook: A Business Guide to Customer Relationship Management, Addison-Wesley Pub Co.*, August 2001.

[14] Fayyad, U. M., *Branching on attribute values in decision tree generalization*, Proceedings of 20th National Conference on Artificial Intelligence, AAAI-94, pp.104-110, 1994.

[15] Feelders, A. J., *Credit scoring and reject inference with mixture models*, International Journal of Intelligent Systems in Accounting, Finance and Management, Vol. 9, pp.1-8, 2000.

[16] Fish, K. E., Barnes, J. H. and Aiken, M. W., *Artificial neural networks - A new methodology for industrial market segmentation*, Industrial Marketing Management, Vol. 24, pp.431-438, 1995.

[17] Hadden, J., Tiwari, A., Roy, R. and Ruta, D., *Computer assisted customer churn management: State-of-the-art and future trends*, Computers & Operations Research, Vol. 34, pp.2902-2917, 2005.

[18] Hornik, K., Stinchcombe, M. and White, H., *Multilayer feedforward networks are universal approximations*, Neural Networks, Vol. 2, pp.336-359, 1989.

[19] Hsieh, N.-C., *Hybrid mining approach in the design of credit scoring models*, Expert Systems with Applications, Vol. 28, pp.655-665, 2005.

[20] Kim, Y. S. and Sohn, S. Y., *Managing loan customers using misclassification patterns of credit scoring model*, Expert Systems with Applications, Vol. 26, pp.567-573, 2004.

[21] Kohonen, T., *Self-organizing maps*, Springer, Berlin, 1995.

[22] Kuo, R. J., An, Y. L., Wang, H. S. and Chung, W. J., *Integration of Self-organizing Feature Maps Neural Network and Genetic K-Means Algorithm for Market Segmentation*, Expert Systems with Applications, Vol. 30, pp.313-324, 2006.

[23] Lancher, R. C., Coats, P. K., Shanker, C. S. and Fant, L. F., *A neural network for classifying the financial health of a firm*, European Journal of Operational Research, Vol. 85, pp.53-65, 1995.

[24] Lee, T. S., Chiu, C. C., Lu, C. J. and Chen, I. F., *Credit scoring using the hybrid neural discriminate technique*, Expert Systems with Applications, Vol. 23, pp.245-254, 2002.

[25] Malhotra, R. and Malhotra, D. K., *Evaluating consumer loans using neural networks*, Omega, Vol. 31, pp.83-96, 2003.

[26] Quinlan, J. R., C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo, California, 1993.

[27] Setiono, R., Thong, J. Y. L. and Yap, C. S., *Symbolic rule extraction from neural networks - An application to identifying organizations adopting IT*, Information & Management, Vol. 34, pp.91-101, 1998.

[28] Tao, Y. H. and Yeh, C. C. R., *Simple database marketing tools in customer analysis and retention*, International Journal of Information Management, Vol. 23, pp.291-301, 2003.

[29] Vellido, A., Lisboa, P. J. G. and Vaughan, J., *Neural networks in business: a survey of applications* (1992∼1998), Expert Systems with Applications, Vol. 17, pp.51-70, 1999.

[30] Zhang, G., Hu, M. Y., Patuwo, B. E. and Indro, D. C., *Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis*, European Journal of Operational Research, Vol. 116, pp.16-32, 1999.

Authors' Information

Nan-Chen Hsieh received the Ph.D. degree in Computer Science and Information Engineering from the Tamkang University, Taiwan, in 1996. He is currently a Professor in the Department of Information Management at National Taipei College of Nursing, Taiwan. His research interests are in the area of data mining, knowledge discovery, fuzzy logic and soft-computing.

Department of Information Management, National Taipei College of Nursing, Taipei, Taiwan 11219, R.O.C.

E-mail: nchsieh@ntcn.edu.tw TEL: +886-2-2822-7101 ext. 2200

Kuo-Chung Chu received the Ph.D. degree in Information Management from the National Taiwan University in 2005. He is currently an Associate Professor in the Department of Information Management at National Taipei College of Nursing, Taiwan. His research interests include decision modeling, simulation, and optimization approach.

Department of Information Management, National Taipei College of Nursing, Taipei, Taiwan 11219, R.O.C.

E-mail: kcchu@ntcn.edu.tw TEL: +886-2-2822-7101 ext. 2200