Applying ANNs into Constructing the CLV Discriminant model

Ching-Cheng Shen
Graduate Program in Travel Industry Management
National Kaohsiung Hospitality College
1, Sung-Ho Rd., Hsiao-Kang, Kaohsiung, Taiwan, R.O.C.

Kun-Lin Hsieh*
Department of Information Management
National Taitung University
684, Sec. 1, Chung-Hua Rd., Taitung 950, Taiwan R.O.C.

Abstract: - Reviewing the marketing environment, the enterprises had met a competitive pressure due to there are many customers with different characteristics. How to achieve the potential customer retention and maintain their core competition will be an important issue. That is, they will meet the issue of measuring the customer lifetime value (CLV). In this study, a CLV discriminant model by using artificial neural networks (ANNs) will be proposed and the key factors can also be determined. Finally, an illustrative case owing to Taiwanese health club will be employed to demonstrating the rationality and feasibility of the propsoed approach.

Key-Words: - Customer lifetime value (CLV), Discriminant analysis, Artificial neural networks (ANNs)

1 Introduction

Reviewing the marketing environment, the enterprises had met a competitive pressure due to there are many customers with different characteristics. Hence, the performance of making the market segmentation by using the social attributes will not be the only one choice. Schmittlein(1987) had mentioned that the segmentation performance from the customers' database is better than that from the social attributes. And, it will lead the customer relationship management (CRM) to be an useful evaluation index to study the customers' loyalty. At the same time, the related researches had also mentioned to create customer value, avoid customer lose, reduce the searching and transaction costs. The customer's requirements can be kept via the data mining techniques and it can provide the active and customerized service to enhance the enterprise's core competition. Data mining techniques had been employed to mine the behind behaviours and make the prediction analysis for the recent years. Therefore, it will let it be an important tool to achieve the CRM. Customer value can be viewed as the infrastructure of CRM (Morgam&Hunt, 1994). Kotler(2000) also pointed out that the importance of relationship management must be focus on how to create the long term relationship with the valued customers and to maintain such relationship for earning profits. Several customer value models were proposed (Sewell & Brown, 1990; Hughes, 1994; Kotler, 2000). Berger & Nasr(1998) also intended to propose a systematic model to compute the customer value, and they summarized the characteristics from Jackson(1985) to form a model with five categories. However, Hughes(1994) proposed a RFM customer value analysis model to address such issue based on recency (R), frequency (F), monetary (M). Basically, the information about RFM model can be mined from the transaction database and it had been frequently used for many enterprises. However, the model construction of RFM model may be difficult and it will limit its applications due to that the assumption being made about the probability distribution of the customer's behaviour.

The history of Taiwan's health clubs is about twenty years. However, the market gradually grew up and made a competitive market environment. Under the strategy of the lower prcining, the cost of creating new customers and the lost rate of club's associators will gradually increase. It will directly react on the profits. Hence, the issue of CRM will play an important role for health clubs, especial for how to provide the equipments or services with high quality to satisfy the customers and decrease the customer lose. However, the buying behaviours of customers are significantly different and it will meet a problem to evaluate the customer value. In the past years, the customer values are frequently evaluated by using RFM model. However, can it directly use for the current status? The health clube will meet such problem. After reviewing the related studies, we will intend to apply the data mining techniques to mine the useful information to construct the customer value prediction model and study the significant factors affecting the customer value.

2 Literature Review

2.1 Customer Lifetime Value (CLV)

Heskett et al. (1994) recognized the customers with re-buying the products or services will lead to a very larger value if the affection of introduction can be included. However, not all customers will be worthy to retain. Hence, the 80/20 rule will be applied in practice due to that about 80% profits will be created via about 20% customers. And, it will lead the enterprises to focus on creating the profitable customers, not promoting their products or services. The concept of customer lifetime value (CLV) will be constructed depending on such thinking. It can be viewed as a concept or a computation result from the viewpoint of the contributions of the customers to the enterprise's profits (Zeithaml and Bitner, 2002).

Levin(1999) recognized the CLV to be the profits from all processes focusing on keeping the relationship between the enterprise and customers. Ranchhod(2002) provided that an important factor affecting the CLV is customer satification and the dynamic environments need to be taken into consideration since evaluating the cutomer's profits. Several studies (Berry,1983；Shain and Chalasain,1992；Morgan and Hunt,1994；Berger and Nasr ,1998； Kotler , 2000) also pointed that the profits of an enterprise can be significantly increased due to keeping the stable and long-term relationship with customers. Kotler (2000) pointed out the core of relationship marketing will be put attention on how to construct the long-term relationship. The CLV had re-defined the related activities for the tranditional marketing due to that the customers will be regarded as an assest. That is, the decision-making of marketing can be viewed as an investment and it will evaluate the future benefits and costs to determine the related activities. Hence, how to mine the valued customers will be an important issue to most enterprises. Kotler & Armstrong(1996) also made the definition about CLV as "The customer making the future profits exceeds the cost spending on it. And, the exception part can be called as CLV". However, several studies (Monroe , 1991；Gale , 1994；Naumann , 1995；Engel, Minird and Blackwell , 1995；Woodruff , 1997； Solomon, 2000) recognized that the CLV to be the value in their mind, and it was called as value of customer recognization. Besides, several studies (Kamakura and Russell1 , 1989；Stone and Bob , 1995；Hughes and Arthur , 1994；Mulhern, 1999 Phillip and Robert , 2000) had also pointed out that the CLV can also be analyzed via the historical transaction data and the CLV will be defined as the predicted value for the current and the future.

Most enterprises frequently apply RFM indexes to quantilize the customer, and the purpose is to quantilize the customer's behaviour and make it obey the marketing formula. The RFM indexes can measure the relationship between customer and enterprise, judge the customer value. Hence, the suitable strategy of custoerm relationship can be determined. Hughes(1994) and Stone(1995) had proposed two different methods to address RFM model. Hughes(1994) considered the same mportance for those three indexes and set the same weight value to them. However, Stone(1995) proposed the different weight value setting for those three indexes via demonstrating a credit card case.

2.2 Customer lose

Customer had known as the keypoint to the grow of an enterprise. Hence, customer lose will lead to the enterprise loss. Besdies, customer lose, customer retention, customer loyalty will represent the same concept. The lower customer lose will denote the higher customer loyalty. Customer retention will be the evaluation of customer loyalty. Hence, the purpose of an enterprise wish to maxmize the customer retention and minimize the customer lose. Customer lose can be viewed as one part of CRM. According to the statistical record, about 85% customers may be retain for an enterprise and about 15% customers will be lose. It will form a rotation theory with inputting and ourputting. Hence, we can apply data analysis or data mining techniques to mine the potential customer lose and mine the possible causes. Then, the enterprise can make the necessary action to avoid customer lose (Li, 1995；Chang and Yuan , 1999；Madden , Savage and Coble-Neal , 1999；Hughes , 2001；Daskaka, Kopanas, Gou and Avouris , 2003).

2.3 Discriminant Analysis

Discriminant Analysis may be used for two objectives: either we want to assess the adequacy of classification, given the group memberships of the objects under study; or we wish to assign objects to one of a number of (known) groups of objects. In both cases, some group assignments must be known before carrying out the Discriminant Analysis. Such group assignments, or labelling, may be arrived at in any way. Hence Discriminant Analysis can be employed as a useful complement to Cluster Analysis (in order to judge the results of the latter) or Principal Components Analysis (SPSS, 2000; Davis, 1986; Weslowsky, 1976). Linear Discriminant Analysis is the 2-group case of MDA. It optimally separates two groups, using the Mahalanobis metric or generalized distance. It also gives the same linear separating decision surface as Bayesian maximum likelihood discrimination in the case of equal class covariance matrices. There is no best discrimination method. A few remarks concerning the advantages and disadvantages of the methods studied are as follows. Analytical simplicity or computational reasons may lead to initial consideration of linear discriminant analysis. Linear discrimination is the most widely used in practice. Often the 2-group method is used repeatedly for the analysis of pairs of multigroup data (yielding k(k-1)/2 decision surfaces for k groups). To estimate the parameters required in quadratic discrimination more computation and data is required than in the case of linear discrimination. If there is not a great difference in the group covariance matrices, then the latter will perform as well as quadratic discrimination.

2.4 Backpropagation neural network

A neural network consists of a number of simple, highly interconnected processing elements or nodes and is a computational algorithm that processes information by a dynamic response of its processing elements and their connections to external inputs. A neural network can model the non-linear relationship between the system's input and system's output. The non-linear relationship or the interaction effect among several variables can be kept in the structure of hidden layer of a neural network model. Restated, it can be learned by passing the training pairs through the network. Among the several conventional supervised learning neural models including the perceptron, backpropagation neural network (BPNN), learning vector quantization (LVQ), and counter propagation network (CPN), the BPNN model is frequently used (Ko et al., 1998; Neural Ware, 1990; Hsieh, 2001; Hsieh, 2006) and, therefore, it will be selected herein. A BPNN consists of three or more layers, including an input layer, one or more hidden layers, and an output layer. Detailed descriptions of the algorithm can be found in various sources (Neural Ware, 1990; Rumelhart et al., 1986). The following is a brief description. To develop a backpropagation neural network, the training and testing data set are firstly collected. The data sets consist of both the input parameters and the resulting output parameters. The backpropagation learning algorithm employs a gradient- or steepest- heuristic that enables a network to self organize in ways that improve its performance over time. The network first uses the input data set to produces its own output. This forward pass through the backpropagation network begins as the input layer receive the input data pattern and passes it to the hidden layer. Each processing element (PE) calculates an activation function in first summing the weighted inputs. This sun is then used by an activation function in each node to determine the activity level of the processing node. The output generated by the network is compared to the known target value. If there is no difference, no learning takes place. If a difference exists, the resulting error term is propagated back through the network, using a gradient- or steepest- descent heuristic to minimize the error term by adjusting the connection weights (Neural Ware, 1990; Rumelhart et al., 1986). As for the training phase, a signal input pattern is presented and the network adjusts the set of weights in all the connecting links such that the desired output is obtained at the output node. On accomplishing the adjustment, the next pair of input and output target value is presented and the network learns that association.

3 Proposed approach and case study

In this section, we will intend to construct a CLV model according to the customer data. In order to demonstrate the rationality and feasibility of the proposed model, we take an illustratve case at Taiwanese health club. In this case, we will also discuss the key factors affecting the CLV and then the comparison of those two models, BPNN and tranditional discriminant analysis, will be made.

3.1 Data analysis

In order to collect the necessary data, we apply the questionaire technique to collect the data from those club associators. The related data structure will be designed as Table 1. And the transferred data structure can be listed in Table 2. Totally, we collect about 160 data. We divide it into two parts according to the 4:1 criterion : the training part is about 128 data and the testing part is about 32 data.

Table 1. Data structure

Item	Variable	Description
1	Number of associator	ch
2	Sex	M：male F：female
3	Birthday	date
4	Time to Apply	date
5	Register fee	numberical
6	Payment fee	numberical
7	Living location	ch

Table 2. Transformation about Data structure

Item	Scale	Variable	Description
1	unuse	Number of associator	ch
2	Nomial	Sex	0：male 1：female
3	Ratio	Age	numberical
4	Ratio	Lifetime	numberical
5	Ratio	Register fee	numberical
6	Ratio	Payment fee	numberical
7	Nomial	Living location	0：outside city 1：inside city

3.2 Discriminant analysis

The following table will represent the results of discriminant analysis for the training data and the key factors affecting discriminant analysis. Herein, we consider thirteen variables as sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location, traffic tool. In order to study the key factors with affecting on discriminant analysis, we will compare the discriminant ratio via disableing necessary variable step by step. The criteria is "if the discriminant ratio will increase since disabling variable, the variable will be decided to disable."

Table 3(1). The comparison table for deleting the variable from the model.

Ratio	Disable variable	Variables in model	Ratio	Disable variable	Variables in model
42.87% (54/128)	-	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location, traffic tool	43.75% (56/128)	Traffic tool time	sex, age, marriage, icomes, associator type, frequency, satisfication, comsumption, way to pay, from which location, living location
46.09% (59/128)	Traffic tool	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location	40.625% (52/128)	Traffic tool frequency	sex, age, marriage, icomes, associator type, time, satisfication, comsumption, way to pay, from which location, living location
39.84% (51/128)	Traffic tool Living location	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location	46.09% (59/128)	Traffic tool Associator type	sex, age, marriage, icomes, frequency, time, satisfication, comsumption, way to pay, from which location, living location
46.09% (59/128)	Traffic tool From which location	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, living location	42.96% (55/128)	Traffic tool incomes	sex, age, marriage, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location

Table 3(2). The comparison table for deleting the variable from the model.

Ratio	Disable variable	Variables in model	Ratio	Disable variable	Variables in model
42.96% (55/128)	Traffic tool Way to pay	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, from which location, living location	39.06% (50/128)	Traffic tool marriage	sex, age, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location
45.31% (58/128)	Traffic tool consumption	sex, age, marriage, icomes, associator type, frequency, time, satisfication, way to pay, from which location, living location	39.06% (50/128)	Traffic tool age	sex, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location
43.75% (56/128)	Traffic tool satisfication	sex, age, marriage, icomes, associator type, frequency, time, comsumption, way to pay, from which location, living location	39.06% (50/128)	Traffic tool sex	age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location

The comparision table will listed in Table 3(1)~3(2). Finally, we can get the Fisher discriminant function for CLV in Table 4 and the corresponding statistical test is listed in Table 5. Next, we apply the test data to make validation for the simiplified Fisher discriminant function (only including sex, age, marriage, frequency, satisfication, from which location, way to pay), and we get the discriminant ratio is about 53.125%. And the distribution for the testing data will be given as Table 6.

Table 4. Fisher discriminant function for CLV.

	CLV					
	1	2	3	4	5	6
sex	8.007	9.128	8.877	7.579	8.443	8.438
Living location	0.187	0.218	0.407	0.228	0.371	0.237
Way to pay	9.284	9.486	11.004	11.662	10.877	11.243
satification	5.551	5.864	5.826	5.600	4.978	5.359
frequency	1.286	0.552	1.068	1.192	1.165	1.067
incomes	1.765	1.542	1.576	1.660	1.946	1.916
marriage	-0.617	-1.363	-0.667	-0.180	-0.913	-0.405
age	1.936	2.575	2.964	2.486	3.231	3.252
consumption	0.159	0.0086	-0.220	-0.174	0.129	-0.0742
From which location	2.818	2.577	2.454	3.153	3.260	3.026
time	7.453	8.409	8.550	7.788	7.896	7.944
type	0.510	0.337	0.343	0.395	0.374	0.404
(constant)	-44.485	-46.374	-51.688	-48.597	-49.675	-50.522

Table 5. The result of statistical test.

discriminant function	eigenvalue	Variation (%)	Accumulated vraiation	Wilk's Lambda	Ch-square	df	significance
1	.516	64.1	64.1	.500	83.081	40	.000
2	.123	15.2	79.4	.759	33.132	28	.231
3	.091	11.4	90.7	.852	19.244	18	.377
4	.056	6.9	97.7	.930	8.748	10	.556

Table 6. Distribution of the testing data for discriminant analysis.

Test / Original Category	data size	Category	1	2	3	4	5	6
		Data size	5	4	3	7	4	9
1	5	CLV	3	1	1	0	0	0
2	4		0	2	0	2	0	0
3	3		1	0	0	1	0	1
4	6		0	0	1	3	1	1
5	3		0	0	0	1	2	0
6	11		1	1	1	0	1	7

4.3 BPNN

The related parameters about BPNN will be given as follows:

1.Processing elements of input layer: It includes 13 variables and the number of PEs will be set as 13.
2.One hidden layer
3.As for the PEs of hidden layer, we will choose 12 to be the number of PEs due to that it is the optimum decision after pilot run.
4.The ratio of traing and testing data is set as 4:1, that is, there are 128 training data and 32 training data.
5.The learning rate will be set as 0.05 and the momentum will be set as 0.8.
6.The number of PEs of the output layer will be set as 1 due to the only one discriminant variable.

Table 7 will represent the results of discriminant analysis for the training data and the key factors affecting discriminant analysis. Herein, we take the same approach mentioned in discriminant analysis for running BPNN. The comparison result will be listed in Table 7. Finally, we can get the key factors affecting discriminant analysis are sex, age, marriage, incomes, consumption and time. Next, we apply the test data to make validation for the simiplified BPNN model (only including sex, age, marriage, incomes, consumption and time), and we get the discriminant ratio is about 96.87% (listed in Table 8). The result significantly exceeds the result obtained from discriminant analysis.

Table 7. The comparison table for deleting the variable from the model.

Ratio	Disable variable	Variable in Model	Ratio	Disable variable	Variable in Model
90.62% (29/32)	-	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location, traffic tool.	96.8% (31/32)	Type. Way to pay. From which location. Living location. Traffic tool.	sex, age, marriage, icomes, frequency, time, satisfication, comsumption.
90.62% (29/32)	Traffic tool.	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location, living location.	96.8% (31/32)	satisfication. Type. Way to pay. From which location. Living location. Traffic tool.	sex, age, marriage, icomes, frequency, time, comsumption.
90.62% (29/32)	Living location. Traffic tool.	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay, from which location.	90.62% (29/32)	time. satisfication. Type. Way to pay. From which location. Living location. Traffic tool.	sex, age, marriage, icomes, frequency, comsumption.
93.75% (30/32)	From which location. Living location. Traffic tool.	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption, way to pay.	96.8% (31/32)	Frequency. satisfication. Type. Way to pay. From which location. Living location. Traffic tool.	sex, age, marriage, icomes, time, comsumption.
96.8% (31/32)	Way to pay. From which location. Living location. Traffic tool.	sex, age, marriage, icomes, associator type, frequency, time, satisfication, comsumption.			

Table 8 Distribution of the testing data for BPNN.

Test / Original Category	data size	Category	1	2	3	4	5	6
		Data size	5	3	4	6	3	11
1	5	CLV	5	0	0	0	0	0
2	4		0	3	1	0	0	0
3	3		0	0	3	0	0	0
4	6		0	0	0	6	0	0
5	3		0	0	0	0	3	0
6	11		0	0	0	0	0	11

5 Concluding Remarks

After demonstrating the illustrative case, we can obtain several concluding remarks as follows:

1. Apply data ming technique, BPNN model, to address the issue of the CLV. Not only the model can be constructed, but the related key factors with affecting on discriminant analysis also can be determined.
2. From the result obtained from BPNN, we founf out that sex, age, marriage, incomes, consumption and time can be taken as the evaluated index to measure the CLV.
3. After comparing with the results obtained from BPNN and tranditional discriminant analysis, the discriminant ratio of BPNN significantly exceeds the discriminant ratio of tranditional method.

References:

[1] Altman, E. I.〈1968〉. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, Journal of Finance, 23, 598-609.

[2] Bhatia,Anil 〈 1999 〉. Customer Relationship Management, 1st ed. U.S.A.;Don Hull.

[3] Berry, M. J. A., & Linoff, G. S.〈1997〉. Data mining techniques: for marketing, sales,and customer support. New York: Wiley Computer.

[4] Bose, S. 〈 1996 〉. Classification using splines. Computational Statistics & Data Analysis, 22, 505-525.

[5] Browne, M. J., & Hoyt, R. E.〈1995〉. Economic and market predictors of insolvencies in the property-liability insurance industry. Journal of Risk and Insurance, 62, 309-327.

[6] Cespedes, Frank V. and H. Jeff Smith 〈1993〉,Database Marketing：New Rules for Policy and Practice,Sloan Management Review, Summer, pp7~22.

[7] Chang, W. L., & Yuan, S. T. (1999, December). Data mining and customer retention/profitability. Paper presented at the meeting of the National Computer Symposium Conference, Taipei, Taiwan.

[8] Dahya, J., Lonie, A. A., & Power, D. M.〈1998〉. Ownership structure, firm performance and top executive change: an analysis of UK firms. Journal of Business Finance & Accounting, 25, 1089-1118.

[9] Davies, P. C.〈1994〉. Design issues in neural network development. NEUROVEST Journal, 21-25.

[10] De Veaux R. D., Gordon, A. L., Comiso, J. C., & Bacherer, N. E.〈1993〉. Modelingof topographic effects on antarctic sea ice using multivariate adaptive regressionsplines. Journal of Geophysical Research, 98(C11), 20307-20319.

[11] Desai, V. S., Crook, J. N., & Overstreet, Jr. G. A. 〈1996〉. A comparison of neuralnetworks and linear scoring models in the credit union environment. EuropeanJournal of Operational Research, 95, 24-37.

[12] Efron, B. 〈 1975 〉. The efficiency of logistic regression compared to normal discriminant function analysis. Journal of the American Statistical Association, 70,892-898.

[13] Fish, K. E., Barnes, J. H., & Aiken, M. W.〈1995〉. Artificial neural networks: a newmethodology for industrial market segmentation. Industrial Marketing Management,24, 431-438.

[14] Freeman, J. A., & Skapura, D. M.〈1992〉. Neural networks algorithms, applications,and programming techniques. MA: Addison-Wesley Publishing Company.

[15] Friedman, J. H., & Roosen, C. B.〈1995〉. An introduction to multivariate adaptive regression splines. Statistical Methods in Medical Research, 4, 197-217.

[16] Hosmer, D. W., & Lemeshow, S.〈1989〉. Applied logistic regression. New York: John Wiley & Sons.

[17] Hughes,Arthur M. (1994) ; Strategic Database Marketing, Chicago：Probus Publishing.

[18] Hughes,Arthur. M. 〈 2001 〉. Strategic database marketing-the masterplan for starting and managing a profitable, customer-based marketing program. New York:McGraw-Hill.

[19] Hwang, D. Y., Lee, C. F., & Liaw K. T.〈1997〉. Forecasting bank failures and deposit insurance premium. International Review of Economics and Finance, 6(3), 317-334.

[20] Jain, A. K., Mao, J., & Mohiuddin, K.〈1996〉. Artificial neural network: A tutorial.IEEE computer, 29, 31-44.

[21] Kamakura, W. A. and G. J. Russell〈1989〉, "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," Journal of Marketing Research, Vol.26, 1989, pp379~390.

[22] Kim, J. C., Kim, D. H., Kim, J. J., Ye, J. S., & Lee, H. S.〈2000〉.Segmenting theKorean housing market using multiple discriminant analysis. Construction Management and Economics, 18, 45-54.

[23] Koch, R. 〈1997〉.The 80/20 principle: the secret of achieving more with less. London:Nicholas Brealey.

[24] Koch,R. 〈1997〉; The 80/20 principle: the secret of achieving more with less. London:Nicholas Brealey.

[25] Lee, H., Jo, H., & Han, I. 〈1997〉. Bankruptcy prediction using case-based reasoning,neural networks, and discriminant analysis. Expert Systems With Applications, 13,97-108.

[26] Little, John D.C.〈1979〉, "Decision Support System For Marketing Managers", Direct Marketing, Vol 157, Nov 1994, pp36~37.

[27] Monroe 〈 1991 〉, "The Value Concept and Relationship Marketing," European Journalof Marketing, Vol. 30,No. 2, 1996, pp.19-30.

[28] Newell,F. 〈 2000 〉 .Loyalty.com: customer relationship management in the new era ofinternet marketing. New York: McGraw-Hill.

[29] Nguyen-Cong, V., Dang, G. V., & Rode, B. M. 〈1996〉. Using multivariate adaptive regression splines to QSAR studies of dihydroartemisinin derivatives. Eur. J Med. Chem., 31(10), 797-803.

[30] Ravald, Annika & Christian Gronroos〈1996〉, "The Value Concept and Relationship Marketing,"

European Journal of Marketing, Vol. 30,No. 2, 1996, pp.19-30.

[31] Stone,Bob 〈 1995 〉 ,Successful Direct Marketing Methods, pp37-59, Lincolnwood, IL ： NTC Business Books.

[32] Shani, David and Sujana Chalasani 〈 1992 〉 , “Exploiting Niches Using Relationship Marketing”, The Journal of Consumer Marketing, Vol 9, Summer 1992.

[33] Sung, T. K., Chang, N., & Lee, G. 〈 1999 〉 . Dynamics of modeling in data mining:interpretive approach to bankruptcy prediction. Journal of Management Information Systems, 16, 63-85.

[34] Trevino, L. J., & Daniels, J. D. (1995). FDI theory and foreign direct investment in the United States: a comparison of investors and non-investors. International Business Review,4, 177-194.

[35] Vellido, A., Lisboa, P. J. G., & Vaughan, J.〈1999〉. Neural networks in business: a survey of applications (1992-1998). Expert Systems With Applications, 17, 51-70.

[36] Vellido, A., Lisboa, P. J. G., & Vaughan, J.〈1999〉. Neural networks in business: a survey of applications(1992-1998).Expert Systems With Applicatio,17,51-70ns.

[37] Werbos, P. J.〈1974〉. Beyond regression: new tools for prediction and analysis in the behavioral sciences. Unpublished doctoral dissertation, Harvard University, MA.

[38] WILLIAM E. SPANGLER，JERROLD H.，LUIS G. VARGAS，1999，『Choosing Data-Mining Methods for Multiple Classification ： Representational and Performance Measurement Implications for Decision Support.

[39] Yuan, S. T., & Chang, W. L.〈2001〉. Mixed-initiative synthesized learning approach for web-based CRM. International Journal of Expert Systems with Applications,20(2), 187-200.

[40] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: the state of the art. International Journal of Forecasting, 14, 35-62.

[41] Zhang, G., Patuwo, B. E., & Hu, M. Y. 〈1998〉. Forecasting with artificial neural networks: the state of the art. International Journal of Forecasting, 14, 35-62.

[42] Davis, J., 1986; *Statistics and Data Analysis in Geology* , (John Wiley & Sons, Toronto, 646p.)

[43] Weslowsky, G., 1976; *Multiple Regression and Analysis of Variance*, (John Wiley & Sons, Toronto, 292p.)