Constructing CLV model by Using BPNN Technique for E-business

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: - For the recent years, market environment had rapidly changed and lead most enterprises gradually to meet the competitive pressure. How to achieve the potential customer retention and maintain their core competition had become an importance issue to them. Reviewing the possible solution in marketing management, the estimation requirement of customer lifetime value (CLV) will be a suitable approach. Artificial neural networks (ANNs) was a well-known modeling technique in practice, hence, we will intend to apply it into our CLV model construction in this study.

Key-Words: - Customer lifetime value (CLV), Artificial neural networks (ANNs), E-business

1 Introduction

The enterprises gradually met a competitive market due to the pressure of rapid change. Therefore, it will lead the enterprise to re-think how to efficiencily make their investment. At the same time, the performance of making the market segmentation by using the social attributes will not be the only one choice. According to the practical experience, the average cost to create a new customer will be five times the cost to retain an old customer. Besdies, from the related reports or studies, we also find out that average loss rate of en enterprose is about 25%. That is, if we can reduce about 5% loss rate, the revenue of an enterprise will increase about 100%. Hence, most enterprise focused their efforts on the current market retention, not on the new market creation. Kotler(2000) also pointed out that the importance of customer relationship management (CRM) must be focus on how to create the long term relationship with the valued customers and to maintain such relationship for earning profits. From the above the summary, we can find out that the customer lifetime value (CLV) gradually become an important issue. 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.

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. Artificial neural networks (ANN) have been used in a wide variety of applications, ranging from classification and pattern recognition to optimization and control (NeuralWare, 1990). Furthermore, ANN has successfully employed to model the complexity structure of a system including linear or non-linear relationship (Funahashi, 1989; NeuralWare, 1990; Hsieh, 2006). Employing the neural networks to model the interrelation among input (or control factors) and output (or response) of a quality system by using the experimental data will be an easy approach than statistical approaches (Ko *et al*., 1998; Su & Miao, 1998; Hsieh, 2001; Hsieh, 2006). In lieu of above circumstances, we can intend to use ANNs to achieve the problem of model construction.

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 & Chalasain,1992 ； Morgan & Hunt,1994 ； Berger & 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 *et al*., 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 & Russell1, 1989; Stone & Bob, 1995; Hughes & 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 Mathematical model

Dwyer (1989) firstly defined the CLV as "the net value between the predicted profit and the related cost". And, he proposed a formula to measure the CLV as Equ (1):

$$CLV = \Sigma\,(\text{sale incomes - cost})/(1+d)^n \qquad (1)$$

Where n denote the periods, and the d will represent as discount rate.

Kotler (1996) recognized that the CLV as "the net profit value that derived from the continuously buying behavior". According to his proposed formula, the CLV can be significantly increased when the gain and the loyalty from customer can be enhanced. The formula can be represented as follows:

$$\text{Average CLV} = \text{Average Gain per annual} \times \text{average loyaltyyears} \times \text{marginal profits} \quad (2)$$

McDonald (1996) will define the CLV as a compound evaluation. Two categories of CLV will be divided as: the first category will include the amount of appliance and the confirmation of loyalty; and the second category will include the commodification of products and the efficiency of public praise. He will construct the CLV model by using the two categories. Such Expanded Customer Lifetime Value can be discussed with three parts including revenue, cost and Opportunity cost.

<u>1.Revenue: the formula will be given as Equ. (3):</u>

$$\text{Revenus} = \text{Years} \times (\text{price} \times \text{amount of sale} \times A_1 \times A_2) \times (\text{inflation rate})^{n-1} \qquad (3)$$

Where A_1 will denote as core relationship degree, A_2 will represented as extended relationship degree and n will denote the years.

<u>2.Cost: the formula will be given as Equ. (4):</u>

$$\text{Cost} = \text{Direct cost} \times (\text{inflation rate})^{n-1} \qquad (4)$$

Where n will denote the year.

<u>3.Opportunity cost: the formula will be given as Equ. (5):</u>

$$\text{Opportunity Cost} = (a+r+L)^n \qquad (5)$$

Where L will denote as the adjustable index ($L \geq 0$), n will denote years and r will represent as discount rate. Hence, McDonald (1996) will incorporate those three elements into his Expanded Customer Lifetime Value as Equ (6):

$$\text{Extended Cost} = \text{Years} \times (\text{revenue-cost}) \times (\text{discount rate})^{n-1}/\text{Opportunity Cost} \qquad (6)$$

where n will denote as years.

2.3 Backpropagation neural 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). 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.

The equation (Neural Ware, 1990; Rumelhart *et al.*, 1986) utilized to adjust the weights following the presentation of an input/output pair for the output layer k is:

$$\Delta W_{kj} = \eta \delta_k O_j \tag{7}$$

where

ΔW_{kj} =the change to be made in the weight from the j-th to k-th unit following the presentation of an input pattern,

δ_k =the error signal for unit k after the presentation of an input pattern,

O_j =the j-th element of the output pattern produced by the presentation of an input pattern,

η =the learning rate that governs how fast the network will encode a set of input/output patterns.

The backpropagation rule for changing weights following the presentation of an input/output pair for the hidden layer j is

$$\Delta W_{ji} = \eta \delta_j O_i \tag{8}$$

where

ΔW_{ji} =the change to be made in the weight from the j-th to i-th unit following the presentation of an input pattern,

δ_j =the error signal for unit j after the presentation of an input pattern,

O_i =the i-th element of the output pattern produced by the presentation of an input pattern,

η =the learning rate that governs how fast the network will encode a set of input/output patterns.

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. The architecture of CLV model construction can be depicted graphically in Figure 1.

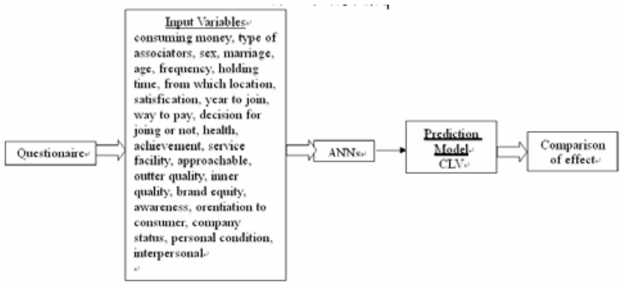

Figure 1. The architecture of CLV model construction.

3.1 Variables definition

The variables to be studied include seven parts: the feature and behaviour, motivation, facility, quality of service, profile of enterprise, barrier of leisure, type of associators. In this case, we initially mail about two hundred and fifty questionnaires and totally have two hundred and eight questionnaires return. The qualified questionnaires is one hundred and ninty eight questionnaires. In order to compute the CLV for this case, we will apply Kolter's CLV model (1996) by incorporating the type of associators, average consuming money and estimated year to keep joining. We can divide ths customer data into five categories in Table 1. The CLV of the first category is from 66000 to 114060, and the CLV of the seond category is about 114061~162000, the third category is from 162001~24900, and the final category is from 249001 to 58500.

Table 1. The customer distribution table.

No	Amount of pepole	Rate	Accumulated rate
1	42	21.2	21.2
2	38	19.2	40.4
3	56	28.3	68.7
4	30	15.2	83.8
5	32	16.2	100.0

3.2 Analysis procedure

The following table will represent the results of discriminant analysis for the training data and the key factors affecting model construction. Herein, we consider twenty-four variables including the basic variables like as consuming money, type of associators, sex, marriage, age, frequency, holding time, from which location, satisfication, year to join, way to pay, decision for joing or not; The other different dimensions like as health, achievement, service facility, approachable, outter quality, inner quality, brand equity, awareness, orentiation to consumer, company status, personal condition, interpersonal. Herein, we will initially determine the significant variables of the basic twelve variables. Hence, the way to achieve such analysis is to disable a basic variable once time and determine the significance of the variable according to the accurate rate. If the accurate rate will reduce since the variable disabling, it will be viewed as a significant variable. The related parameters about BPNN will be given as follows:

1.Processing elements of input layer: It includes 12 variables and the number of PEs will be set as 12.
2.One hidden layer. As for the PEs of hidden layer, we will choose 8 to be the number of PEs due to that it is the optimum decision after pilot run.
3.The ratio of traing and testing data is set as 4:1, that is, there are 156 training data and 52 training data.
4.The learning rate will be set as 0.1 and the momentum will be set as 0.7.
5.The number of PEs of the output layer will be set as 1 due to the only one variable.

After performing such procedure, we get consuming money, type of associators, sex, age, frequency, holding time, from which location, satisfication, year to join, way to pay to be the significant variables. Next, we will separately add one dimension (variable) into BPNN and determine the significance of dimension. Hence, the related parameters about BPNN will be given as follows:

1.Processing elements of input layer: It includes 11 variables and the number of PEs will be set as 11.
2.One hidden layer. As for the PEs of hidden layer, we will choose 9 to be the number of PEs due to that it is the optimum decision after pilot run.
3.The ratio of traing and testing data is set as 4:1, that is, there are 156 training data and 52 training data.
4.The learning rate will be set as 0.1 and the momentum will be set as 0.7.
5.The number of PEs of the output layer will be set as 1 due to the only one variable.

After performing such procedure, we get the other three dimensions to be the significant variables: health, service facility, inner quality. Hence, we get thirteen variables to construct the CLV model. Next, 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. As for the PEs of hidden layer, we will choose 10 to be the number of PEs due to that it is the optimum decision after pilot run.
3.The ratio of traing and testing data is set as 4:1, that is, there are 156 training data and 52 training data.
4.The learning rate will be set as 0.1 and the momentum will be set as 0.7.
5.The number of PEs of the output layer will be set as 1 due to the only one variable.

Table 3. Distribution of the validation data for BPNN.

Predicted data / Real data		Type	1	2	3	4	5
		Amont	4	6	7	7	6
Type	Amount		CLV testing sample				
1	4	CLV	4	0	0	0	0
2	5		0	4	1	0	0
3	7		0	1	5	1	0
4	7		0	0	1	6	0
5	7		0	1	0	0	6

We will apply the test data to make validation for the simiplified BPNN model (only including consuming money, type of associators, sex, age, frequency, holding time, from which location, satisfication, year to join, way to pay, health, service facility, inner quality), and we get the accurate ratio is about 83.3% (listed in Table 2). The rationality and feasibility can be demonstrated well in this case study.

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 found out that consuming money, type of associators, sex, age, frequency, holding time, from which location, satisfication, year to join, way to pay, health, service facility, inner quality can be taken as the evaluated index to measure the CLV.

References:

[1]Bhatia,Anil（1999）. Customer Relationship Management, 1st ed. U.S.A.;Don Hull.

[2]Blattberg,.C,&Thomas,J.S(2000).Valuing,Analyzing And Managing The Marketing Function Using Customer EQUITY Principles. Unpublish manuscript, Northwestern University, Evanston,IL

[3]Berger, Paul D.,Nasr, N.I.Customer Lifetime Value: MarketingModels and Application.Journal of Interactive Marketing,12,winter 1998,pp17-30

[4].Brown, L. G. (1989), “The Strategic and Tactical Implications of Conveniencein Consumer Product Marketing,”Journal of Consumer Marketing, 6(3),pp.13-19.

[5]Fish, K. E., Barnes, J. H., & Aiken, M. W.（1995）. Artificial neural networks:a newmethodology for industrial market segmentation. Industrial MarketingManagement, 24, 431-438.

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

[7]Griffin, J.（1995）,“The Internet’s expanding role in building Customer Loyalty”, Direct Marketing, Vol. 59, pp.50-53.

[8]Hsieh, K. L., Process Improvement in the Presence of Qualitative Response by Combining Fuzzy Sets and Neural Networks, *Integrated Manufacturing Systems*, 12(6-7), pp. 449-462, 2001.

[9]Hsieh, K. –L., (2006), “Parameter Optimization of a Multi-response Process for Lead frame Manufacturing by Employing Artificial Neural Networks”, International Journal of Advanced Manufacturing Technology, 28(5-6), 584-591

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

[11]Hughes,Arthur. M.（2001）. Strategic database marketing-the masterplan for starting and managing a profitable, customer-based marketing program.

NewYork: McGraw-Hill.

[12]Ko, D. C., Kim, D. H., Kim, B. M. and Choi, J. C., Methodology of perform design considering workability in metal forming by the artificial neural network and Taguchi method, *Journal of Materals Processing Technology*, 80-81, pp. 487-492, 1998.

[13]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.

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

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

[16]Monroe(1991), "The Value Concept and Relationship Marketing," European Journal of mMarketing, Vol. 30, No. 2, 1996, pp.19-30.

[17]Neural Ware, Inc., *Neural Works Professional II/Plus and NeuralWorks Exporer*, Penn Center West: Neural Ware, Inc, 1990.

[18]Parasuraman, A., Zeithaml, V.A., and Berry, L.L.(1985)," A Conceptual Model of Service Quality and Its Implication for Furture Research,"Journal of Marketing,Vol.49,P41-50.

[19]Ravald, Annika & Christian Gronroo(1996), "The Value Concept and Relationship Marketing," European Journal of Marketing, Vol. 30,No. 2, 1996, pp.19-30.

[20]Rumelhart, D. E., Hinton, G. E. and Williams, R. J., Learning internal representations by error propagation, *Parallel Distributed Processing: Explorations in the Microstructure of cognition,* Edited by Rumelhart, D. E., McClelland, J. L., MIT Press, Cambridge, MA, 1, pp.318-362, 1986.

[21]Stone, B.(1995),Successful Direct Marketing Methods, pp37-59, Lincolnwood, IL：NTC Business Books.

[22]Su, C. T. and Miao, C. L., Neural network procedure for experimental analysis with censored data, *International Journal of Quality Science,* 3(3), pp.239-253, 1998.