The Intelligent Method for Identifying Customer Purchasing Pattern Changed Over a Period in Time

Choong-Soo Lee

Gwangju University, Republic of Korea

cslee@gwangju.ac.kr

Gye-Hang Hong

Dongbu Co., Republic of Korea

kaistduck@dongbu.com

ABSTRACT

As the global business environment is beginning to change from the product-centered stage to the customer centered stage, companies can not survive from the global business competition without understanding the customers. The general Customer Relationship Management (CRM) method has focused on marketing and Voice of Customer (VOC) like campaign, call center, etc. However, we have an interest in how to design the customer-centered production system. Therefore, we will measure how customer satisfaction and sales revenue are increased or decreased according to the change in the factors, quality, quantity, price and delivery frequency. To accomplish that in this paper, we identified the main customers, who have an effect on sales quantity, through the RFM analysis. We then designed a neural networks model and feature weighting model to discover the knowledge of buying pattern for valuable customers.

Keywords: Feature Weighting, Customer Relationship Management, RFM Method, Neural Network, Voice of Customer

INTRODUCTION

In this paper, we suggested a method of how to identify main customers and how to analyze the key factors, which have an effect on buying behaviors of the main customers, in each period of time.

As the global business environment is beginning to change from the product-centered stage to the customer centered stage, companies can not survive from the global competition without understanding the customers.

The Customer Relationship Management (CRM) research is one of the methods for understanding the customer. In general, the CRM issue is how each company identifies his or her main customers and finds the buying patterns of the main customers changed over a time period. Therefore, the customer relationship management system has focused on marketing and Voice of Customer (VOC) like campaign, call center, etc.

However, we have an interest in how to design the customer-centered production system. Therefore, we will measure how customer satisfaction and sales revenue are increased or decreased according to the change in the factors: quality, quantity, price and delivery frequency.

LITERATURE REVIEW

We applied the following methods to develop an intelligent method for identifying purchasing pattern of valuable customers.

Firstly, we applied the Self Organizing Maps (SOM) and – Recency, Frequency, Monetary- (RFM) method to distinguish valuable customers into overall customers.

SOM (Kohonen, 1982) is an unsupervised learning scheme to train the neural network. Unsupervised learning comprises those techniques for which the resulting actions or desired outputs for the training sequences are not known. The network is only told the input vectors, and it then self-organizes these inputs into categories.

The RFM clustering method is one of the analyzing methods for discovering customer patterns (Ha and Park ,1998). The RFM is defined as follows:

- Recency (R): the time period of last purchase during an analyzing time period
- Frequency (F): the number of purchases during an analyzing time period
- Monetary (M): the amount of spent money during an analyzing time period

Secondly, we applied the Neural Network to identify the key buying pattern of valuable customer.

The back propagation algorithm (Mitchell, 1997), one of the neural networks models, is employed for discovering the customer buying behaviors. In general, the neural networks model is applied to predict a coming environment or classify entities from history patterns.

PROPOSED METHOD

Distinguishing Valuable Customers from Valueless Customers

The CRM system identifies the main customers who have an effect on sales quantity from sale history database and customer profile database and then, it discovers the knowledge of their consuming pattern. The system extracts the Recency, Frequency, Monetary information from the sales data for identifying the main customers.

The Recency of the RFM metrics is redefined as the average purchase period of time during 1 year because we consider the case of seasonal buying trends. The system segments all customers into some customer groups which have a similar consuming pattern with the SOM. Each input (RFM) can fall into one of following categories: above the average or below the average. So, the number of customer groups is eight, possible combination of inputs, 23. Because the output nodes of the SOM are set by two dimensions, nine nodes are needed to design the model.

After clustering groups, the system classifies the inputs of the each group into the two categorical values, and then merges the groups which have the same categorical values about all inputs. Table 1 shows the result of merged clusters. We represent above the average of each input as ↑ and below the average of each input as ↓.

Table 1. Segmenting All Customers Into Four Groups In Terms Of RFM

	Recency	Frequency	Monetary	Number of customers
Cluster 1	R↑ (0.66)	F↑ (0.42)	M↑ (0.54)	20
Cluster 2	R↓ (0.38)	F↑ (0.88)	M↑ (0.72)	4
Cluster 3	R↓ (0.22)	F↓ (0.17)	M↓ (0.08)	46
Cluster 4	R↑ (0.76)	F↓ (0.09)	M↓ (0.05)	70
Average	0.51	0.39	0.35	

As shown in Table 1, cluster 1 is superior to the others in terms of all inputs, R, F, and M. Cluster 2 is superior to the others in terms of the F and M inputs, however, is not in term of the R input. It implies that customers, who belong to the cluster 2, are very valuable customers during the little period of 1 year. Therefore, customers of the two clusters are very important customers. The cluster 3 is the group of volatile customers because customers of the cluster sometimes buy small quantities of our items during a little period. The cluster 4 steadily buys our items, however, consumes a little quantity. Therefore, the cluster 3 and cluster 4 aren't the manageable groups.

Identifying Behaviors Of The Valuable Customers

The CRM system analyzes buying behaviors of the main customers by each Period for Assessment (PFA). Unlike the general CRM techniques, the system identifies the important factors, which have an effect on the reason that the main customers buy the items. Then, it measures the degree of contribution/risk of the factors, which represents the ratio of change of revenue and customer satisfaction to change of value of the factor. The system employs a neural network model of the Figure 1 to discover the knowledge about the important factors and the degree of contribution of them.

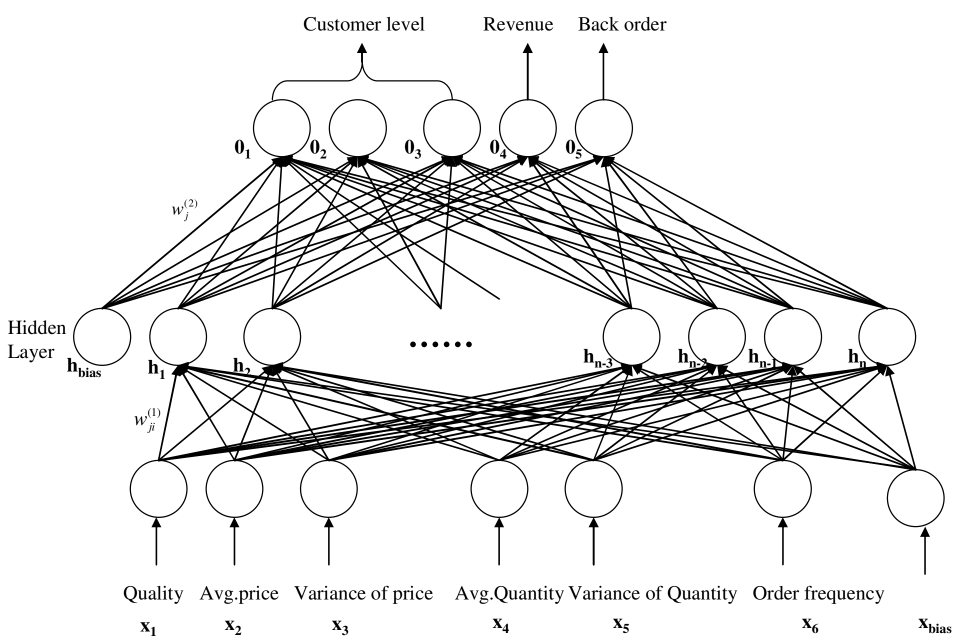

Figure 1: The Neural Network Model For Discovering The Knowledge Of Customer Behaviors

In generally, neural network models are applied to classification problem or prediction problem. However, we designed the neural network model to operate both classification and prediction. As shown in Figure 1, output nodes of the model have 3 nodes to classify the customer level and 2 nodes to predict revenue and customer satisfaction, back order. The model has one hidden layer and two or three times nodes of the number of input nodes and one bias node. And the model has input nodes which consist of six nodes and one bias node. In this paper, the input nodes represent vector x, x = { x0, x1, x2, x3, x4, x5, x6 }, hidden layer does vector h, h = { h0, h1, h2, ……, hn }, and output does vector o, o = { o1, o2, o3, o4, o5 }. $w_{ji}^{(1)}$ denotes a connecting weight from xi to hj and $w_{kj}^{(2)}$ is also the connecting weight between ok and hj. The bias nodes set h0= 1, x0 = 1. Each node of the hidden layer is activated by a weighted linear combination of the connected inputs (aj) and each output node is also activated by a weighted linear combination of the connected hidden nodes (bk). That is,

$$a_j = \sum_{i=0}^{6} w_{ji}^{(1)} x_i , b_k = \sum_{j=0}^{n} w_{kj}^{(2)} h_j$$

All nodes of the model use the Sigmoid activation function (g(a) = 1/(1+e-a)).

Then, each output of the model is calculated as follow:

$$o = g(\sum_{j=0}^{n} w_{kj}^{(2)} g(\sum_{i=0}^{6} w_{ji}^{(1)} x_i))$$

For the training the model, we used six factors, quality, average price, variance of price, average quantity, variance of quantity, and order frequency as the input data. The factors have an effect on buying behavior of the main customers. Output data of the customer level for classification used the group number described in the previous section. That is, the best customer group, the cluster 1 (see Table 1), denotes (1, 0, 0), good customer group, the cluster 2, does (0, 1, 0), and the other groups, the cluster 3 and the cluster 4, do (0, 0, 1). The other output data use the profit and the number of back order from each customer during each PFA. We retrieve some data sets randomly from sales history database and divide the data sets into the leaning set and the validation set. And then, the system trains the classification and prediction NN model and discovers the important knowledge of main customers from the model.

As mentioned above, the system discovers two kinds of important knowledge from the neural networks model : The Important Factors (IF) and the Degree of Contribution/Risk (DCR) of each IF. In general, NN model are not so commonly used in real data mining problems as other learning strategies, such as decision tree technique. This is partly due to its shortcoming of being "Black Box", meaning the neural network provides end users with little comprehensible knowledge about how it arrived at a given result (Benitez, Castro, and Requena,1997). For overcoming the shortcoming, the system uses the sensitivity and the activity, one of the feature weighting methods.

Firstly, the system uses the sensitivity method (Shin, Yun, Kim and Park, 2000) to identify the important factors. The sensitivity of input node is calculated by removing the input node from the trained neural network. That is, an input node might be removed by setting all the connected weights to zero. A measure of the sensitivity of an input factor is the difference in the prediction value between when the feature is removed and when it is left in place. That is,

$$w_i = \frac{\left(\sum_L \frac{|p^0 - p^i|}{p^0} \right)}{n}$$

Where P0 is the normal prediction value for each training instance after training, and Pi is the modified prediction value when the input node i is removed. L is the set of training data, and n is the number of training data. Because our model has the classification output and prediction output, the method is modified. That is, the classification output is calculated as follow:

$$w_i = \frac{\left(\sum_L difference \right)}{n}$$

difference (0, i) = 0 if factor i is $o_1^0 = o_1^i \wedge o_2^0 = o_2^i \wedge o_3^0 = o_3^i$

1 otherwise

Where the sensitivity (wi) is how well the input i classifies customers. The ratio of the number of exactly classification to the total number of training data is importance of the input i.

Secondly, the system uses the activity method (Shin, Yun, Kim and Park, 2000) to measure the DCR of each IF. The activity method measures the variance of activation level of an input node. When the activation value of a node varies much according to its input value, the activity of the node may be considered to be high. On the contrary, if the activation value of a node remains constant over the training data, the activity of the node become zero.

$$R_i = \sum_{j=1}^{n}((w_{ji}^{(1)})^2 \times (w_j^{(2)})^2 \times \mathrm{var}(sigm(\sum_{i=1}^{6} w_{ji}^{(1)} \times x_i))$$

Because activity can measure the change of output to variance of an input, we may obtain the knowledge of how revenue or customer satisfaction can be increased if we improve value of the input factor (CONTRIBUTION). On the contrary, we can know how revenue or customer satisfaction can be decreased if we do not manage the input factor well (RISK).

Table 2 Consuming Pattern of Main Customers Over Periods in Time

Factor PFA		T1		T2		T3		T4	
		P	C	P	C	P	C	P	C
Quality	Level	1st – 2nd level		4th – 7th level		1st – 2nd level		1st level	
	IF	0.652	0.176	0.473	0.085	0.345	0.042	0.486	0.112
	DCR	0.329		0.426		0.289		0.019	
Frequency	Level	High		Average		Average		Above Average	
	IF	0.072	0.206	0.024	0.026	0.002	0.116	0.574	0.418
	DCR	0.010		0.042		0.042		0.356	
Price	Level	Above Average		Average		Above Average		High	
	IF	0.432	0.176	0.517	0.341	0.365	0.304	0.045	0.074
	DCR	0.044		0.506		0.427		0.001	
Quantity	Level	High		Average		Average		Below Average	
	IF	0.334	0.059	0.023	0.286	0.195	0.019	0.462	0.007
	DCR	0.254		0.001		0.174		0.316	
Key Factors		Quality, Quantity		Quality, Price		Quality, Price		Frequency, Quantity	

* P : Prediction, C : Classification

Table 2 shows values of IFs and their DCR by each PFA and Figure 2 shows a correlation graph, represents a correlation between the IF and the DCR. The graph helps the companies to support the knowledge of what factors they should manage intensively. The horizontal axis of the graph is a value of an IF and the vertical axis is the DCR value of the IF. The larger values of both IF and its DCR is, upper position of right in the graph, the more important the IF is. The priority of IF is determined from the upper position of right in the graph to the lower position of left in the graph. The results are written in key factors of Table 2.

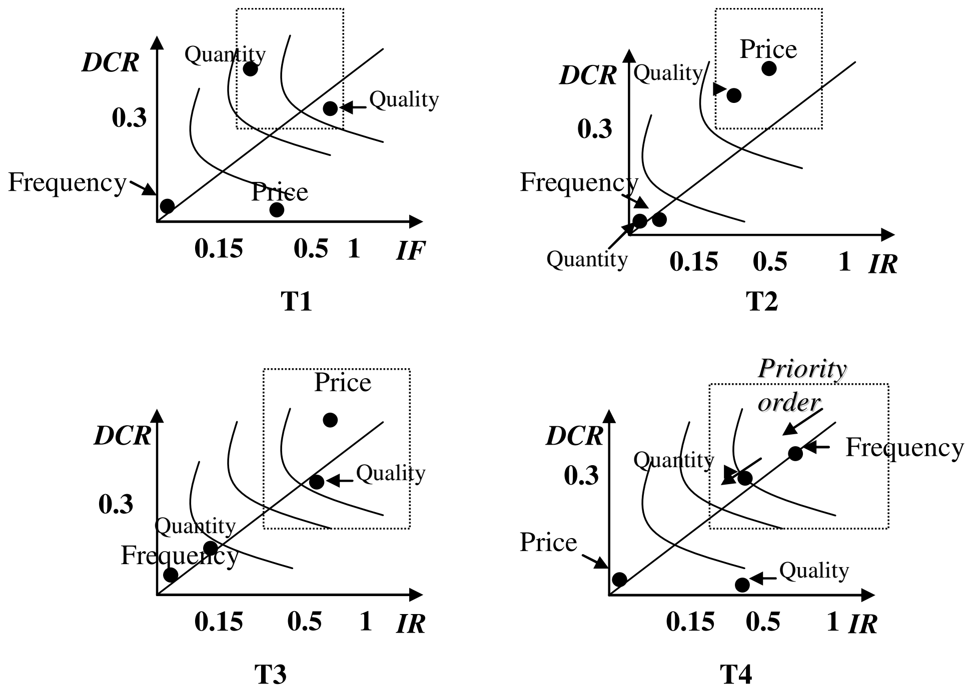

Figure 2: The Correlation Graph Between the IF and the DCR

CONCLUSION

In this paper, we identified the main customers, who have an effect on sales quantity, through the RFM analysis. And then, we designed a neural networks model and feature weighting model to discover the knowledge of buying pattern for valuable customers.

That is, we identified two kinds of important knowledge from the neural networks model: IF(Important Factors) and DCR(the Degree of Contribution/Risk) of each IF.

We applied the model to the customers who buy the agricultural product having feature of seasonal trend in Korea. In the result, the same customer may have different needs over the period in time: quality and quantity in T1, quality and price in T2 and T3, and frequency and quantity in T4.

Because our research has the limitation to be applied to customers who buy overall types of product, we should verify our model to the products in various industries.

REFERENCES

T. Kohonen (1982), Self-oranized formation of topologically correct feature maps, Biolog. Cybern., vol.43, pp59-69.

S. H. Ha, S. C. Park (1998), Application of data mining tools to hotel data mart on the Intranet for database marketing, Expert System With Applications, vol. 15, pp. 1-31.

Mitchell Tom M. (1997), Machine Learning, McGraw-Hill.

J.M.Benitez, J.L.Castro, and I. Requena (1997), Are neural networks black boxes?, IEEE Trans. Neural Networks, Vol. 8.

C.K. Shin, U.T Yun, H.K. Kim and S.C. Park (2000), A Hybrid Approach of Neural Network and Memory-Based Learning to Data Mining. Int J IEEE Trans. on Neural Networks. 2000. Vol.11, No.3

AUTHOR BIOGRAPHIES

	**Dr. Choong_Soo Lee** is an Associate Professor in the Department of *e*-business, Gwangju University in Korea. He hold a BS and MS degrees from Hanyang University and Ph. D. from Korea University, all in Industrial Engineering. His main research themes are Customer Relationship Management(CRM), Enterprise Resource Planning(ERP), Supply Chain Management (SCM), Service Quality and Scheduling
	**Dr. Gye-Hang Hong** is a manager in the Dongbu Financial Service Co. in Korea. He hold a BS degree from DongGuk University and MS and Ph. D from Korea University, all in Industrial Engineering. His main research themes are Customer Relationship Management(CRM), Datamining, and Supply Chain Management (SCM).