韓國經營科學會誌
第28卷 第2號
2003年 6月

데이터 마이닝 기법의 기업도산예측 실증분석*

이 기 동**

A Study of Data Mining Techniques in Bankruptcy Prediction*

Kidong Lee**

■ Abstract ■

In this paper, four different data mining techniques, two neural networks and two statistical modeling techniques, are compared in terms of prediction accuracy in the context of bankruptcy prediction. In business setting, how to accurately detect the condition of a firm has been an important event in the literature. In neural networks, Backpropagation (BP) network and the Kohonen self-organizing feature map, are selected and compared each other while in statistical modeling techniques, discriminant analysis and logistic regression are also performed to provide performance benchmarks for the neural network experiment. The findings suggest that the BP network is a better choice among the data mining tools compared. This paper also identified some distinctive characteristics of Kohonen self-organizing feature map.

Keyword : Data Mining, Kohonen Self-Organizing Feature Map, BP, Classification.

1. Introduction

Bankruptcy prediction has long been a interesting subject in business classification. In the mid 1980's to the early 1990's, the savings and loan (S&L) debacles led to the closing of many S&Ls, resulting in a significant economic and financial loss to the U.S. economy. Recently, the incident of Asian financial crisis gives us the lesson that the capability of accurately predict-

논문접수일 : 2003년 5월 16일 논문게재확정일 : 2003년 5월 28일
* 이 논문은 2003년도 인천대학교 교내연구비에 의하여 연구되었음.
** 인천대학교 경영학과 전임강사

ing financial condition, whether it is a country or a firm, is an important activity. Further, breakthroughs in information and computer technologies today trigger more intense competition among corporations. To maintain a competitive edge, companies must adjust, or adapt their organizational behavior faster than their competitors according to environmental changes. Thus, business research communities attempt to understand these nonlinear adaptive behavioral patterns of today's companies whose main purpose is to compete with the changes of competitors and of environments constantly.

Artificial neural networks (ANNs), a branch of artificial intelligence, as alternative classification technologies to the statistical modeling, have often been used in bankruptcy prediction. In fact, backpropagation (BP) network has been the most widely used neural network method in the bankruptcy prediction (Tam and Kiang, 1992 ; Wilson and Sharda, 1994 ; O'Leary, 1998 ; Zhang, 1999). Recently, however, researchers try to expand their methodological choices experimenting with different types of ANNs. In particular, Kohonen Self-Organizing Feature Map (KSOFM) has been frequently appeared in business and information system literature (Martin-del-Brio and Serrano-Cinca ; 1995, Kiviluto, 1998 ; Kohonen, 1997 ; Alam *et al.*, 2000).

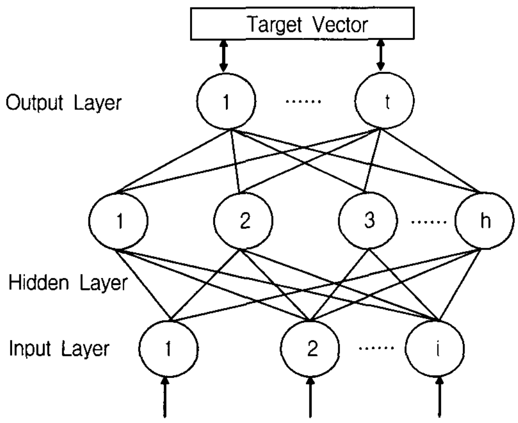
Results and performance of neural network experiments are largely depending on the available size of data sets and the selection of modeling techniques. In applying neural network modeling techniques to nonlinear patterns such as business failure, there is no assumption made in advance. It means that by and large, the pattern of a business event may be simulated and modelled by each experiment until we confirm some underlying patterns. In other words, the performance of these two popular data modelling techniques, BP and KSOFM, might be different in the context of one particular application area, here bankruptcy prediction. Thus, it is interesting to contrast these two different neural network types to see how they behave, more specifically, the prediction accuracy of their behavioral aspect in bankruptcy prediction which is a typical two group classification problem. The purpose of this paper is then to compare the prediction accuracy of four different data mining techniques : BP, KSOFM, discriminant analysis, logistic regression. Note that the first two data mining techniques come from neural network, while the latter two are from statistic modelling techniques which served as performance benchmarks.

The rest of this paper is organized as follows. Section 2 illustrates the prior research on the BP network, the Kohonen self-organizing feature map, and bankruptcy prediction. Section 3 provides the research design and methodologies in terms of data, variables, cross-validation scheme, and classification technologies used for the study. Experimental results are shown and discussed in section 4 where the results of each of the four individual classification techniques are compared. The limitations, the future research directions, and the conclusions for the study are given in the section 5.

2. Prior Literature

2.1 Back-Propagation (BP) Network

The BP algorithm, a systematic training method for a Multi-Layer Perceptron (MLP), has been the most widely used in bankruptcy prediction tasks. [Figure 1] shows MLP with one hidden layer whose architecture shows i nodes in the input layer, h nodes in the hidden layer, and t nodes in the output layer.



[Figure 1] MLP with one hidden layer

The main idea of the BP network is to reduce this error to some desirable levels by way of adjusting weight vector (Wasserman, 1989 ; O'Leary, 1998 ; Zhang *et al.*, 1999). The process of finding a desirable solution in the BP network is briefly described below.

The actual output value, **Y**, in the output node of the BP network is computed as in equation (1).

$$\mathbf{Y} = f(\mathbf{X} \times \mathbf{W}) \tag{1}$$

where **Y** stands for the output vector, **X** the input (row) vector, **W** the weight vector (including bias), and $f(\,)$ denotes an activation function. The activation function, $f(\,)$, transforms the sum of input values into output values of the node. Typical choices of the activation function consist of the logistic, the tangent, the sign, and the linear. The logistic function is used in this study, shown in equation (2).

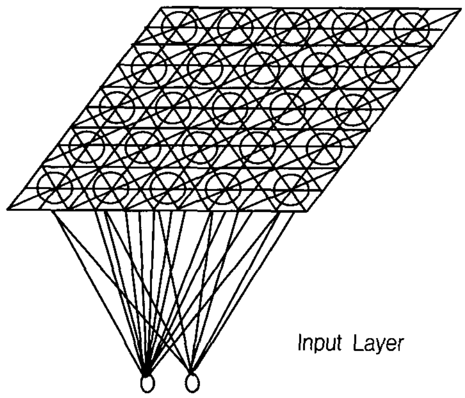
$$\mathbf{Y} = \mathrm{f}\ (\mathbf{X} \times \mathbf{W}) = 1/(1 + \mathrm{e}^{-(\mathbf{X} \times \mathbf{W})}) \tag{2}$$

Then, these actual output values, **Y**, compare with the target (or desired) values shown in [Figure 1], and the difference (or error) between the target values and the actual network outputs (**Y**-Target), measured by some error measures such as sum of squares, are computed. By reducing this error to some desirable levels by adjusting weight vector, we can reach the solution for a problem. Researchers must find an optimal architecture for the BP network with its associated parameters before they use it. This process is called training. This task involves determining a number of design variables for a BP network (i.e., activation functions, error measure functions, MLP architecture, and a particular training method). After training done, researchers can use the network for testing purposes.

2.2 Kohonen self-organizing neural network

Another popular neural network model is the Kohonen self-organizing feature map (Kohonen, 1997 ; 1982). The Kohonen self-organizing neural networks have appeared in many fields, for example, classification (Corridoni *et al.*, 1996 ; Schonweiler *et al.*, 1996 ; Deschenes and Noonan, 1995), pattern recognition (Xin-Hua and Hopke, 1996), clustering (Martin-del-Brio and Serrano-Cinca, 1995 ; Kiviluoto, 1998), and forecasting (Der Voort *et al.*, 1996). In the Kohonen network, input vectors are presented to the outer space (feature map).

Some researcher's argue that this Kohonen self-organizing feature map is more plausible in describing pattern in an ever-changing today's business environment (Waldrop, 1992 ; Kelly, 1994 ; Hinton and Sejnowski, 1999). [Figure 2] shows the typical architecture of the Kohonen self-organizing feature map with two input nodes and 5 × 5 output nodes. As is seen, the input vector space of the object is projected into the two dimensional Kohonen output space.



[Figure 2] Typical architecture of the Kohonen self-organizing feature map

The Kohonen training process succinctly summarized by Nour (1994) is reproduced below as :

1. Initialize $\mathbf{W}_i(t)$ to random values, and set $t = 0$.
2. Present an input vector $\mathbf{X}$ to the network, and compute the distance (similarity) $\mathbf{D}$ using the Euclidean metric to find the closest matching unit c, to each input vector.

$$\mathbf{D}_i = \| \mathbf{X}(t) - \mathbf{W}_i(t) \|, \ \forall\ i$$
$$N_c = \min\{\mathbf{D}_i\}, \ \forall\ i$$

3. Update the weight vector according to the following rule.

$$\mathbf{W}_i(t) = \mathbf{W}_i(t) + lr(t)^{*-}\ h(t,r)^{*}\{\mathbf{X}(t) - \mathbf{W}_i(t)\} \qquad i \in N_c$$
$$\mathbf{W}_i(t) = \mathbf{W}_i(t) \qquad i \notin N_c$$

4. If $t > T$ stop, else increment t and go to Step 2.

where t is the iteration step, T is a number of iterations predefined, and $lr(t)$ is a learning rate which ranges in [0, 1]. The term $h(t, r)$ is a neighborhood function, which decreases over the iteration step and the topological distance, $r = (r_i - r_c)$, between unit i, and unit c (the winner), where r_i and r_c are the coordinates of units i and c, respectively.

The most basic self-organizing learning rule is a winner-take-all approach whereby only the winning node adapts its weight (no neighborhood function). This winner-take-all approach is often called competitive learning. The Kohonen's self-organizing learning method is a variation of the above competitive, winner-take-all approach in that not only the winner but also its neighbors can update their weights together.

2.3 Bankruptcy Prediction

Bankruptcy prediction is a typical case of binary decision-making process. Improvement of this bankruptcy prediction area comes from the incessant effort of past researchers that have developed ratio analysis to linear modeling to nonlinear modeling including the neural network approach.

Beaver (1966) was one of the first researchers to study bankruptcy prediction by testing the efficacy of several financial ratios in their classification and prediction capabilities. Altman (1968) introduced a class of models based on discriminant analysis in classifying bankruptcy prediction using the following well-known five variables as working capital/total assets, retained earnings/total assets, earnings before interest and taxes/total assets, market value of equity/total debt, and sales/total assets. Ohlson (1980), with the use of a logistic regression to estimate the probabilities of a bankruptcy, reported a much higher prediction on corporate failure.

Neural networks as classification data mining tools were not used as a bankruptcy classification technology until the early 1990's. Odom and Sharda (1990) were the first researchers to investigate the feasibility of neural networks in firm failure prediction. They found that BP networks are at least as accurate as discriminant analysis. After this first neural experiment, a significant volume of neural network research followed (Tam and Kiang, 1992 ; Salchenberger *et al.,* 1992 ; Udo, 1993 ; Tsukuda and Baba, 1994 ; Wilson and Sharda, 1994 ; Sharda and Wilson, 1996 ; Martin-del-Brio and Serrano-Cinca, 1995 ; Jo, Han, and Lee, 1997 ; O'Leary, 1998 ; Kiviluoto, 1998 ; Zhang *et al.*, 1999 ; Alam *et al.*, 2000) due to their nonlinear mapping capabilities, thus, improving prediction accuracy.

Tam and Kiang (1992) compared a BP network experiment with a linear classifier, a logistic regression, *k*NN, and ID3. The 19 financial ratios were used in classification task of the 59 matched pairs of Texas banks that failed in the period of 1985～1987. They found that the BP network approach outperforms the other classification techniques.

Salchenberger *et al.* (1992) initially selected 29 variables and then performed a stepwise regression to reduce the number of variables into the final five variables. With the bank data set in the period January 1986 to December 1987, their experiment was to test the possible performance difference of BP networks over a logistic regression.

Udo's study (1993) compared the effectiveness of a BP network with a multiple regression in bankruptcy prediction. Udo's findings confirmed that the BP network is as accurate as or more accurate than a multiple regression model.

Tsukuda and Baba (1994) compared the effectiveness of a BP network versus discriminant analysis in bankruptcy prediction using financial data for 1 and 3 years prior to failure for two listed and unlisted company sets of Japanese corporation. The results showed that the BP network approach seems to work rather well with noisy data than statistical counterparts.

Wilson and Sharda (1994) and Sharda and Wilson (1996) used an experimental design of training and test sets to test BP's effectiveness compared to many statistical classification methods. With use of Monte Carlo resampling techniques, they confirmed BP's prediction accuracy over conventional statistical methods.

Martin-del-Brio and Serrano-Cinca (1995) applied the self-organizing neural networks (Kohonen) to two financial data sets taken from the state of Spanish economy : the Spanish banking crisis of 1977～1985 and the financial state of Spanish companies in 1990～1991. Their results were interesting in that the input feature spaces were projected into the natural clustering (or delimiting) of regions of interest on the output maps.

Jo, Han, and Lee (1997) compared three different techniques in bankruptcy prediction : discriminant analysis, case-based forecasting, and BP network. In classifying Korean firms during 1991～1993, matched by industry and average credit rating within industry, they found that the BP network was better than the two other techniques. Their experiments also showed that experiments with raw data produced a better result than with normalized data.

O'Leary (1998) launched a meta study comparing 15 prior neural network studies on bankruptcy prediction in terms of sampling, impact of different ratios of failed and non-failed firms, software used, input variables, the number of hidden layers and hidden nodes, performance measures, and misclassification costs. The findings of this comparative study confirmed that in general the BP approach outperforms other statistical classifiers, but often with a high cost of time and effort of experiments.

Kiviluoto (1998) applied the self-organizing map to 1,137 Finnish industrial enterprises. In Kiviluoto's study, the self-organizing map is used for a way to indicate the *bankruptcy zone* in the output space.

Using a 5-fold cross-validation scheme, Zhang *et al.* (1999) provided a comprehensive review of a neural network approach on firm failure. They used 6 input variables (Altman's variables plus the current ratio) with a data set covering a 12 year-period and confirmed that BP network outperforms a logistic regression.

Alam *et al.* (2000) compared three different algorithm's a fuzzy clustering algorithm and two self-organizing neural network approache's in a data set representing a real bankruptcy proportion in the real world. Their findings were interesting in that they identified some gray area between healthy and bankruptcy firms. Firms in the gray area could be possible bankrupt candidates that need to be watched closely.

In sum, the early ANN researchers heavily relied on the BP network but recently we see the pattern of neural network researchers that have tried to experiment with the Kohonen self-organizing feature maps more frequently. Once again, direct comparison between these two different learning styles is somewhat difficult, but we try to contrast each other so that, it is hoped, the advantages and disadvantages of the Kohonen self-organizing feature map to the BP network becomes apparent. Next is the detailed research design for this study.

3. Research Design and Methodology

Data sets and variables, the cross-validation scheme as well as the detailed specifications of data mining techniques are illustrated below.

3.1 Data and Variables

The data sample for this bankruptcy prediction study consists of Korean firms that have filed for bankruptcy in the period 1995~1998, selected from a pool of the listed companies on the Korean Stock Market. An initial search of bankrupt firms is made through the Security and Exchange Commission (SEC) filings stored in an on-line database (commercially available in a electronic format) of the Korea Investors Service, Inc., which is a strategic partner with Moody's Investors Service in the Korean security market.

Financial institutions such as commercial banks or investment bankers are excluded in this data set since in the Korean market, the fates of such financial intermediaries seem to be much more affected by the government policies and decision, not from their own financial strength. Therefore, including such financial institutions might deteriorate the quality of this study by obscuring the prediction accuracy of the applied methods.

Searching for failed firms resulted in 113 non-financial failed firms among the listed companies in the Korean stock market. Then, the availability of the financial ratios for the failed firms further reduced the final bankrupt sample size to 84 since some of them seemed not to report their financial status on their bankruptcy filings.

Each failed firm is matched with a non-failed firm in terms of (1) asset size and (2) a two-digit Standard Industrial Classification (SIC) code as control measures. The asset size of a non-failed firm is matched with that of a failed firm using the three-year period prior to bankruptcy filings. As a result, we have a matched sample of 168 firms ; 84 failed firms and 84 non-failed firms. Two time-framed financial data sets, the two-year and the three-year prior to bankruptcy filings, are prepared for this experiment in order to see if any of these classification tools can detect any discrepancy of the financial condition of a firm between this time-difference of the data. As noted, this selected period closely resembles the outbreak of Asian financial crisis. Thus, we test the early warning capability of the classification tools using only two or three years ahead of bankruptcy filing, rather than using the much preceding year.

Each firm is described by Altman's five variables since the prediction capabilities of these ratios are well documented in the previous literature (Zhang *et al.*, 1999 ; Boritz and Kennedy, 1995 ; Odom and Sharda, 1990 ; Altman, 1968) :

1. WCTA = working capital/total assets as a measure of the net liquid assets of the firm to the total capitalization.
2. RETA = retained earnings/total assets as a measure of cumulative profitability.
3. EBITTA = earnings before interest and taxes/total assets as a measure of true productivity of the firm's assets subtracting any tax or leverage factors.
4. MEDEBT = market value of equity/book value of total debt as a measure how much the firm's assets can decline in value before the liabilities exceed the assets and the firm becomes insolvent.
5. SALETA = sales/total assets as a measure of the sales generating ability of the firm assets.

For data normalization issues, this study does not use a data normalization method since many researchers (*i.e.*, Jo *et al.*, 1997 ; Zhang *et al.*, 1999) suggest that data normalization does not seem to improve the effectiveness of neural networks.

3.2 Cross-Validation Scheme

Any bias due to changing data set composition could have a detrimental impact on determining neural network architecture and its parameters. A cross-validation technique is introduced to investigate the classification performance of neural networks in terms of sampling variation. The cross-validation technique enables us to use a whole data set so that it is hoped, any bias effect would be minimized (Zhang *et al.*, 1999 ; Tam and Kiang, 1992).

In this study, a four-fold cross validation technique is used. The total data set that contains 84 matched firms (84 failed firms and 84 non-failed firms) is divided into four equal and mutually exclusive subsets, each of which thus contains 21 matched objects. <Table 1> shows the details of this four-fold cross-validation scheme.

〈Table 1〉 Four-fold Cross-Validation

	Whole set of data			
	First (One-fourth)	Second (One-fourth)	Third (One-fourth)	Fourth (One-fourth)
Subset 1	Training set		Validation set	Testing set
Subset 2	Testing set	Training set		Validation set
Subset 3	Validation set	Testing set	Training set	
Subset 4	Training set	Validation set	Testing set	Training set

Note) That each of one-fourth of whole data set contains 21 matched objects.

Training is conducted on any two of the four subsets while the remaining two sets are used for validation and testing purposes, respectively. The validation set is introduced as an early stopping technique, to improve generalization (MathWorks, 1997). Using this early stopping technique, the validation process is embedded into the training results, which would prevent a possible upward bias of the classification accuracy of a training set.

3.3 The BP Network

For the BP experiment, a BP network with a three-layer model (also called a one-hidden layer model) is considered for this study.

The previous studies proved that as long as there are sufficient numbers of hidden nodes provided, this architectural BP network is able to approximate any arbitrary function well (Funahashi, 1989 ; Hornik *et al.*, 1989).

Five input nodes are used. In a cross-sectional study, the number of input nodes is usually the same as the number of independent variables used in a study (Zhang *et al.*, 1999). For the output nodes, only one output node is sufficient for a binary classification problem. A logistic transfer function is used for the hidden nodes and the output node. So, the range of an output value from the BP network is in [0, 1]. If the output value exceeds. 5, a firm is classified as bankrupt and non-bankrupt, otherwise.

The number of hidden nodes is used as a major experimental factor : there is no definite rule to follow in this hidden node decision. Thus, one - ten hidden nodes are tested in this study.

The Levenberg-Marquardt algorithm for training is used since it provides the fastest convergence and is specifically designed for the square error cost function. Its algorithm works in a way that when error size is large, the algorithm approximates gradient decent, whereas error size become smaller, the algorithm becomes the Gauss-Newton method which is more faster and efficient (MathWorks, 1997). The mean square error (MSE) function is used for the error function. It has been a popular choice in the past literature for theoretical consideration and provides a consistent error function (Berardi, 1998 ; Tam and Kiang, 1992).

3.4 Kohonen Self-Organizing Feature Map

A two-dimensional Kohonen output layer is used to help provide a visual presentation. Five input nodes are used corresponding to the five financial ratios. However, selecting the appropriate number of output nodes is a quite difficult and this is usually experiment-dependent. There is no consensus among researchers about the subject. For example, Serrano-Cinca (1995) used a 14 by 14 grid for 74 training vectors. As a result, some output nodes may not win for the given input vector. Nour (1994) suggested that to obtain good mapping results, the number of output nodes in the Kohonen neural network should be at least 10~20% of training vectors (or objects). However, using too few output nodes may cause the congestion of input vectors over an output node, which may make it difficult to distinguish the characteristics of the output space. Thus, it seems that it had better use a large number of output nodes and this study thus uses 200 output nodes.

The weight of each connection between an input node and an output node is initialized in a random value. The four subsets (generated by the cross-validation technique) are randomly sequenced to minimize any variation due to any possible input sequence pattern. So, with a combination of initial weight randomization and input sequence randomization, it is hoped that any bias effect can be minimized.

For the training cycle decision, there is no definitive stopping point. A heuristic is to use enough training cycles so that a network approaches a stable state. A preliminary experiment showed that the self-organizing feature map usually adjust its weights quickly to their inputs. Thus it may not be necessary that the Kohonen self-organizing feature map needs too many training cycles.

The performance results of the Kohonen self-organizing map are also shown in a tabular form so that they can be easily compared with the outcomes of the BP and of other statistical classifiers.

3.5 Statistical Models

Two widely used statistical techniques, discriminant analysis and logistic regression, are conducted to provide performance benchmarks. For discriminant analysis, a quadratic discriminant analysis (QDA) is used rather than a linear function since the covariance matrices are different (Hair *et al.,* 1995). The range of the expected output value of the logistic regression is [0, 1], so it is usually interpreted as the probability of class belonging. Further, it has been suggested in practice that a logistic regression approach is often preferred, especially when the basic assumption of normality of the variables is not met (Hair *et al.,* 1995).

In sum, the four classification technologies compared are : (1) the BP network, (2) the Kohonen self-organizing feature map, (3) logistic regression analysis, and (4) discriminant analysis. For statistical analysis, the SAS program is used. For neural network experiments, the Matlab Neural Network Toolbox 5.3 is used for the BP network while the Viscovery Som 3.0 is used for the Kohonen self-organizing feature map. Neural network experiments were done on three Pentium III personal computers. Note that for the BP experiment, five different random seeds of initial weights were generated and then, five different runs were done for each factor.

The prediction accuracy of the test sets obtained by each run was ranked by total misclassifications, and the median run was taken as the reported result. Performance outcomes of each classifier are measured in terms of : 1) the number and percentage of correct classification/prediction, and the number and percentage of misclassification : 2) Type I error (bankrupt firms being classified as non-bankrupt) versus Type II error (non-bankrupt firms being classified as bankrupt).

4. Experimental Results

4.1 The BP Network

Appendix 1 and 2 summarize the hidden node effects on the BP neural network performance when applied to the four-fold cross-validation scheme to the two data sets, 2-year (Appendix 1) or 3-year (Appendix 2) prior to a bankrupt filing. Four subsets of the whole data were iteratively tested with one to ten hidden nodes of the architectural design of the BP networks. So, the four-fold cross-validation scheme with the ten different hidden-node combinations resulted in 40 cells explored in each table.

In both appendix, each of the hidden node experiments shows the number (and percentage) of the correct *classification* for the four training sets (Panel A) as well as of the correct *prediction* for the four test sets (Panel B). Note that the validation process using the early stopping is embedded in the training period so that this embedded training would, it is hoped, minimize an upward bias. Appendix 1 and 2 also give the Type I error and Type II error. Since we did not incorporate the misclassification cost difference between Type I and Type II errors within the training procedure itself, only thing matter in this study is the total misclassification number, rather than each separate misclassification number or cost.

Following procedure identifies the best model (**bold**-faced) for a subset by exercising two criteria : (1) the highest *prediction* rate in terms of the test set results and (2) if there were a tie among the model results, the parsimonious model, the one with the least number of hidden nodes, was chosen as the best model. In Appendix 1, for example, where the 2-year data sets were used, the best model for Subset 1 was the one-hidden node model. In fact, the correct prediction rate for Subset 1 is highest at 73.81% in the one-, the three-, or the five- hidden node models (when applied the highest prediction rate criterion). But with the use of the parsimony criterion, the one-hidden node model is one that shows not only the highest correct prediction rate but is also the most economic model.

It is interesting to see that in Appendix 1 and 2, the best model for each subset is either the one- or the two-hidden node model except the eight-hidden node model for Subset 4 in Appendix 2. There have been neural network studies focusing on the issue of tradeoff between sample size and model selection (Berardi, 1998). The general opinion seems to be that as the amount of data increases, the algorithm tends to use more complex models such as adding more hidden nodes. In other words, for a small sample size, the simpler models are usually preferred (Berardi, 1998). Though research findings in this area may not be conclusive at the present time, the findings of this study also support the above argument that the model selection is constrained by the sample size being used. Training objects of this study are less than 90 and thus, simpler models such as one-or two-hidden node models were selected as the best models shown above.

Prediction rates varied considerably across the four subsets. In Panel B in Appendix 1, the correct prediction rates for the test sets were 73.81% (Subset 1), 71.43% (Subset 2), 78.57% (Subset3) and 59.52% (Subset 4). In Panel B in Appendix 2, the correct prediction rates for the test set were 71.43% (Subset 1 thru Subset 3) and 66.67% (Subset 4). The prediction accuracy for the two-year data set is greater for two subsets, equal for one subset, and lower for the remaining subset. From these limited results, we may not say that the BP network identifies and differentiates the 2-year data sets from the 3-year data sets.

4.2 Kohonen Self-Organizing Feature Map

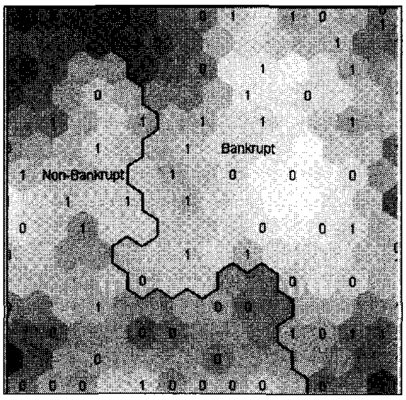
[Figures 3] to [Figure 4] gave some portions of the performance results of the Kohonen self-organizing feature map. Note first that each of these Kohonen self-organizing maps is given in two-dimensional space.

The 2-year data set was used to draw [Figure 3] (classification result) and [Figure 4] (prediction results).

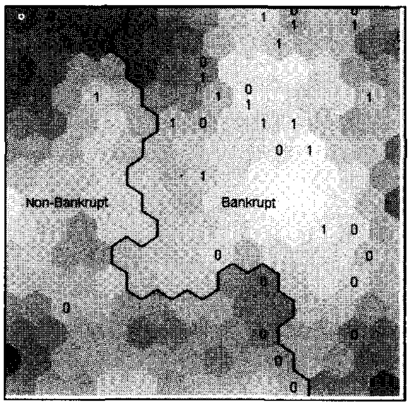
In [Figures 3]~[Figure 4], the label 1 denoted a bankrupt firm and the label 0 a non-bankrupt firm. The naming of a cluster is as follows. A cluster is named bankrupt if it contained more bankrupt firm than non-bankrupt ones. Otherwise it was named a non-bankrupt group.

<Table 2> gave the summary of the experimental results of the Kohonen self-organizing feature maps in tabular form.

When the 2-year data sets were used, the correct classification rates for the training sets ranged from 64.29% to 73.81% while the correct predication rates for the test sets ran from 54.76% to 66.67%. When the 3-year data sets were used, the correct classification rates of the Kohonen self-organizing feature maps for the training sets were in between 66.67% and 70.24% while the correct prediction rates for the test sets were in the range of 52.38% to 76.19%.



[Figure 3] Performance results of the Kohonen network for the Subset 2 (training set) in 2-year



[Figure 4] Performance results of the Kohonen network for the Subset 2 (test set) in 2-year

In a bankruptcy prediction study, the early detection is very important. That is why we test the 2-year versus the 3-year data sets using neural networks or statistical classifiers, to see if either methodology can detect any early distress call. In this Kohonen experiment, like the BP, it is difficult to say that the Kohonen network differentiates the 2-year data set from the 3-year.

4.3 Statistical Models

Discriminant analysis (DA) was performed. Since the assumption of the equal covariance is not met in this paired bankruptcy study, quadratic DA was conducted and the performance results were presented in <Table 3>. Discrimi-

〈Table 2〉 Performance Results of the Kohonen Self-Organizing Feature Map

Year	Type		Subset 1		Subset 2		Subset 3		Subset 4	
			#	%	#	%	#	%	#	%
2	Training	Correct Classification	58	69.05%	54	64.29%	56	66.67%	62	73.81%
		Type I error	12	14.29%	17	20.24%	22	26.19%	13	15.48%
		Type II error	14	16.67%	13	15.48%	6	7.14%	9	10.71%
	Test	Correct Prediction	28	66.67%	27	64.29%	28	66.67%	23	54.76%
		Type I error	1	2.38%	14	33.33%	13	30.95%	7	16.67%
		Type II error	13	30.95%	1	2.38%	0	0.00%	12	28.57%
3	Training	Correct Classification	59	70.24%	58	69.05%	56	66.67%	59	70.24%
		Type I error	12	14.29%	14	16.67%	15	17.86%	14	16.67%
		Type II error	13	15.48%	12	14.29%	13	15.48%	11	13.10%
	Test	Correct Prediction	32	76.19%	25	59.52%	28	66.67%	22	52.38%
		Type I error	4	9.52%	15	35.71%	5	11.90%	9	21.43%
		Type II error	6	14.29%	2	4.76%	9	21.43%	11	26.19%

〈Table 3〉 Performance Results of QDA

Year	Type		Subset 1		Subset 2		Subset 3		Subset 4	
			#	%	#	%	#	%	#	%
2	Training	Correct Classification	59	70.24%	56	66.67%	54	64.29%	56	66.67%
		Type I error	16	19.05%	19	22.62%	17	20.24%	20	23.81%
		Type II error	9	10.71%	9	10.71%	13	15.48%	8	9.52%
	Test	Correct Prediction	25	59.52%	28	66.67%	28	66.67%	23	54.76%
		Type I error	8	19.05%	12	28.57%	10	23.81%	12	28.57%
		Type II error	9	21.43%	2	4.76%	4	9.52%	7	16.67%
3	Training	Correct Classification	62	73.81%	51	60.71%	51	60.71%	54	64.29%
		Type I error	15	17.86%	24	28.57%	26	30.95%	25	29.76%
		Type II error	7	8.33%	9	10.71%	7	8.33%	5	5.95%
	Test	Correct Prediction	26	61.90%	27	64.29%	27	64.29%	26	61.90%
		Type I error	10	23.81%	14	33.33%	13	30.95%	14	33.33%
		Type II error	6	14.29%	1	2.38%	2	4.76%	2	4.76%

<Table 4> Performance Results of Logistic Regression

Year	Type		Subset 1 #	Subset 1 %	Subset 2 #	Subset 2 %	Subset 3 #	Subset 3 %	Subset 4 #	Subset 4 %
2	Training	Correct Classification	64	76.19%	56	66.67%	56	66.67%	62	73.81%
		Type I error	8	9.52%	12	14.29%	13	15.48%	9	10.71%
		Type II error	12	14.29%	16	19.05%	15	17.86%	15	17.86%
	Test	Correct Prediction	24	57.14%	29	69.05%	33	78.57%	27	64.29%
		Type I error	8	19.05%	10	23.81%	8	19.05%	6	14.29%
		Type II error	10	23.81%	3	7.14%	1	2.38%	9	21.43%
3	Training	Correct Classification	59	70.24%	52	61.90%	52	61.90%	56	66.67%
		Type I error	12	14.29%	15	17.86%	17	20.24%	10	11.90%
		Type II error	13	15.48%	17	20.24%	15	17.86%	18	21.43%
	Test	Correct Prediction	26	61.90%	26	61.90%	28	66.67%	26	61.90%
		Type I error	5	11.90%	10	23.81%	9	21.43%	8	19.05%
		Type II error	11	26.19%	6	14.29%	5	11.90%	8	19.05%

nant analysis was implemented using SAS procedure DISCRIM. In all cases the prior probability proportional to group size option was used.

Whereas the 2-year data sets were used, the correct classification rates for the training sets ranged from 64.29% to 70.24%. The correct prediction rates for the test sets were in between 54.76% and 66.67%. Whereas the 3-year data sets were used, the correct classification rates for the training sets went from 60.71% to 73.81% and the correct prediction rates for the test sets are in between 61.90% and 64.29%. Second, the performance results of logistic regression were shown in Table.

Logistic regression was implemented using SAS procedure LOGISTIC. For the 2-year data sets, the correct classification rates for the training sets ranged 66.67% to 76.19% while the correct prediction rates for the test sets are in between 57.14% and 78.57% in <Table 4>. For the 3-year data sets being used, the range of the correct classification rates for the training sets gave 61.90% to 70.24% while the correct prediction rates for the testing sets went between 61.90% and 66.67% in <Table 4>.

4.4 Performance Comparisons Among Classification Techniques

<Table 5> provides a summary of performance comparison among the four methodological choices used in this bankruptcy prediction study : the two neural networks, the BP networks and the Kohonen self-organizing feature maps, and of the two statistical classifiers, QDA and logistic regression.

In <Table 5>, the best *classification* model (for training) for each subset was identified by the <u>underlined</u> character. As expected, the BP networks showed the highest classification accuracy.

Among the total eight sets, seven from the BP network, one from the Kohonen network, and another from QDA were selected as the best classification models. Note that there was a tie

〈Table 5〉 Summary of Performance Result

Year	Subset	Type	BP		Kohonen		QDA		Logistic	
			#	%	#	%	#	%	#	%
2	Subset 1	Training	66	78.57%	58	69.05%	59	70.24%	64	76.19%
		Test	**31**	**73.81%**	28	66.67%	25	59.52%	24	57.14%
	Subset 2	Training	64	76.19%	54	64.29%	56	66.67%	56	66.67%
		Test	30	71.43%	27	64.29%	28	66.67%	29	69.05%
	Subset 3	Training	59	70.24%	56	66.67%	54	64.29%	56	66.67%
		Test	**33**	**78.57%**	28	66.67%	28	66.67%	**33**	**78.57%**
	Subset 4	Training	69	82.14%	62	73.81%	56	66.67%	62	73.81%
		Test	25	59.52%	23	54.76%	23	54.87%	**27**	**64.29%**
3	Subset 1	Training	61	72.62%	59	70.24%	62	73.81%	59	70.24%
		Test	30	71.43%	**32**	**76.19%**	26	61.90%	26	61.90%
	Subset 2	Training	59	70.24%	58	69.05%	51	60.71%	52	61.90%
		Test	**30**	**71.43%**	25	59.52%	27	64.29%	26	61.90%
	Subset 3	Training	56	66.67%	56	66.67%	51	60.71%	52	61.90%
		Test	30	71.43%	28	66.67%	27	64.29%	28	66.67%
	Subset 4	Training	62	73.81%	59	70.24%	54	64.29%	56	66.67%
		Test	**28**	**66.67%**	22	52.38%	26	61.90%	26	61.90%

in the classification result between the BP and the Kohonen in Subset 3 in the 3-year data set being used.

Again, the best *prediction* model (for test) for each subset was identified in the **bold**-faced character. The BP networks showed the best prediction capabilities across the sample variability. Among the total eight sets, six from the BP network, two from the logistic regression, and one from the Kohonen network were recognized as the best prediction models. Note again that there was a tie between the BP network and the logistic model in Subset 3 when the 2-year data set was used.

The effect of data set year was tested in <Table 6>. It gave the SAS ANOVA results for both training and test sets.

This test was performed to see whether there was a performance difference in the two different data sets, namely, the 2-year or the 3-year data sets prior to bankruptcy filing.

〈Table 6〉 Effect of Data Set Year on Performance Accuracy

Type	2-year Mean	3-year Mean	F-value	P-value
Training	59.44	56.69	3.45	0.0729
Test	27.63	27.31	0.10	0.7544

In the training sets, there was a performance difference at 10% significant level.

However in test sets, ANOVA results indicated that there was no significant difference using either these two different timed data sets. In other words, the year-effect did not seem to be a critical factor influencing the prediction capabilities of the classification techniques compared here. One speculation for no year effect presented was that for most bankrupt firms in these particular data sets, financial deterioration had already begun well before the three fiscal years, and thus, no financial strength difference was detected between the 2-year or 3-year data sets being used.

Effectiveness of each individual technique should also be tested, which is shown in <Table 7>. It gave the SAS ANOVA results with the Duncan Option for both training and test sets. Letters in the Duncan grouping columns showed the group to which the mean of each group belongs. Different letters in the Duncan Option indicated that groups are significantly different at a 5% level. In other words, same letters in the Duncan grouping mean that they can be grouped together statistically.

〈Table 7〉 Results of Duncan's multi-group test for both training and test sets

Set	Technique	Means	Duncan Grouping
Training	BP	62.000	A
	Logistic	57.125	B
	Kohonen	57.750	B
	DA	55.375	B
Test	BP	29.625	A
	Logistic	27.375	A B
	Kohonen	26.625	B
	DA	26.250	B

<Table 7> showed that the BP network seemed to be the best accuracy model. The runner-up was the logistic regression as shown in the test section of the Duncan grouping. The performance accuracy of the Kohonen network and DA showed that they were not as accurate as the BP network for both training and test. But their prediction rates were comparable to the logistic regression. Back in <Table 5>, discriminant analysis did not seem to show a comparable performance to the remaining three other techniques, and thus its presence might obscure some statistical tests.

For this reason, we dropped the discriminant analysis method in the following ANOVA test, which focused on the prediction (test) performance of the remaining three individual classification techniques : BP, Kohonen, and logistic regression. A one-way ANOVA with one repeated-measure design was used to compare the possible performance (test) rate differences among the three classification techniques.

The null hypothesis was that there is no difference in the prediction rates of the three classification techniques. Since the p value is .0434 ($F_{2,14}$ = 3.96) of the one-way ANOVA with one repeated-measure design, we have the test result that rejected the null hypothesis of no differences in the mean levels of prediction rates among the three classification techniques.

Thus, we went on the paired comparison between methods which were performed three times as : (1) the BP versus the Kohonen, (2) the Kohonen versus the logistic regression, and finally (3) the logistic regression versus the BP networks in <Table 8>.

〈Table 8〉 Paired Comparison among the three Classifiers with Prediction Rate

	BP vs. Kohonen	Kohonen vs. Logistic	Logistic vs. BP
F-value	11.45	0.29	5.30
p-value	0.0117	0.6083	0.0548

<Table 8> showed that there was a prediction rate difference between the BP network and the Kohonen network at almost a 1% level. Also, the performance rates between the BP network and

the logistic regression differed nearly at a 5% significant level. It means that the BP networks showed the best performance results among the three classification tools. It was reassuring that the results of this study confirmed the findings of the previous literature that the BP networks provide a good mapping function for bankruptcy indication (Zhang *et al.*, 1999 ; Berardi, 1998 ; Tam and Kiang, 1992). The performance results of the Kohonen self-organizing feature maps were not as good as the BP networks, but they were comparable to logistic regression in this study.

5. Conclusions

The main purpose of this study was to investigate four different data mining techniques : BP, Kohonen self-organizing network, logistic regression, and discriminant analysis in bankruptcy prediction context. Two statistical modeling techniques, discriminant analysis and logistic regression, are provided to give some performance benchmarks for the neural network classifiers.

The tested for this study is the Korean listed companies. The findings of this study can be summarized as :

The impact of data set size in neural network experiment is particularly important. It should be noted that training data sets (84 objects) used in this study is rather a small one. Usually, BP networks provide a good posterior probability when they have enough objects to be learned. It is because the neural network paradigm is, in essence, a data driven non-parametric approach. However, we show that even with the small sample size, the BP networks consistently outperform the Kohonen self-organizing feature map and other statistical modeling techniques.

Though having observed in the discrepancy of these classification tools, we noticed that the Kohonen self-organizing feature map could be as used an alternative classification tool. The main advantage of Kohonen self-organizing feature map is in its visual presentation. Thus, researchers can draw and see the difference between bankruptcy groups and non-bankruptcy groups.

Another way of using the Kohonen self-organizing neural network is the tracking the financial condition of a firm over time. That is, each of a sample or population is plotted to the Kohonen map initially, and over a certain period of time, the position of a firm in the map is kept in track while the characteristic of each of the clusters is evaluated. The experimental prototype of this approach was given by Martin-del-Brio and Serrano-Cinca (1995). A major advantage of this approach is that we can examine the comparative financial condition of a firm in a real-time base. This method does not provide a high prediction rate, but it is more practical to detect a failing firm in a more timely fashion, the trait that is especially important in a fast changing business environment.

There are a variety of neural networks available for pattern classification tasks. Each of these neural network classifiers has its own advantages and disadvantages based on its algorithm and architecture. In this study, only two neural network types, the BP and Kohonen networks, along with two statistical data mining techniques, are compared in the bankruptcy prediction. In this study, as the first step to explore the potential of different types of networks, we contrasted the BP with the Kohonen networks in bankruptcy prediction. We found that even with some limitations such as inaccuracy, the Kohonen self-organizing feature map could be used as a promising classification technique in the fast-changing business environment.

Appendix

〈Appendix 1〉 Panel A : Effects of Hidden Nodes on BP experiment (2-year) for Training Sets

# of Hidden Nodes	Type	Subset 1		Subset 2		Subset 3		Subset 4	
		#	%	#	%	#	%	#	%
1	Correct Classification	**66**	**78.57%**	**64**	**76.19%**	59	70.24%	**69**	**82.14%**
	Type Ⅰ error	**4**	**4.76%**	**6**	**7.14%**	15	17.86%	**3**	**3.57%**
	Type Ⅱ error	**14**	**16.67%**	**14**	**16.67%**	10	11.90%	**12**	**14.29%**
2	Correct Classification	65	77.38%	63	75.00%	**59**	**70.24%**	68	80.95%
	Type Ⅰ error	4	4.76%	6	7.14%	**15**	**17.86%**	4	4.76%
	Type Ⅱ error	15	17.86%	15	17.86%	**10**	**11.90%**	12	14.29%
3	Correct Classification	63	75.00%	65	77.38%	56	66.67%	65	77.38%
	Type Ⅰ error	3	3.57%	5	5.95%	11	13.10%	6	7.14%
	Type Ⅱ error	18	21.43%	14	16.67%	17	20.24%	13	15.48%
4	Correct Classification	63	75.00%	68	80.95%	58	69.05%	67	79.76%
	Type Ⅰ error	9	10.71%	7	8.33%	10	11.90%	5	5.95%
	Type Ⅱ error	12	14.29%	9	10.71%	16	19.05%	12	14.29%
5	Correct Classification	64	76.19%	68	80.95%	62	73.81%	63	75.00%
	Type Ⅰ error	10	11.90%	8	9.52%	15	17.86%	8	9.52%
	Type Ⅱ error	10	11.90%	8	9.52%	7	8.33%	13	15.48%
6	Correct Classification	59	70.24%	63	75.00%	59	70.24%	65	77.38%
	Type Ⅰ error	3	3.57%	6	7.14%	10	11.90%	10	11.90%
	Type Ⅱ error	21	25.00%	15	17.86%	15	17.86%	9	10.71%
7	Correct Classification	61	72.62%	64	76.19%	57	67.86%	64	76.19%
	Type Ⅰ error	3	3.57%	6	7.14%	12	14.29%	6	7.14%
	Type Ⅱ error	19	22.62%	14	16.67%	15	17.86%	14	16.67%
8	Correct Classification	64	76.19%	57	67.86%	58	69.05%	62	73.81%
	Type Ⅰ error	5	5.95%	3	3.57%	6	7.14%	15	17.86%
	Type Ⅱ error	15	17.86%	24	28.57%	20	23.81%	7	8.33%
9	Correct Classification	66	78.57%	69	82.14%	50	59.52%	70	83.33%
	Type Ⅰ error	4	4.76%	6	7.14%	25	29.76%	6	7.14%
	Type Ⅱ error	14	16.67%	9	10.71%	9	10.71%	8	9.52%
10	Correct Classification	64	76.19%	57	67.86%	64	76.19%	66	78.57%
	Type Ⅰ error	9	10.71%	3	3.57%	9	10.71%	6	7.14%
	Type Ⅱ error	11	13.10%	24	28.57%	11	13.10%	12	14.29%

Note) That # denotes the number of objects classified into each category and that % shows the percent of objects classified into each category. Type Ⅰ error occurs when a bankrupt firm is classified as a non-bankrupt firm while Type Ⅱ error occurs when a non-bankrupt firm is classified as a bankrupt firm.

<Appendix 1> Panel B : Effects of Hidden Nodes on BP experiment (2-year) for Test Sets

# of Hidden Nodes	Type	Subset 1		Subset 2		Subset 3		Subset 4	
		#	%	#	%	#	%	#	%
1	Correct Prediction	**31**	**73.81%**	**30**	**71.43%**	31	73.81%	**25**	**59.52%**
	Type I error	**8**	**19.05%**	**4**	**9.52%**	6	14.29%	**5**	**11.90%**
	Type II error	**3**	**7.14%**	**8**	**19.05%**	5	11.90%	**12**	**28.57%**
2	Correct Prediction	30	71.43%	25	59.52%	**33**	**78.57%**	25	59.52%
	Type I error	6	14.29%	5	11.90%	**2**	**4.76%**	6	14.29%
	Type II error	6	14.29%	12	28.57%	**7**	**16.67%**	11	26.19%
3	Correct Prediction	31	73.81%	30	71.43%	30	71.43%	25	59.52%
	Type I error	5	11.90%	3	7.14%	4	9.52%	6	14.29%
	Type II error	6	14.29%	9	21.43%	8	19.05%	11	26.19%
4	Correct Prediction	27	64.29%	25	59.52%	32	76.19%	25	59.52%
	Type I error	11	26.19%	4	9.52%	5	11.90%	6	14.29%
	Type II error	4	9.52%	13	30.95%	5	11.90%	11	26.19%
5	Correct Prediction	31	73.81%	28	66.67%	33	78.57%	25	59.52%
	Type I error	7	16.67%	3	7.14%	6	14.29%	7	16.67%
	Type II error	4	9.52%	11	26.19%	3	7.14%	10	23.81%
6	Correct Prediction	29	69.05%	28	66.67%	32	76.19%	25	59.52%
	Type I error	5	11.90%	1	2.38%	4	9.52%	7	16.67%
	Type II error	8	19.05%	13	30.95%	6	14.29%	10	23.81%
7	Correct Prediction	29	69.05%	28	66.67%	33	78.57%	24	57.14%
	Type I error	8	19.05%	1	2.38%	3	7.14%	8	19.05%
	Type II error	5	11.90%	13	30.95%	6	14.29%	10	23.81%
8	Correct Prediction	28	66.67%	28	66.67%	32	76.19%	24	57.14%
	Type I error	12	28.57%	0	0.00%	4	9.52%	11	26.19%
	Type II error	2	4.76%	14	33.33%	6	14.29%	7	16.67%
9	Correct Prediction	29	69.05%	27	64.29%	32	76.19%	24	57.14%
	Type I error	11	26.19%	4	9.52%	4	9.52%	8	19.05%
	Type II error	2	4.76%	11	26.19%	6	14.29%	10	23.81%
10	Correct Prediction	29	69.05%	26	61.90%	33	78.57%	23	54.76%
	Type I error	10	23.81%	2	4.76%	4	9.52%	5	11.90%
	Type II error	3	7.14%	14	33.33%	5	11.90%	14	33.33%

Note) That # denotes the number of objects classified into each category and that % shows the percent of objects classified into each category. Type I error occurs when a bankrupt firm is classified as a non-bankrupt firm while Type II error occurs when a non-bankrupt firm is classified as a bankrupt firm.

〈Appendix 2〉 Panel A : Effects of Hidden Nodes on BP experiment (3-year) for Training Sets

# of Hidden Nodes	Type	Subset 1		Subset 2		Subset 3		Subset 4	
		#	%	#	%	#	%	#	%
1	Correct Classification	61	72.62%	**59**	**70.24%**	**56**	**66.67%**	61	72.62%
	Type Ⅰ error	10	11.90%	**11**	**13.10%**	**16**	**19.05%**	11	13.10%
	Type Ⅱ error	13	15.48%	**14**	**16.67%**	**12**	**14.29%**	12	14.29%
2	Correct Classification	**61**	**72.62%**	65	77.38%	56	66.67%	59	70.24%
	Type Ⅰ error	**9**	**10.71%**	6	7.14%	17	20.24%	18	21.43%
	Type Ⅱ error	**14**	**16.67%**	13	15.48%	11	13.10%	7	8.33%
3	Correct Classification	61	72.62%	65	77.38%	59	70.24%	58	69.05%
	Type Ⅰ error	13	15.48%	6	7.14%	12	14.29%	14	16.67%
	Type Ⅱ error	10	11.90%	13	15.48%	13	15.48%	12	14.29%
4	Correct Classification	59	70.24%	59	70.24%	58	69.05%	57	67.86%
	Type Ⅰ error	9	10.71%	18	21.43%	9	10.71%	19	22.62%
	Type Ⅱ error	16	19.05%	7	8.33%	17	20.24%	8	9.52%
5	Correct Classification	62	73.81%	66	78.57%	55	65.48%	65	77.38%
	Type Ⅰ error	6	7.14%	8	9.52%	8	9.52%	6	7.14%
	Type Ⅱ error	16	19.05%	10	11.90%	21	25.00%	13	15.48%
6	Correct Classification	60	71.43%	73	86.90%	62	73.81%	64	76.19%
	Type Ⅰ error	18	21.43%	5	5.95%	14	16.67%	7	8.33%
	Type Ⅱ error	6	7.14%	6	7.14%	8	9.52%	13	15.48%
7	Correct Classification	62	73.81%	64	76.19%	64	76.19%	70	83.33%
	Type Ⅰ error	9	10.71%	6	7.14%	9	10.71%	2	2.38%
	Type Ⅱ error	13	15.48%	14	16.67%	21	25.00%	12	14.29%
8	Correct Classification	66	78.57%	71	84.52%	65	77.38%	**62**	**73.81%**
	Type Ⅰ error	4	4.76%	5	5.95%	10	11.90%	**8**	**9.52%**
	Type Ⅱ error	14	16.67%	8	9.52%	9	10.71%	**10**	**11.90%**
9	Correct Classification	68	80.95%	64	76.19%	62	73.81%	68	80.95%
	Type Ⅰ error	4	4.76%	6	7.14%	13	15.48%	5	5.95%
	Type Ⅱ error	12	14.29%	14	16.67%	9	10.71%	11	13.10%
10	Correct Classification	70	83.33%	61	72.62%	62	73.81%	65	77.38%
	Type Ⅰ error	7	8.33%	10	11.90%	7	8.33%	9	10.71%
	Type Ⅱ error	7	8.33%	13	15.48%	15	17.86%	10	11.90%

〈Appendix 2〉 Panel B : Effects of Hidden Nodes on BP experiment (3-year) for Test sets

# of Hidden Nodes	Type	Subset 1		Subset 2		Subset 3		Subset 4	
		#	%	#	%	#	%	#	%
1	Correct Prediction	29	69.05%	**30**	**71.43%**	**30**	**71.43%**	21	50.00%
	Type I error	10	23.81%	**2**	**4.76%**	**7**	**16.67%**	12	28.57%
	Type II error	3	7.14%	**10**	**23.81%**	**5**	**11.90%**	9	21.43%
2	Correct Prediction	**30**	**71.43%**	26	61.90%	24	57.14%	25	59.52%
	Type I error	**7**	**16.67%**	2	4.76%	11	26.19%	7	16.67%
	Type II error	**5**	**11.90%**	14	33.33%	7	16.67%	10	23.81%
3	Correct Prediction	30	71.43%	26	61.90%	30	71.43%	25	59.52%
	Type I error	10	23.81%	2	4.76%	5	11.90%	7	16.67%
	Type II error	2	4.76%	14	33.33%	7	16.67%	10	23.81%
4	Correct Prediction	29	69.05%	27	64.29%	24	57.14%	24	57.14%
	Type I error	7	16.67%	2	4.76%	9	21.43%	7	16.67%
	Type II error	6	14.29%	13	30.95%	9	21.43%	11	26.19%
5	Correct Prediction	30	71.43%	27	64.29%	28	66.67%	25	59.52%
	Type I error	8	19.05%	4	9.52%	6	14.29%	7	16.67%
	Type II error	4	9.52%	11	26.19%	8	19.05%	10	23.81%
6	Correct Prediction	28	66.67%	26	61.90%	30	71.43%	27	64.29%
	Type I error	13	30.95%	3	7.14%	8	19.05%	7	16.67%
	Type II error	1	2.38%	13	30.95%	4	9.52%	8	19.05%
7	Correct Prediction	29	69.05%	28	66.67%	30	71.43%	25	59.52%
	Type I error	9	21.43%	4	9.52%	3	7.14%	6	14.29%
	Type II error	4	9.52%	10	23.81%	9	21.43%	11	26.19%
8	Correct Prediction	29	69.05%	25	59.52%	29	69.05%	**28**	**66.67%**
	Type I error	8	19.05%	6	14.29%	8	19.05%	**7**	**16.67%**
	Type II error	5	11.90%	11	26.19%	5	11.90%	**9**	**21.43%**
9	Correct Prediction	30	71.43%	28	66.67%	29	69.05%	25	59.52%
	Type I error	9	21.43%	2	4.76%	7	16.67%	6	14.29%
	Type II error	3	7.14%	12	28.57%	6	14.29%	11	26.19%
10	Correct Prediction	29	69.05%	26	61.90%	28	66.67%	28	66.67%
	Type I error	10	23.81%	7	16.67%	4	9.52%	7	16.67%
	Type II error	3	7.14%	9	21.43%	10	23.81%	7	16.67%

Note) That # denotes the number of objects classified into each category and that % shows the percent of objects classified into each category. Type I error occurs when a bankrupt firm is classified as a non-bankrupt firm while Type II error occurs when a non-bankrupt firm is classified as a bankrupt firm.

REFERENCES

[1] Alam, P, Booth, D, Lee, K, Thordarson, T., "The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks : an experimental study," *Expert Systems With Applications*, 18(2000), pp.185-199.

[2] Altman, EL., "Financial ratios, discriminate analysis and the prediction of corporate bankruptcy," *Journal of Finance*, 23(3) (1968), pp.589-609.

[3] Beaver, W., "Financial ratios as predictors of failure, empirical research in accounting," *selected studies 1966, Journal of Accounting Research*, 4(1966), pp.71-111.

[4] Berardi, V., *An Investigation of neural network ensemble methods for posterior probability estimation*, Unpublished Ph.D. dissertation, Kent State University, 1998.

[5] Corridoni, JM, Del Bimbo, A, Landi, L., "3D object classification using multi-object Kohonen networks," *Pattern Recognition*, 29(6)(1996), pp.919-935.

[6] Der Voort, M, Dougherty, M, Watson, S., "Combining Kohonen maps with arima time series models to forecast traffic flow," *Transportation Research Part C : Emerging Technologies*, 4(5)(1996), pp.307-318.

[7] Deschenes, CJ, Noonan, J., "Fuzzy Kohonen network for the classification of transients using the wavelet transform for feature extraction," *Information Sciences*, 87(4) (1995), pp.247-266.

[8] Funahashi, K., "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, 2(1989), pp. 189-192.

[9] Hinton, G Sejnowski, T. (eds.), *Unsupervised learning : foundations of neural computation.* The MIT Press Cambridge Massachusetts, 1999.

[10] Hornik, K, Stinchcombe, M, White, H., "Multi-layer feedforward networks are universal approximators," *Neural Networks*, 2(1989), pp.359-366.

[11] Jo, HY, Han, IG, Lee, HY., "Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis," *Expert Systems With Applications*, 13(2)(1997), pp.97-108.

[12] Kelly, K., *Out of Control, The New Biology of Machines, Social Systems, and the Economic World*, Perseus Books, Reading Massachusetts, 1994.

[13] Kiviluoto, K., "Predicting bankruptcies with the self-organizing map," *Neurocomputing*, 21(1998), pp.191-201.

[14] Kohonen, T., "Correlation matrix memories," *IEEE Trans. on Computers*, C-21 (1972), pp.353-359.

[15] Kohonen, T., "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, 43(1982), pp.59-69.

[16] Kohonen, T., *Self-organizing Maps.* Second Ed, Springer-Verlag, Berlin, 1997.

[17] MathWorks., *Matlab : Neural Network Toolbox, User's Guide*, The MathWorks, Inc., 24 Prime Park Way, Natick, MA, 01760, 1997.

[18] Martin-del-Brio, B Serrano-Cinca, C., "Self-organizing neural networks : the financial state of spanish companies," *Neural Network in the Capital Markets* Refenes (eds). (1995), pp.341-357.

[19] Nour, MA., *Improved clustering and*

classification algorithms for the Kohonen self-organizing neural network, Unpublished Ph.D. dissertation, Kent State University, 1994.

[20] Odom, M Sharda, R., "A neural network model for bankruptcy prediction," *Proceedings of the IEEE International Conference on Neural networks*, (1990) pp. 163-168.

[21] Ohlson, J., "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, 18(1)(1980), pp.109-131.

[22] O'Leary, DE., "Using neural networks to predict corporate failure," *International Journal of Intelligent Systems in Accounting, Finance and Management*, 7 (1998), pp.187-197.

[23] Schonweiler, R, Kaese, S, Moller, S, Rinscheid, A, Ptok, M., "Classification of spectrographic voice patterns using self-organizing neuronal networks (Kohonen maps) in the evaluation of the infant cry with and without time-delayed feedback," *International Journal of Pediatric Otorhinolaryngology*, 38(2)(1996).

[24] Serrano-Cinca, C., "Self organizing neural networks for financial diagnosis. *Decision Support Systems*, 17(1996), pp.227-238.

[25] Sharda, R Wilson, RL., "Neural network experiments in business-failure forecasting : predictive performance measurement issues," *International Journal of Computational Intelligence and Organization*, 1 (2)(1996), pp.107-117

[26] Tam, KY Kiang, MY., "Managerial applications of neural networks : the case of bank failure predictions," *Management Science*, 38(7)(1992), pp.926-947.

[27] Tsukuda, J Baba, SI., "Predicting Japanese corporate bankruptcy in terms of financial date using neural network," *Computers and Industrial Eng.*, (1994), pp.445-448.

[28] Udo, G., "Neural network performance on the bankruptcy classification problem," *Computers and Industrial Engineering*, 25 (1993), pp.377-380.

[29] Waldrop, MM., *Complexity.* Touchstone, Rochefeller Center, 1230 Avenue of the Americas, New York, NY, 10020, Simon & Schuster, Inc, 1992.

[30] Wasserman, PD., *Neural computing : theory and practice*, Van Nostrand Reinhold, NY, 1989.

[31] Wilson, RL Sharda, R., "Bankruptcy prediction using neural networks," *Decision Support Systems*, 11(1994), pp.545-557.

[32] Xin-Hua, S Hopke, PK., "Kohonen neural network as a pattern recognition method based on the weight interpretation," *Analytica Chimica Acta*, 336(1996), pp.57-66.

[33] Zhang, G, Hu, MY, Patuwo, BE, Indro, DC., "Artificial neural networks in bankruptcy prediction : general framework and cross-validation analysis," *European Journal of Operational Research*, 116(1999), pp.16-32.