Bankruptcy Predictions for Korea Medium-Sized Firms using Neural Networks and Case Based Reasoning

Ingoo Han, Cheolsoo Park, Chulhong Kim

Graduate School of Management
Korea Advanced Institute of Science and Technology
Telephone : 02-958-3613   E-mail : ingoohan@kaist.msd.ac.kr

ABSTRACT

Prediction of firm bankruptcy have been extensively studied in accounting, as all stockholders in a firm have a vested interest in monitoring its financial performance. The objective of this paper is to develop the hybrid models for bankruptcy prediction. The proposed hybrid models are two phase. Phase one are (a) DA-assisted neural network, (b) Logit-assisted neural network, and (c) Genetic-assisted neural network. And, phase two are (a) DA-assisted Case based reasoning, and (b) Genetic-assisted Case based reasoning. In the variables selection, We are focusing on three alternative methods - linear discriminant analysis, logit analysis and genetic algorithms - that can be used empirically select predictors for hybrid model in bankruptcy prediction. Empirical results using Korean medium-sized firms data show that hybrid models are very promising neural network models and case based reasoning for bankruptcy prediction in terms of predictive accuracy and adaptability.

Key Words : Bankruptcy Prediction, Neural Networks, Genetic Algorithms, Case Based Reasoning, Discriminant Analysis, Logit Analysis, Hybrid Model

INTRODUCTION

Firm bankruptcy worldwide have occurred at rates higher than at any time since the early 1960s. During the past two decades, economic figures indicated a significant upsurge in the number of bankruptcies. The Korea data suggest that the annual bankruptcy rate of medium-sized firms is steadily increasing. Higher interest rates, recessionary economic conditions, and heavy debt burdens have made medium-sized companies increasingly vulnerable. The rate of bankruptcy is an important economic indicator and is the focus of much public attention.

Bankruptcy prediction has been one of the most challenging tasks in accounting since the study of Fitzpatrick in 1930's and during the last 60 years an impressive body of both theoretical and empirical research concerning this topic has evolved. From 1966 until 1980's discriminant analysis was the dominant method in bankruptcy prediction. During the 1980's the method was replaced by logistic analysis and in the 1990's artificial neural networks have produced very promising results in predicting bankruptcies. Artificial neural networks have performed well in business classifications including bankruptcy prediction[1]. A neural network model has a good ability in modeling and forecasting. The success of the neural network model is attributable to its generalization capability to predict the output for new data after the neural network was trained[8]. Previous empirical results show that neural network models provide higher predictive accuracy than statistical methods. However, it has not been resolved yet how to design a neural network model suitable for a specific problem. It has been proposed that the hybrid model combing two or more models has a potential to achieve a high predictive performance.

Odom and Sharda[10] compared the performance of neural network and discriminant analysis. The accuracy level was 81.48% for the neural network and 59.26% for DA. They used traditional ratios, but also absolute numbers from financial statements as prediction variables.

Jo and Han[7] compared DA, neural network and case based reasoning. The accuracy level was as high as 83.79% for neural network, 82.22% for DA and 81.52% for case based reasoning.

Lee, Han & Kwon[6] showed hybrid neural network model for performance. They proposed hybrid neural network either integrate different kinds of neural network (like SOFM-assisted neural network) or combine neural network with other statistical or AI models(like MDA-assisted or ID3-assisted neural network). The accuracy level was 68.57% for MDA, 75.24% for MDA-assisted neural network and 80.48 for SOFM assisted neural network

Back et al.[5] used neural network and the back propagation algorithm and received an

accuracy level of 95% in bankruptcy prediction. In a recent study back et al. Compared discriminant analysis, logit analysis and neural network with each other. They used different variable selection methods for each prediction method. For discriminant and logit analysis they used stepwise variables selection methods. The best prediction result - an accuracy level of 97% - was received with neural network combined with genetic algorithms.

This paper propose the five kinds of hybrid neural network Models: (a) DA-assisted neural network Models (b) Logit-assisted neural network Models, (c) Genetic Algorithm-assisted neural network Models, (d) DA-assisted case based reasoning, and (e) GA-assisted case based reasoning. DA, logit analysis and genetic algorithms all have different assumptions concerning the relationships between the independent variables. The purpose of this paper are three (1) to explore the possibility of the proposed hybrid models, (2) to compare the accuracy in bankruptcy prediction, and (3) to study if the variable selection methods used in discriminant analysis and logit analysis just give lower bounds of the hybrid models predictive performance and if the genetic algorithm is the algorithm that improves the results.

RESEARCH DATA AND VARIABLES SELECTION
Research Data

The data sample consists of Korean Medium-sized firms that failed in the 1993-1995. They were selected from the bankrupt companies listed in the Korea Credit Guarantee Fund Database. The experiment case consists of same number of bankruptcy and non-bankruptcy case. We have included the data of 1993 which includes the financial value of 242 bankruptcy firm. The data of 1994 and 1995 consists of 293 and 786 cases respectively. Bankruptcy firms are distributed in various industries. To make experimental sets, we adopted the pair matching method. Therefore, 2548 firms in total were selected as data set for experiment(Sound : 1274, Failure : 1274).

Variable Selection

There are two approaches in bankruptcy prediction: quantitative and qualitative analysis. Quantitative analysis use financial index or ratios evaluated by financial statement, and qualitative analysis use company additive information, for example, growing possibility, history, and executive's leadership.

This study adopts two main principles, on is to use ratio which were regarded as important in

previous research and the other is to use ratios in the practical area. Initially 65 financial variables and 2 non-financial variables(History, and Type) are selected for the prediction model.

Discriminant Analysis. In this research, we used the stepwise method in SPSS using Wilk's Lambda to sequentially select important variables from 65 variables. Table 1 show the DA results. In the DA results, 8 financial variables were selected as important input variables for predicting bankruptcy.

Logit Analysis. The logit analysis was derived using the logistic regression procedures available in SPSS. AS in DA models, we used stepwise selection method. In the logit analysis results, 10 financial variables were selected as important input variables for predicting bankruptcy. Table 2 show the logit analysis results.

Genetic Algorithms. The genetic algorithm that performs input variable selection is looking for sets of variables that act in a synergistic manner as good predictors of the output data[7], [9]. We used *NeuralWorks* Professional V5.0 Predict(Neural Ware Inc., 1994) to select input variables for predicting bankruptcy. Table 3 show the GA results.

HYBRID MODELS FOR NEURAL NETWORKS
: Phase I
This study has constructed 2-layer neural networks model showed in Table 4 The data was arbitrarily split into a training set, consisting of 1544 companies, a test set consisting of 502 companies, and a validation set consisting of 502 companies. The input layer consisted of the variables from each model. The output layer had only one node in each network, with the output value 0 for bankruptcy, and 1 for non-bankruptcy. Several different hidden layer architectures for each model were tested keeping track on the error in order to be able to find the best generalizing network in each experiment. We used *Neuro Shell* V2.0 to implement the neural network models. We are proposed the (a) DA - Assisted Neural Networks, (b) Logit - Assisted Neural Networks, and (c) GA - Assisted Neural Networks.

HYBRID MODELS FOR CASE BASED REASONING : Phase II
Case based reasoning(CBR) makes data more accessible by organizing it as a set of examples from past experience that can be generalized and applied to current problems. CBR provides a technique for capturing the experience in a large set of historical cases. This paper compared to the accuracy in bankruptcy prediction using case

based reasoning. This CBR approach based on discrete logic usually seek the single best instance, or a weighted combination of a small number of neighbors in the observation space. The key to this approach depends on the determination of the most effective set of weights to use in order to construct the neighbor[11]. The input data are the firm's financial ratios. The CBR approach selected DA and GA method so that we compare with neural network model. The in put data was randomly split into a training set, consisting of 750 companies, and a test set consisting of 250 companies.

In phase II, We are proposed the (a) DA-assisted Case based reasoning, and (b) Genetic-assisted Case based reasoning.

EMPIRICAL RESULTS

The Prediction accuracy by DA is 73.2 %. And the prediction result by logit is 74.4%.

DA-assisted neural network uses the 8 financial variables, selected by the stepwise method as input variables. The prediction accuracy is 85.26 %.

Logit-assisted neural network uses the 10 financial variables, selected by the stepwise method as input variables. The prediction accuracy is 78.29 %.

GA-assisted neural network uses the 9 financial variables, selected by the genetic algorithm as input variables. The prediction accuracy is 86.78 %. The results are summarized in Table 5. Therefore, this study presented hybrid neural network outperform the DA and logit. DA-assisted neural network and GA-assisted neural network shows increase in the prediction performance.

And, DA-assisted Case based reasoning uses the 8 financial variables, selected by the stepwise method as input variables. The prediction accuracy is 86.8 %. The results are summarized in Table 6.

Genetic-assisted case based reasoning uses the 9 financial variables, selected by the genetic algorithm as input variables. The prediction accuracy is 88.4 %. The results are summarized in Table 7.

GA-assisted case based reasoning perform the best. This model shows an outstanding prediction accuracy 88.4%. Therefore, this study presented hybrid case based reasoning better than the neural networks model.

CONCLUSION

This study has suggested the hybrid models which perform very well in the bankruptcy prediction tasks. This study proposed hybrid

model combine neural network model with other statistical or AI models(like DA-assisted or GA-assisted neural network models). DA-assisted or Logit-assisted neural network model uses DA or Logit models as a preprocessor for selecting the appropriate input variables which are to be used by the supervised neural network model.

This study experimented DA, Logit, and hybrid models using Korea medium-sized firm bankruptcy data. Experimental results showed that GA-assisted CBR performs the best. This is an exploratory study to propose and test hybrid models. Future studies are expected to present improved hybrid model and to apply monthly prediction.

REFERENCES
[1] Agarwal, A. (1993). Neural Networks and Their Extensions for Business Decision Making. Ph. D Dissertation, College of Business, The Ohio State University.
[2] Hecht-Nielsen, R. (1990). Neurocomputing. Reading, Massachusetts: Addison-Wesley.
[3] Back, B., T. Laitinen and K. Sere(1996), Neural network and Genetic algorithms for Bankruptcy Predictions", Proceedings of the third World congress on Expert System, 123-130.
[4] Lee, K., I. Han, and Y. Kwon(1995), Hybrid Neural network Models for Bankruptcy Predictions", Decision Support System.
[5] Jo, H., I. Han, and H. Lee (1995), Comparative Analysis of Bankruptcy Prediction Accuracy : Using Case-based forecasting, NN and DA," Proceedings of '95 Pacific-Asian Conference on Expert System, China.
[6] Francis, W., and D. Lee, (1993), Hybrid Neural network for Stock Selection," The second annual Informational Conference on Artificial Intelligence Application on Wall Street, NY USA.
[7] Deboeck D. J.(1994), Trading on the Edge, John Wiley & Sons, Inc.
[8] Apostolos-Paul Refenes(1995) *Neural Networks in the Capital Markets*, John Wily & Sons.
[9] Goldberg, D.E.(1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, MA.
[10] Odom, M.D. and R. Sharda(1989), "A Neural network model for Bankruptcy prediction," Proceedings of the IEEE International Conference on Neural Network, San Diego, CA., pp.163-168.
[11] Steven H. Kim and D.S. Kang(1996), "Implicit versus Explicit Learning for Forcasting: Case Study in Intraday Stock Index Prediction," KAIST Working Paper.

Table4 Neural Network Model
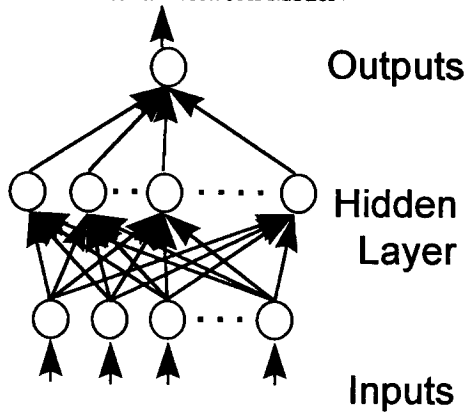
Table 1 Variables of DA Selection

Selected Financial Ratios	Wilk' λ
Ordinary Income to Total Assets	0.8556
Equity	0.7839
Growth Rate of Total Assets	0.7436
Fixed Assets Turnover	0.6947
Net Financial Expenses	0.6761
Ordinary income to Business capital	0.6619
Break-Even Point Ratio	0.6364
Total Asset	0.6309

Table 2 Variables of Logit Analysis Selection
- Operating Income to Financial Expenses
- Growth Rate of Total Assets
- Ordinary Income to Business Capital
- Financial Expenses to Sales
- Quick Ratio
- Current Ratio
- Cost of Goods Sold Ratio
- Receivables Turnover Ratio
- CF to Debt Ratio after Operating
- Lending Efficiency

Table 3 Variables of GA Selection
- Net Income to Stock Holder Equity
- Ordinary Income to Sales
- Quick Ratio
- Sales
- Equity
- Growth Rate of Total Assets
- Fixed Assets Turnover
- Ordinary income to Business capital
- CF to Debt Ratio after Operating

Table5 Prediction Accuracy of Neural networks Model

	DA	Logit	DA/NN	Logit/NN	GA/NN
Test	75.8	76.1	84.86	80.48	84.31
Validation	73.2	74.4	85.26	78.29	86.78

Table 6 DA-Assisted CBR

Nearest Number	Number of Correct Classification	Result(%)
2	217	86.8
5	211	84.4
7	202	80.8
10	199	79.6

Table 7 GA-Assisted CBR

Nearest Number	Number of Correct Classification	Result(%)
2	221	88.4
5	213	85.2
7	215	86.0
10	203	81.2