Neural Networks Versus Logit Regression Models For Predicting Financial Distress Response Variables

Jozef M. Zurada, (E-mail: jmzura01@homer.louisville.edu), University of Louisville
Benjamin P. Foster, (E-mail: bpfost01@homer.louisville.edu), University of Louisville
Terry J. Ward, (E-mail: tward@mtsu.edu), Middle Tennessee State University
Robert M. Barker, (E-mail: rmbark01@homer.louisville.edu), University of Louisville

Abstract

Neural networks are designed to detect complex relationships among variables better than traditional statistical methods. Our study examined whether the complexity of the response measure impacts whether logistic regression or a neural network produces the highest classification accuracy for financially distressed firms. We compared results obtained from the two methods for a four state response variable and a dichotomous response variable. Our results suggest that neural networks are not superior to logistic regression models for the traditional dichotomous response variable, but are superior for the more complex financial distress response variable.

Introduction

Our study examines the impact that the complexity of the financial distress response variable has on the relative predictive ability of logistic regression models and neural networks. We examine this impact by comparing the relative ability of neural networks and logistic models to distinguish between (1) a dichotomous response variable and (2) a multi-state response variable.

As recommended by Greenstein and Welsh (1996), we conduct our analysis on an unbalanced sample. We use Ward and Foster's (1996) unbalanced sample, but collapse their three distressed states into one distressed state to create a dichotomous response variable. Neural networks and logistic regression models produce comparable classification accuracy rates for the dichotomous response variable. In contrast, with the same sample, Zurada *et al.* (1997) found that their neural network better classified the firms into four states of distress than did Ward and Foster's (1996) logistic regression models. Comparing our results to Zurada *et al.*'s (1997) results provides evidence that neural networks provide higher predictive accuracy than logistic models for a more complex response variable. However, logistic models may perform as well or better than neural networks in predicting a dichotomous response variable.

The rest of this paper is organized as follows. The next section more fully discusses

Readers with comments or questions are encouraged to contact the authors via e-mail.

the motivation for the paper and reviews prior literature. We then describe our research methods and discuss the results of our analysis. The final section contains our conclusions.

Motivation and Literature Review

Some researchers believe neural networks can capture more complex relationships between and among variables than can traditional statistical analysis. Explaining a multi-state response variable should require more complex relationships among and between independent variables than would explaining a two-state response variable. Consequently, while logistic models may perform as well as neural networks in predicting a dichotomous financial distress response variable, neural network literature suggests that neural networks should better predict companies in a multi-state financial distress context than logistic models. To examine this issue we compared the classification accuracy rates for neural networks to those of logistic regression models, on the same sample data. Unique to this study, we compared the classification accuracy rates of the two techniques for both a dichotomous and multi-state financial distress response variable.

Comparison of logistic regression and neural networks for dichotomous financial distress response variables

Several studies compared the predictive ability of neural networks and traditional statistical models to distinguish between financially troubled companies and healthy companies. Some studies found neural networks were somewhat better at distinguishing between the firms. For example, Coats and Fant (1993) found that a neural network was better at distinguishing the companies than a multiple discriminant model while Lenard *et al.* (1995), Fletcher and Goss (1993), and Salchenberger *et al.* (1992) found a neural network superior to a logistic regression model. Limitations of these studies are that they conducted their analysis on a sample that matched one healthy firm with one distressed firm and that they use a dichotomous response variable.

Not all prior research supports the superiority of neural networks over traditional statistical techniques. For example, in a study of healthy and financially vulnerable or unsound Italian industrial firms, Altman *et al.* (1994) found that a neural network produced classification results comparable and sometimes superior to linear multiple discriminant analysis (LDA). However, they concluded that LDA produced better results overall because of concern about the neural network producing illogical weightings of some indicators and overfitting in the training stage.

Greenstein and Welsh (1996) compared the ability of a neural network and logistic analysis to predict bankrupt and healthy U.S. firms. They used a training sample in which 20% of the firms had declared bankruptcy and a testing data set that included less than 1% bankrupt firms. Greenstein and Welsh found that, overall, a logistic model outperformed the neural network in predicting the testing data set. Thus, they noted their results should generate caution when interpreting the results of prior studies that used matched-pair data sets.

Research with multi-state distress response variables

Some researchers have examined the usefulness of different accounting information in traditional multi-state statistical models. (For example, see Bahnson and Bartley (1992), Ward (1994), and Ward and Foster (1996)). Bahnson and Bartley (1992) cautioned that when a multi-state definition of financial distress was used rather than the traditional bankrupt vs. nonbankrupt definition, logistic regression models produce different research results. Ward and Foster (1996) used ordinal logistic models to analyze a four-state response variable from an unbalanced sample of healthy and distressed firms. (Over

70% of the firms included in their sample were healthy.) Zurada *et al.* (1997) performed a neural network analysis on Ward and Foster's (1996) to compare classification accuracy rates on the multi-state holdout sample for the neural network and logistic models. Zurada *et al.* concluded that neural networks generally produced better overall classification rates than logistic models.

Current study's goals

Prior literature suggests that the relative predictive ability of neural networks and logistic regression models can be impacted by sample proportions and the complexity of the response variable. We want to examine this issue. Thus, we compare the differences between neural network and logistic model classification rates for a dichotomous response variable to the differences between the classification rates from the same sample for a multi-state response variable.

Research Methods

Sample

We use Ward and Foster's (1996, p. 140) sample for the analyses reported in this study. The sample was derived by using *Wall Street Journal Index* and *Compact Disk Disclosure* information from 1988 and 1989 to identify companies that became bankrupt or defaulted on loan interest or principal payments or received a favorable debt accommodation. Also, a search of *Compustat* tapes identified companies that had, after paying a dividend for each of the three prior years, cut their dividend by at least 40%. Ward and Foster then randomly matched the distressed companies with healthy companies across the same industries. (See Ward and Foster (1996, p. 140) for a complete description of the sample and sampling technique employed.)

Nineteen eighty-eight firms were used as the training sample and 1989 firms as a holdout sample. The 1988 sample totals 204 companies of which 150 were healthy and 54 were distressed firms (16 cut dividends, 21 loan defaulted or received a favorable debt accommodation, and 17 went bankrupt). The 1989 sample serves as a test sample and includes 141 companies of which 103 were healthy and 38 were distressed (12 cut dividends, 13 loan defaulted or received a favorable debt accommodation, and 13 went bankrupt). Unlike many samples used in previous studies, distressed companies are not as numerous as the healthy companies included in the sample.

Variables and Time Period Investigated

Because we compare neural network and logistic regression model results for a dichotomous and multi-state response variable from their sample, we examine the same independent variables used by Ward and Foster (1996, 139-141). The independent variables are traditional accrual accounting ratios or traditional ratios adjusted for accounting allocations: (1) SALESCA = sales/current assets, (2) CACL = current assets/current liabilities, (3) OETL = owners' equity/total liabilities, (4) CATA = current assets/total assets, (5) CASHTA = cash plus marketable securities/total assets, (6) SIZE = log (total assets), (7) CFFF = estimated cash flow from financing activities/total liabilities, and (8) CFFI = estimated cash flow from investing activities/total liabilities.

Ward and Foster (1996) developed four different logistic models with these variables. The ninth variable varied depending on the model used. Each model included the independent variables described above and one of the following variables: (a) NITA = net income/ total assets, (b) DPDTADJ = depreciation and amortization and deferred tax allocations adjusted operating flow/total assets, (c) NQA FLOW = net-quick-assets operating flow/total assets, or (d) CFFO = estimated cash flow from operating activities/total assets. Ward and Foster (1996) focused on the impact of accounting allocations on predictive models. Consequently, for the models including DPDTADJ, NQA FLOW, and CFFO, they removed the impact of depreciation and deferred taxes from the scaling measures of total assets, total liabilities, and owners' equity. (See Ward and Foster, 1996, p. 141 for a more complete description of the variables.)

Ward and Foster (1996) and Zurada *et al.* (1997) examined an ordinally scaled response variable with companies coded as: 0 = healthy, 1 = dividend cut, 2 = loan default or favorable debt accommodation, or 3 = bankruptcy. To conduct our comparison of the relative predictive ability of neural networks and logistic models, we collapse the three distressed categories into one category to produce the following dichotomous response variable: 0 = healthy and 1 = distressed. Following Ward and Foster (1996), we analyze data from 1987, 1986, and 1985 for the 1988 companies. A neural network and logistic regression model are developed for each year's data. We then apply the results from these analyses to data from 1986, 1987, and 1988 to the holdout sample of 1989 companies. Like Ward and Foster (1996, 148), based on the output probabilities produced for the 1989 companies, we use the holdout sample proportions as the cutoff classification point when classifying firms.

Neural Network Model

In this study we use the same neural network structure as Zurada *et al.* (1997). This is one of the most popular neural networks used in financial applications, a two-layer feed-forward neural network with error back-propagation. Zurada *et al.*'s network architecture included four neurons in the output layer and the number of neurons in the hidden layer varied from 12 to 30. We used the same neural network architecture in this study, except the output layer included only two output neurons. Figure 1 provides a graphical representation of the neural network used by Zurada *et al.* (1997) while Figure 2 provides a graphical representation of our neural network. The figures illustrate the additional complexity caused by including four rather than two output neurons.

We train and test the neural network on the nine input variables for the years discussed above. Each neuron uses a sigmoidal unipolar activation function that produces a value within the range (0, 1). Before training begins, the network's weights are initialized at random with values within the range (-0.1, 0.1). Also, to prevent network's saturation, the nine variables are normalized to values from within the interval (-1, 1). During training, a nine-tuple (nine variables) that constitutes a training pattern is submitted to the network. The pattern flows through the network's layers and for the dichotomous response variable appears as a couple on the output. The network's output is confronted with the true response submitted by a teacher. In this case the true responses, (1, 0) and (0, 1) represent a healthy firm or a distressed firm, respectively. In Zurada *et al.*'s (1997) study, for the multi-state response variable, the pattern appears as a four-tuple on the output and the true responses, (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), and (0, 0, 0, 1) represent: a healthy firm, a firm that experienced dividend cuts, a firm that experienced a loan default, and a firm that filed for bankruptcy, respectively.

The differences between the network's output and the true output are passed back to the neural network to modify its weights. This process, called a training step, is repeated for all 204 training patterns (the 204 firms) in the training set. The training set is shown to the neural network 30,000 times. (This number is chosen arbitrarily and it guarantees that the maximum allowable error value will be reduced to a small value.) After training is finished, the network's performance is tested for 141 test patterns (the 141 holdout firms) which the network has not seen during training. The network's decision is determined by the larger of its two outputs for the dichotomous response variable and was determined by the largest of its

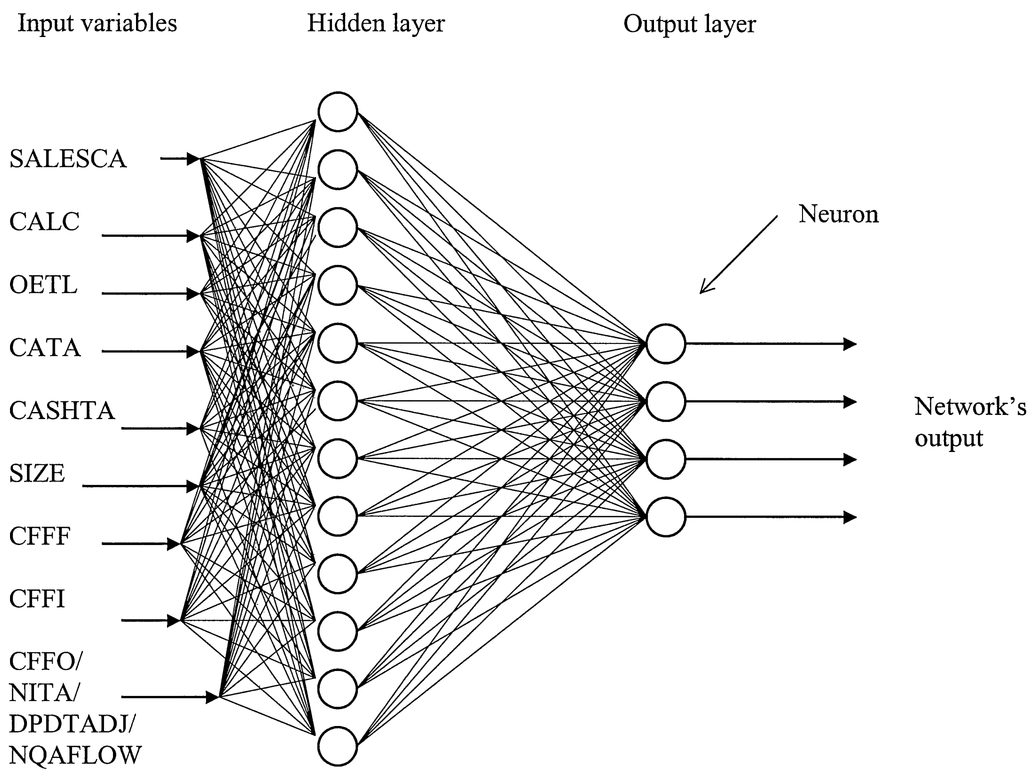

Figure 1. The neural network used for testing financial distress for four-state models.

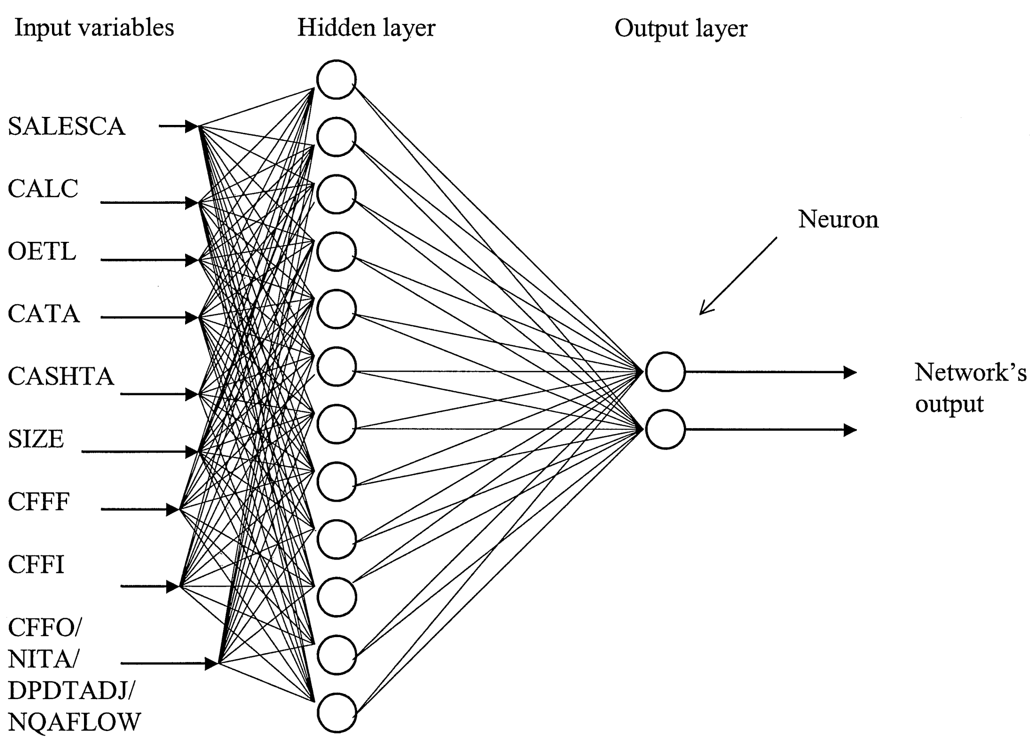

Figure 2. The neural network used for testing financial distress for two-state models.

four outputs for the multi-state response variable.

Comparing Logistic Predictions and Neural Network Predictions

The main objective of this study is to compare the relative classification accuracy rates of logistic models and neural networks for dichotomous and multi-state response variables. Thus, after obtaining classification rates on the holdout sample for the dichotomous response variable, we test for significant differences between the overall accuracy rates for the two methods. To test for significant differences between the two methods' classification ability we use the Z-test for differences between binomial proportions described by Bhattacharyya and Johnson (1977, p. 310). For the four-state response variable, we also test for significant differences between the overall classification accuracy rates for the logistic models and neural networks reported by Ward and Foster (1996) and Zurada *et al.* (1997), respectively.

Discussion of the Results

Table 1 reports the classification accuracy rates for the two-state logistic regression models and neural networks. The logistic models better classified the distressed firms, while the neural networks better classified the healthy firms. Overall, the logistic regression models correctly classify the companies slightly better than do the neural networks. We test for significant differences in the overall classification rates using a two-tailed binomial proportional Z-test because conflicting prior research with two-state response variables did not lead to any definite expectations that one method would properly classify more firms than the other.

The only significant difference between overall classification rates for the methods is for the analyses that included DPDTADJ two years prior to distress; the logistic model's overall classification rate is significantly higher (at p-value $< .10$) than the neural network. These results are comparable to prior research with a two-state financial distress response variable when the sample of distressed firms was not matched one for one with a healthy firm (Greenstein and Welsh 1996, Table 6).

Table 2 presents classification rates obtained for Ward and Foster's (1996, p. 148) four-state ordinal logistic regression models and Zurada *et al.*'s (1997) neural networks for the data. We test for significant differences between these classification accuracy rates. Because we expect neural networks to perform better than logistic models when the response variable is more complex, we use a one-tailed binomial proportional Z-test to test for significant differences. For all twelve models (four models for each of three years), the neural network's overall classification accuracy is higher than the logistic regression models' accuracy - significantly higher for nine of the twelve models.

Neural networks' superiority is due to a higher classification rate for the healthy companies (class 0) for all models in all years. However, comparing classification accuracy within classes 1 through 3 for the four models for the three years (36 classes), the neural network outperforms the regression analysis only 15 out of 36 times and the two methods classify firms equally well two times. No clear pattern of superiority by logistic models or the neural networks is evident among the three distress classes for any particular year prior to distress.

Conclusions

Proponents of neural networks claim that the networks can better capture and analyze complex relationships than traditional statistical analysis. A dichotomous financial distress response variable may lack the complexity necessary for a neural network to outperform logistic models when analyzing an unbalanced sample (a sample that more realistically reflects the true population). However, a multi-state financial

Table 1
Classification Rates for Dichotomous Response Variable by Logistic Regression Models and Neural Networks

Percentage of firms classified correctly

	Year-1		Year-2		Year-3	
Model	Regression	NN	Regression	NN	Regression	NN
NITA:						
All firms	81.56	79.43	76.60	76.60	68.09	64.54
State 0	82.52	87.38	80.58	90.29	67.96	75.73
State 1	78.95	57.89	65.79	39.47	68.42	34.21
DPDTADJ:						
All firms	83.69	81.58	81.56*	73.05	68.79	68.79
State 0	86.41	87.38	84.47	85.44	68.93	78.64
State 1	76.32	65.79	73.68	39.47	68.42	42.11
NQAFLOW:						
All firms	80.14	80.85	75.89	74.47	67.38	70.21
State 0	83.50	90.29	78.64	84.47	65.05	83.50
State 1	71.05	55.26	68.42	47.37	73.68	34.21
CFFO:						
All firms	80.85	80.85	77.31	73.76	68.09	66.67
State 0	84.47	92.23	77.67	87.38	67.96	84.47
State 1	71.05	50.00	76.32	36.84	68.42	18.42

Results of 2-tailed proportional z-tests that neural network and logistic regression models' overall classification rates not equal: *significant at $\leq$ .10. The models and neural networks were developed for a dichotomous response variable where 0 = healthy and 1 = distressed. Each model and neural network included the following variables: SALESCA = sales/current assets, CACL = current assets/current liabilities, OETL = owners' equity/total liabilities, CATA = current assets/total assets, CASHTA = cash plus marketable securities/total assets, SIZE = log (total assets), CFFF = estimated cash flow from financing activities/total liabilities, and CFFI = estimated cash flow from investing activities/total liabilities. The models as listed in the table included the above variables and either: NITA = net income/total assets, DPDTADJ = depreciation and amortization and deferred tax allocations adjusted operating flow/total assets, NQAFLOW = net-quick-assets operating flow/total assets, or CFFO = estimated cash flow from operating activities/total assets. For the analyses including DPDTADJ, NQAFLOW, and CFFO, the impact of depreciation and deferred taxes have been removed from the scaling measures of total assets, total liabilities, and owners' equity.

distress response variable may require a more complex relationship structure among the predictor variables than can be adequately modeled with logistic regression. Consequently, we anticipated that a neural network would more likely outperform a traditional logistic model in classifying companies into a four-state financial distress response than into a two-state financial distress response. Our findings provide some evidence of the superiority of neural networks in financial distress analysis with a more complex response variable.

Our analysis of a two-state distress variable reveals little difference in the overall classification rates by neural networks and logistic models. Our results and prior research results (Greenstein and Welsh 1996) question the superiority of neural networks over logistic regression models when analyzing a dichotomous financial distress response variable with more realistic sample proportions. In contrast, our

Table 2
Classification Rates for Multi-state Response Variable
by Ordinal Logistic Regression Models and Neural Networks

Percentage of firms classified correctly

Model	Year-1 Regression (Ward & Foster 1996)	Year-1 NN (Zurada *et al.* 1997)	Year-2 Regression (Ward & Foster 1996)	Year-2 NN (Zurada *et al.* 1997)	Year-3 Regression (Ward & Foster 1996)	Year-3 NN (Zurada *et al.* 1997)
NITA:						
All firms	63.31	74.47**	57.45	68.79**	53.19	62.41*
State 0	77.67	89.32	70.87	86.41	66.02	78.64
State 1	8.33	16.67	0.00	16.67	0.00	0.00
State 2	7.69	57.14	38.46	21.43	46.15	50.00
State 3	53.85	25.00	23.08	25.00	7.69	0.00
DPDTADJ:						
All firms	70.92	74.47	64.54	69.50	56.74	65.25*
State 0	80.58	90.29	74.76	85.44	67.96	83.50
State 1	41.67	16.67	33.33	0.00	8.33	0.00
State 2	61.54	50.00	53.85	28.57	46.15	42.86
State 3	30.77	25.00	23.08	50.00	15.38	0.00
NQAFLOW:						
All firms	69.50	75.18	61.70	71.63**	51.06	66.67***
State 0	79.61	93.20	72.82	85.44	64.08	84.47
State 1	66.67	16.67	0.00	33.33	0.00	8.33
State 2	46.15	42.86	84.61	21.43	38.46	28.57
State 3	15.38	16.67	7.79	50.00	7.69	16.67
CFFO:						
All firms	65.96	75.18**	60.28	70.21**	53.90	66.67**
State 0	80.58	94.17	72.62	87.32	67.96	83.50
State 1	33.33	16.67	25.00	8.33	8.33	8.33
State 2	23.08	35.71	30.77	28.57	30.77	42.86
State 3	23.08	16.67	23.08	33.33	7.69	8.33

Results of 1-tailed proportional z-tests that neural network overall classification rates were higher than the logistic regression models' overall classification rate: ***significant at $\leq .01$, **significant at $\leq .05$, and *significant at $\leq .10$. The response variable was an ordinal scale with companies coded as: 0 = healthy, 1 = dividend cut, 2 = loan default or favorable debt accommodation, or 3 = bankruptcy. See Table I for a description of the independent variables included in each analysis.

comparison of differences between the ability of ordinal logistic regression models and neural networks to classify companies in a multi-state distress response variable show that neural networks provide higher (significantly higher in nine of sixteen comparisons) overall classification accuracy rates than logistic models. However, neural networks' superiority mainly results from better classification of the healthy firms in the holdout sample.

Suggestions for Future Research

Neural networks do not provide statistics on individual variables included in the network. Thus, they are not well suited for testing the usefulness of specific information (such as a specific accounting ratio) to predict financial distress. In contrast, logistic regression output provides statistics on each variable included in the model; researchers can analyze these statistics to test the usefulness of specific information. Thus, for unbalanced samples of financially distressed and healthy firms, our results lead to these suggestions. When predictive accuracy is the most important goal, (a) a neural network analysis should be used for a multi-state financial distress response variable and (b) logistic regression and neural networks produce comparable results for a two-state financial distress response variable. When testing the usefulness of specific information is the most important goal, researchers should use logistic regression analysis. Future research studies should test whether these suggestions are valid for research with other samples of healthy and distressed firms and for other response variables.

References

1. Altman, E. I., Marco, G., and Varetto, F, "Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience)," *Journal of Banking and Finance,* Vol. 18, No. 3, pp. 505-529, 1994.
2. Bahnson, P. R. and Bartley, J. W, "The Sensitivity of Failure Prediction Models to AlternativeDefinitions of Failure," *Advances in Accounting,* 10, pp. 255-278, 1992.
3. Bhattacharyya, G. K. and R. A. Johnson, *Statistic Concepts and Methods*. New York: John Wiley & Sons, 1977.
4. Coats, P. K., and Fant, F. L, "Recognizing Financial Distress Patterns Using a Neural Network Tool," *Financial Management*, Vol. 22, September, pp. 142-150, 1993.
5. Fletcher, D., and Goss, E, "Forecasting with Neural Networks: An Application Using Bankruptcy Data," *Information & Management*, Vol. 24, No. 3, pp. 159-167, 1993.
6. Greenstein, M. M., and Welsh M. J., "Bankruptcy Prediction Using *Ex Ante* Neural Networks and Realistically Proportioned Testing Sets," Working Paper, Lehigh University, 1996.
7. Lenard, M. J., Alam, P., and Madey, G. R., "The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty Decision," *Decision Sciences*. Vol. 26, No. 2, pp. 209-227, 1995.
8. Salchenberger, L. M., Cinar, E. M., and Lash, N. A., "Neural Networks: A New Tool for Predicting Thrift Failures," *Decision Sciences*. Vol. 23, No. 4, pp. 899-916, 1992.
9. Ward, T. J., "An Empirical Study of the Incremental Predictive Ability of Beaver's Naive Operating Flow Measure Using Four-State Ordinal Models of Financial Distress," *Journal of Business Finance and Accounting*, June, pp. 547-561, 1994.
10. Ward, T. J., and Foster, B. P., "An Empirical Analysis of Thomas's Financial Accounting Allocation Fallacy Theory in a Financial Distress Context," *Accounting and Business Research*, Vol. 26, No. 2, pp. 137-152, 1996.
11. Zurada, J. M., B. P. Foster, T. J. Ward, and R. M. Barker. "A Comparison of the Ability of Neural Networks and Logit Regression Models to Predict Levels of Financial Distress," in *System Development Methods for the Next Century,* G. Wojtkowski and W. Wojtkowski (Eds.), Plenum Press, New York, NY, pp. 291-295, 1997.