Improving Prediction of Neural Networks: A Study of Two Financial Prediction Tasks

TARUN K. SEN   tksen@vt.edu

Department of Accounting and Information Systems, Pamplin College of Business, Virginia Tech, Falls Church, VA, 22043, USA

PARVIZ GHANDFOROUSH[†]   parviz@vt.edu

Department of Business Information Technology, Pamplin College of Business, Virginia Tech, Falls Church, VA 22043, USA

CHARLES T. STIVASON

Department of Accounting, Christopher Newport University,Newport News, VA 23606, USA

Abstract. Neural networks are excellent mapping tools for complex financial data. Their mapping capabilities however do not always result in good generalizability for financial prediction models. Increasing the number of nodes and hidden layers in a neural network model produces better mapping of the data since the number of parameters available to the model increases. This is detrimental to generalizability of the model since the model memorizes idiosyncratic patterns in the data. A neural network model can be expected to be more generalizable if the model architecture is made less complex by using fewer input nodes. In this study we simplify the neural network by eliminating input nodes that have the least contribution to the prediction of a desired outcome. We also provide a theoretical relationship of the sensitivity of output variables to the input variables under certain conditions. This research initiates an effort in identifying methods that would improve the generalizability of neural networks in financial prediction tasks by using mergers and bankruptcy models. The result indicates that incorporating more variables that appear relevant in a model does not necessarily improve prediction performance.

Keywords: Neural networks, Bankruptcy Prediction, Merger Prediction

1. Introduction

Models that use adaptive algorithms like neural networks are excellent tools for accurately mapping a data set, but have often been criticized for their complexity resulting in poor generalizability in financial prediction tasks [21]. Researchers have had difficulty in sifting through the maze of weights

---

that build a neural network model in order to interpret it. However, several researchers have used neural networks to successfully predict financial phenomena, such as, stock market trends [31] and [29], firm bankruptcy [20, 23, 28], gold prices [8, 9], and foreign exchange prices [19]. Wider acceptance of neural networks in the financial world can be promoted by adopting simpler network architectures that have more generalizability and provide some explanatory facilities. This research evaluates the hypothesis that generalizability of neural networks can be increased in financial prediction tasks by simplifying neural networks using techniques that induce knowledge from the task domain. These techniques reduce the search space for neural networks enabling them to ignore noise that is inherent in financial data. The noise in financial data often makes it difficult to obtain a generalizable model.

The financial worlds of the stock market, corporate mergers, and firm bankruptcy are complex. They are difficult to understand and more difficult to predict. There is growing emphasis in the economics literature to treat this world as inherently complex, and not simplify it using mathematical postulations [4, 30]. Neural network training and other adaptive algorithms like genetic algorithms are being considered as potential tools in explaining and predicting such market phenomena. Most research effort in this direction attempts to combine several methods to provide better explanation capabilities [24, 7].

An interesting question is why use neural networks for financial models instead of statistical models or symbolic rule-based models? Shavlik [26] argues that neural networks have at least comparable accuracies to symbolic learning algorithms. Neural networks have been found to be successful in several tasks of inherent complexity like speech recognition, character recognition, language learning, etc. (see [26] for several applications). The advantage of neural networks over symbolic learning algorithms is that they use a more natural learning representation, much like the human brain. Thus, they can arrive at more precise models albeit complex [25]. However, like neurological interactions in the human brain, these models are hard to interpret. To simplify human decision-making, decision theorists propose different decision-making models. Combining symbolic approaches with neural networks has precisely the same objective that decision theorists have for their models, i.e., to provide a heuristic that explains complex connection mechanisms in a neural network model.

In this paper we use neural networks to predict two financial tasks. The two applications are prediction of targets for corporate mergers and prediction of firms that face impending bankruptcy. Both applications have significant financial implications. Being able to predict firms that are po-

tential targets for acquisition could result in financial gains since there is
typically an increase in the stock price of targeted firms. Targets are consid-
ered as good investments by firms that acquire them. Bankruptcy predic-
tion enables auditors to make going-concern judgments, banks to predict
potential loan defaulters, and stock analysts to identify poor investment
opportunities.

Prediction of targets for mergers is a challenging task. Prior researchers
have tried to predict mergers using logistic regression. Their results are
reflective of the difficulties associated with predicting targets. Although
Simkowitz and Monroe[27] and Dietrich and Sorensen [6] show that their
models are capable of predicting targets and non-targets (firms that are
not taken over) with a degree of accuracy of about 63

Prediction of bankruptcy is critical to the financial auditor. If a firm
goes bankrupt, the independent accounting firm that performed the audit
prior to bankruptcy is often held liable for damages to investors from the
bankruptcy. A number of studies have tried to develop bankruptcy mod-
els for auditors. Altman and McGough [1] compared the accuracy of a
bankruptcy prediction model to auditors' going-concern opinions one year
prior to the bankruptcy event. For the years 1970 to 1973, their model
correctly predicted bankruptcy for 82 percent of the cases while auditors
issued the going-concern opinion in only 46 percent of the cases. Their sam-
ple, however, consisted only of bankrupt firms. Other authors have found
similar results [2, 15, 14]. Tam and Kiang [28] showed that neural networks
are as good a predictor as logistic regression when predicting bank failures.

2. Complexity Reducing Mechanisms

Complexity in neural network architectures increases with the number of
input nodes, the hidden nodes, and the connections between them. In-
creased complexity provides better mapping of the training data as the
number of parameters (weights in the case of neural networks) increases
[11]. This, however, reduces generalizability as the network memorizes the
idiosyncrasies in training data. The optimal neural network configuration
attempts to balance network complexity with generalizability. In this study
we will attempt to show that if network complexity is reduced using tech-
niques which reduce the search space, generalization and thus, prediction,
can be improved.

A neural network model is composed of a set of input nodes U_i and a set
of output nodes Y_j. U_i are predictor variables and the Y_j are the desired
output. The output variables Y_j are mapped to the independent variables
U_i using intermediate variables V_k and X_i in the hidden layers (Figure 1).
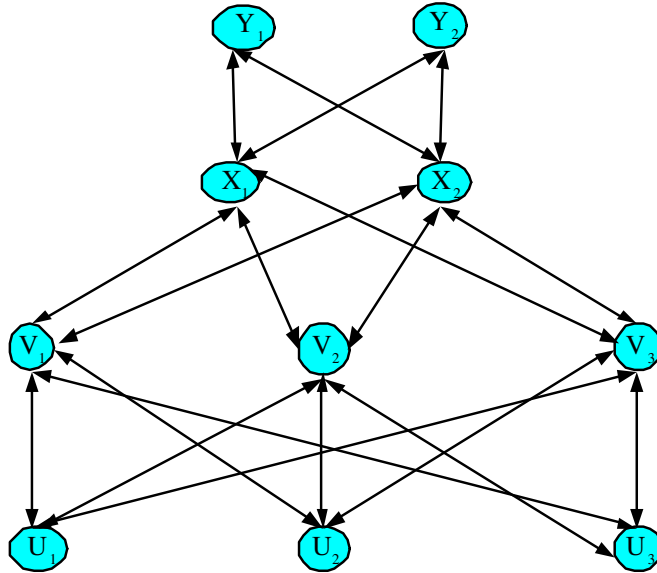
Figure 1. Neural Network Model.

An intermediate node variable V_k is computed using a linearly weighted regression like model and a sigmoidal transformation function, where

$$V_k = \frac{1}{1 + e^{\left\{-\sum_i W_i U_i\right\}}} \qquad (1)$$

W_i represents the weight of the input U_i mapped onto intermediate variable V_k. V_k becomes the input layer variable for the next intermediate or hidden layer, and the next layer input variables X_i are computed using a formula similar to (1). The process is repeated until the Y_j are estimated.

The total error of the estimated outputs can be computed using Rumelhart, et. al. [22] as

$$E = \frac{1}{2} \sum_j \left(y_i - y_i'\right)^2$$

where y_i' is the estimated output and y_i is the desired output. Based on Rumelhart's back propagation algorithm, the weights can be changed so that convergence in the output estimates occurs. This can be accomplished by minimizing the error function E with respect to the weights.

This leads to a weight change $\Delta W = - \in \sum \frac{\partial E}{\partial W}$, where $\in$ is called a learning coefficient.

These weight changes are then back propagated to the input layer. This process is repeated until convergence of the output variables occur.

Although the back propagation algorithm works very well for mapping a given input-output data set, it does not do a very good job of predicting financial phenomena. This is partly due to having unnecessary nodes, resulting in additional parameters to be estimated. In section 5.2 we demonstrate that by removing some of the input variables, prediction performance does not deteriorate. It is more likely that performance will improve.

2.1. A heuristic for node reduction

To reduce complexity, we develop a heuristic by which input nodes can be eliminated. Consider the following neural network model, where nodes are marked 1,2,3,4,5,6 for convenience.
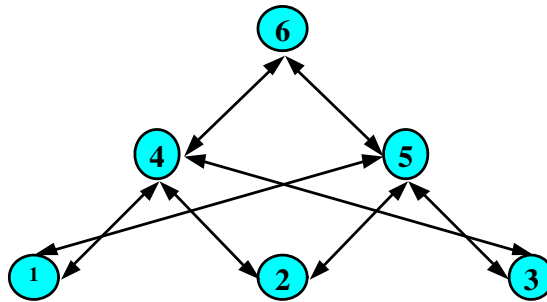

Figure 2. Neural Network Model with One Dependent Variable.

X_1, X_2, X_3 are the input variables and Y_6 is the output variable (subscripts refer to the respective node numbers). Linear weighted regression functions for the intermediate variables X_4 and X_5 may be written as:

$$X_4 = W_{14}X_1 + W_{24}X_2 + W_{34}X_3 \tag{2}$$

and

$$X_5 = W_{15}X_1 + W_{25}X_2 + W_{35}X_3 \tag{3}$$

The transformed variables Y_4, Y_5 for the intermediate layer are:

$$Y_4 = \frac{1}{1 + e^{-X_4}} \; and \; Y_5 = \frac{1}{1 + e^{-X_5}} \tag{4}$$

Finally,

$$X_6 = W_{46}Y_4 + W_{56}Y_5 \tag{5}$$

and

$$Y_6 = \frac{1}{1 + e^{-X_6}} \tag{6}$$

These completely specify all node variables in the neural network model. To find the impact of each input variable X_1, X_2, X_3, on the output Y_6, we find $\frac{\partial X_6}{\partial X_1}$.

From (2), we know that

$$\frac{\partial X_4}{\partial X_1} = W_{14} \; and \; \frac{\partial X_5}{\partial X_1} = W_{15} \tag{7}$$

Using the chain rule for derivatives,

$$\frac{\partial Y_4}{\partial X_1} = \frac{\partial Y_4}{\partial X_4} * \frac{\partial X_4}{\partial X_1} = W_{14}[\frac{\partial Y_4}{\partial X_4}] \tag{8}$$

and

$$\frac{\partial Y_5}{\partial X_1} = \frac{\partial Y_5}{\partial X_5} * \frac{\partial X_5}{\partial X_1} = W_{15}[\frac{\partial Y_5}{\partial X_5}] \tag{9}$$

We take the derivative for (5) with respect to X_1:

$$\frac{\partial X_6}{\partial X_1} = W_{46} * \frac{\partial Y_4}{\partial X_1} + W_{56} * \frac{\partial Y_5}{\partial X_1} \tag{10}$$

Substituting (8) and (9) in (10) we obtain

$$\frac{\partial X_6}{\partial X_1} = W_{46} * W_{14} * [\frac{\partial Y_4}{\partial X_4}] + W_{56} * W_{15} * [\frac{\partial Y_5}{\partial X_5}] \qquad (11)$$

We can similarly obtain

$$\frac{\partial X_6}{\partial X_2} = W_{46} * W_{24} * [\frac{\partial Y_4}{\partial X_4}] + W_{56} * W_{25} * [\frac{\partial Y_5}{\partial X_5}] \qquad (12)$$

Let us define $K_1 = W_{46} * [\frac{\partial Y_4}{\partial X_4}]$ and $K_2 = W_{56} * [\frac{\partial Y_5}{\partial X_5}]$ We can rewrite (12) in terms of weights K_1 and K_2 as:

$$\frac{\partial X_6}{\partial X_1} = W_{14} * K_1 + W_{15} * K_2 \qquad (13)$$

and

$$\frac{\partial X_6}{\partial X_2} = W_{24} * K_1 + W_{25} * K_2$$

From (11) and (12) we see that the sensitivity of the output variable Y_6 to X_1, and X_2 differ only on the input layer weights. If K_1 and K_2 were known, then the impact of X_1 and X_2 on the output variable Y_6 can be derived from (13). However, if we make some assumptions on K_1 and K_2, knowing the input layer weights, we can determine the impact of the input layer variables on the output. A simple heuristic that can be followed for eliminating input variables is to sum the input layer weights and eliminate the variables with the lowest weights. We test the impact of this heuristic on prediction performance.

3.   The Neural Network Models

Two neural network models are developed. One is a complete model with all the input nodes. The second is a reduced model, where input variables are eliminated by using the sum of the input layer weights described in the previous section.

3.1.   The fully connected neural network models (model 1)

The first model is a fully connected neural network that was developed with a single hidden layer (with $2n + 1$ nodes where n is the number of input nodes) for both the mergers and bankruptcy data sets. The number of nodes in the hidden layer is based on Kolmogrovs theorem [11], which

suggests that there exists a continuous mapping function f for a neural network with one hidden layer having $2n + 1$ nodes. This model included all the input variables for each of the data sets as in any traditional neural network model. The transfer function used in the networks was the sigmoidal function: $\frac{1}{1+e^{-x}}$. Both networks (for the two financial data sets) consisted of twelve input nodes, one hidden layer of 25 nodes and two output nodes. No momentum was used, and a learning constant of .2 was assumed.

To control for overfitting, the neural network model's training is stopped based on prediction performance on a test sample. Training is stopped periodically to observe the prediction performance using the test sample. Training is continued until prediction performance keeps on improving on the test sample. At the point where prediction performance starts to deteriorate, training is stopped. This method for controlling overfitting of neural network models is explained in [11].

3.2.   Reduced neural network based on weight analysis (model 2)

The weights in a neural network topology can be interpreted to a limited extent. It was shown in section 2 that the partial derivative of an output variable (at the final layer of a neural network model) with respect to an input variable (at the first layer) is a function of the sum of the product of weights (connected to the input node in the first hidden layer) and a function K which is independent of any parameters in the first hidden layer. This function K is the same for all partial derivatives between an output layer node and an input node. This implies that the sum of the weights connected to an input node in the first layer determines the impact on the output variable, since the function K is common to all input nodes. This assumes that all inputs are standardized and positive. The inputs to our models were standardized on a 0 to 1 scale.

After the networks are trained, the absolute values of the weights for each input node at the first layer are averaged. A high value indicates that this input variable is more important in distinguishing targets and failed firms from non-targets and non-failed firms. Four variables that have the strongest impact on the outcome in each model are identified. We eliminated all other input parameters from the networks. This simplified the network making it easier to be interpreted. It is expected that the mapping accuracy of the network on the training set will be reduced by eliminating some parameters. This may not be a problem since a high mapping accuracy may suggest overfitting of the model [11]. We expect that this reduced model will either generalize better or at least not do any worse than model 1.

Table 1. Distribution of Sample Sizes.

Task	Type of Firm	Training Sample Size	Holdout Sample Size
Merger Prediction	Targets	37	10
	Non-Targets	37	140
Bankruptcy Prediction	Bankrupt Firms	80	49
	Non-Bankrupt Firms	80	80

4. Data and Variables

The merger data were collected for all firms that were merger targets (and non-targets) in 1984-85 from Compustat and CRSP tapes. There were 37 target firms and 1260 non-target-firms that were identified for 1984 that had complete data. This unequal distribution of target and non-target firms gives rise to a state-based sampling problem [18]. To avoid a bias in the training sample, the training data set was comprised of an equal number of target and non-target firms (37 each) obtained from the 1984 data set. If the whole sample was used as the training data set, then the model could simply predict all firms as non-targets and obtain excellent prediction performance. Thus, learning would be impaired. The hold out sample was randomly selected from all firms in the 1985 Compustat and CRSP tapes. The hold out sample was comprised of 10 target firms and 140 non-target firms. This ratio is similar to the population percentages of targets and non-targets for 1984 and 1985. The distribution of the sample data is shown in Table 1.

The independent variables selected for predicting mergers were based on three factors:

1. The inefficient management hypothesis for takeovers [10] suggests that firms that are poorly managed are likely targets for takeover. Evidence of management performance can be obtained by (1) average excess stock return (AER), (2) return on equity (ROE), and (3) average sales to assets turnover (AVTURNOV).

2. The undervaluation hypothesis implies that a firms stock price does not reflect the true economic value of the firm. The firm is considered a good investment opportunity. Some firms have high growth and low resources and vice versa. These firms are identified as good investment opportunities. This can be proxied by a dummy variable (GRDUMMY) that indicates the growth-resource mismatch (Palepu 1986). Other indicators of good investment opportunities are average dividend payout

(AVPAYOUT), market to book ratio (MTB), and price earnings (PE) ratio [18, 6, 5].

3. Demographic and environmental factors [6, 5] may also be indicators of potential targets. A dummy variable (IDUMMY) was included to indicate whether a firm in the same industry category was merged in the year prior to the observation year. Total net book value of assets as an indicator of the size (SIZE) of the company was also included as an independent variable.

The data were obtained from Computstat and CRSP tapes. For simplicity most of the data items were averaged over three to five years prior to the two observation years 1984-85. Targets were identified by locating delistments of firms from the New York Stock Exchange. This information was obtained from CRSP tapes and later verified by the *The Wall Street Journal Index.*

Bankruptcy data were collected for "failed" firms that have declared Chapter 11 or Chapter 7 bankruptcy or were involuntarily liquidated during the years 1968-1990, with all data available on COMPUSTAT. A randomly selected sample of 160 "non-failed" firms, along with 129 "failed" firms, from the same period was used for training and testing purposes. The training set had 80 failed and non-failed firms each. The holdout sample had 80 non-failed firms and 49 failed firms. The firms were listed on either the NYSE, AMEX or NASDAQ stock exchanges. The accounting data for one year prior to the failure event were obtained using the COMPUSTAT tapes. A distribution of the data is shown in Table 1.

Prior research in the bankruptcy literature has focused on the distinctions between "stressed" and "non-stressed" firms [3, 17, 16, 12, 13] show that in their holdout period the actual bankruptcy proportion was approximately .0016 for non-stressed firms and .01406 for stressed firms. They had a total of 134 bankrupt firms of which 118 were stressed and only 16 were non-stressed. A stressed firm is one which satisfies one of the following conditions: (1) negative working capital in the current year, (2) a loss from operations in any of the three years prior to bankruptcy, and (3) a retained earnings deficit three years prior to bankruptcy.

Hopwood, et. al. [12] question the validity of prior studies because of the lack of control for stressed firms. For this reason 80 of the "non-failed" firms are from each classification "stressed" and "non-stressed", which follows the approach used by [12]. Stress was included in the model as a binary variable (STRESS).

Zmijewski [32] identified three important dimensions of financial bankruptcy prediction, return on assets, leverage, and liquidity. The independent

variables for the bankruptcy model were selected based on these dimensions. Leverage can be measured by long term liabilities / total assets (LTLTA). Liquidity measures that were used in this study were current assets / total assets (CATA), current assets / current liabilities (CR), and cash / total assets (CASHTA). Based on Ohlsons (1980) study another independent variable that was included was the natural logarithm of firm size (LSIZE). A firms asset management capability is a strong indicator of its ability to continue in existence. The ratios current assets / sales (CASALES), inventory / sales (INVSALE), sales / net plant (SALESSPPE), and accounts receivable / inventory (ARINV) were used as indicators of a firm's asset management capability.

5. Results and Analysis

Neural networks were built using the 2 models described earlier. To compare the results of these neural network models, a logistic regression model was built using the same data sets. The logistic regression prediction results are used for comparison only. The prediction performances of the two models and the logistic regression model are summarized in Table 2. A non-parametric chi-square test of homogeneity (using SAS) was run to see if there was a statistical difference in prediction between the logistic regression model and the individual neural network models. The p-values from this test are shown for the holdout sample only in Table 2. The following section describes the results for each of the two models.

5.1. Neural network model (model 1)

Recall that model 1 is a fully connected neural network model with one hidden layer with 2n+1 nodes. For both the bankruptcy prediction task and the merger prediction task, there were no significant differences between model 1s predictions and the logistic regression models predictions. For the bankruptcy data the logistic regression model predicted 40 out of 49 bankrupt firms correctly and 58 out of 80 non-bankrupt firms correctly. The neural network (NN) model 1 predicted 39 out of 49 bankrupt firms correctly and 62 out of 80 non-bankrupt firms correctly. The overall prediction effectiveness for each model was high.

The logistic regression model for the merger data predicted 2 out of 10 target firms correctly and 75 out of 140 non-target firms correctly. The NN model 1 predicted 4 out of 10 targets correctly and only 68 out of 140 non-targets correctly. Both models had very poor prediction performance.

Table 2. Prediction performances for models 1, 2 and logistic regression.

			Logistic Regression	Neural Networks	
				Model 1 (Fully Connected)	Model 2 (Pruned using weights)
Mergers	Training	Targets	27/37	32/37	18/37
		Non-Tragets	24/37	35/37	27/37
	Holdout	Targets	2/10	4/10 (.329)#	3/10 (.606)
		Non-Tragets	75/140	68/140 (.403)	77/140 (.81)
Bankruptcy	Training	Bankrupt	66/80	66/80	66/80
		Non-Bankrupt	56/80	76/80	55/80
	Holdout	Bankrupt	40/49	39/49 (.798)#	40/49 (1.0)
		Non-Bankrupt	58/80	62/80 (.465)	57/80 (.860)

The values in the cells indicate the proportion of firms predicted correctly. # The figures in parentheses for the hold out sample are the p values for a chi-square for the hold out sample are the p values for a chi-square test of difference in prediction between the logistic regression model and the neural network (NN) models.

This is consistent with past research, where prediction of mergers, due to lack of availability of data, has been extremely difficult while bankruptcy prediction has been comparatively easier.

5.2.   Reduced neural network using weight analysis (model 2)

Model 1 was pruned to arrive at a reduced model. The pruning was done by observing the first layer weights of the fully connected network (model 1). The nodes with the highest average weights in the first layer were retained in the reduced model. All other nodes were eliminated to obtain better generalization of the neural network model.

The four highest average weights for input nodes were for AVPAYOUT, AVGROWTH, PE, and AER. Thus model 2 for mergers had only these four variables. We attempted to use more than 4 variables in the model in terms of the average weights. These models did not perform very differently. For the bankruptcy model, the highest average weights for input nodes were for CR, STRESS, CASALES, and LTLTA. These variables were then used in model 2 for the bankruptcy data.

Model 2 also did not perform significantly better than the logistic regression in predicting mergers or bankruptcy. However, contrary to popular

expectation, these results indicate that incorporating more variables that appear relevant in a model does not necessarily improve prediction performance. Future research could attempt to refine this method to obtain better generalizability.

6. Conclusion

In this study, neural network models were simplified to improve generalizability by preprocessing methods used to identify those variables that were most likely to improve prediction, and remove those that did not seem to have much effect. Although, by eliminating input variables prediction performance for firm mergers and bankruptcies did not improve significantly, there is evidence to indicate that simplification does not deteriorate performance

The results also indicate that removing input variables in neural networks by using the weights (Model 2) does not improve generalizability in the two tasks considered here. It would appear that further simplification of the data using binary variables and decision trees may provide better generalizability as well as explanatory capability.

This research points to a direction aimed at simplifying data sets with knowledge based techniques before applying neural networks to financial data. This appears to have an impact on generalizability of financial models. Future research should be directed towards further validating this approach by identifying techniques that improve the generalizability of neural networks using knowledge-based techniques. One such method would be genetic algorithms.

This research also indicates that the first layer weights of a neural network model have a significant impact on the strength of the relationship between the output variables and the independent variables. The nature of that relationship is subject to further research. This research shows that by using simple heuristics based on first layer weights prediction performance can potentially be improved.

References

1. E. I. Altman and T. P. McGough. Evaluation of a Company as a Going-Concern. *Journal of Accountancy*, pp. 50-57, 1974.
2. E. I. Altman. Accounting Implications of Failure Prediction Models. *Journal of Accounting, Auditing and Finance*, pp. 4-19, 1982.
3. J. Argenti. Corporate Collapse: The Cause and Symptoms. *John Wiley & Sons*, New York, NY, 1976.

4. B. Arthur. Designing Economic Agents that Act Like Human Agents: A Behavioral Approach to Bounded Rationality. *The American Economic Review*, pp. 353-359, 1991.

5. P. A. Brous, and O. Kini. A Reexamination of Analyst's Earnings Forecasts for Takeover Targets. *Journal of Financial Economics*, v. 33, p. 201-225, 1993.

6. J. K. Dietrich, and E. Sorensen. An Application of Logit Analysis to Prediction of Merger Targets. *Journal of Business Research*, v. 12, pp. 393-402, 1984.

7. R. C. Eberhart. The Role of Genetic Algorithms in Neural Network Query Based Learning and Explanation Facilities. *Proceedings of the International Workshop on Combinations of Genetic Algorithms and Neural Networks*, IEEE computer Society Press, Los Alamitos, CA, pp 169-183, 1992.

8. G. Grudnitski. Important Factors in Neural Networks-Forecasts of Gold Future Prices. *First International Workshop on Neural Networks in the Capital Markets*, London Business School, 1993.

9. G. Grudnitski and L. Osburn. Forecasting S&P and Gold Future Prices: An Application of Neural Networks. *Journal of Futures Markets*, Vol. 13, 6, pp. 631-643, 1993.

10. P. M. Healy, K. G. Palepu, and R. S. Ruback. Does Corporate Performance Improve After Mergers? *Journal of Financial Economics*, Vol. 31, pp. 133-175, 1992.

11. Nielsen R. Hecht. *Neurocomputing*. Addison-Wesley Publishing Company, Reading, MA, 1990.

12. W. Hopwood, J. McKeown and J. Mutchler. A Reexamination of Auditor Versus Model Accuracy Within the Context of the Going Concern Opinion Decision. *Contemporary Accounting Research*, pp. 409-431, 1994.

13. G. D. Kane, F. M. Richardson, and P. Graybeal. Recession-Induced Stress and the Prediction of Corporate Failure. *Working Paper*, University of Delaware, 1994.

14. H. C. Koh and L. N. Killough. The Use of Multiple Discriminant Analysis in the Assessment of the Going-Concern Status of an Audit Client. *Journal of Business, Finance and Accounting*, pp. 179-192, 1990.

15. A. S Levitan and J. A. Knoblett. Indicators of Exceptions to the Going Concern Assumption. *Auditing: A Journal of Practice and Theory*, pp. 26-39, 1985.

16. K. Menon, and K. Schwartz. An Empirical Investigation of Audit Qualification Decisions in the Presence of Going Concern Uncertainties. *Contemporary Accounting Research*, v. 4, pp. 302-315, 1987.

17. J. Mutchler. A Multivariate Analysis of the Auditor's Going-Concern Opinion Decision. *Journal of Accounting Research*, v. 23, pp. 668- 682, 1985.

18. K. G. Palepu. Predicting Takeover Targets: A Methodological and Empirical Analysis. *Journal of Accounting and Economics*, v. 8, pp. 3-35, 1986.

19. H. Pi. Dependency Analysis and Neural Network Modeling of Currency Exchange Rates. *First International Workshop on Neural Networks in the Capital Markets*, London Business School, 1993.

20. W. Raghupati, L. S. Schkade, and B. S. Raju. A Neural Network Application for Bankruptcy Prediction. *Proceedings of the 24th Annual International Conference on Systems Sciences*, pp. 147-155, 1991.

21. A. Refenes. *Neural Networks in the Capital Markets*. John Wiley and Sons, Chichester, England, 1995.

22. D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning Representation by Back-Propagating Error. *Nature*, Vol. 323, pp. 533-536, 1986.

23. L. Salchenberger, M. Cinar, and N. Lash. Predicting Savings and Loans Failures Using Neural Networks. *Decision Sciences*, Vol. 23, 4, pp. 899-916, 1986.

24. J. D. Schaffer, D. Whitley, and L. J. Eshelman. Combinations of Genetic Algo-
rithms and Neural Networks: A survey of the state of the art. *Proceedings of
the International Workshop on Combinations of Genetic Algorithms and Neural
Networks*, IEEE Computer Society Press, Los Alamitos, CA, pp. 1-37, 1992.

25. S. Sestito and T. Dillon. Knowledge Acquisition of Conjunctive Rules Using Multi-
layered Neural Networks. *International Journal of Intelligent Systems*, Vol. 8, pp.
779-805, 1993.

26. J. W. Shavlik. Combining Symbolic and Neural Learning. *Machine Learning*,Vol.
14, pp. 321-331, 1994.

27. M. Simkowitz and R. J. Monroe. A Discriminant Analysis Function for Conglom-
erate Targets. *Southern Journal of Business*, 6, 3-16, 1971.

28. K. Y. Tam and M. Y. Kiang. Managerial Applications of Neural Networks: The
case of bank failure predictions. *Management Science*, pp. 926-947, 1992, .

29. Fidelity's Bradford Lewis takes aim at indexes with his neural network computer
program. *The Wall Street Journal 1992*, October 27, 1992.

30. M. M. Waldrop. Complexity. *Simon and Schuster*, New York, NY, 1992.

31. Y. Yoon and G. Swales. Predicting Stock Price Performance: A Neural Network
Approach. *Proceedings of the 24th Annual International Conference on Systems
Sciences*, pp. 156-162, 1991 .

32. M. E. Zmijewski. The Prediction of Corporate Failure: The State of the Art.
Journal of Accounting Research, (Supplement), pp. 59-82, 1984.