Thai Journal of Development Administration Vol.36, No.3, July-September 1996

A Use of Untransformed Financial Accounts in a Neural Network Model for Predicting Bankruptcy

*Aekkachai Nittayagasetwat**

The significance of this study is to develop a bankruptcy-prediction model by using artificial neural network (ANN) technology and untransformed financial accounts as predictors. An artificial neural network (ANN) is the newest approach to studies of bankruptcy or financial distress predictions. The previous studies of bankruptcy prediction models, including the ANN technology-based models, usually employed financial ratios obtained from balance sheets and income statements as explanatory variables in distinguishing between bankrupt and nonbankrupt firms. This initiates the question whether transformations of data really matter in an ANN model

The use of untransformed financial accounts is not common for studies in finance because untransformed financial accounts are not standardized by firm size. For most econometric models, transformations of data may be necessary because untransformed accounts are vulnerable to multicollinearity and are likely to violate a priori assumptions of an underlying methodology (e.g., the Classical Linear Regression Model). The violation of those assumptions makes biased, inefficient, and/or inconsistent estimators. Also, proper transformations of data are required to prevent the possibility of spurious regressions. (Granger and Newbold, 1974). The spurious regressions reflect the existence of autocorrelated errors in regression analysis. Therefore, transformations of data, such as the natural logarithm of accounts, the ratio constructed from two or more items of accounts, or the first difference of independent variables (as suggested by Granger and Newbold) are definately required for econometric models.

However, the need for data transformation should not be the case for an ANN model. Hornik, Stinchcombe, and White (1989) proved that an ANN model has the property of a universal approximator; an ANN model is capable of approximating any measurable function. Based on the fact that transformed accounts are derived from untransformed

*Lecturer, Graduate School of Business Administration, National Institute of Development Administration.

accounts, both untransformed and transformed accounts initially use the same information. ANN can probably approximate the function based on the original information. Untransformed accounts should provide ANN with the freedom of approximation without restrictions generated by the assumptions of transformation techniques.

Neither does ANN require the unjustified a priori assumptions of traditional econometrics, nor should it require the unjustified restrictions of input transformations. Therefore, an ANN model should be adjustable without the necessity of transforming the input. Additionally, the performance of an ANN model using untransformed accounts should be as good as that of an ANN model using transformed accounts. The use of untransformed accounts as predictors for bankruptcy-prediction models will simplify research by eliminating the search for proper transformations.

The results of the ANN model are compared with those of the Logit model. The sign tests are used for statistical comparisons between Logit and ANN since both are probability models and can be compared observation by observation. (Note that the sign test is nonparametric and distribution-free, so it is quite suitable for ANN in which the distribution is unknown.)

The paper continues in five sections. Section 1 presents a review of the literature on predicting bankruptcy, the primary contribution and null hypotheses of this study. Section 2 describes the data set, data sources, and steps in data collection. Section 3 describes the methodology. Section 4 presents the empirical results. Finally, Section 5 summarizes the results and significance of this study.

1. Literature, Primary Contribution, and Null Hypotheses

The use of untransformed financial accounts is not common in the finance literature. There is no previous literature referring to the use of untransformed financial accounts as predictors in bankruptcy-prediction models. As later discussed, studies in finance consistently use financial ratios in the development of bankruptcy-prediction models.

The development of bankruptcy-prediction models started from univariate analysis with Beaver (1996), but the best-known technique became the multivariate discriminant analysis (MDA) of Altman (1968). MDA dominated the studies of bankrupty-prediction models for two decades after it was developed. The studies which used MDA as the tool for bankruptcy prediction included Edmister (1971), Deakin (1972), Sinkey (1975), and Altman, Haldeman and Narayanan (1977). The studies of bankruptcy prediction did not stop at MDA. Logit (e.g., Ohlson 1980, Zavgren 1982), Probit, and the recursive partitioning algorithm (RPA) of Frydman, Altman and Kao (1985) were developed to supercede MDA.

An artificial neural network (ANN) is the newest approach in predicting bankruptcy and has gotten great attention from many researchers. Bell, Ribar and Verchio (1990) used an ANN model trained by the backpropagation algorithm to predict bank failures. They found that ANN performed better than Logit.

Singleton and Surkan (1991) used ANN models with two hidden layers (10 and 5 hidden nodes) to classify bond ratings. They found that performance of ANN was relatively accurate, compared to MDA and Probit.

Salchenberger, Cinar and Lash (1992) developed a neural network model using the backpropagation algorithm to predict the failure of thrift institutions. As usual, the authors used financial ratios as input variables and used stepwise regression to select the significant variables from the CAMEL categories. Then, they compared the results of ANN with those of Logit and found the results in favor of the ANN model. Also, they observed that decrease of type I errors was accompanied by greater increase of type II errors for the Logit model than for the ANN model. So an ANN model was quite a promising tool in terms of prediction ability and sensitivity between type I and type II errors.

Coats and Fant (1993) studied financial distress classification by using neural network technology. The authors used the Cascade-Correlation (or CASCOR) of Fahlman and Lebiere (1990) as a training algorithm for the neural network. As claimed by Fahlman and Lebiere, CASCOR has many advantages over other neural network algorithms. For example, it learns fast, it determines its own size and design (e.g., hidden nodes), and it does not need feedback of results. In their paper, Coats and Fant used the set of financial ratios from Altman (1968) and also compared the results with Altman's MDA model (or Z-score model). The concluded that CASCOR is more accurate, effective, and robust than MDA.

Boose, Dorsey and Huang (1992) developed an ANN model trained by the genetic adaptive neural network training (GANNT) algorithm to predict insolvency of life insurers. The authors found that, given the same misclassification costs of type I and type II errors, the performance of ANN was the best compared to the performances of MDA, Logit,

---

[1]The use of untransformed financial accounts as predictors in ANN is an attempt to test empirically the theory of the universal approximator of an ANN model (Hornik, Stinchcombe, and White, 1989). The idea of using untransformed accounts is not applied to Logit because there is no theoretical support. Especially, the use of untransformed accounts in Logit is subject to multicollinearity, spurious regressions, and violations of a priori assumptions of underlying methodology.

k-NN, and the naive classification model.

Dorsey, Edmister and Johnson (1993) compared ANN with RPA. Using the data set from Frydman, Altman and Kao (1985), they found that ANN was superior to RPA.

Luther (1993) used an ANN model to predict the outcomes of bankruptcy filings under Chapter 11. Unlike most studies which predicted whether a firn would be bankrupt in the future, this research focused on whether a firm already filing for bankruptcy would afterward be able to reorganize or would subsequently go for liquidation. Compared to Logit, ANN had high prediction ability and was not excessively sensitive to changes in cutoff points.

Barney (1993) used an ANN model to predict failure of the Farmers Home Administration (FHA) borrowers to make scheduled FHA debt payments. The neural network model used in this study provided a satisfying prediction with relatively low error rates compared to Logit and OLS.

The previous studies of bankruptcy prediction models, including the ANN technology-based models, consistently used financial ratios as predictors. There is no previous literature referring to the use of untransformed financial accounts as explanatory variables for classification and prediction in financial models. The primary contribution of this study is to use untransformed financial accounts as explanatory variables in an ANN model for predicting bankruptcy.[1]

If the theory of the universal approximator of ANN is supported, untransformed financial accounts should be at least as good as financial ratios as predictors in the bankruptcy models. Then the use of untransformed accounts can simplify the research and eliminate the search for proper transformations.

In addition to the primary contribution described above, the objectives of this study are to improve the performance of the bankruptcy-prediction model and to compare the performances of the models using different approaches. The basic techniques are provided for developing bankruptcy-prediction models as follows:

1. Basic approaches: 1) artificial neural network (ANN) using the genetic adaptive neural network training (GANNT) algorithm, and 2) Logit.[2]

---

[2]The modified version of the GANNT program (Dorsey, Johnson, and Mayer, 1992) was used to run ANN; and the logit procedure in SAS version 6.0 (SAS institute, Inc., 1990) was used to run Logit. ANN was run on the Cray Y-MP Supercomputer of the Mississippi Center for Supercomputing Research. Logit was run on the IBM 3084 VM/ XA of the University of Mississippi.

2. Basic predictors: 1) untransformed financial accounts, and 2) financial ratios.

3. Complementary techniques: the sign test for the statistical comparison of ANN and Logit.

The null hypotheses of this study are as follows:

H1: ANN does not predict bankruptcy as accurately as Logit.

H2: Financial ratios are better predictors than untransformed financial accounts.

The expected results are such that both of the above hypotheses are rejected.

2. Data Set

The data used in this study are entirely extracted from 10-K Compact Disclosure. The learning sample (or the estimation sample) consists of the firms filing in the years 1991 and 1992. The information (weights, parameter estimates, etc.) developed from the learning sample is used to predict the outcomes for the holdout sample consisting of the firms filing in 1993. The sample set consists of two groups of firms: a group of bankrupt firms and a group of nonbankrupt firms.

The bankrupt firms are the firms filing for bankruptcy under Chapter 7 or Chapter 11 of the U.S. Bankruptcy Code during the years 1991 through 1993. To obtain the bankrupt companies from 10-K Compact Disclosure, the selection is achieved by including all bankrupt companies which file for 8-K-3 filings (8-K-3 reports are filed on corporate bankruptcy events), and, at the same time, report bankruptcy in "Status", "Corporate Exhibits", or "Other Corporate Events" fields.[3]

The other group of the sample, the group of nonbankrupt firms, is chosen from 10-K Compact Disclosure with the exclusion of 8-K filings (8-K reports are filed on any unscheduled corporate events, such as mergers, acquisitions, dispositions of assets, or bankruptcy). In addition, the nonbankrupt firms must not mention bankruptcy in text or

[3]The criterion of using the intersection between the companies which file for 8-K-3 and the companies which report bankruptcy in text is to eliminate the companies with unclear bankruptcy. For example, a parent company which is not entering bankruptcy but has bankrupt subsidiaries also files for 8-K-3, so does a subsidiary of a bankrupt parent company. Companies which are thinking about but not entering bankruptcy report bankruptcy in text, so these firms have to be eliminated. Another example of sample selection is that some companies have filed for bankruptcy in the previous year(s) but are able to exit bankruptcy when the sample is selected.

in "Status" fields. The nonbankrupt group is matched about 9-to-1 to the bankrupt group.Table 1 shows the number of bankrupt and nonbankrupt firms in the learning and holdout samples.

Table 1 Learning Sample and Holdout Sample of Bankrupt and Nonbankrupt Firms

Sample	Filing Year	Number of Firms		Percentage of Bankrupt Firms
		Bankrupt	Nonbankrupt	
Learning	1991	97	884	9.89 %
Learning	1992	76	694	9.87 %
Holdout	1993	60	607	9.88 %

This study uses cross-validation tests which are extreme out-of-sample tests. Described in Dorsey, Edmister and Johnson (1993), cross-validation tests use the weights (or any information such as the parameter estimates) developed in the previous year(s) to predict the outcomes in the subsequent year. Cross-validation tests are useful because they are practical in reality. For example in reality, if we want to predict the ex ante outcomes for the year 1994 (as the current year), we can only use the ex post information of the year 1993 or earlier (as the past years) to develop the prediction model. In this study cross-validation of 1993 predictions is based on 1991-1992 weights. The firms filing in 1991-1992 are used as the learning (in-sample) observations while the firms filing in 1993 are used as the holdout (out-of-sample) observations.

Untransformed financial accounts are original data from balance sheets and income statements. Financial ratios are calculated from untransformed financial accounts. The financial data used are reported in the year before the firms file for bankruptcy. Table 2 shows the filing years and the years used to collect the data.

Table 2 Bankruptcy Filing Years and Data Collection Years

Bankruptcy Filing Year	Data Collection Year
1993	1992
1992	1991
1991	1990

The independent variables are selected from important accounting information which was used to compute the financial ratios in previous studies. The input variables used in this study are based on the study by Dorsey, Edmister and Johnson (1993), in which the variables were derived from the financial ratios used in the paper by Frydman, Altman and Kao (1985). Table 3 shows the untransformed financial accounts used in this study.

Table 3 Untransformed Financial Accounts

Variable Code	10-K CD Code	Definition
CH	CH	Cash
IV	IV	Inventories
CA	CA	Current Assets
TA	TA	Total Assets
CL	LI	Current Liabilities
TL	TL	Total Liabilities
NS	SA	Net Sales
IE	IF	Interest Expenses
IB	IB	Income Before Taxes
NI	NI	Net Income
CF	KA	Cash Flow
RE	RT	Retained Earnings
SE	SE	Shareholder Equity

When data are missing, Compact Disclosure reports code information as "NA"or leaves the cells blank in both text and numeric fields. As suggested in Dorsey, Edmister, and Johnson (1993), "NA" and blank cell can be changed to zero (0) because they represent data that do not exist. For example, a company which does not have preferred stock will report the code "NA" in this field.

Some of the variables described above are very likely to have correlation with others. A beauty of ANN technology is that it eliminates the noise so that the system can be fitted without transforming the data. Traditional econometric models have to be greatly concerned for correlation among independent variables because it can cause the problem of multicollinearity.

The financial ratios, used in Dorsey, Edmister and Johnson (1993), are computed using the untransformed accounts described above. Table 4 shows the ratio variables, represented in the form of untransformed account variable codes.

Table 4 Financial Ratios

Variable Code	Definition
CHTA	Cash/Total Assets
CHNS	Cash/Net Sales
CFTL	Cash Flow/Total Liabilities
CACL	Current Assets/Current Liabilities
CATA	Current Assets/Total Assets
CANS	Current Assets/Net Sales
EBITTA	EBIT/Total Assets[a]
LOGINT	Log (Interest Expenses + 15)
LOGTA	Log (Total Assets)
SETK	Shareholder Equity/(Total Assets-Current Liabilities)
NITA	Net Income/Total Assets
QACL	Quick Assets/Current Liabilities[b]
QATA	Quick Assets/Total Assets[b]
QANS	Quick Assets/Net Sales[b]
RETA	Retained Earnings/Total Assets
TLTA	Total Liabilities/Total Assets
NSTA	Net Sales/Total Assets
WKTA	Working Capital/Total Assets[c]
WKNS	Working Capital/Net Sales[c]

[a]EBIT = Income Before Taxes + Interest Expenses

[b]Quick Assets = Current Assets-Inventories

[c]Working Capital = Current Assets-Current Liabilities

Dorsey, Edmister and Johnson(1993) also presented an argument about undefined values caused by transformations. When the denominator of a ratio is zero or when we try to compute the natural logarithm of a value less than or equal to zero, the final transformed value is undefined. Therefore, some observations with undefined values have to be deleted. This surely reduces the sample size by 96 observations in the learning sample and 34 observations in the holdout sample, and underlines an advantage of using untransformed accounts over using financial ratios as predictors because some data can be lost during the process of transformation. Table 5 shows the number of bankrupt and nonbankrupt firms in the learning and holdout samples, which have been left after deleting undefined values.

Table 5 Learning Sample and Holdout Sample of Bankrupt and Nonbankrupt Firms for Financial Ratio Analysis

Sample	Filing Year	Number of Firms		Percentage of Bankrupt Firms
		Bankrupt	Nonbankrupt	
Learning	1991	96	827	10.40 %
Learning	1992	73	659	9.97 %
Holdout	1993	58	515	10.12 %

3. Methodology

The general idea of classification is based on the statistical framework of pattern recognition of two distinct groups and strategy to discriminate these distinct groups. Each classification methodology provides strategies of optimal decision rules for classification with a specific objective of either maximizing prediction accuracy or minimizing misclassification costs. Bayes' optimal decision rule is applied in order to partition the sample space into solution spaces Ω_i, i = 1,2,...,K where K is the number of groups. In this study firm x, which actually belongs to either the bankrupt group (G_1) or the nonbankrupt group (G_2), has the financial characteristic set X in which it can be classified into either the solution space of bankrupt firms (Ω_1) or the solution space of nonbankrupt firms (Ω_2).

For firm x which has the financial characteristic set X, the Bayes optimal decision rule is used such that:

$$\frac{P\,(X \backslash G_1)}{P\,(X \backslash G_2)} < \frac{P\,(\,G_2)}{P\,(\,G_1)} \Rightarrow \begin{array}{l} \chi \in \Omega_1 \\ \chi \in \Omega_2 \end{array} \qquad (1)$$

where $P(X \backslash G_i)$ is the class-conditional probability density function (pdf) of group i, and $P(G_i)$ is the prior probability of group i.

The primary approach used in this study is ANN. Then we compare ANN with Logit. Logit is a nonlinear specification and it uses the logistic function which is the dominant candidate among activation functions for an ANN model. As a parametric technique, Logit requires a distributional assumption between input and output. ANN is also considered as a parametric technique because its predicting ability is based on its estimated weight matrix. Although it is a parametric technique, ANN captures a most dominant characteristic of nonparametric techniques in that ANN does not require an assumption of the functional relationship between input and output.

3.1 Logit

Logit is a parametric technique used in classification and regression. It is a nonlinear probability model in which the dependent variable is not continuous but represents a discrete choice, such as bankruptcy or nonbankruptcy in this study. The nature of a nonlinear specification would be an S-shaped curve bounded in the interval (0,1). One such curve represents the logistic function used in a Logit model. The estimation of Logit is obtained through the maximum likelihood technique. Logit does not need the assumptions which are required in MDA, such as the condition that two groups of sample need to have the same variance-covariance matrix.[4]

As the dependent variable is assigned to be 1 for the bankrupt case and 0 for the nonbankrupt case, the probability with which a firm is classified as a bankrupt firm is given by:

$$Prob\,(Y = 1) \quad = \quad \frac{1}{1 + \mathrm{e\,x\,P}^{-\beta' x}} \quad = \; F\,(\beta' x) \qquad (2)$$

where Prob (Y = 1) = probability of bankruptcy
X = financial characteristics of the firm
β = the matrix of coefficients.

The parameters, β_{ig} can be estimated by using the maximum likelihood technique. Since each observation can be treated as a single draw from a Bernoulli distribution, the model with probabilities F ($\beta' x_i$) and independent observations leads to the likelihood function for a sample of n observations:

$$L = \prod_{i=1}^{n} [\ F(\beta' x_i)\ ]^{y_i} [\ 1 - F(\beta' x_i)]^{1-y_i} \qquad (3)$$

Therefore, the log likelihood can be obained as:

$$LnL = \sum_{i=1}^{n} [\ Y_i \ \text{ln} F(\beta' x_i) + (1-y_i) \ \text{ln} \ (\ 1-F(\beta' x_i)\ )] \qquad (4)$$

Now the estimators are obtained by maximizing the log likelihood function in equation(4) and then hypothesis testings of a Logit model can be conducted with either the Wald statistic or the likelihood ratio (LR) statistic, both chi-square distributed.

3.2 Artificial Neural Network (ANN)

An artificial neural network (ANN) has recently been viewed as a new approach of artificial intelligence (AI) tools applied to financial decision-making tasks, and one of the most successful applications of ANN in finance is to tasks involving classification and prediction. ANN does not need distributional assumptions or any specification of functional relationship between input variables and classification results. In fact, the most fascinating characteristic of ANN is that it provides a flexible mapping between inputs and outputs, with an appropriate functional relationship and without unjustified a priori assumptions required in econometric methodology. The idea of flexibility can be referred to in the neural net literature starting from Kolomogorov (1957) and followed by Sprecher (1965), Lorentz (1976), Hecht-Nielsen (1987), Irie and Miyake (1988), Hornik, Stinchcombe, and White (1989), and Funahashi (1989), among others.

Wasserman and Schwartz (1987) defined neural nets as highly simplified models of the human nervous system, exhibiting abilities such as learning, generalization,and abstraction.

---

[4]This definition section for Logit is derived from Kmenta (1990), and Greene (1990).

For financial situations which appear to be either unstructured or impractical to solve quantitatively, an ANN is considered as an effective and helpful tool for supporting financial decision-making tasks (Hawley, Johnson and Rains, 1990). The superiority of ANN is basically built on self-learning and self-maintaining abilities, based on new information and existing experience. ANN also offers a high degree of prediction accuracy, consistency, and robustness.

The artificial neuron structure of multilayer feedforward networks consists of three highly interconnected layers: 1) input layer, 2) hidden layer, and 3) output layer. Each interconnection of a network has a scalar weight which is computed by each processing element through a nonlinear function. A natural candidate among nonlinear functions would be an S-shaped curve bounded in the interval (0,1). One such curve is the sigmoid logistic function shown in Equation 7, also used in the Logit model.

$$g(\alpha) = \frac{1}{1 + e^{-\alpha}} \qquad (7)$$

This interconnected structure of nodes in a multilayer feedforward neural network provides a flexible mapping between inputs and outputs with an appropriate functional relationship regardless of its "true" functional form. The idea of flexibility has been bonded into the neural net literature starting from the study by Kolmogorov (1957) and followed by the subsequent studies by Sprecher(1965), Lorentz (1976), Hecht-Nielsen (1987), Irie and Miyake (1988), Hornik, Stinchcombe, and White (1989), and Funahashi (1989), among others. The importance of having a general mapping between the input and output vectors, as stated in Hawley, Johnson, and Raina (1990), is that it eliminates the need for the unjustified a priori restrictions so commonly used to facilitate estimation.

In addition, one of the most important advantages of an ANN is its ability to learn from historical data by automatically adjusting the weights when the problem environment is changed. This property of an ANN is particularly useful in an environment of sensitive information.

In this research, the genetic algorithm technique is proposed as an appropriate learning algorithm for the ANN model. This learning algorithm is explained below.

3.2.1. Genetic Algorithm

ANN literature typically uses backpropagation as the training algorithm for an ANN model. However, backpropagation, which is a point-to-point gradient-search technique, has a tendency for the network to become trapped in local optima (Hecht-Nielsen, 1990).

The genetic algorithm, on the other hand, is a global-search technique which iterates a random search until it reaches the optimal solution of the objective function (Dorsey, Johnson and Mayer, 1991). Since the genetic algorithm does not use a gradient method which requires the derivative of outputs to adjust the weights, the network trained by the genetic algorithm can use any kind of objective function, not necessarily only differentiable functions (Dorsey and Mayer, 1992). Therefore, this study uses the genetic adaptive neural network training (GANNT) algorithm of Dorsey, Johnson and Mayer(1991) as the training algorithm for the ANN model.

The GANNT algorithm starts with the **selection** process in which a random population of solutions or a set of weight matrices is initiated.The objective function is then calculated and surviving factors are obtained. A subset of population is selected as the strongest surviving solution which optimizes the objective function. Then the selected weight matrices will be used in the reproduction process which produces "offspring" for the next generation. The network can be accelerated to obtain optimization by using the crossover technique in which the vector elements of weight matrices are swopped among each weight matrix. Also, a mutation decides whether any element of the weight matrices should be changed. The process iterates until the optimum solution of the objective function can be reached.

3.2.2. Hidden Units, Weight Matrices, and Transformations of Data

The preliminary tests of this paper were conducted by assuming the number of input nodes equal to the number of hidden nodes for every model. The results did indeed imply that a model with more weights (more input and hidden nodes) performed better than a model with less weights. This intuition is actually presented in the paper "An Additional Hidden Unit Test for Neglected Nonlinearity in Multilayer Feedforward Networks" by Halbert White (1990). White proved that multilayer feedforward networks exactly represent some unknown mappings between input and outputs and that the functions can be estimated to any degree of accuracy by adding hidden units to the networks.

The preliminary study of this paper could not reject the hypothesis that financial ratios were better predictors than untransformed accounts. One explanation was observed from the fact that thirteen untransformed accounts produced as many as nineteen financial ratios. Transformations of data may be able to represent the implied hidden nodes; and the more the hidden nodes, the fitter the model. As stated in Hornik, Stinchcombe, and White (1989), a neural network can approximate the relationship between inputs and outputs to any degree of accuracy provided that a sufficient number of hidden units are given. So the number of hidden nodes of the model using untransformed accounts has to be increased to offset the implied and actual weights in the model using financial ratios.

Dorsey (1992) conducted some experiments on transformations of data, such as the inverse function (1/X) and the multiplication (XY) of data.[5] He observed the empirical result that the networks could be trained very fast to produce the transformed data. Consequently, this study will avoid the problem of implied hidden units by assuming equal weights for every model.

3.3. Prior Probabilities and Misclassification Costs

Mentioned in Luther (1993), the cutoff point of 0.5 is optimum if and only if prior probabilities of both groups in the population are the same and costs of type I and type II errors are also the same.[6] However, the prior probability of the nonbankrupt group is obviously much bigger than that of the bankrupt group and a type I error (misclassifying bankrupt firms as nonbankrupt firms) is certainly much more costly than a type II error (misclassifying nonbankrupt firms as bankrupt firms). Therefore, the cutoff point of 0.5 is clearly not suitable. In this study, Prob (Y = 1) is defined as the probability of failure. When the cutoff point is greater than or equal to 0.5, firms are likely to be classified as nonbankrupt firms because the nonbankrupt group dominates the bankrupt group in the sample.

Due to the facts that a type I error is much more costly than a type II error and that the prior probability of the bankrupt group is much less than that of the nonbankrupt group, bankruptcy-prediction models have to be adjusted with the type-I to type-II cost ratio and the prior probability. Zavgren (1983) mentioned Diamond's[7] study which concluded that the prior probability of 0.1 for failure with a type-I to type-II cost ratio of 20 gave the best results. The Manski-Lerman (1977) weights incorporate prior probabilities and error cost ratios for Logit models. The log-likelihood function can be defined as

$$L = \sum_{i=1}^{n} [\ W_1\ Y_i\ \ln\pi_i + W_2\ (1-Y_i) \ln(\ 1 - \pi_i)\ ] \qquad (8)$$

where the weights, W1 and W2, serve to adjust the misclassification costs and prior probabilities. Mathematically,

$$\frac{W_1}{W_2} = \frac{C_{12}\,(P_1/S_1)}{C_{21}\,(P_2/S_2)} \qquad (9)$$

[5]The experiments are reported from private communication with Dr. Dorsey.

[6]This idea can also be referred to in Neter, Wasserman and Kutner (1989).

where C_{12} is defined as the cost of misclassifying a bankrupt firm as a nonbankrupt firm,

C_{21} is defined as the cost of misclassifying a nonbankrupt firm as a bankrupt firm,

P_1 is the proportion of the bankrupt group in the population,

P_2 is the proportion of the nonbankrupt group in the population,

S_1 is the proportion of the bankrupt group in the sample, and

S_2 is the proportion of the nonbankrupt group in the sample.

If the sample proportions (S1/S2) equal the population proportions (P1/P2), the weights reflect only the factor of misclassification costs (C12/C21). This study assumes the prior probability of 10% bankrupt firms as Diamond's study did, so that S1/S2 equals P1/P2. The log-likelihood function can be defined as

$$LnL = \sum_{i=1}^{n} [\; C_{12}\, y_i \;\; \ln \pi_i + C_{21} \;(1-Y_i) \ln (\; 1 - \pi_i)\;] \qquad (10)$$

The costs of type I and type II errors are analyzed by experimenting with various type-I to type-II cost ratios. The best ratio, which yields the best classification accuracy subject to the constraint that the type I error is less than or equal to the type II error, is selected.

Total misclassification costs can also be calculated as:

$$MC = P_1 C_{12} TypeI_e + P_2 C_{21} TypeII_e \qquad (11)$$

where $TypeI_e$ is the probability of a bankrupt firm misclassified as a nonbankrupt firm, and

$TypeII_e$ is the probability of a nonbankrupt firm misclassified as a bankrupt firm. Mathematically,

$$TypeI_e \;=\; \frac{n_1}{N_1} \qquad (12.1)$$

and

$$TypeII_e \;=\; \frac{n_2}{N_2} \qquad (12.2)$$

---

[7]Diamond, H., Jr.,"Pattern Recognition and the Detection of Corporate Failure," 1976.

where n1 is the number of misclassified "bankrupt" firms,

N1 is the number of bankrupt firms in the sample,

n2 is the number of misclassified "nonbankrupt" firms, and

N2 is the number of nonbankrupt firms in the sample.

3.4. The Sign Test

The sign test is used to compare the performance of Logit and that of ANN. Both Logit and ANN are probability and parametric models. Although ANN is parametric, we do not have knowledge of how the basic variables in ANN are distributed. Therefore, nonparametric methods are quite suitable for comparison of ANN and Logit, and comparison of ANN models using different predictors. Also, the conflict of results due to the different cutoff values can be eliminated. As discussed in Hoel (1971, pp. 309-315), the sign test is not only a nonparameric but also a distribution-free method.

The hypothesis for testing the performance of Logit and that of ANN is that there is no superiority of either Logit or ANN over the other. In other words, the probability that Logit is better than ANN equals the probability that ANN is better than Logit.

First, the median, which has the property that the probability is 1/2, is defined for testing hypothetical values of the median without knowing the form of the distribution.

Let Y be a dichotomous variable representing 1 for a bankrupt firm and 0 for a nonbankrupt firm. Let P_L be the probability of failure classified by Logit and P_{NN} be the probability of failure classified by ANN.

Let $X_1, X_2, \ldots, X_n$ be a random sample of X and

$$X = \text{abs}\,(Y-P_L) - \text{abs}(Y-P_{nn}) \qquad (13)$$

where abs (PB) is the absolute value.

The hypothesis tests whether Logit is as good as ANN. If X is negative (<0), Logit is better than ANN. If X equals zero (=0), Logit is as good as ANN. If X is positive (>0), ANN is better than Logit. Therefore, if the hypothesis is rejected, either ANN or Logit is better than the other.

Assume that f(x) is a continuous density function and assume that its median, denoted by ξ, is uniquely defined by[8]

$$\int_{\xi} f(x)\,dx = \frac{1}{2} \qquad (14)$$

for testing the hypothesis

$$H_0 : \xi = 0 \quad against\ H_1 : \xi \neq 0 \qquad (15)$$

From the definition of the median it follows that when H_0 is true, Prob[X>0] = 1/2 and therefore Prob[X_i>0] = 1/2, i = 1,..., n.

Let

$$Z_i = 1, \quad if\ X_i > 0;\ 0,\ if\ X_i < 0 \qquad (16)$$

When $X_i = 0$, the observations are cumulated as m.[9] Then the variable Z_i is a binomial variable corresponding to a single trial of an experiment for which probability equals 1/2. Therefore,

$$E\,[\,Z_i\,] \quad = \quad \frac{1}{2} \qquad (17)$$

Since the Z_i are independent, their sum (U) will be a binomial variable corresponding to (n - m) independent trials of an experiment for which probability equals 1/2. Thus, when H_0 is true,

$$E\,[U] = \sum_{i=1}^{n-m} E\,[Z_i] = \sum_{i=1}^{n-m} \frac{1}{2} \; = \; \frac{n-m}{2} \qquad (18)$$

For the probability equal to 1/2, the binomial distribution is approximated well by its normal approximation. Therefore, let τ be a standard normal variable where

$$\tau \; = \; \frac{U \pm \frac{1}{2} - \frac{n-m}{2}}{\sqrt{\frac{n-m}{4}}} \qquad (19)$$

The correction +1/2 is used for a left tail critical region and -1/2 for a right tail region.

[8]This section is derived from Hoel (1971) pp. 309-315.

[9]Since we want to test the superiorities of the models, the inclusion of the equalities of the models will bias the distribution.

4. Results and Analysis

Table 6 shows the prediction performances of the Logit and ANN models at the cutoff point of 0.5. The table incorporates the models using untransformed accounts and financial ratios. Cross-validation tests are conducted for the learning sample consisting of 1991-1992 data and are verified by the holdout sample consisting of 1993 data.

Table 6 Learning and Holdout Sample Classification Results (The cutoff point = 0.5)

Untransformed Accounts						
Model	Type I Error[a]		Type II Error[b]		Overall Accuracy	
	Learn	Holdout	Learn	Holdout	Learn	Holdout
Logit**	15.03 %	15.00 %	19.33 %	20.29 %	81.10 %	80.23 %
ANN*	12.14 %	16.67 %	19.01 %	20.48 %	81.67 %	79.90 %
Financial Ratios						
Model	Type I Error[a]		Type II Error[b]		Overall Accuracy	
	Learn	Holdout	Learn	Holdout	Learn	Holdout
Logit**	18.34 %	15.52 %	26.65 %	25.24 %	74.20 %	75.74 %
ANN*	10.65 %	17.24 %	15.21 %	16.70 %	85.26 %	83.25 %

a A type-I error is defined as classifying a bankrupt firm as a nonbankrupt firm.

b A type-II error is defined as classifying a nonbankrupt firm as a bankrupt firm.

* The type-I to type-II cost ratio equals 10.

** The type-I to type-II cost ratio equals 11.

ANN gives the best in-sample performance with the highest prediction accuracy and the lowest type I and type II errors in almost every case. For the out-of-sample performances, ANN still offers the best prediction accuracies in the models using financial ratios as predictors. For the models using untransformed accounts as predictors, ANN is slightly less accurate than Logit (79.90% vs. 80.23%). The type I errors of the ANN models are somewhat higher than those of Logit and RPA, but still no more than 20%. The type II errors of the ANN models are lower than those of Logit in almost every case. For the models using untransformed accounts as predictors, the type II error of ANN is 20.48% while that of Logit is 20.29%.

Compared with financial ratios, untransformed accounts are seemingly less preferable predictors for ANN. In the training sample, The ANN model using financial ratios as predictors offers lower type I and type II errors and a higher prediction accuracy than does the ANN model using untransformed accounts as predictors. In the holdout sample, although the ANN model using untransformed accounts yields a slightly lower type I error (16.67% vs. 17.24%), the ANN model using financial ratios as predictors yields a lower type II error and a higher prediction accuracy.

The performances of the bankruptcy models at the cutoff point of 0.5 can be summarized as follows: 1) ANN offers a better performance than does Logit in most cases; and 2) financial ratios are better than untransformed accounts as predictors in the ANN models. The prediction accuracy of each model is over 70%; for a type I error under 30%. Especially, accuracies of the ANN models are very close to and over 80%; while type I and II errors are close to and under 20%. Prediction accuracies of out-of-sample models are not much different from those of in-sample models. In other words, the in-sample performance is comparable to the out-of-sample performance. Therefore, the bankruptcy models developed in this study are quite practical.

When financial ratios are used as predictors, ANN apparently performs better both in- and out-of-sample than does Logit. However, when untransformed accounts are used, ANN can only perform well in the training sample. ANN is not superior to Logit in predicting the holdout sample when untransformed accounts are used as predictors.

As discussed earlier, econometric models which use untransformed accounts are more likely to be jeopardized by the problems of spurious regressions and multicollinearity. These problems also seem serious in ANN. Data transformations are empirically confirmed to improve the predictive performance of ANN.

Statistical comparisons of the models are performed by the sign tests. Table 7 shows the sign tests for Logit and ANN in the learning sample, and Table 8 shows those in the holdout sample. It can be seen that ANN is better than Logit in every case.

Table 7 The Sign Tests for Logit and ANN (Learning Sample)

Untransformed Accounts						
Sample	Superiority			U	n-m	τ
	Logit	ANN	Neither			
Bankrupt	45.09 %	51.45 %	3.47 %	89	167	0.77
Nonbankrupt	35.42 %	64.32 %	0.25 %	1015	1574	11.47***
Overall	36.38 %	63.05 %	0.57 %	1104	1741	11.17***
Financial Ratios						
Sample	Superiority			U	n-m	τ
	Logit	ANN	Neither			
Bankrupt	26.04 %	73.37 %	0.59 %	124	168	6.09***
Nonbankrupt	25.77 %	74.23 %	0.00 %	1103	1486	18.65***
Overall	25.80 %	74.14 %	0.06 %	1227	1654	19.65***

*** Significant at the 1% level of confidence.
** Significant at the 5% level of confidence.
* Significant at the 10% level of confidence.

Superiority = Percentage that a model is better than the other.

U = Sum of Z_i

Z_i = 1 when a model is better than the other;
0 when the other is better.

n = Number of observations in the sample.

m = Number of observations with no superiority.

τ = Standard normal variable.

When untransformed accounts are used as predictors, ANN performs significantly better than Logit (at the 1% level of confidence) in both learning and holdout samples. The superiority of performance of ANN comes from the classification accuracy in the nonbankrupt group. ANN performs marginally better than Logit in the bankrupt group, but not significantly at the 10% level of confidence.

When financial ratios are used as predictors, ANN is approximately twice as good as Logit (ANN is better than Logit in 73-74% of the learning sample and in 62-69% of

Table 8 The Sign Tests for Logit and ANN (Holdout Sample)

Untransformed Accounts						
Sample	Superiority			U	n-m	τ
	Logit	ANN	Neither			
Bankrupt	41.67 %	53.33 %	5.00 %	32	57	0.79
Nonbankrupt	33.64 %	66.00 %	0.37 %	361	545	7.54***
Overall	34.43 %	64.74 %	0.82 %	393	602	7.46***
Financial Ratios						
Sample	Superiority			U	n-m	τ
	Logit	ANN	Neither			
Bankrupt	37.93 %	62.07 %	0.00 %	36	58	1.71**
Nonbankrupt	29.90 %	69.90 %	0.19 %	360	514	9.04***
Overall	30.72 %	69.11 %	0.17 %	396	572	9.16***

*** Significant at the 1% level of confidence.
** Significant at the 5% level of confidence.
* Significant at the 10% level of confidence.

Superiority = Percentage that a model is better than the other.
U = Sum of Z_i
Z_i = 1 when a model is better than the other;
0 when the other is better.
n = Number of observations in the sample.
m = Number of observations with no superiority.
τ = Standard normal variable.

the holdout sample). The superiorities of ANN over Logit are statistically significant at the 5% level of confidence in both learning and holdout samples and both bankrupt and nonbankrupt groups.

The sign tests shown in Table 9 indicate statistical comparisons of the models using different predictors. Comparing between untransformed accounts and financial ratios, financial ratios are significantly better in both learning and holdout samples (the bankrupt and nonbankrupt groups are not separated). In the nonbankrupt group, the superiorities of financial ratios as predictors are statistically significant at the 1% level of confidence in both learning and holdout sample. In the bankrupt group, financial ratios are significantly better at the 10% level of confidence for the learning sample, and untransformed accounts are marginally, but not significantly, better for the holdout sample.

Table 9 The Sign Tests for Statistical Comparison of the Predictors for ANN

The Learning Sample						
Sample	Superiority			U	n-m	τ
	UA	FR	Neither			
Bankrupt	40.83 %	52.07 %	7.10 %	88	157	1.44^{*}
Nonbankrupt	40.92 %	51.88 %	7.20 %	771	1379	4.36^{***}
Overall	40.91 %	51.90 %	7.19 %	859	1536	4.62^{***}
The Holdout Sample						
Sample	Superiority			U	n-m	τ
	UA	FR	Neither			
Bankrupt	48.28 %	41.38 %	10.34 %	28	52	0.42
Nonbankrupt	39.81 %	50.49 %	9.71 %	260	465	2.50^{***}
Overall	40.66 %	49.56 %	9.77 %	248	517	2.20^{**}

*** Significant at the 1% level of confidence.
** Significant at the 5% level of confidence.
* Significant at the 10% level of confidence.
UA = Untransformed Accounts.
FR = Financial Ratios

Two "brief" conclusions are implied: 1) ANN is significantly better than Logit; and 2) financial ratios are significantly better than untransformed accounts as predictors for ANN.

5. Conclusions

The use of untransformed financial accounts is not common in the finance literature because untransformed financial accounts are not standardized by firm size. In addition, the use of untransformed accounts in econometric models has a possibility of encountering the problems of multicollinearity, spurious regressions, and violation of a priori assumptions of econometric methodology. However, with the property of a universal approximator, the ANN models should be indifferent to using either.

The primary contribution of this study is to use untransformed financial accounts and artificial neural network technology to develop a bankruptcy-prediction model with an attempt to achieve two properties: 1) the model should be convenient to develop; and 2) the model should be accurate and efficient. This study uses an artificial neural network (ANN) trained by the genetic algorithm and uses untransformed accounts as predictors.

Also, cross-validation tests are conducted to verify the bankruptcy models with the extremely out-of-sample data. The data used in this study are entirely extracted from 10-K Compact Disclosure. The sample consists of bankrupt firms which file for bankruptcy under Chapter 7 or Chapter 11 of the U.S. Bankruptcy Code, and nonbankrupt firms which are randomly selected and matched about 9-to-1 to bankrupt firms. This ratio is suggested in Diamond's study as a proper proportion of bankrupt firms in the population.

The performances of the bankruptcy models are measured by prediction accuracy and prediction errors (type I and type II errors). The performances of bankruptcy models are evaluated at the cutoff point of 0.5. The results are as follows: 1) ANN is a better predictor than Logit; and 2) financial ratios are better than untransformed accounts as predictors in the ANN models. Statistical comparisons of the ANN and Logit models are performed by the sign tests. The results of the sign tests indicate that ANN is significantly better than Logit, and that financial ratios are significantly better than untransformed accounts as predictors for ANN.

All results indicate that financial ratios are better than untransformed accounts as predictors. The ANN model using financial ratios as predictors can be trained very fast and very accurately, compared to the ANN model using untransformed accounts as predictors. Therefore, the use of untransformed financial accounts as predictors in an ANN model still has serious problems.

As discussed, untransformed financial accounts are not standardized by firm size. In addition, the spurious regressions and multicollinearity which are problems in econometric models also jeopardize the ANN model when untransformed accounts are used as predictors.

It may need more time to train an ANN model which uses untransformed accounts as predictors, and time consumption will be a major weakness of using untransformed accounts as predictors. So untransformed accounts may offer convenience, but not accuracy.

REFERENCES

1. Altman, E.I.,"Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, September 1968.
2. Altman, E.I., "Commercial Bank Lending: Process, Credit Scoring and Costs of Errors in Lending," *Journal of Financial and Quantitative Analysis*, Proceedings Issue, Vol.XV, No.4, November 1980, 813-833
3. Altman, E.I., "Bankruptcy and Reorganization," in Altman (ed), *Handbook of Corporate Finance*, John Wiley & Sons, Inc., 1986.
4. Altman, E.I., "Evaluating the Chapter 11 Bankruptcy-Reorganization Process," *Columbia Business Law Review*, Vol. 1993, No.1.
5. Altman, E.I., Avery, R., Eisenbeis, R., and Sinkey, J., *Application of Classification Techniques in Business, Banking and Finance*, JAI Press, 1981.
6. Altman, E.I., Haldeman, R., and Narayanan, P.,"ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations," *Journal of Banking & Finance*, June 1977.
7. Anderson, J.A., "Cognitive and Psychological Computation with Neural Models." *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-13, No.5, September/ October 1983, 799-815.
8. Amemiya, T., "Qualitative Response Models: A Survey," *Journal of Economic Literature*, 19, 1981, 1483-1536.
9. Barney, D.K., "Modeling Farm Debt Failure: The Farmers Home Administration," *Ph.D. Dissertation*, The University of Mississippi, 1993.
10. Beaver, W.H., "Financial Ratios as Predictors of Failure, Empirical Research in Accounting: Selected Studies." Supplement to *Journal of Accounting Research 4*, 1966, 71-111.
11. Bell, T.B., Ribar, G.S., and Verchio, J.R., "Neural Nets VS. Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Banks' Failures." Presented at *Deloitte & Touche/University of Kansas Auditing Symposium*, 1990.

12. Blum, M., "Failing Company Discriminant Analysis," *Journal of Accounting Research*, Spring, 1974.

13. Boose, M.A., Dorsey, R.E., and Huang, C.S., "Neural Networks for NAIC Solvency Surveillance of Life Insurers," Working Paper, *University of Mississippi*, 1992.

14. Caporaletti, L., Dorsey, R.E., Johnson, J.D., and Powell, W.A., "Artificial Neural Systems as In-Sample Simultaneous Equation Systems Forecasters," forthcoming, *Decision Support Systems.*

15. Caporaletti, L.E. and Seydel, J., "Using Artificial Neural Networks to Assist in Determining Optimal Bids," The University of Mississippi.

16. Casey, C.J., McGee, V.E., and Stickney, C.P., "Discriminating Between Reorganized and Liquidated Firms in Bankruptcy," *Accounting Review*, April 1986, 249-263.

17. Caudill, M., "Using Neural Nets: Representing Knowledge, Part 1," *AI Expert*, December 1989, 34-41.

18. Chester, D.L., "Why Two Hidden Layers Are Better Than One," *Proc. IJCNN* 90, Washingon D.C., January 15-19, 1990.

19. Coats, P.K., and Fant, L.F., "Recognizing Financial Distress Patterns Using a Neural Network Tool," *Financial Management*, Vol. 22, No.3, 1993, 142-155.

20. Cowan, J.D. and Sharp, D.H., "Neural Nets," *Quarterly Review of Biophysics*, Vol. 21, No.3, 1988, 305-427.

21. Deakin, E.B., "A Discriminant Analysis of Predictors of Business Failure," *Journal of Accounting Research*, March 1972.

22. Diamond, H., Jr., "Pattern Recognition and the Detection of Corporate Failure," *Ph.D. Dissertation*, New York University, 1976 (Mentioned in Zavgren 1983).

23. Dorsey, R.E., Edmister, R.O., and Johnson, J.D., "Financial Distress Prediction: A Neural Net Model for Large Corporations," *Working Paper*, University of Mississippi, 1993.

24. Dorsey, R.E., Johnson, J.D, and Mayer, W.J., "The Genetic Adaptive Neural Network Training (GANNT) Algorithm for Generic Feed Forward Artificial Neural Systems," 1991.

25. Dorsey, R.E., and Mayer, W.J., "Optimization Using Genetic Algorithms," *Working Paper*, University of Mississippi, 1992.

26. Edmister, R.O., "Financial Ratios and Credit Scoring for Small Business Loans," *Journal of Commercial Bank Lending*, September 1971.

27. Edmister, R.O., "An Empirical Test of Financial Ratio Analysis for Small Business Failure Predictions," *Journal of Financial and Quantitative Analysis*, March 1972.

28. Fahlman, S.E., and Lebiere, C., "The Cascade-Correlation Learning Architecture," *Technical Report: CMU-CS-90-100*, Carnegie Mellon University, February 1990.

29. Frydman, H., Altman, E.I., and Kao, D.L., "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress," *Journal of Finance*, 1985.

30. Funahashi, K., "On the Approximate Realization of Continuous Mappings by Neural Networks," *Neural Networks*, Vol. 2, 1989, 183-192.

31. Granger, C.W.J., and Newbold, P., "Spurious Regressions in Econometrics," *Journal of Econometrics 2*, 1974, 11-120.

32. Greene, W.H., *Econometric Analysis*, Macmillan Publishing Company, New York, 1990.

33. Hansen, J.V., McDonald, J.B., and Stice, J.D. "Artificial Intelligence and Generalized Qualitative-Response Models: An Empirical Test on Two Audit Decision-Making Domains," *Decision Sciences*, Vol.23, No.3, May/June 1992, 708-723.

34. Hawley, D.D., Johnson, J.D., and Raina, D., "Artificial Neural Systems; A New Tool for Financial Decision Making," *Financial Analysts Journal*, November/December 1990, 63-72.

35. Hecht-Nielsen, R., "Kolmogorov's Mapping Neural Network Existence Theorem," *IEEE First International Conference on Neural Networks*, June 21-24, Vol.3, 1987, 11-14.

36. Hecht-Nielsen, R., *Neurocomputing*, Addison-Wesley Publishing Company, Massachusetts, 1990.

37. Hoel, P.G., *Introduction to Mathematical Statistics*, Fourth Edition, John Wiley & Sons, Inc., 1971.

38. Holland, J.H., "Genetic Algorithms," *Scientific American*, July 1992.

39. Hopfield, J.J. and Tank, D.W., "Neural Computation of Decisions in Opinizaion Problems," *Biological Cybernetics*, Vol. 52, 1985, 141-152.

40. Hornik, K., Stinchcombe, M., and White, H., "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, Vol.2, 1989, 359-366.

41. Hornik, K., Stinchcombe, M., and White, H., "Universal Approximation of An Unknown Mapping and Its Derivatives Using Multilayer Feedforword Networks," *Economics*, University of California, San Diego, February 1990.

42. Irie, B. and Miyake, S., "Capabilities of Three-Layered Perceptrons," *IEEE International Conference on Neural Networks*, 1988, 641-648.

43. Johnston, J., *Econometric Methods*, McGraw-Hill Book Company, New York, 1984.

44. Jorgensen, C. and Matheus, C., "Catching Knowledge in Neural Networks, " *AI Expert*, December 1986, 31-41.

45. Kennedy, P., *A Guide to Econometrics,* Second Edition, The MIT Press, Cambridge, Massachusetts, 1985.

46. Kmenta, J., *Elements of Econometrics,* Second Edition, Macmillan Publishing Company, New York, 1990.

47. Kohonen, T., "State of the Art in Neural Computing," *International Conference on Neural Networks.*

48. Kolmogorov, A.N., "On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition," *Dokl. Akad. Nauk USSR,* 114, 1957, 953-956.

49. Lapedes, A. and Farber, R., "How Neural Nets Work," *IEEE Second International Conference on Neural Networks,* Vol. 2, 442-456.

50. Lippman, R.P., "An Introduction to Computing with Neural Nets," *IEEEASSP Magazine,* April 1987, 4-22.

51. Lorentz, G.G., "The 13^{th} Problem of Hilbert." Proc. of Symposia in Pure Math, *American Mathematical Society,* 28, 1976.

52. Luther, R.K., "Predicting the Outcome of Chapter 11 Bankruptcy: An Artificial Neural Network Approach," *Ph.D. Dissertation,* The University of Mississippi, December 1993.

53. Manski, C.F. and Lerman, S.R., "The Estimation of Choice Probabilities From Choice Based Samples," *Econometrica 45,* November 1977.

54. Martin, D., "Logit Analysis and Early Warning Systems for Bank Supervision," *Journal of Banking and Finance I,* November 1977, 249-276.

55. McIntyre, J.E. Jr., "Predicting Bank Failures: Bankruptcy Theories, Classification Models, and Empirical Evidence," *Ph.D. Dissertation,* The University of Georgia, 1990.

56. Neter, J., Wasserman, W., and Kutner, M.H., *Applied Linear Regression Models,* Second edition, Richard D. Irwin, Inc.

57. Ohlson, J.A., "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research,* 1980.

58. Parzinger, M.J., "Artificial Neural Network Model for the Prediction of Bank Failures," Memphis State University.

59. Reece, P., "Perceptrons and Neural Nets," *AI Expert,* January 1987, 50-57.

60. Rosenblatt, F., "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review,* Vol.65, No.6, 1958, 386-408.

61. Rumelhart, D.E., Hinton, G.E., and Williams, R.J., "Learning Internal Representation

by Error Propagation," in D.E. Rumelhart and J.L. McClelland, Eds., *Parallel Distributed Processing: Exploration in the Microstructures of Cognition*, MIT Press, MA, 1986, 318-362.

62. Rumelhart, D.E., Hinton, G.E., and Williams, R.J., "Learning Representation by Propagating Errors," *Nature*, 323, 1986, 533-536.
63. 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, July/August 1992.
64. Singleton, J.C., and Surkan, A.J., "Modeling the Judgement of Bond Rating Agencies: Artificial Intelligence Applied to Finance," *Journal of Midwest Finance Association*, 1991, 72-80.
65. Sinkey, J.F., Jr., "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," *Journal of Finance 30*, March 1975, 21-36.
66. Sinkey, J.F., Jr., Terza, J.V., and Dince, R.R. "A Zeta Analysis of Failed Commercial Banks," *Quarterly Journal of Business and Economics 26*, Autumn, 1987, 35-49.
67. Sprecher, D.A., "On the Structure of Continuous Functions of Several Variables," *Trans. Amer. Math. Soc.*, 115, March 1965, 340-355.
68. Wasserman, P.D., and Schwartz, T., "Neural Network, Part I," *IEEE Expert*, Winter, 1987, 10-12.
69. Wasserman, P.D., and Schwartz, T., "Neural Network, Part II," *IEEE Expert*, Spring, 1988, 10-15.
70. White, H., "Economic Prediction Using Neural Networks: the Case of IBM Daily Stock Returns," *IEEE Conference on Neural Networks*, Vol.2, 1988, 451-458.
71. White, H., "An Additional Hidden Unit Test for Neglected Nonlinearity in Multilayered Feedforward Networks," *Proceedings of the International Joint Conference on Neural Networks*, Wash. D.C., Vol.2, 1989, 451-455.
72. White, H., "Consequences and Detection of Misspecified Nonlinear Regression Model," *Journal of the American Statistical Association*, Vol. 76, No. 374, June 1981.
73. White, H., "Neural Network Learning and Statistics," *AI Expert*, December 1989, 49-52.
74. White, H., "Some Asymptotic Results for Bank-Propagation," *IEEE Second International Conference on Neural Networks*, Vol. 3, 1990, 261-266.
75. Werbos, P.J., "Backpropagation: Past and Future," IEEE *Second International Conference on Neural Networks*, Vol. 1, 343-353.
76. Zavgren, C.V., "The Prediction of Corporate Failure: The State of the Art," *Journal of Accounting Literature*, 1983