December 1993

Intelligent Credit Rating System

Ingoo Han
Korea Advanced Institute of Science and Technology

Won Jhee
Hong IK University

INTELLIGENT CREDIT RATING SYSTEM

INGOO HAN, Department of Management Information Systems,
Korea Advanced Institute of Science and Technology, Seoul, Korea

WON CHUL JHEE, Department of Industrial Engineering,
Hong Ik University, Seoul, Korea

ABSTRACTS

Credit rating represents an assessment of the relative level of risk associated with the timely payments required by the debt obligations. Business researchers have traditionally used statistical techniques for classification. In late 1980's, inductive learning started to be used for business classification. Recently, neural network began to be applied for business classification. Korea has a short history of credit rating. This study applies the neural network approach to Korean commercial paper rating data and compare the performance by neural network algorithm with that by the discriminant analysis. The results show that neural network has a superior predictve potential.

INTRODUCTION

Classification refers to separating distinct sets of objects or observations and allocating new objects or observations into previously defined groups. Classification needs an algorithm to separate and allocate objects or observations. This algorithm is called a classification technique. The ultimate goal of a classification method is to provide the relevant outcome or to replicate the expert's judgment. The relative performance of different classification techniques may depend on data conditions.

Business researchers have traditionally used statistical techniques for classification. Statistical methods frequently employed in business classification research include multiple discriminant analysis(MDA), logit, and probit methods.

Recently, inductive learning, a subfield of artificial intelligence(AI), began to be applied to the classification research in business. Examples include stock market prediction (Braun and Chandler, 1987) and scholarship and fellowship grant cases (Garrison and Michaelsen, 1989). Inductive learning uses a data set of examples and determines a relationship between these examples via inductive inference. The induced rules can then be used to predict outcomes or to replicate judgments. ID3 is an inductive learning algorithm which has been used most widely.

Quite recently, the neural network approach, another field of AI, has started to be applied to the classification research in business. Neural network model is based on how human brain cells and their interactions are able to perform complex tasks. A neural network model consists of many processing elements(PE). These PEs are grouped into linear arrays called layers. A neural network model has an input layer and an output layer, and may or may not have hidden layers. Each PE computes the linear combination of input signals and applies the transfer function to calculate the output value. Examples using the neural network approach include accounting inventory method choice(Liang, Chandler, Han, and Roan, 1992), bankruptcy prediction(Tam and Kiang, 1992), and bond rating(Surkan and Singleton, 1990).

Areas of business classification research include bankruptcy prediction, accounting method choice, audit opinion decisions, credit rating prediction, and bank loan classification. The credit rating is an important task for the capital market to work efficiently. The two largest rating agencies are Standard and Poor's, and Moody's both of which have long histories. The credit rating in Korea has started in 1985 when the Korea Credit Rating Corporation was founded. This paper applies the neural network approach to the Korean commercial paper rating data and compares the results using the neural net approach with those using the statistical approach.

STATISTICAL CLASSIFICATION TECHNIQUES

The statistical methods frequently employed in business classification research are MDA, logit, and probit. Discriminant analysis, evolved as a variant of univariate analysis of variance, is generally concerned with comparisons of the distribution of one or more variables across different groups or populations (Altman et al., 1981). Discriminant analysis can also be used to test for differences in variable mean vectors and/or covariance structures across groups.

Discriminant analysis assumes that the explanatory variables are distributed with a multivariate normal distribution. There are two types of MDA according to functional form: LDA(linear discriminant analysis) and QDA(quadratic discriminant analysis). LDA has been far more frequently used than QDA in business classification research.

Two years after Beaver's(1966) univariate approach, Altman (1968) introduced the LDA to bankruptcy prediction. As expected, the accuracy achieved by Altman's multivariate approach was higher than that by Beaver's univariate approach. Studies using LDA include Altman, Haldeman, and Narayanan (1977) and Mutchler (1985). LDA assumes that the variance-covariance matrices of groups are the same. Violation of these assumption may reduce the power of the test.

Probit and logit models evolved from the traditional regression model. Probit and logit (generally qualitative response (QR) models) postulate a line of causality running from exogenously determined independent variables and stochastic errors to a discrete dependent variable. QR models are used to estimate the conditional probability of an event given explanatory variables.

The probit model assumes that the conditional probability of category membership follows the standard normal distribution. The underlying assumption of the normal distribution for the probit or MDA model is sometimes justified by applying the central limit theorem and considering the large number of factors that may influence the probabilities. Studies using the probit method include Kaplan and Urwitz (1979), Zmijewski and Hagerman (1981), and Zmijewski (1984).

The logit model is identical to the probit model except that the conditional probability of category membership follows the standard logistic distribution. The distribution and density functions of the standard logistic are very similar in shape to the standard normal distribution. It is known that the use of probit or logit makes little difference except when data are heavily concentrated in the tails. Studies using the logit method include Ohlson (1980) and Gentry, Newbold, and Whitford (1985).

Recursive partitioning, a recently developed statistical technique, was applied to commercial bank loan classifications by Marais, Pattel, and Wolfson (1984) and financial distress prediction by Frydman, Altman, and Kao (1985). The recursive partitioning procedure is a nonparametric method using recursive binary partitioning of the explanatory variables to classify observations. This method provides a treelike structure for classifying observations. Recursive partitioning selects and partitions the independent variable or linear combination of variables that most improves the homogeneity of class assignments based on misclassification costs and prior probabilities.

ARTIFICIAL INTELLIGENCE CLASSIFICATION TECHNIQUES

Learning is defined as changes in a system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the population more efficiently and effectively the next time (Simon, 1983). AI researchers have devoted much effort to implanting learning capabilities in computer software. The computer modeling of learning processes constitutes the field of machine learning (Carbonell, Michalski, and Mitchell, 1983).

The ability of human beings to make generalizations from scattered observations or to discover structures in collections of observations has been a long-standing issue of interest. The understanding of this process, called inductive inference, may be the key to an improvement of methods by which the computer can acquire knowledge. Inductive learning, a subfield of machine learning, is viewed as a heuristic search through a space of symbolic descriptions, generated by an application of various inference rules to the observational statements (Michalski, 1983).

Various inductive learning algorithms have been developed in AI domains. ID3 has been the most frequently used in inductive learning applications. ID3 (generally inductive learning algorithms) was developed for concepts which are qualitative in nature. Hence, it is expected that the ID3 method can deal with qualitative variables better than statistical methods. ID3 is a nonparametric method which may be preferred to statistical techniques based on specific assumptions when the data severely violate those assumptions. The ID3 method generates a decision tree, where each leaf node contains examples which are of the same class, based on Shannon's information theory.

Inductive learning methods have limitations in several ways. Messier and Hansen (1988) discussed three sources of limitations: the difficulty to apply to very large problem domains, the potential error introduced into a production system from omitting important instances or diagnostic attributes, and conflicting instances.

Scientists began to apply inductive learning to classification in early 1970's. Buchanan et al. (1976) developed a computer program called Meta-DENDRAL that assisted chemists in the discovery of rules from empirical data on mass-spectra. Meta-DENDRAL provided qualitative explanations of the characteristics of fragmentations and rearrangements among a set of molecules. Buchanan et al. showed that the inductive program was capable of rationalizing the mass-spectra data and suggested that it offered a powerful and useful complement to traditional methods for finding a structural relationship from spectral data.

It has been only a few years since business researchers began to apply the inductive learning approach to business domains. Braun and Chandler(1987) applied the inductive learning approach to predict stock market movements. Braun and Chandler suggested that inductive learning can be beneficial in developing a decision support system for market analysts or in developing their own decision processes.

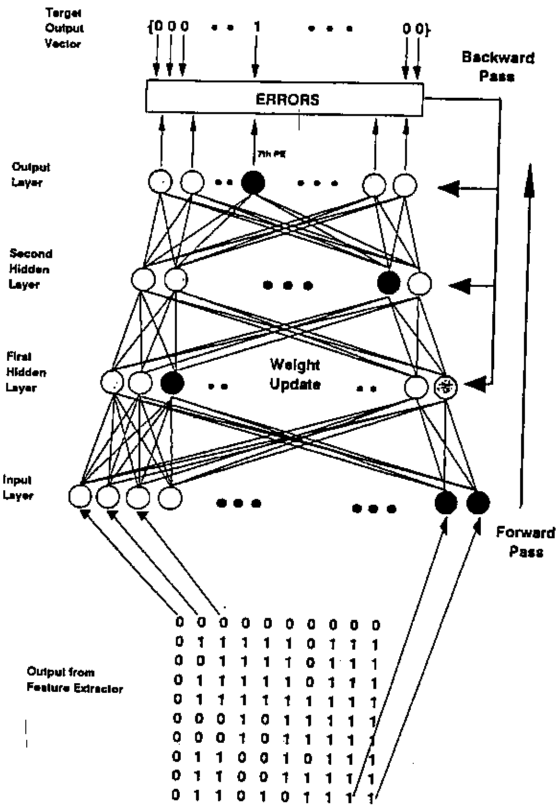
Garrison and Michaelsen (1989) applied inductive learning to analyze Tax Court cases in determining scholarship or fellowship grant status. They chose fourteen attributes which are all nominal. The results using the holdout technique indicated that the rules from ID3 were more accurate than MDA or logit at predicting the outcome of Tax Court decisions. The authors suggested that inductive learning algorithms may be more applicable to qualitative measurement situations such as tax studies than statistical techniques such as MDA or logit, which are based on interval parametric statistical theory.

Quite recently, the neural network approach has started to be applied to business classification. A neural network system consists of many simple interconnected PEs. Each PE calculate the wight of input value and applies the transfer function to generate the output value. The transfer function remains unchanged, but the weights for the linear combination can dynamically be adjusted to produce a desirable output. The PEs are grouped into linear arrays called layers. A neural network model has an input layer and an output layer, and may or may not have hidden layers.

A widely used learning algorithm is the back-propagation. The back-propagation algorithm can be applied to multilayer neural network. The application of back-propagation involves two phases. In the forward stage, the input is propagated forward through the network to compute the output value and the error. In the backward stage, the recursive computation of the error term is performed in a backward direction. The back-propagation algorithm converges very slowly and may get stuck at local minimum.

The following figure represent a network model with two hidden layers using back-propagation algorithm(Jhee and Lee, 1993).

Figure 1: Neural Network Model



The studies applying neural network models show that the neural network approach is very powerful tool for business applications. Liang, Chandler, Han, and Roan(1992) compares the performance of neural network, ID3, and probit models under various data conditions. The classification models include eight numeric variables (financial ratios and accounting numbers) and one nominal variable (industry classification). They shows that the neural network model mostly outperform the probit and ID3 methods in predicting accounting inventory method choice.

Tam and Kiang(1992), using bank failure data, compares a neural network approach with linear discriminant function, logit model, k nearest neighbor, and ID3. They shows that the neural net is a promising method of evaluating bank conditions in terms of predictive accuracy, adaptability, and robustness.

Surkan and Singleton(1990) applied neural networks with single and multiple hidden layers for bond rating. The results show that neural networks models with single or multiple hidden layers outperform discriminant analysis. They also show that networks trained with two hidden layers outperform a network having only one hidden layer containing a comparable number of PEs.

As discussed above, there exist various classification techniques available to researchers. The previous studies shows that neural network model is a very promising classification tool. The neural network contains nonparametric advantages like ID3 and has provided better results than ID3. The neural network model with two hidden layers using the back-propagation algorithm will be applied to Korean commercial paper rating data. The previous studies show that ID3 is not good at predicting numerical data though it could be very powerful when dealing with qualitative data. The benchmark for comparison will be the discriminant analysis which has been used mostly widely in business classification research.

CREDIT RATING

Moody's started rating railroad bonds in 1909 and began to issue rating for utility and industrial debts in 1914. Poor's began to issue ratings in 1922 and Standard Statistics started issuing rating in 1923. The two companies merged into Standard and Poor's in 1941. The credit rating in Korea has started in 1985 when the Korea Credit Rating Corporation was founded. Now, there are three professional credit rating agencies in Korea which rate the corporate bond and commercial papers. Korean history of credit rating is quite short compared to that of the United States.

The previous studies show that about two-thirds of bond rating can be predicted with simple model using financial information(see Kaplan and Urwitz(1979) for a review of these studies). It will be interesting to test whether the credit rating in Korea can be predicted like those in the United States or other developed countries. In Korea, the number of companies whose bond have been rated is not large enough for an empirical analysis while relatively large number commercial papers have been rated. In this study, the rating of commercial papers published by Korea Credit Rating Corporation will be used for the analysis.

Korea Credit Rating Corporation provides 6 grades for commercial papers:

Table 1: Commercial Paper Grade

Grade	Implications
A1	The ability of timely payment and its safety are the best.
A2	The ability of timely payment is excellent but its safety is worse than A1.
A3	The ability of timely payment and its safety are good but worse than A2.
B	The ability of timely payment is fine but its safety is volatile depending on the situation.
C	The ability of timely payment and its safety are very volatile.
D	no grade

The job of a professional credit agency is to determine the ability of an industrial company with regard to the timely payment of principal and interests. The Standard & Poor's rating methodology profile(Standard & Poors, 1986) addresses the following nine criteria categories: industry risk, issuer's industry position-market position, issuer's industry position-operating efficiency, management evaluation, accounting quality, earnings protection, leverage and asset protection, cash flow adequacy, financial flexibility. The first four categories are oriented to business analysis, the remainder to financial analysis.

The business analysis data are the output of time-consuming work of credit analysts and may be confidential. Thus, it is difficult to use these data for research. The financial analysis data are mostly public and available for a research. It is why the previous studies use financial data for credit analysis. This study is not an exception. However, the future research should incorporate business analysis to develop a practicable credit rating system.

The total sample available includes 216 companies whose commercial papers were rated in 1985, 1986, and 1987. None of commercial papers have been rated C or D. Therefore, only four grades (A1, A2, A3, and B) are included in the study as the followings:

Table 2: Sample

Grade	Number of companies
A1	24
A2	47
A3	65
B	80
Total	216

When classification accuracy is estimated from the same sample used for model specification, the estimate of classification accuracy is biased upward because the classification model is tailored to the data. Techniques to avoid the overfitting problem include the holdout technique, jackknife procedure, and bootstrapping. The holdout technique that is the most frequently used in classification research is used in this study. The testing sample is a set of 80 companies selected randomly. The rest is used as a training sample. The performance measure used in this study is the classification accuracy which is defined as the number of holdout cases correctly predicted divided by the total number of holdout cases.

It is ideal for a researcher to use a theory in choosing financial variables that will predict credit rating. However, the lack of theory in this issue is a long-standing problem for financial analysts. The 26 financial variables to be adopted in this study are those used by credit analysts in the Korea Credit Rating Corporation.

DISCRIMINANT ANALYSIS APPLICATIONS

The lack of theory led researchers to consider a multitude of variables. A model that includes too many variables may be overfitted to the training sample. Though the overfitted model may be highly successful in classifying the training sample, it is expected to be less effective in predicting the testing sample. In addition, inclusion of overlapped ratios results in multicollinearity among ratios and distorts the relationship between independent and dependent variables.

The stepwise selection method selects independent variables for entry into analysis on the basis of their discriminating power. In many instances

Table 3: Financial Variables

SA	sales
TA	total assets
SE	stockholders' equity
SAGR	sales growth
TAGR	total asset growth
SEGR	stockholders' equity growth
NITA	net income/total assets
NIBC	net income/business capital
NISE	net income/stockholders' equity
ORTA	ordinary income/total assets
OPSA	operating income/sales
ORSA	ordinary income/sales
FESA	financial expenses/sales
TLSE	total liabilities/stockholders' equity
SETA	stockholders' equity/total assets
CLSE	current liabilities/stockholders' equity
FALC	fixed assets, investments, and other assets/fixed capital
LTTA	long-term liabilities/total assets
EBIN	earnings before interests and taxes/Interests
FASE	fixed assets, investments, and other assets/stockholders' equity
DRCF	debt repayment/cash flows
CFTL	cash flows/total liabilities
QACL	quick assets/current liabilities
CACL	current assets/current liabilities
SLAR	sales/accounts and notes receivable
SLIV	sales/inventories

the full set of independent variables contains excess information about the group differences, or perhaps some of the independent variables may not be very useful in discriminating among groups. By sequentially selecting the next best discriminator at each step, a reduced set of variables will be found which is almost as good as, and sometimes better than, the full set. In this study, the stepwise selection method using Rao's V_r, a generalized distance measure, is adopted. The variables selected is the one which contributes the largest increase in V when added to the previous variables.

Factor analysis is a statistical tool designed to group variables into a few factors that retain a maximum of information contained in the data. The factor analysis is performed to check the timeseries stability and pattern of financial ratios. The results of factor analysis shows that the financial ratios are stable in the reasonable pattern during the three year periods.

The results of discriminant analysis for the testing sample is as the following:

Table 4: Prediction Results Using Discriminant Analysis

<1987>

predicted / actual	A1	A2	A3	B
A1	5	2	1	0
A2	1	9	6	1
A3	1	3	15	6
B	0	2	4	24

<1986>

predicted / actual	A1	A2	A3	B
A1	2	4	2	0
A2	1	7	7	2
A3	2	3	16	4
B	2	2	5	21

<1985>

predicted / actual	A1	A2	A3	B
A1	3	4	1	0
A2	2	9	4	2
A3	2	6	13	4
B	0	3	9	18

The predictive accuracy is 66.25% for 1987, 57.50% for 1986, and 53.75% for 1985. The predictive power of financial variables are improving as time goes by from 1985 to 1987. The credit rating has begun in 1985. In the first year, the credit rating may have been less reliable. The credit rating system will be improved gradually. The financial variables will be more creditable as the economy advances. the financial variables which have large coefficients in the standardized discriminant function are stockholders' equity, total assets, total liabilities/stockholders' equity, ling-term liabilities/total assets, financial expenses/sales, earnings before interests and taxes/interests, etc.

NEURAL NETWORK APPLICATIONS

The design of input pattern, selection of PE, the number of PE in the hidden layer, the design of output pattern, and learning strategy are important factors in using the connectionist network algorithm. The financial variable is divided into five classes using the average and standard deviation of financial variable.

Table 5: Input Pattern

input pattern	rule
1 0 0 0 0	if $X \le EX-2S$
1 1 0 0 0	if $EX-2S<X \le EX-S$
1 1 1 0 0	if $EX-S<X \le EX+S$
1 1 1 1 0	if $EX+S<X \le EX+2S$
1 1 1 1 1	if $EX+2S<X$

EX is the industry average and S is the standard deviation of financial variable. The processing element include the summation function of inputs and the transformation function. The Sigmoidal nonlinear function is used in all layers but the input layer. The values of the function is greater than 0 and less than 1.

$$net_j = W_{ij} \Sigma X_i - \theta$$

where θ is threshold
X_i is the input from prior layer.

$$f(net_j) = 1/[1+\exp(-net_j)].$$

The number of PE in the input layer is 130(26 financial variables times 5 classes). There is no general rule on how many PE are the most appropriate in the hidden layers. It is known that the input pattern is decomposed if the number of PE in the hidden layer is greater than that in the input layer and that the input patter is summarized if the number of PE in the hidden layer is smaller than that in the input layer. In this study, there are two hidden layers and the number of PE in the hidden layer get smaller as approaching the output layer. The number of PE in the output layer is four because there are four grades of credit rating in this study. The value of the output layer is 1 for the PE representing the corresponding grade and 0 otherwise. For example, the A1 grade has the output value, (1 0 0 0).

In this study, the error back-propagation method, which is widely used, is adopted as the learning algorithm of connectionist network.

$$\Delta W_{ij} = \eta \delta_k o_j$$

$$\delta_k = (t_k - o_k) o_k (1 - o_k) \text{ for the output layer}$$

$$\delta_k = o_j (1 - o_j) \Sigma \delta_k W_{kj} \text{ otherwise}$$

where η=learning coefficient
δ=error signal
o=the output value of PE
t=the desired output value of output layer PE

The convergence of this learning algorithm can be explained by the Gradient Descent Algorithm(Jones and Hopkins, 1987).

The learning strategy consists of two steps. First, the five cases for each of four grades are selected. The twenty patterns are used for learning until all the patterns are correctly classified. Second, the whole training cases are added for learning to the network resulting from the first step. The reason for the two-step process is to minimize the error and expedite the convergence of network since the back-propagation algorithm does not guarantee these.

There are three kinds of network used in this paper. Network A use 80 PEs in the first hidden layer and 60 PEs in the second hidden layer, network B 60 PEs and 30 PEs, and network C 50 PEs and 40 PEs. Network B show the best predictive accuracy. It implies that the use of more PEs does not always lead to the better performance but only takes more time. The predictive accuracy of network B for the testing sample is as the following:

Table 6: Prediction Results Using Neural Network

<1987>

predicted / actual	A1	A2	A3	B
A1	8	0	0	0
A2	1	15	1	0
A3	0	2	22	1
B	0	1	1	28

<1986>

predicted / actual	A1	A2	A3	B
A1	6	2	0	0
A2	1	16	0	0
A3	0	3	21	1
B	0	1	2	27

<1985>

predicted / actual	A1	A2	A3	B
A1	7	1	0	0
A2	2	14	1	0
A3	1	4	21	3
B	0	1	4	25

The predictive accuracy is 91.25% for 1987, 87.5% for 1986, and 83.75% for 1985. The predictive performance has gradually improved like when using the discriminant functions. The predictive performance of discriminant analysis(DA) and neural network(NN) can be compared as the following:

Table 7: Comparison of Results between DA and NN

	1987	1986	1985
DA	66.25%	57.50%	53.75%
NN	91.25%	87.50%	83.75%

The difference in the predictive accuracy between discriminant analysis and neural network is more than statistically significant. The much better performance of connectionist network implies that the bond rating cannot be well explained by a simple linear model. The linear model can be represented as a network model without hidden layers. Therefore, the hidden layers in the network model make difference in the predictive performance.

CONCLUSION

The discriminant analysis using korean CP rating data shows the similar performance as the previous studies using the data of the US or developed countries. The neural network model shows much better performance than the discriminant analysis. This implies that the neural network approach is a powerful tool for bond rating research. However, the neural network model does not provide an logical explanation to the bond rating unlike the regression analysis. There is no general method to determine the relative importance of an input from the weights of a neural net. This is an important research issue to be studies in the future.

References

Altman, E. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." Journal of Finance, September 1968, 589-609.

Altman, E., R. Haldeman, and P, Narayanan. "ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations." Journal of Banking and Finance, 1977, 29-54.

Beaver, W., "Financial Ratios as Predictors of Failure." Journal of Accounting Research, supplement, 1966, 71-111.

Braun, H. and J. Chandler, "Predicting Stock Market Behavior Through Rule Induction: An Application of the Learning-From-Example Approach." Decision Science, 1987, 415-29.

Buchanan, B., D. Smith, W. White, R. Gritter, E. Feigenbaum, J. Lederberg, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference. 22. Automatic Rule Formation in Mass Spectrometry by means of the Meta-DENDRAL program." Journal of American Chemical Society, September 1976, 6168-6178.

Carbonell, J. G., R. S. Michalski, and T. M. Mitchell, "An Overview of Machine Learning." In R. S. Michalski, J. G., Carbonell, and T. M. Mitchell (Ed.), Machine Learning: An Artificial Intelligence Approach. Tioga, Palo Alto, Calif., 1983.

Frydman, H., E. I. Altman, and D. Kao, "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress." Journal of Finance, March 1985, 269-91.

Garrison, L. and R. Michaelson, "Symbolic Concept Acquisition: A New Approach to Determining Underlying Tax Law Constructs." Journal of American Taxation Association, Fall 1989, 77-91,

Gentry, J., P. Newbold, and D. Whitford, "Classifying Bankrupt Firms with Fund Flow Components." Journal of Accounting Research, Spring 1985, 146-60.

Jhee, W. and J. Lee, "Performance of Neural Networks in Managerial Forecasting." Intelligent Systems in Accounting, Finance, and Management, 1993, 55-71.

Jones, W. and J. Hopkins, "Back-propagation: A Generalized Delta Learning Rule." BYTE Magazine, October 1987.

Kaplan, R. S. and G. Urwitz, "Statistical Models of Bond Ratings: A Methodological Inquiry." Journal of Business, April 1979, 231- 62.

Liang, T., J. Chandler, I. Han, and J. Roan, "An Empirical Investigation of Some Data Effects on the Classification Accuracy of Probit, ID3, and Neural Networks." Contemporary Accounting Research, Fall 1992, 306-328.

Marais, M., J. Patell, and M. Wolfson, "The Experimental Design of Classification Models: An Application of Recursive Partitioning and Bootstrapping to Commercial Bank Loan Classifications." Journal of Accounting Research, supplement, 1984, 87-114.

Messier, W., Jr., and J. Hansen, "Inducing Rules for Expert System Development: An Example Using Default and Bankruptcy Data." Management Science, December 1988, 1403-1415.

Michalski, R. S., "A Theory and Methodology of Inductive Learning." in R. Michalski, J. Carbonell, T. Mitchell, (Ed.), Machine Learning: An Artificial Intelligence Approach. Tioga, Palo Alto, CA., 1983.

Mutchler, J. "A Multivariate Analysis of the Auditor's Going-Concern Opinion Decision." Journal of Accounting Research, Fall 1985, 668-82.

Ohlson, J. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." Journal of Accounting Research, Spring 1980, 109-31.

Simon, H. A., "Why Should Machine Learn." in R. S. Michalski, J. G. Carbonell, and T. M. Mitchell (Ed.), Machine Learning: An Artificial Intelligence Approach. 1983.

Standard and Poor's, Credit Overview: Industrial Rating, New york, 1986.

Surkan, A. and J. Singleton, "Neural Networks For Bond Rating Improved by Multiple Hidden Layers." Proceedings of the IEEE International Conference on Neural Network, San Diego, 1990, 157-162.

Tam, K. and M. Kiang, "Managerial Applications of Neural Networks: The Case of Bank Failure Predictions." Management Science, July 1992, 926-947.

Zmijewski, M., "Methodological Issues related to the estimation of financial distress prediction Models." Supplement to Journal of Accounting Research, 1984, 59-82.

Zmijewski, M. and R. Hagerman, "An Income Strategy Approach to the Positive Theory of Accounting Standard Setting/Choice." Journal of Accounting and Economics, August 1981, 129-149.