DEVELOPING FINANCIAL SOUNDNESS PREDICTION MODEL USING NEURAL NETWORK

V.Bindu Neeharika, VVS Kesava Rao, Ch. Maheswara Rao, I. Indra Kiran,
Asst.Professor, Department of Mechanical Engineering, ANITS, Sangivalasa-531162, VSP
Professor, Department of Mechanical Engineering, Andhra University, VSP
Asst.Professor, Department of Mechanical Engineering, ANITS, Sangivalasa-531162, VSP
Professor, Department of Mechanical Engineering, ANITS, Sangivalasa-531162, VSP

ABSTRACT

One interesting area for the use of neural networks is in event prediction. This study develops a neural network model for prediction of financial soundness of business organizations and tests it using financial data from these organizations. Multivariate discriminant analysis is a more traditional method of prediction of financial soundness of organizations. The same set of data is analyzed using neural network method. A comparison of the predictive abilities of both the neural network and the discriminant analysis method is presented. The results show that neural networks might also be applicable to predict financial soundness of business organizations.

Keywords: artificial intelligence, artificial neural network, multi-layer perception, radial basis function, multivariate discriminant analysis, alpha error, beta error.

1. INTRODUCTION

Paula M. Weller (2010) has conducted a test of several models of artificial networks and multiple and probit analysis to predict bankruptcy of United States of America’s textile industry. In this study, financial data of 47 bankrupt companies and 104 public shares non-bankrupt companies of textile industry has been examined during the years of 1998 to 2004 which includes Asia’s currency crisis and competition increase on behalf of China. Results show that in case of bankrupt companies, Altman model (1968) and neural network model based on its variables has the most prediction ability for one or two years prior to bankruptcy; and neural network model based on Zmijewski’s variables (1984) and also neural network model based on Altman variables (1983) show the best classification results in case of non-bankrupt companies during the entire period of research. Hung & Chen’s study (2009) deals with the prediction of companies’ bankruptcy probability using Decision Tree Algorithm, back propagation neural network and Support Vector techniques and shows that back propagation neural network has had a better performance than the other two methods. In this study, bankruptcy prediction model based on decision tree and support vector techniques has had an accuracy equivalent to 70% and 70.89% respectively, while back propagation neural network model has had an accuracy equivalent to 72.37%. Sungbincho, Kim and Kwon Bae’s study (2009) compares logit bankruptcy prediction model, artificial neural network, combinatory multiple discriminant analysis, decision tree, support vector technique and combinatory neural network to each other and presents a combinatory model of bankruptcy prediction using neural network learning. Its results show that artificial neural network has a better performance than the other models in case of classification accuracy. Average prediction accuracy of the aforementioned models are listed as 78.04%, 78.01%, 78.15%, 72.38%, 78.01%, and 78.92% respectively. Using fuzzy logic techniques, Slim (2007) presents a bankruptcy prediction model in his study and believes that this prediction method has a better predicting performance than the other methods like back propagation neural network and linear discriminant analysis. In this study 17 financial ratios of 68 companies from 2000 to 2005 have been used. The result of using five layer fuzzy neural networks confirms researcher’s reasoning. In average, bankruptcy prediction accuracy of linear discriminant analysis, back propagation neural network, and fuzzy neural network models in the training sample have been 70.83%, 81.92%, and 97.92% respectively, and in the test sample have been 60%, 70%, and 90% respectively. This shows that fuzzy neural network has had a better performance in case of prediction accuracy in both training and test samples. Kim and Gu's study (2006) deals with the comparison of performance accuracy to predict bankruptcy of two models of multiple discriminant analysis and logit analysis using financial data of 36 bankrupt and non-bankrupt companies from 1986 to 1998. 12 independent variables have been used including financial ratios and the results show that logit model's performance accuracy to predict bankruptcy has been 93% which is the same as that of multiple discriminant analysis and they both have had similar performance.

2. METHODOLOGY

2.1. Neural Network Analysis Through SPSS

The methodology of neural network analysis through SPSS is explained in the following steps:
Step1: Consideration of dependent variable
The dependent variable may be i) Nominal. ii) Ordinal. iii) Scale

Step2: Consideration of Predictors.
These are independent variables to predict the dependent variable. In this study, 20 financial ratios under 4 four categories are considered to develop the prediction model for financial soundness of steel manufacturing organizations.

Step 3: Partitioning the data set
Partitioning of the dataset into training, testing, and holdout samples is necessary in neural network analysis. The training sample comprises the data records used to train the neural network.

Step4: Specify the Architecture
The structure of the network is need to be specified. The procedure can select the "best" architecture automatically, or you can specify a custom architecture. Automatic architecture selection builds a network with one hidden layer. Specify the minimum and maximum number of units allowed in the hidden layer, and the automatic architecture selection computes the "best" number of units in the hidden layer. Automatic architecture selection uses the default activation functions for the hidden and output layers. Custom architecture selection gives you expert control over the hidden and output layers and can be most useful when you know in advance what architecture you want or when you need to tweak the results of the Automatic architecture selection.

Step5: Training the data set

Step6: Specify the Optimization Algorithm: This is the method used to estimate the synaptic weights. There are the following methods for optimization algorithm in training the data set

- Scaled conjugate gradient: Conjugate gradient methods apply only to batch training types, so this method is not available for online or mini-batch training.
- Gradient descent. This method must be used with online or mini-batch training. it can also be used with batch training.

Step 7: Specify the Training Options: The training options allow to fine-tune the optimization algorithm. Training options for the scaled conjugate gradient algorithm include the following values.

- Initial Lambda: The initial value of the lambda parameter for the scaled conjugate gradient algorithm. Specify a number greater than 0 and less than 0.000001.
- Initial Sigma: The initial value of the sigma parameter for the scaled conjugate gradient algorithm. Specify a number greater than 0 and less than 0.0001.

- Interval Center and Interval Offset: The interval center (a, 0) and interval offset (a) define the interval [a 0−a, a 0+a], in which weight vectors are randomly generated

Step 8: Record the Output.

3. RESULTS AND DISCUSSION

The aim of this study was to examine whether a MLP and RBF neural networks can help to correctly predict financial soundness (High or Low), by analyzing data obtained from the annual reports from FY 2102-13 to FY 2016-17 of the 5 steel manufacturing organizations. Table * gives information about the datasets used to build the ANN model. From the table it is observed that the training dataset contains in 92% of the sample and testing dataset contains 8% of the sample.

Multi-layer Perceptron Neural Network:

Table -1: Case Processing Summary

		N	Percent
Sample	Training	25	100.0%
Valid		25	100.0%
Excluded		0	
Total		25	

The Table -1 shows the number of neurons in every layer and one independent variable (Financial Soundness group) denoted as cluster number (CLN). Automatic architecture selection chose 8 nodes for the hidden layer, while the output layer had 2 nodes to code the depended variable financial soundness. For the hidden layer the activation function was the hyperbolic tangent, while for the output layer also the softmax function is used.

Table-2 : Network Information

Input Layer	Factors	1	PSR1
		2	PSR4
		3	PSR6
		4	PSR7
		5	PR1
		6	PR3
		7	PR5
		8	PR8
		9	PR9
		10	LR1
		11	LR2
		12	LR3
		13	LR4
		14	LR6
		15	VR1
		16	VR2
		17	VR3
		18	VR4
		19	VR5

		20	VR8
	Number of Units[a]		446
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		5
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	CLN
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy
a. Excluding the bias unit			

The model summary, shown in Table -2, provides information related to the results of training and testing sample. Cross entropy error is given for both training and testing sample since is the error function that network minimizes during the training phase. The small value (2.279 E-6) of this error indicates the power of the model to predict financial soundness. The cross entropy error (8.695 E-8) is also very less for the testing data set, meaning that the network model has not been overfitted to the training data. The result justifies the role of testing sample which is to prevent overtraining. From the results, it is observed that, there are no incorrect predictions based on training and testing sample.

Table-3: Model Summary

Training	Cross Entropy Error	.016
	Percent Incorrect Predictions	0.0%
	Stopping Rule Used	Training error ratio criterion (.001) achieved
	Training Time	0:00:00.05

Dependent Variable: CLN

Table 3 displays classification for categorical dependent variable (financial Soundness), by partition and overall. For each case, the predicted outcome is defined as success if the predicted probability is greater than 0.5.

As can be seen, the MLP network correctly classified 23 steel manufacturing organizations, out of 23, in the training sample and three out of three in testing sample. Overall 100.0% of the training cases were correctly classified.

Table -4: Classification

		Predicted		
Sample	Observed	1	2	Percent Correct
Training	1	12	0	100.0%
	2	0	13	100.0%
	Overall Percent	48.0%	52.0%	100.0%

Dependent Variable: CLN

1-Low Financial Soundness; 2-High Financial Soundness;
The ROC curve is a diagram of sensitivity versus specificity that shows the classification performance for all possible cutoffs. Figure -1 gives the sensitivity and specificity chart, based on the combined training and testing samples. The 45-degree line from the upper right corner of the chart to the lower left represents

the scenario of randomly guessing the class. The more the curve moves away the 45-degree baseline, the more accurate is the classification. In this study, these curves are at maximum distance from the 45-degree line.

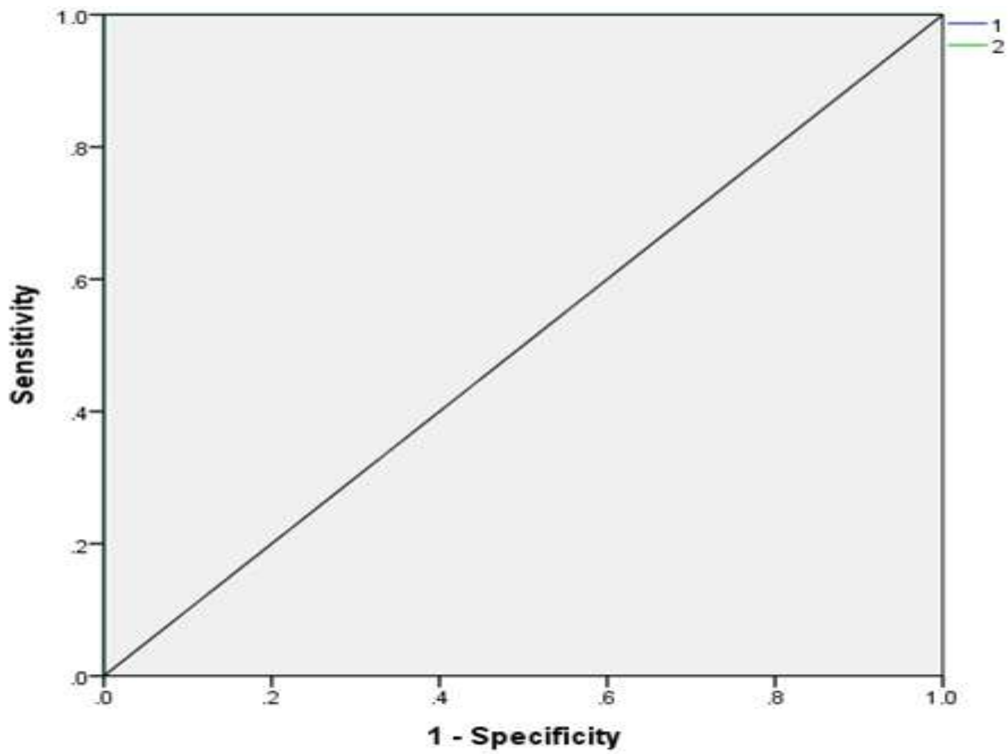

Figure -1:ROC Curve

Table-4 gives the area under the ROC curve. The area value 0.097 show that, if a steel manufacturing organization from the high financial soundness category and another from the low financial soundness category are randomly selected, there is complete certainty for these steel manufacturing organizations of being in the respective categories.

Table 5: Financial Soundness Prediction through MLP

Steel Manufacturing Organization	Year	Actual Category	Predicted Category
SMO1	2013	1	1
SMO2	2013	1	1
SMO3	2013	2	2
SMO4	2013	1	1
SMO5	2013	2	2
SMO1	2014	1	1
SMO2	2014	1	1
SMO3	2014	2	2
SMO4	2014	1	1
SMO5	2014	2	2
SMO1	2015	1	1
SMO2	2015	1	1

SMO3	2015	2	2
SMO4	2015	2	2
SMO5	2015	2	2
SMO1	2016	1	1
SMO2	2016	2	2
SMO3	2016	2	2
SMO4	2016	1	1
SMO5	2016	2	2
SMO1	2017	1	1
SMO2	2017	2	2
SMO3	2017	2	2
SMO4	2017	1	1
SMO5	2017	2	2

Radial Basis Function (RBF) Neural Network

RBF neural networks is also adopted to predict financial soundness (High or Low), by analyzing data obtained from the annual reports from FY 2102-13 to FY 2016-17 of the 5 steel manufacturing organizations. Table * gives information about the datasets used to build the RBF ANN model. From the table it is observed that the training dataset contains in 88% of the sample and testing dataset contains 12% of the sample

Table 6:Case Processing Summary

		N	Percent
Sample	Training	25	100.0%
Valid		25	100.0%
Excluded		0	
Total		25	

The Table 6 shows the number of neurons in every layer and one independent variable (Financial Soundness group) denoted as cluster number (CLN). Automatic architecture selection chose 8 nodes for the hidden layer, while the output layer had 2 nodes to code the depended variable financial soundness. For the hidden layer the activation function was the softmax, while for the output layer also the identity function is used.

Table 7:Network Information

Input Layer	Factors	1	PSR1
		2	PSR4
		3	PSR6
		4	PSR7
		5	PR1
		6	PR3
		7	PR5
		8	PR8
		9	PR9
		10	LR1
		11	LR2
		12	LR3
		13	LR4

		14	LR6
		15	VR1
		16	VR2
		17	VR3
		18	VR4
		19	VR5
		20	VR8
	Number of Units		446
Hidden Layer	Number of Units		2[a]
	Activation Function		Softmax
Output Layer	Dependent Variables	1	CLN
	Number of Units		2
	Activation Function		Identity
	Error Function		Sum of Squares
a. Determined by the Bayesian Information Criterion: The "best" number of hidden units is the one that yields the smallest BIC in the training data.			

The model summary, shown in Table 7, provides information related to the results of training and testing sample. Cross entropy error is given for both training and testing sample since is the error function that network minimizes during the training phase. The small value (1.000) of this error indicates the power of the model to predict financial soundness. The sum of squares error (0.0669) is also very less for the testing data set, meaning that the network model has not been overfitted to the training data. The result justifies the role of testing sample which is to prevent overtraining. From the results, it is observed that, there are no incorrect predictions based on training and testing sample.

Table 8:Model Summary

Training	Sum of Squares Error	5.936
	Percent Incorrect Predictions	48.0%
	Bayesian Information Criterion (BIC)	2818.657[a]
	Training Time	0:00:00.19

Dependent Variable: CLN

a. The number of hidden units is determined by the Bayesian Information Criterion: The "best" number of hidden units is the one that yields the smallest BIC in the training data.

Table 9: Classification

		Predicted		
Sample	Observed	1	2	Percent Correct
Training	1	8	4	66.7%
	2	8	5	38.5%
	Overall Percent	64.0%	36.0%	52.0%

Dependent Variable: CLN

ROC Curve

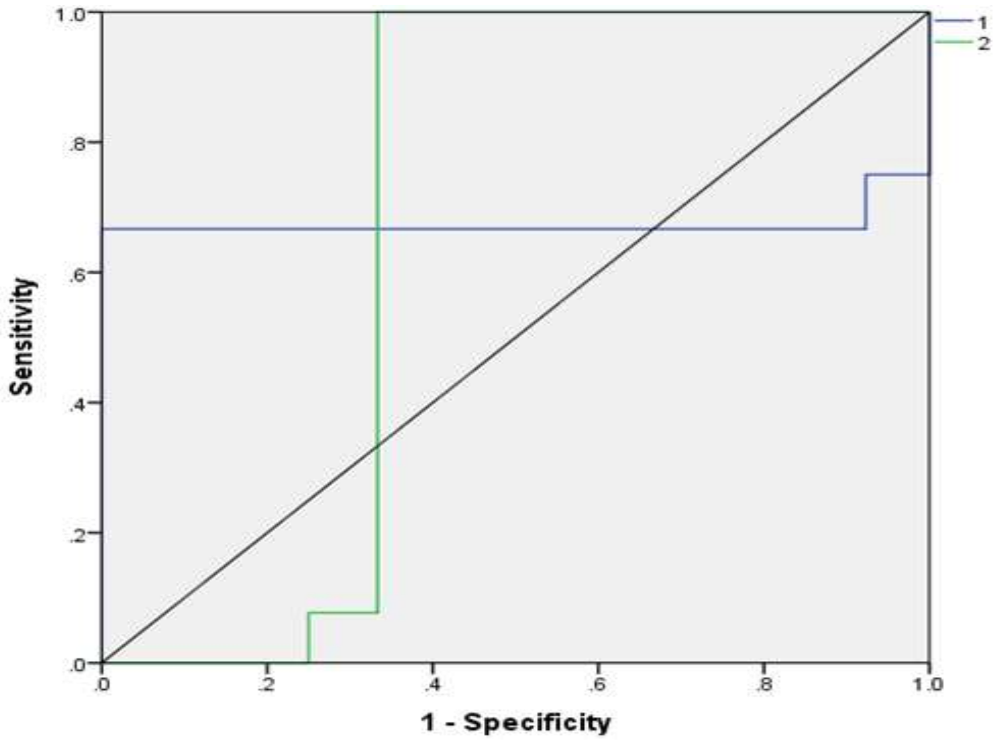

Figure 2: ROC Curve

Table 10:Financial Soundness Prediction using RBF

Steel Manufacturing Organization	Year	Actual Category	Predicted Category
SMO1	2013	1	1
SMO2	2013	1	1
SMO3	2013	2	1
SMO4	2013	1	1
SMO5	2013	2	1
SMO1	2014	1	1
SMO2	2014	1	1
SMO3	2014	2	2
SMO4	2014	1	1
SMO5	2014	2	1
SMO1	2015	1	1
SMO2	2015	1	1
SMO3	2015	2	1
SMO4	2015	2	2
SMO5	2015	2	1
SMO1	2016	1	2
SMO2	2016	2	2
SMO3	2016	2	1
SMO4	2016	1	2
SMO5	2016	2	1
SMO1	2017	1	2

SMO2	2017	2	2
SMO3	2017	2	2
SMO4	2017	1	2
SMO5	2017	2	1

COMPARISON OF MLP AND RBF

The comparison of the two models (MLP and RBF) in terms of predictability shows that the MLP shows good prediction.

Table 11: COMPARISON OF MLP AND RBF

	MLP-NNA	RBF-NNA
Area Under ROC Curve	1.00	0.673
Correct Classification(%)	100%	52%
Sum of Squares Error	0.016	5.936

Thus, artificial neural networks appear to be a powerful tool for the prediction of the financial distress of companies. This work goes hand in hand with the empirical studies already established (Kerling & Podding, 1994; Oden & Sharada, 1990; Abdou et al., 2008; Khashman, 2011).

Artificial neural network models are increasingly used in scoring with varying success. According to some statisticians, although these new methods are interesting and sometimes more efficient than traditional statistical techniques, they are also less robust and less well founded. Furthermore, neural networks are unable to explain the results they provide. Finally, they are as black boxes with unknown operating rules. They create their own representation in learning. In terms of interpretation of weights, discriminates analysis seems to be more efficient.

4. Concluding Remarks

The aim of this paper is to determine the effectiveness of artificial neural networks in predicting financial soundness, based on financial ratios data collected from annual reports of five steel manufacturing organizations during FY 2012-13 to FY 2016-17. Also, the results of the neural network analysis.

The literature review indicated that neural networks outperform all other classifiers, regarding prediction accuracy. A multilayer perceptron and Radial Basis function neural networks were trained, to predict financial soundness. The classification accuracy rate of multilayer perceptron and Radial Basis function neural networks was very high, with 100% and 95.5% respectively. The results also showed that the most powerful predictors of financial soundness. Although future work will need to validate these findings in larger and more diverse samples, there is strong evidence that the proposed model can be used effectively to predict financial soundness of business organizations in general and steel manufacturing organizations in particular and to help the management to design interventions that increase the financial soundness.

References

[1] Altman, E. I. (1983). Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing With Bankruptcy. New York: John Wiley & Sons, Inc.

[2] Beaver. W. H. (1966). Financial ratios as predictors of bankruptcy. Journal of Accounting Research, 4, 71-111. http://dx.doi.org/10.2307/2490171

[3] Chen, W. S., & Du, Y. K. (2009). Using neural networks and data mining techniques for the financial distress prediction model.

Expert systems with applications, http://dx.doi.org/10.1016/j.eswa.2008.03.020
[4] Gordon, M. J. (1971). Towards a Theory of Financial Distress. The Journal of Finance, 26, 347-356. http://dx.doi.org/10.1016/j.eswa.2008.03.020
[5] Haykin, S. (1999). Neural Networks: A Comprehensive Foundation (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.
[6] Hung, C., & Chen, J. H. (2009). A selective ensemble based on expected probabilities for bankruptcy prediction. Expert Systems with Applications, 36(3), 5297-5303. http://dx.doi.org/10.1016/j.eswa.2008.06.068
[7] Jones, F. J. (1987). Current Techniques in Bankruptcy Prediction. Journal of Accounting Literature, 6, 131-164. Kim, H., & Gu, Z. (2006). Predicting Restaurant bankruptcy: A Logit Model in comparison with a Discriminant Model. Journal of Hospitality Tourism research, 30, 474-793. http://dx.doi.org/10.1177/1096348006290114
[8] McCulloch, W. S., & Pitts, E. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 5, 115-133. http://dx.doi.org/10.1007/BF02478259
[9] Norusis, M. J. (2005). SPSS 13.0 Guide To Data Analysis. Paula, M. W. (2010). The Application of Altman, Zmijewski and Neural Network Bankruptcy Prediction Models to Domestic Textile-Related Manufacturing Firms: A Comparative Analysis. A dissertation of H. Wayne Huizenga School of Business and Entrepreneurship Nova Southeastern University.
[10] Slim, C. (2007). Fuzzy Neural Model for Bankruptcy Prediction. The Journal of Business Review Cambridge, 8(2), 117-122.
[11] Sungbin, C., Jinhwa, K., & Jae, K. B. (2009). An integrative model with subject weight based on neural network learning for bankruptcy prediction. http://dx.doi.org/10.1016/j.eswa.2007.09.060. Expert Syst. Appl., 36(1), 403-410.
[12] Whitaker, R. (1999). The Early Stage of Financial Distress. Journal of Economics and Finance, 23(2), 123-133. http://dx.doi.org/10.1007/BF02745946
[13] Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research, 24, 59-82. http://dx.doi.org/10.2307/2490859