Evaluation of Banks Insolvency Using Artificial Neural Networks

QEETHARA KADHIM AL-SHAYEA
MIS Department
Al Zaytoonah University of Jordan
Amman - Jordan
kit_alshayeh@yahoo.com

GHALEB AWAD EL-REFAE
Finance and Banking Department
Al Ain University of Science and Technology
U.A.E.
ghalebelrefae@yahoo.com

Abstract: - Bankruptcy prediction has been an important and widely studied topic. The goal of this study is to predict bank insolvency before the bankruptcy using artificial neural networks, to enable all parties to take remedial action. Artificial neural networks are widely used in finance and insurance problems. Generalized Regression Neural Network (GRNN) is used to evaluate the predictor variable used to predict the insolvency. The most important predictor variable influencing insolvency is consistently having the largest regression. Results showed that the most affecting factor in banks insolvency evaluation is the net income, total equity capital, cost of sales, sales, cash flows and loans. The Feed-forward back propagation neural network is used to predict the bankruptcy. The results of applying Feed-forward back propagation neural network methodology to predict financial distress based upon selected financial ratios show abilities of the network to learn the patterns corresponding to financial distress of the bank. The percent correctly classified in the training sample by the feed-forward back propagation network is approximately 91 percent. Artificial neural networks show significant signs for providing early warning signals and solvency monitoring. The proposed neural network is evaluated using confusion matrices.

Key-Words: - Artificial Neural Networks, General Regression Network, Feed-forward Back Propagation Neural Network, Regression, Financial Distress Analysis, and Receiver Operating Characteristic (ROC) Curves.

1 Introduction

Bankruptcy prediction has been an important and widely studied topic. The prediction of the likelihood of failure of a company given a number of financial measures, how soon an “ill” business can be identified, possibility of identifying the factors that put a business at risk — these are of main interest in bank lending [1]. Recently researchers have used neural networks as a bankruptcy classification models. Artificial neural networks showed accurate results as discriminant analysis to early detect bank failures [2].

Beaver [3], Altman [4], Williams and Goodman [5], Sinkey [6], and Altman, Haldeman, and Narayanan [7] have used discriminant analysis to solve bankruptcy prediction problem.

Kaski, Sinkkonen, and Peltonen [8] introduced a method for deriving a metric, locally based on the Fisher information matrix, into the data space. A self-organizing map (SOM) is computed in the new metric to explore financial statements of enterprises. The metric measures local distances in terms of changes in the distribution of an auxiliary random variable that reflects what is important in the data.

Pramodh and Ravi [9] proposed and implemented a variant for Baek and Cho’s [10] neural network and named it Modified Great Deluge Algorithm based Auto Associative Neural Network (MGDAAANN), wherein a metaheuristic is used to train the auto associative neural network.

Kim and Kang [11] proposed an ensemble with neural network for improving the performance of traditional neural networks on bankruptcy prediction tasks.

Nachev, Hill and Barry [12] showed the potential of neural networks based on the Adaptive Resonance Theory as tools that generate warning signals when bankruptcy of a company is expected (bankruptcy prediction problem). Using that class of neural networks is still unexplored to date. They examined four of the most popular networks of the class fuzzy, distributed, instance counting, and default ARTMAP.

Ravisankar [13] presented three hitherto unused neural network architectures for bankruptcy prediction in banks. These networks are Group Method of Data Handling (GMDH), Counter Propagation Neural Network (CPNN) and fuzzy Adaptive Resonance Theory Map (fuzzy ARTMAP). Efficacy of each of these techniques is

ISBN: 978-1-61804-068-8

tested by using four different datasets pertaining to Spanish banks, Turkish banks, UK banks and US banks. Further t-statistic, f-statistic and GMDH are used for feature selection purpose and the features so selected are fed as input to GMDH, CPNN and fuzzy ARTMAP for classification purposed.

2 Artificial Neural Networks

An artificial neural network (ANN) is a computational model that attempts to account for the parallel nature of the human brain. An (ANN) is a network of highly interconnecting processing elements (neurons) operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layer(s) of units, called hidden layer(s). Fig. 1 represents the typical neural network. You can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements.

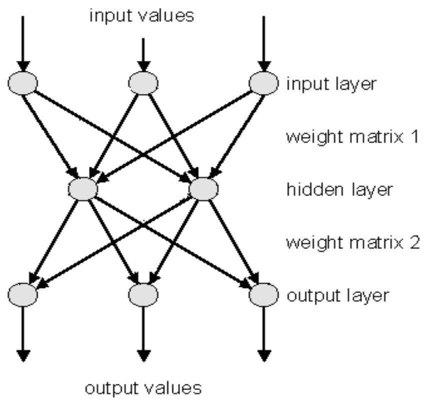

Fig. 1 A typical neural network

For the researcher and the financial analyst, the main advantage of ANNs is that there is no need to specify the functional relation between variables. Since they are connectionist-learning machines, the knowledge is directly imbedded in a set of weights through the linking arcs among the processing nodes. In order to train a neural network properly one needs a large set of representative 'good quality' examples. In the case of bankruptcy problems, the researcher should be cautious when drawing conclusions from neural networks trained with only one or two hundred cases, as observed in most previous studies [14].

2.1 Generalized Regression Neural Network

The GRNN was applied to solve a variety of problems like prediction, control, plant process modeling or general mapping problems [15].

General regression neural network Specht [16] and Nadaraya [17] , does not require an iterative training procedure as in back-propagation method.

The GRNN is used for estimation of continuous variables, as in standard regression techniques. It is related to the radial basis function network and is based on a standard statistical technique called kernel regression. By definition, the regression of a dependent variable y on an independent x estimates the most probable value for y, given x and a training set. The regression method will produce the estimated value of y, which minimizes the mean-squared error. GRNN is a method for estimating the joint probability density function (pdf) of x and y, given only a training set. Because the pdf is derived from the data with no preconceptions about its form, the system is perfectly general. Furthermore, it is consistent; that is, as the training set size becomes large, the estimation error approaches zero, with only mild restrictions on the function. In GRNN, instead of training the weights, one simply assigns to w_{ij} the target value directly from the training set associated with input training vector i and component j of its corresponding output vector [18]. GRNN architecture is given in Fig. 2.

GRNN is based on the following formula [19]:

$$E[y \mid x] = \frac{\int_{-\infty}^{\infty} y.f(x,y).dy}{\int_{-\infty}^{\infty} f(x,y).dy} \quad (1)$$

where y is the output of the estimator, x is the estimator input vector, $E[y|x]$ is the expected output value, given the input vector x and $f(x,y)$ is the joint probability density function (pdf) of x and y.

The function value is estimated optimally as follows:

$$y_j = \frac{\sum_{i=1}^{n} h_i.w_{ij}}{\sum_{i-1}^{n} h_i} \quad (2)$$

where w_{ij}= the target output corresponding to input training vector x_i,

$h_i = e^{\frac{-D_i^2}{2.spread^2}}$, the output of the hidden layer neuron, $D_i^2 = (x-u_i)^T(x-u_i)$, the squared distance

ISBN: 978-1-61804-068-8

between the input vector x and the training vector u,
x= the input vector,
u_i=training vector i, the center of neuron i, spread=a constant controlling the size of the receptive region.

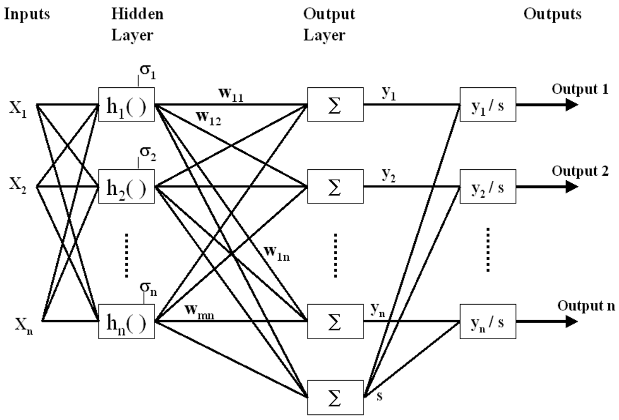

Fig. 2 Generalized Regression Neural Network (GRNN) Architecture

2.2 Feed-Forward Back Propagation Neural Network

Feed-forward neural networks are widely and successfully used models for forecasting and problem solving. A typical feed-forward back propagation neural network consists of three layers: the input layer, a hidden layer, and the output layer. A feed-forward fully connected network is trained in supervised manner. Based on the way they learn, all artificial neural networks are divided into two learning categories: supervised and unsupervised. In supervised learning, the network is trained by providing it with input and output patterns. During this phase, the neural network is able to adjust the connection weights to match its output with the actual output in an iterative process until a desirable result is reached.

It allows signals to travel one-way only; from source to destination; there is no feedback. The hidden neurons are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs. Each neuron in the hidden layer uses a transfer function to process data it receives from input layer and then transfers the processed information to the output neurons for further processing using a transfer function in each neuron

The output of the hidden layer can be represented by

$$YNx1 = f(WNxM\ XM,1+ bN,1) \quad (3)$$

where Y is a vector containing the output from each of the N neurons in a given layer, W is a matrix containing the weights for each of the M inputs for all N neurons, X is a vector containing the inputs, b is a vector containing the biases and $f(\cdot)$ is the activation function [20].

3 Experimental Results

3.1 Data

Spanish banking industry suffered the worst crisis during 1977-1985 resulting in total cost 12 billion dollars. The Spanish bank's data is obtained from Olmeda and Fernandez [21]. This dataset contains 66 banks where 37 of them went bankrupt and the rest solvent. Table 1 presents the financial ratios which are considered as predictor variables used in the study. The first three predictor variables are liquidity ratios, whilst the fourth measures the self-financing capacity of the bank. Ratios five, six and seven relate profit to various items on the balance sheet. Ratio eight relates the cost of sales to sales and ratio nine relates the cash flow of the bank to the debts.

Table 1 Predictor variable of datasets

Spanish banks data	
S. No.	Predictor Variable Name
1	Current assets/total assets
2	Current assets-cash/total assets
3	Current assets/loans
4	Reserves/loans
5	Net income/total assets
6	Net income/total equity capital
7	Net income/loans
8	Cost of sales/sales
9	Cash flow/loans

3.2 Results Analysis

Generalized regression neural networks (GRNN) are a kind of radial basis network that is often used for function approximation. The GRNN with a radial basis layer and a special linear layer and linear output neurons was created using the neural network toolbox from Matlab 7.9. The insolvency output is used as a target to the network and the nine predictor variable mentioned in table 1 are used as an input to the network. A spread slightly lower than the distance between input values is used, in order, to get a function that fits individual data points fairly closely. The spread value was chosen 0.1.

Table 2 shows the results of regression values. It is obvious, that the most affecting factor in banks insolvency evaluation is the net income, total equity capital, cost of sales, sales, cash flows and loans. The multiple regressions for all the affecting factors show that percent correctly predicted in the simulation sample are approximately 80 percent.

ISBN: 978-1-61804-068-8

Table 2

Regression for Spanish banks data		
No.	Predictor Variable Name	Regression Value
1	Current assets/total assets	0.23513
2	Current assets-cash/total assets	0.13947
3	Current assets/loans	0.26166
4	Reserves/loans	0.23377
5	Net income/total assets	0.39355
6	Net income/total equity capital	0.69428
7	Net income/loans	0.39377
8	Cost of sales/sales	0.63899
9	Cash flow/loans	0.63897

A two-layer feed-forward network with 9 inputs and 20 sigmoid hidden neurons and linear output neurons was created as shown in Fig. 3. The dataset contains 66 samples. Training is done using scaled conjugate gradient back propagation network. The scaled conjugate gradient algorithm (SCG) developed by Moller [9] was designed to avoid the time-consuming line search. This algorithm combines the model-trust region approach with the conjugate gradient approach.

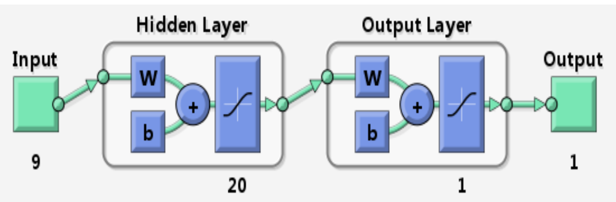

Fig. 3

The results of applying the proposed neural network to predict financial distress based upon selected financial ratios showed very good abilities of the network to learn the patterns corresponding to financial distress of the bank. The results were good; the network was able to classify approximately 94% of the cases in the training set as shown in Fig. 4. Fig. 4 shows the confusion matrices for training, testing, and validation, and the three kinds of data combined. The diagonal cells in each table show the number of cases that were correctly classified, and the off-diagonal cells show the misclassified cases. The blue cell in the bottom right shows the total percent of correctly classified cases (in green) and the total percent of misclassified cases (in red).

Fig.5 shows the training state values. Best validation performance is 0.020337 at epoch 19 as shown in Fig.6. The mean squared error (MSE) is the average squared difference between outputs and targets.

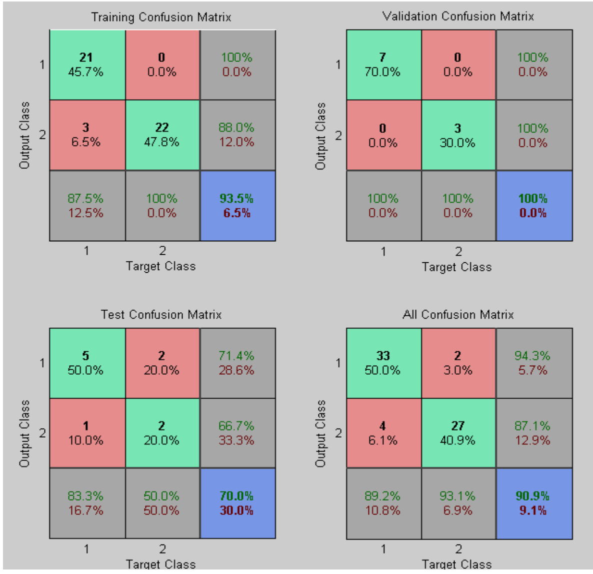

Fig. 4

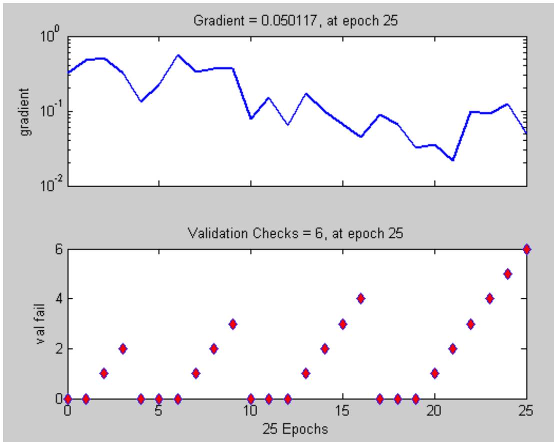

Fig. 5

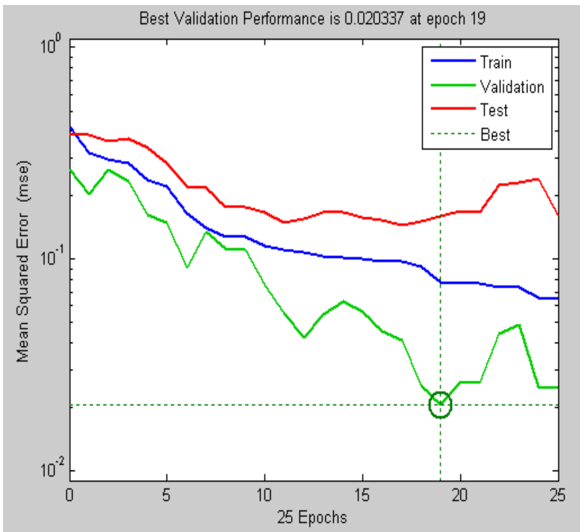

Fig. 6

ISBN: 978-1-61804-068-8

The Receiver Operating Characteristic (ROC) curve is used to inspect the classifier performance more closely as shown in Fig. 7. By definition, a ROC curve shows true positive rate versus false positive rate (equivalently, sensitivity versus 1–specificity) for different thresholds of the classifier output. You can use it, for example, to find the threshold that maximizes the classification accuracy or to assess, in more broad terms, how the classifier performs in the regions of high sensitivity and high specificity.

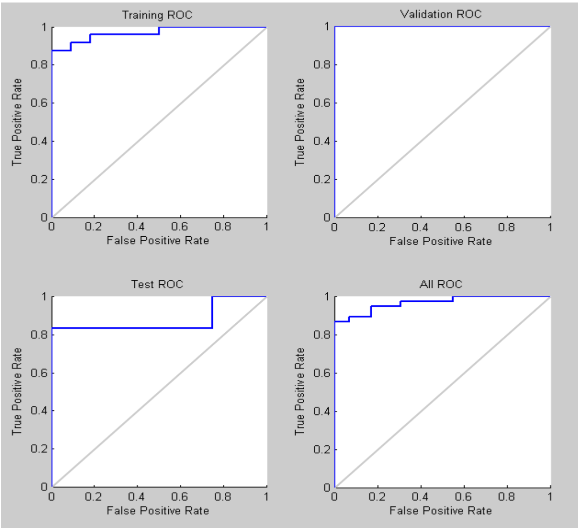

Fig. 7

The percent correctly classified in the simulation sample by the feed-forward back propagation network is approximately 91 percent as shown in Fig.8. Fig.9 shows the ROC curve for simulation sample.

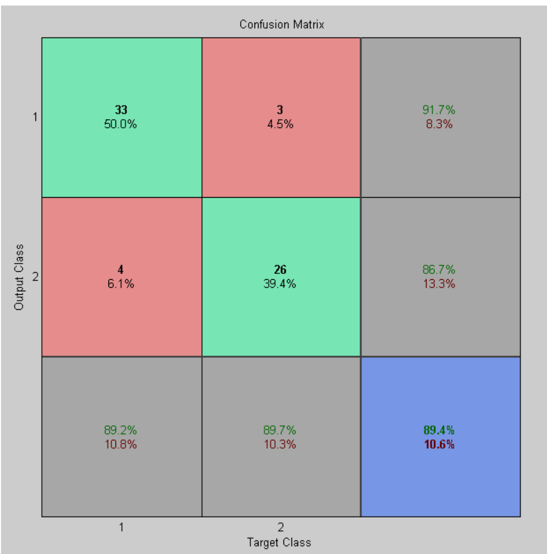

Fig. 8

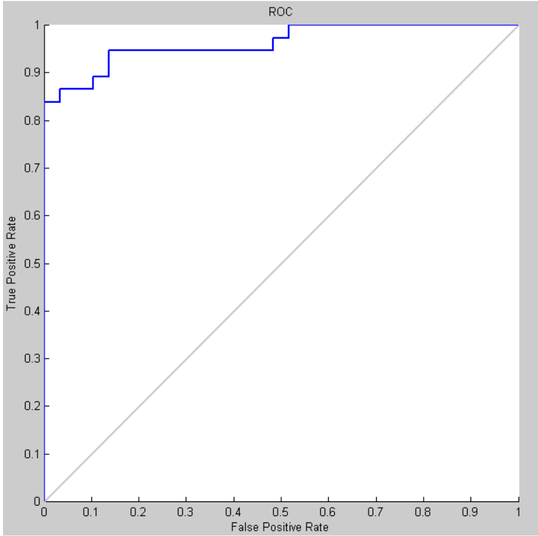

Fig. 9

4 Conclusion

The goal of this paper is to evaluate the predictor variable used to predict the insolvency using GRNN. Results showed that the most affecting factor in banks insolvency evaluation is the net income, total equity capital, cost of sales, sales, cash flows and loans. It is also aimed to predict bank insolvency before the bankruptcy using feed-forward back propagation neural network. The results of applying the supervised neural network in classification of financial distress based upon the selected financial variables showed that artificial neural networks are able to learn the patterns corresponding to financial distress of the bank. Even with the limited data used in this study, artificial neural networks showed significant signs for providing early warning signals and solvency monitoring.

References:

[1] J Kamruzzaman, R. A. Sarker and R. Begg, Artificial Neural Networks: Applications in Finance and Manufacturing, Idea Group Publishing, 2006.
[2] Q. K. Al-Shayea, G. A. El-Refae and S. F. El-Itter, Neural Networks in Bank Insolvency Prediction, International Journal of Computer Science and Network Security, Vol. 10, No. 5, May 2010, pp. 240-245.
[3] W. H. Beaver, Financial ratios as predictors of failures. Empirical research in accounting: selected studies, Journal of Accounting Research, 5(Suppl.), 1966, pp. 71–102.
[4] Altman, E. I., Financial Ratios, Discriminant Analysis and the Prediction of Corporate

ISBN: 978-1-61804-068-8

Bankruptcy, Journal of Finance, Vol. 23, No. 4, 1968, pp. 589-609.

[5] W. H. Williams and M. L. Goodman, A Statistical Grouping of Corporations by Their Financial Characteristics, Journal of Financial and Quantitative Analysis, Vol. 6, No. 4, 1971, pp. 1095-1104.

[6] J. Sinkey, A Multivariate Statistical Analysis of the Characteristics of Problem Banks, Journal of Finance, Vol. 30, No. 1, 1975, pp. 21-36.

[7] E. I. Altman, R. G. Haldeman, and P. Narayanan, ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations, Journal of Banking and Finance, Vol. 1, No. 1, 1977, pp. 29-54.

[8] S. Kaski,J. Sinkkonen and J. Peltonen, Bankruptcy Analysis with Self-Organizing Maps in Learning Metrics, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 12, NO. 4, JULY 2001, pp. 936-947.

[9] C. Pramodh and V. Ravi, Modified Great Deluge Algorithm based Auto Associative Neural Network for Bankruptcy Prediction in Banks, International Journal of Computational Intelligence Research, Vol.3, No.4, 2007, pp. 363–370.

[10] J. Baek and S. Cho, Bankruptcy Prediction for credit Risk Using an Auto- Associative Neural Network in Korean Firms, The proceedings of the CIFEr Hong Kong, 2003.

[11] M. Kim and D. Kang, Ensemble with neural networks for bankruptcy prediction, Expert Systems with Applications, Vol. 37, Issue 4, April, 2010, pp. 3373-3379.

[12] A. Nachev, S. Hill and C. Barry, "Fuzzy, Distributed, Instance Counting, and Default ARTMAP Neural Networks for Financial Diagnosis", International Journal of Information Technology & Decision Making, Vol. 9, No. 6, 2010, pp. 959–978.

[13] P. Ravisankar, Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP, Knowledge-Based Systems, Vol. 23, Issue 8, Dec., 2010, pp. 823-831.

[14] J. C. Neves and A. Vieira, Improving Bankruptcy Prediction with Hidden Layer Learning Vector Quantization, European Accounting Review, Vol. 15, No. 2, 2006, pp. 253–271.

[15] D. W. Patterson, Artificial Neural Networks, Theory, and Applications, Englewood Cliffs, NJ: Prentice-Hall, 1995.

[16] D. F. Specht, A general regression neural network, IEEE Trans. Neural Networks, Vol. 2, 1991, pp. 568–576.

[17] E. A. Nadaraya, On estimating regression, Theory of Probability Application, Vol. 9, 1964, pp. 141–142.

[18] M. T. Hagan, H. B. Demuth, M. Beale, Neural network design, PWS Publishing Company, Boston, 1996.

[19] K. Kayaer and T. Yildirim, Medical Diagnosis on Pima Indian Diabetes Using General Regression Neural Networks, web page available at: www.yildiz.edu.tr/~tulay/publications/Icann-Iconip2003-2.pdf.

[20] J. A. Freeman and D. M. Skapura, Neural networks: algorithms, applications and programming techniques, Addison Wesley Longman, 1991.

[21] I. Olmeda and E. Ferandez, Hybrid Classifiers for Financial Multicriteria Decision Making: The Case of Bankruptcy Prediction, Computational Economics, 1997, pp. 317-335.