Neural Networks: Is it hermeneutic?

Rama Prasad Kanungo[1]

[1] Rama Kanungo is the Research Co-ordinator of the Asian Accounting, Finance and Business Research Unit, Cardiff Business School, Colum Drive, Cardiff, CF10 3EU, UK.Email:Kanungomba@hotmail.com

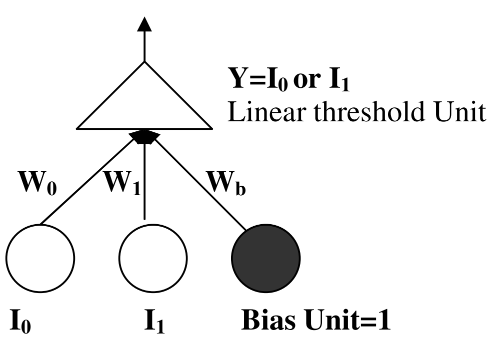
Abstract

This paper proposes a synoptic methodology to evaluate the determinants of audit fees by utilising Neural Networks. First, a brief discussion is presented to highlight the significant application of Neural Network in the areas of financial management; second the framework of proposed methodology has been outlined to examine the implication of audit fees on target sample. The underlying rational of this paper is to establish NNs as a diagnostic tool to assess the effect of audit fees on firms, which indeed warrants further empirical investigation. The importance of NNs emerges from the fact that if external and internal audit fees can be disseminated by employing this methodology which is perceived more significantly robust than other econometric models, then accounting standards can be improved.

Overview of Neural Networks: Efficacy in Finance

The applied process of Neural Networks and the cognitive competence of our brain are functionally coterminous. Both the processes learn through different patterns and predict most probable outcomes. However, the function of NNs is very much limited with reference to our real synaptic versatility. The origin of neural networks dates back to the 1940s. McCulloch and Pitts (1943) and Hebb (1949) researched networks of simple computing devices that could model neurological activity and learning within these networks, respectively. Later, the work of Rosenblatt (1962) focused on computational ability in perceptrons, or single-layer feed-forward networks. Proofs showing that perceptrons, trained with the Perceptron Rule on linearly separable pattern class data, could correctly separate the classes truly generated excitement among researchers and practitioners.

NNs are a form of multiprocessor computing system with simple processing elements, high degree of interconnectedness, simple scalar message and adaptive interaction between each element (Smith, 1996). Each simple element is like our cerebral neurons competent of propagating systematic reception to next one following a layered matrix. This matrix is highly connected and conduits message in a scalar form. They adapt the message and process it to identify an optimal outcome. The main subjectivity of such process minimises the error through a sequential elimination approach. Haykin(1994) states that NNs resemble brain in two ways, first, knowledge is acquired by the network through a learning process and second, interneuron connection strengths known as synaptic weights are used to store the knowledge . Smith (1996) presented single unit adaptive network in this following diagram



This network has two inputs and one output. All are binary by structural mode. The output is outlined as

1 if $W_0 * I_0 + W_1 * I_1 + W_b > 0$
0 if $W_0 * I_0 + W_1 * I_1 + W_b <= 0$

For the purpose of simple presentation we say output is 1 if either I_0 or I_1 is 1

The NNs identify the change by an amount proportional to the difference between the desired output and the actual output. This is referred as Perceptron learning rule, which is presented below in the Eq. (1).

Recently NNs are in vogue as a diagnostic tool to examine the topology of financial applications. Fadlalla and Lin (2004) state that some major characteristics of such applications are to investigate data intensity, unstructured and noisy data, high uncertainty scenarios and hidden relationship between parameters. So far many studies have utilised NNs to examine the predictive aspects of such applications. Essentially, during 1993-1995 NNs were increasingly been used in the financial studies. However, so far NNs have mainly been unutilised for forecasting and asymmetric prediction, so this paper intends to propose NNs in a different context to measure the binary effects of certain parameters such as audit fees.

The following table illustrates some key studies where NNs are employed to explore studies in financial applications:

Study	Domain	Year
Altman, Marco, and Varetto	Bankruptcy prediction	1994
Coleman, Graettinger and Lawrence		1991
Klemic		1993
Odom and Sharda		1990
Raghupati, Schkade and Raju		1991
Rahimiana et al		1993
Salchenberger, Cinar and Lash		1992
Tam		1991
Tam and Kiang		1992
Wilson and Sharda		1994
Ahmadi	Stock-market forecasting	1990
Baba et al		1993
Barker		1990
Barr and Mani		1994
Bergerson and Wunsch		1991
Berry and Trigueiros		1993
Bosarge		1991
Cheng, Wagner and Lin		1996
Chenoweth and Obradovic		1995
Chiang, Urban and Blabridge		1996
Chuah		1992
Collard		1993
Ganesh		1994
Kamijo and Tanigawa		1990

Kimoto and Asakawa		1990
Kryzanowski, Gellar and Wright		1993
March		1995
Quah, Tan and Heng		1993
Refenes		1993
White		1988
Yoon and Swales		1991
Morase	Credit analysis	1990
Reilly et al		1991
Roy and Cosset		1994
Van Eyen and Cronjie		1994
Vishwakrama	Business-cycle recognition	1994
Dutta and Shekhar	Bond rating	1988

Source: Fadlalla, A. and Lin, C. (2004)

Outline of Proposed methodology: Audit Fees

Following Kutsurelis's (1998) study to examine the financial market this methodology proposes a similar approach. In the light of previous researches undertaken it can be envisaged that in the place of assessing the accuracy of financial forecasting, it can be potentially employed to investigate the effect of audit fee on public or private sector firms. In this discussion it has been proposed to use the network utilising two adaptive functions, one is a Logistic function $f(x)=1/(1+\exp(-X))$ and other one is the Gaussian function $f(x)=\exp(-X^2)$. Both the functions are utilised in the study undertaken by Kutsurelis. Also these methods were used by both Kartalopoulos (1996) and Haykin (1994) to examine the financial forecasting accuracy. After calculating the functions, each one will be employed as input nodes in the final NNs model. The prior assumption in this instance is drawn upon the concept that NNs is highly effective to identify signals and eliminate noisy data flow. Particularly over a long period and lagged time series data where unstructured nature and hidden relationship in variables are not correctly identified. Hence it is presumed that such approach would be more appropriate where a sampling 5-10 years data is compiled and collapsed on target firms. Furthermore, least square approximation and random sampling do not always provide a robust calculation. Therefore, a log linearisation approach is more practical in this instance to establish the maxima threshold of parameters.

It would be reasonable to focus these types of research on the backpropagation algorithm learning method. This method has often been employed in many previous researches. Backpropagation algorithm focuses to minimise residual error between NNs output and actual desired output value (Kutsurelis, 1998). A brief view of backpropagation model is presented below following Dhar and Stein (1996) and

Kutsurelis (1998). The BPN model has been selected for this proposed methodology for two reasons, first, it converges on the error minimisation problem; second, it provides optimal solution to high frequency momentum and learning rates. Following NNs it is rational to theorise that output node corresponds to 1= positive value, 0= negative value in terms of the effect of audit fees on the sample firms. The input nodes represent financial ratios that the model will use to predict the effect of audit fees as a whole on the sample pool. The audit fee would be categories into external and internal fees to conclude each ones effect on firms. This would differentiate apparent effects of the each category on the firms. In this approach the ratios used by Altman *et. al*. (1994) with some variations are feasible and often been investigated, thus seem more adaptable in this context. The ratios proposed are:

Working Capital/Total Assets
Total Liabilities/Total Assets
External Audit Fees/ Total Assets
Internal Audit Fees/Total Assets
ROI/Total Assets
Cash Flow/ Total Debt
Cash/ Current Liabilities
Current Assets/Current Liabilities
EBIT/Total Interest Payments
Working Capital/Total Assets

Initially, it is expected that these ratios should be calculated by utilising above two functions and subsequently each output would serve as inputs in the final model. The basic concepts of BPN, which is sometimes referred as multi-layered perceptrons (MLPs) and radial Basic Function Networks (RBF) is described in the following section.

Usually BP uses a gradient descent algorithm which follows the error curvature by moving down the slope. In this study it is imperative to minimise the total squared errors by repeated error elimination process, such as using BNP. The following BP error minimisation models are adapted from Kaastra and Boyd (1996)

$$\mathbf{E} = \tfrac{1}{2} \sum_{n}^{M} \mathbf{E}_n = \sum_{n}^{M} \sum_{i}^{N} (\mathbf{t}_{ni} - \mathbf{o}_{ni}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{(1)}$$

Where E is the total error of all neural patterns, E_n represents the error pattern n, the index n ranges over the set of input patterns, and I refers to the ith output of neuron. The variable t_{ni} is the desired output for the ith output neuron when the nth pattern is presented, and o_{ni} is the actual output of the ith output neuron when pattern n is presented. The adaptive interface as learning rule to adjust the weight between neuron I and j is defined as:

$$\boldsymbol{\delta_{ni}= (t_{ni}-o_{ni})\, o_{ni}\, (1-o_{ni})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{(2)}$$

$$\boldsymbol{\delta_{ni}= o_{ni}\, (1- o_{ni}) \sum_{k}^{N} \delta_{nk}\, w_{jk}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{(3)}$$

$$\boldsymbol{\Delta w_{ij}\, (h+1)= \varepsilon\, (\delta_{ni}\, o_{ni})} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{(4)}$$

Where h is the presentation number, δ_{ni} is the error signal neuron I for pattern n, and ε is the learning rate. The learning rate is a constant of proportionality which determines the size of the weight changes. Essentially the weight changes of neurons are proportional to the impact of the weight from the neurons on the error. The error signal for an output neuron and a hidden neuron are calculated by Eq. (3) and (4) by using simplified partial differential equations. Kaastra and Boyd (1996) suggests that to increase the learning rate, thus to enhance the backpropagation training time without leading to oscillation, a momentum term needs to be included in BP learning rule. They further add that this momentum term determines how past weight changes affect current weight changes. The modified training rule as defined by them is

$$\boldsymbol{\Delta w_{ij}\, (h+1)= \varepsilon\, (\delta_{ni}\, o_{ni})+ \alpha\, \Delta w_{ij}\, (h)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{(5)}$$

However NNs software programme usually provides default values for learning rate and momentum that mostly adjust the high frequency oscillation.

Mainly running the NNs consists of in two fold approaches. First, forwardpass when the outputs are calculated and the error at the output units are calculated; second backpropagating or backwardpass, the output error is used to alter the weights on the output units and the error at the hidden nodes is calculated, subsequently the weights on the hidden nodes are altered using these values. Thus the model optimises error to a minimal threshold. Dhar and Stein (1996) state that by recursive backward propagation such errors are minimised and all the weights are adjusted to converge on error minimisation solution.

Essentially, it is secure to save the data in SPSS and subsequently import the files into NNs software package, Neuroshell 2 as patterned files. The ratios would be converted using SPSS commands to calculate logistic function and Gaussian function values, respectively, i.e., $f(x) = 1/(1+\exp(-X))$ and $f(x)=\exp(-X^2)$. Furthermore, a Ward Network would be used to train the Network. Ward Network is a BPN with one input layer, three hidden layers and one output layer. Each layer is known as a slab. Each slab is customarily consists of one to sixteen individual neurons depending upon the location within the networks. Each slab has a different activation function. In this research aforesaid two functions are recommended only to restrict the complicated computational procedure.

However the application of Neural Networks in this context is subjective to further rigour of empirical estimation. The essence of robustness can only be established upon critical examination of such rigour, which requires diligence in pooling the data and layering it with necessary commands.

Bibliography

Altman,E. I., Marco, G. and Veretto, F. (1994) 'Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks', *Journal of Banking and Finance*, Vol.18, No.3, pp 505-529.

Dhar, V. and Stein, R. (1996) 'Intelligent Decision Support Methods: The Science of Knowledge Work (Ed.), Prentice Hall Business Publishing, NJ, USA.

Fadlalla, A. and Lin, C. (2004) 'An Analysis of the Application of Neural Networks in Finance', Working paper, Department of Computer and Information Services, Cleveland State University, Ohio, USA.

Haykin, S. (1994) 'Neural Networks: A Comprehensive Foundation', MacMillan College Publishing Company, NY.

Hebb, D. O. (1949) 'The Organization of Behavior' New York: John Wiley.

Kaastra, I. and Boyd, M.(1996) 'Designing a Neural Network for Forecasting Financial and Economic Time Series' *Neurocomputing*, Vol. 10, pp 215-236.

Kartalopoulos, S. V. (1996) 'Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications, IEEE Press, NY.

Kutsurelis, J. E. (1998) 'Forecasting Financial Markets using Neural Networks: An Analysis of Methods and Accuracy', PhD Thesis, Naval Postgraduate School, Carlifornia, USA.

McCulloch, W. S. and Pitts, W. H. (1943) 'A Logical Calculus of the ideas immanent in Nervous Activity', *Bulletin of Mathematical Biophysics*, Vol. 5, pp 115-133.

Rosenblatt, F. (1962) 'Principles of Neurodynamics', Spartan, New York.

Smith, L. (1996) 'An Introduction to Neural Networks', Presented in NSYN meeting Edinburgh, Scotland, February.