A Comparative study of the forecasting ability of Backpropagation Artificial Neural Network Models with Learning Rate Adaptation for Colombo Stock Market.

EJKP Nandani[1] and Janak R Wedagedara[2]

Department of Mathematics, University of Ruhuna, Matara, Sri Lanka.

[1]ejkpnandani@yahoo.com, [2]janak@maths.ruh.ac.lk

I. BACKGROUND

A stock market is one of the fundamental types of financial markets. The stock market can also be thought of as a highly complex and adaptive system. Financial indices defined on stock prices are used as indicators of the economical trend of a country.Thus, forecasting the behaviour of the stock market is at primary concern not only of the business community but also of the policy makers of a country.

We present here the designing and analysis of a few Artificial Neural Network (ANN) models to predict future trend of the two major financial indices used in the CSE namely, the Milanka Price Index (MPI) and, the All Share Price Index (ASPI) using the financial time-series data of the CSE.

In so doing we consider here two computational methods of input selection for the Training algorithms for the ANN - the *Gradient Descent* and, the *Gradient Descent with Learning Rate Adaptation*[2].

A number of papers have addressed forecasting problems pertaining to the Sri Lankan financial/stock market: for instance, in [3] authors report of a methodology developed to predict the direction of tomorrow's closing price of the ASPI of Sri Lankan stock market using a genetic algorithm as well as an ANN model. In [4] the author reports about stock market prediction system which is based on the quantitative and qualitative input vectors where they consider three different types of ANNs in the analysis.

The results we report in the present paper significantly different from these earlier studies in the sense that the use of the learning rate adaptation against constant learning rate used in the above works. Moreover, we shall use the information theoretic statistics - Mean Square Error (MSE), Akaike Information Criteria (AIC) [5] and the Bayesian Information Criteria (BIC) [5] for comaring the ANN models that we employ in the forecasts.

II. MATERIALS AND METHODS

A. Collection and Preprocessing of Data:

We use the CSE market data for the period 1985 – 2009 (available for the general public in electronic form in the official CSE Data CD), of which the first $1/3^{rd}$ has been used in training of the ANN, the second one-third for validation (in parallel with training) and the final (unseen) part to test against the forecasts made by the ANN.

The input range required for the network must be determined. Let's assume that the input range is from I_{min} to I_{max}. The formula for transforming each data value x to an input value I is [1]: $x' = I_{min} + (I_{max} - I_{min})\left(\dfrac{x - x_{min}}{x_{max} - x_{min}}\right)$

B.Input Selection Methods:

As with any forecasting model, the selection of appropriate input is extremely important for generalization ability [6].Based on the way the input is presented, we propose two different models for the ANN:

- **Model-A**: The (historical) daily values of MPI, ASPI and their stocks have been considered as inputs and the next day value of MPI and ASPI have been considered as the desired output.

For example, in the vector

$$\{[(x_1^n, x_2^n, x_3^n, \ldots, x_{14}^n), (x_1^{n+1}, x_2^{n-1})]\}.$$

$(x_1^n, x_2^n, x_3^n, \ldots, x_{14}^n)$ is the n^{th} day input vector, for which the desired output is (x_1^{n+1}, x_2^{n+1}). Where x_1 and x_2 are the MPI & ASPI and the reamining $x_i^n, (i = 3, 4, \ldots, 14)$ are underlying 12 industry sector values.

- **Model-B**: The daily MPI and ASPI data which collected from the CSE has been grouped into 30 consecutive non-overlapping blocks. Thus, each block shall serve as an input vector whence the MPI (ASPI) of the 31^{st} day has been taken as the desired output corresponding to those 30 inputs. For example, in the vector,

$$\{[(x_1, x_2, \ldots, x_{30}), (x_{31})], [(x_2, x_3, \ldots, x_{31}), (x_{32})], \ldots, [(x_{n+1}, x_{n+2}, \ldots, x_{n+30}), (x_{n+31})]\},$$

$(x_{n+1}, x_{n+2}, \ldots, x_{n+30})$ is the $(n+1)^{th}$ input vector, for which the desired output is x_{n+31} where x_n is the MPI(ASPI) on the n^{th} recorded day.

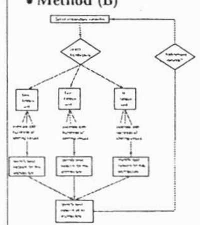
C. Determination of the Network Architecture:

Different network architectures have been tested considering two computational methodologies employed for minimization of the Mean Square Error (MSE) resulted due to the errors made in the output produced by the network and the actual (desired) output, namely:

- **Method (A)**

The network architecture mainly depends on the number of hidden nodes in the network. We first chose arbitrary values (between 0 and 1) for both the learning rate (η) and the momentum term (α) to train the network varying number of hidden node units $m = 1, \ldots, m', \ldots, 20$. For each m the MSE has been computed. The optimal value of m is m' for which $MSE_{m'} \leq MSE_m$ for all $m > m'$. This way one can expect a faster algorithm as this essentially imposes an *early stopping criterion* for the training algorithm.

- **Method (B)**



Network model selection procedure using trial and error.

D. Performance Measurements:

We compare the forecasting ability of each model via Mean Square Error (MSE), Akaike Information Criterion (AIC) [5] and Bayesian Information Criterion (BIC) [5] via which we examine the best network topology for each model.

III. RESULTS: Forecasts for MPI and ASPI

Figure 1: Red and Blue curves correspond to gradient descent backpropagation learning algorithm with and without learning rate adaptation respectively

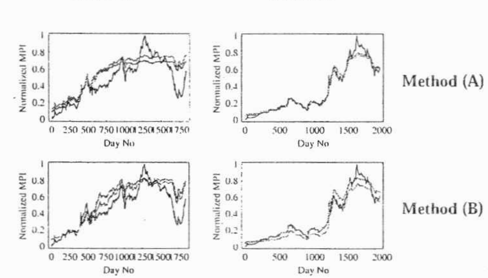
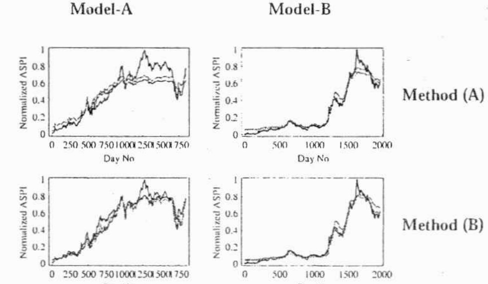


Figure 2: Red and Blue lines correspond to gradient descent backpropagation learning algorithm with and without learning rate adaptation respectively



III. CONCLUSIONS:

– The convergence rate of error in the algorithm with learning rate adaptation is found to be faster than in the one without learning rate adaptation. Performance of forecasting has degraded with increasing number of future days in the learning algorithms in both models Model-A and Model-B.

– Our results show that an ANN which uses historicalmarket data as input with $30-10-n$ architecture for MPI and $30-7-n$ architecture for ASPI (where $n = 1, 2, 3, 7$ represents the next n days for which the forecast is required) and which use learning rate adaptation together with the Method (B) (trial and error model selection approach) outperforms all the other models in forecasting indicating the knowledge transferred by the biological neural network (the human brain) to the ANN in the study improves the performance of the ANN.

– As it can be seen from the Bottom-Right of the Figure 2, the forecast made by this ANN for unseen data is in excellent agreement with the actual behaviour and the forecasting ability of the ANN outperforms most of the other standard (regression) statistical techniques.

References:

[1] Roberto Battiti, *Accelerated backpropagation learning: Two optimization methods*, Complex System 3:331-342, [1989] . [2] Dassanayake M.M.K. and Tilakaratne C. *Technological Developments in Networking, Education and Automation*, Springer, Chapter Predicting Trading Signals of Sri LAnkan Stock Market Using genetic ALgorithms and Neural Networks:269-273 [2010] . [3] Robert E. Kass and Adrian E. Raftery *Bayes Factors*, Journal of the American Statistical Association 90:773-795, [Jun 1995] . Addison Wesley longman (Singapore) pvt Ltd . [1999] . [4] A. Senanayake, *Automated neural-ware system for stock market prediction*, Cybernetics and Intelligent Systems,IEEE Conference 2:1166-1171, [2005] . [5] Iiuo Klevecka and Janis Lelis, *Preprocessing of input data of neural networks: The case of forecasting telecommunication network traffic* , ISSN 0085-710, [2008] [6] Anders Krogh and John A. Hertz, *A simple weight decay can improve generalization*, In Neural Information Processing System 4:950-957 [1995]