Artificial Neural Network: s-Empirical Analysis of Predictive Accuracy in the Indian Stock Market

Mr. Vaibhav Sahu*, Mr. Gagan K Patra* and Mr. Khagesh Agarwal**

*Associate Consultant, SI-BI, Infosys Technologies Ltd.
Email: sahu.vaibhav@gmail.com, gagan.patra@gmail.com, agarwalkhagesh@gmail.com

Abstract: Stock market movement is driven by numerous factors, both at national and international levels, and because of the multiplicative effect of these factors, the market movement has been majorly random and very less predictable. Any model, which can predict the stock market movement would be helpful to investors to reduce their risk exposure, increase hedging effectiveness and maximize returns.

In the context of behavior modeling, it would be noteworthy to understand the complex structure of the human brain, which learns by experience. The learning is facilitated by an arrangement of neurons, which store the information about a particular event and its outcome. These neurons get trained in due course of time by encountering the events and recording the respective outcome, and by predicting the outcome when any similar event is encountered again. A similar logic could be applied to model the stock market movement by designing a neural network and developing a structure of neurons using the historic data, which can be used to predict the market movement. This research is an application of the concept of neural networking, wherein historical data is analyzed using the software MATLAB and designing an optimal network with the minimum root-mean-squared-error, using the NARX (Non-linear Autoregressive Exogenous) model. The network so obtained was applied on the current data, and an attempt was made to predict the closing price of the script under consideration for the next day. It was found that using the model, an accuracy rate of 80% was achieved in predicting the closing price of the script for the next day.

Keywords: Hedging, Neural Networks, NARX (Non-linear Autoregressive Exogenous), Risk.

Accepted On: 26.11.2012

1 Introduction

Stock market forecasting has always been a topic of great interest, both in the academic arena as well as the real-life markets. The increased attention towards this area of study can be attributed to the fact that if the direction of the stock markets can be successfully predicted, the investors could then be better guided. The pay-off from any investment decision and trading activity depends on the market predictability to a great extent. If a model can be so developed which is capable of predicting the trends of the dynamic stock market movements effectively, it can be used to make profits and generate wealth. Moreover, the predictive capability of the model can enable the market regulators take corrective measures and keep the system healthy.

Forecasting of stock market movements and its predictive accuracy comes with numerous challenges, both academic as well as experimental. A notable mention in this front could be the Efficient Market Hypothesis (EMH), which was promulgated by Eugene Fama [10] in Theory of Capital Market.

It was argued that stock prices fully reflect all the available information in an efficient market, and that there is no way to model inefficiencies in the market movements whatsoever to make supernormal returns. Even in the weak form of market efficiency, it was argued that the stock prices reflect all the past information, and that the past information could not be used to derive results which can help to make supernormal profits. If we go by the premise laid down under the EMH, then it can be arguably said that there is no system which can accurately yet consistently outperform the markets. This can very well lead
us to the conclusion that modeling of market movements under the assumption of EMH would only be possible on the speculative, stochastic component not on the changes in value or other fundamental factors [25]. Another yet reference worth mentioning is the Random Walk Hypothesis, which is consistent with the EMH. It also essentially states that the stock price movements are random departure from the previous prices, and evolve according to a random walk. Given that there has been a lot of debate to ascertain the validity and reliability of the EMH and the random walk theory, however, with the advent of computational and behavioral finance, economists have tried to propose an alternate to the EMH, or rather say, an opposite hypothesis which may be collectively referred to as the Inefficient Market Hypothesis (IMH). This essentially states that the markets are not always efficient, that the stock price movements don’t always follow a random walk leading us to conclude the presence of inefficiencies or anomalies in the market. A notable reference could be made to the work of Pan Heping [25] where the Swing Market Hypothesis was postulated. It stated that the markets are sometimes efficient while sometimes inefficient, and tends to swing between these two modes intermittently. The theory also promulgates that the market movements can be decomposed into four different types of components, viz., dynamical swings, physical swings, abrupt movements and random walks. This can lead us to the conclusion that we may try to identify these inefficiencies in the market, and try to model the market movements when inefficiencies do exist.

Past researches in the field of modeling stock market movements have tried to propose various models by making use of fundamental analysis, technical analysis and a number of other analytical techniques for making an attempt to analyze and predict the market behavior. Fundamental analysis involves an in-depth analysis of the change in stock process with respect to exogenous macro-economic variables. It is based under the assumption that the market price of a stock depends on its intrinsic value and the expected return of the investors. However, given the fact that the expected return by the investors is different for different class of investors, and the same is also subjected to change as new information pertaining to the stock is available in the market, leading to a change in the stock price. Moreover, the analysis of various macro-economic factors is subjective to the interpretation of the analyst. Technical analysis, on the other hand, is carried out by interpretation of the data pertaining to price, volume and open-interest in the form of statistical charts, so as to predict future movements. The underlying premise behind technical analysis is that all of the internal and external factors that affect a market at any given point of time are already factored into that market’s price [21].

In addition to these common tools used to analyzing and charting stock price movements, a number of traditional time series forecasting tools are used for the similar purpose. In a typical time series forecasting technique, the past data of the prediction variable is analyzed and modeled, so as to capture the patterns of the historical changes in these variables. Based on the model so developed, we try to provide current information so as to predict future price movements. There are mainly two approaches of time series modeling and forecasting: linear approach and the non-linear approach. Commonly used linear approach includes simple/weighted moving average, exponential moving average, exponential smoothing, regression analysis etc. One of the most popular and widely used linear approaches is the Autoregressive Integrated Moving Average (ARIMA) model, proposed by Box & Jenkins [3]. This model, also sometimes referred to as the Box & Jenkins Approach, presumes a linear model but is quite flexible because of its ability to represent different types of time series, viz., the Autoregressive (AR), Moving Average (MA) and combined AR & MA (ARMA) series.

Given the fact that there is very less empirical evidence of perfect linearity of the stock market returns as the residual variance between the predicted and actual returns is high [1], the existence of non-linearity in the stock market returns is generally accepted, and have also been propounded by financial researchers and analysts with empirical evidences. Some of the widely used parametric non-linear forecasting models such as Autoregressive Conditional Heteroskedasticity (ARCH) and the General Autoregressive Conditional Heteroskedasticity (GARCH) have been in use for the purpose of
financial forecasting. However, most of these non-linear statistical techniques require that the non-linear model must be specified before the estimation of the parameters is done.

As many researchers have claimed that the stock market is a chaos system, i.e., a non-linear deterministic system which appears to be random because of the irregular functions, they are therefore difficult to deal with using normal analytical methods given the dynamism in these systems and high sensitivity to the initial set of conditions. Neural networks are very effective in learning the characteristic behavior of such non-linear chaotic systems because they make very few assumptions about the functional form of the underlying dynamic dependencies and their initial conditions. A neural network is a massively parallel distributed processor made up of simple processing units which has a natural propensity for storing experiential knowledge and making the same available for use [12]. They have a remarkable ability to derive meaning from complicated or imprecise data, and can be used to extract patterns and identify trends, which otherwise is too complex to be noticed either by humans or other computer techniques. From a statistical perspective, neural networks are analogous to non-parametric, non-linear, regression models. However, the traditional statistical models have limitations in understanding the relationship between the input and the output of the system because of the complex and chaos nature of the system. The past few years have witnessed tremendous advancements in the application of neural networks for forecasting stock price movements with a hope that the patterns could be extracted. The superior resolution of ANN lies in its predictive ability to discover and interpret non-linear relationships in the input data set without any a-priori assumption of the knowledge of the existence of any such relationships [32]. The input variables are mapped to the output variables by transformation using a special function known as activation function, which independently discovers and adapts to the inherent relationships existing between the variables from a set of labeled training example and therefore involves a modification in the network parameters. Neural networks also have a built-in capacity to adapt the network parameters to reflect the changes in the system. Any neural network which is trained to a particular input data set corresponding to a particular environment can be easily retained to a new environment to predict at the same level of the environment. Moreover, when the system under study is non-stationary and dynamic in nature, the network so designed can change its parameters (synaptic weights) in real-time basis to adapt to the changes. Thus the use of Artificial Neural Networks (ANN) can provide an edge over the traditional forecasting tools and identify characteristic inefficiencies in the forms of predictable patterns to forecast future price movements, making ANNs a very promising tool for security price modeling. The objective of this research is to empirically study the architecture of the artificial neural networks and ascertain its predictive ability in the context of the Indian stock markets.

2 Review of Literature

Forecasting of stock market returns and price movements have gained importance for the past two decades, where researchers tried to model a linear relationship between the input macro-economic variables and the resultant stock returns. However, post the empirical findings of the non-linearity in the stock market returns [1], we have witnessed a paradigm shift in the focus of the researchers towards non-linear statistical modeling for predicting stock market returns. There have been many researches in the past in the context of non-linear modeling of the stock market returns, but most of them have required the estimation parameters of the model to be specified before the estimation is carried out. However, given the characteristic nature of the stock market returns being noisy, uncertain, chaotic and non-linear in nature, ANN has over the time, evolved out to be a better technique in identifying and analyzing the stock’s performance and price movements, viz-a-vie its determinant factors more accurately than other statistical techniques. A noteworthy reference could be made to the implementation of a neural network model by Chan, Wong & Lam.[1], using the technical analysis variables for the companies listed in the Shanghai Stock Exchange. The findings of this paper reported that the prediction of stock market
movements and returns thereof are quite possible with the algorithm and initialization methods. In the context of financial theory and practice, another noteworthy reference could be made to the empirical work of Kuan & Liu [33], to model the forecasting and predictive ability of exchange rates using the ANN’s. It was established that a properly designed model has a lower out-of-sample mean squared prediction error relative to the random walk model. In a yet another research in the context of stock price movements, Jasic & Wood [34] had studied the profitability of the trading signals generated from the out-of-sample short term predictions for the daily returns of S&P 500, DAX, TOPIX and FTSE stock market indices evaluated over the period 1965-99. The results were compared with a benchmark linear autoregressive model, and it was found out that the buy and sell signals derived from neural network predictions were significantly different from the unconditional one-day mean returns, and have the possibility of generating significant net profits for the reasonable decision rules and transaction cost assumptions.

Research by Cao et. al. [35] provides an interesting comparison between the Fama & French model and the ANN model, in the context of prediction of the returns by the Chinese stock markets, where it was reported and empirically established that the ANNs are superior to other linear models in terms of its predictive ability. Another research where ANN’s were used to forecast the Canadian GDP growth reported that the neural network models yield statistically lower forecast errors for the year-on-year growth rates of the real GDP in comparison to other linear and univariate models [36]. Similarly, predictive accuracy of forecasting market returns have been compared across the results of linear regression with that of the neural networks [6].

In many past researches, the ARIMA model has been used as a benchmark model and results so obtained have been compared to ascertain the superiority of modeling using ANN. It was found that the prediction of ANN was better than that of ARIMA, and that useful prediction can also be made even without the use of extensive data or knowledge [14]. ANNs are being extensively used in the financial domain for credit risk and bankruptcy prediction [37], derivative pricing [38], asset allocation and portfolio management [27], risk management, interest rate modeling, insurance, stock valuation, exchange rate forecasting, predicting stock market prices [42], generate buy, sell signals, design of trading system [41], prediction of option and futures prices [42], etc., apart from its use in the financial market and corporate arena (companies such as General Electric, American Express etc. are using neural network for screening credit applications, spot stolen credit cards, detect patterns which may indicate fraud and predict commodity and stock prices, bond ratings and currency trading trends. Financial institutions have carried out many projects on financial forecasting with neural networks).

In the Indian context, ANN has been extensively used for the purpose of research. One reference could be made to a research on modeling the Sensex weekly closing values by using ANN, by developing two networks with three hidden layers [10]. In a yet another research, ANN was extensively used to forecast the weekly INR-USD exchange rate and the results were found to be superior when compared with that of autoregressive and random walk models [26]. Similarly, the forecasting ability of the ANN was tested on the Nifty index returns by applying three-layer architecture and it was found that the predictions were efficient for the period under investigation.

The previous studies have used various forecasting techniques in order to predict the stock market trends. Some attempted to forecast the daily returns where others developed forecasting models to predict the rate of returns of individual stocks. In many papers it was also found that researchers have attempted to compare their results with other statistical tools. And these findings provide strong motivation for modeling forecasting tools for stock market prediction. Given that most of the past researches have been based on ascertaining the predictive accuracy of the forecasting ability of the ANN, not much of research has been into the development and deployment of a network and ascertaining whether the network output confirms to the market behavior. Under our research, a network has been developed to forecast the closing price of the stock under consideration, rather than checking the lag structure of the actual.
3 Research Methodology

For the better performance of the model care have been taken to various steps of modeling. First the performance measurement statistics were decided to judge the performance and then other steps and issues such as input variable selection, data pre-processing, neural network architecture design, selection of optimal number of neurons were handled by Trial and Error Method.

In this research work 5 input parameters are mainly used to find the pattern of the closing price of Infosys in BSE Sensex and also the value of the closing price of the t+1 day. The 1st and foremost parameter was the closing price of the scrip i.e. Infosys for 2 years (1st Jan 2009 – 24th Feb 2011). The reason behind choosing the parameter is the non-linearity in the scrip value. We have chosen the closing price as we are trying to forecast the closing price and it also reflects the best trade price of the day, as all information are also absorbed for the day. The 2nd input parameter taken is the Crude Oil Prices, as the Indian stock market reacted to the variation in the crude oil pricing at many times. The recent Egypt crisis also made an impact on the BSE/NSE values. The Iraq war and the inflation effect also can be identified by the same. Statistically the correlation between the stock prices and the crude oil prices is 0.88 which shows a strong positive correlation between the two variables. The USD-INR exchange rate comes into play whenever any international transaction takes place. Infosys being an IT company, its stock prices are very much affected by the USD-INR exchange rates. The correlation between the stock prices and exchange rate is -0.865 which shows a strong negative correlation.

As we considered the Crude Oil price, we have also taken the USD-INR exchange rate to convert into INR value of the International crude oil price. The change in the rate of exchange will also help in identifying the relation, if any. The market index all the information available in the market. As we are analyzing stock value in BSE Sensex, we selected the BSE Sensex closing prices as one of the input variables, whose correlation with the stock under study is 0.95, a strong positive correlation. The last variable is IT Index in BSE Sensex as the stock under study is Infosys which is an IT company and its correlation with the stock is 0.99. IT Index is considered to absorbing all the market information available related to IT market and IT companies. With these 5 input parameters we tried to make a model to find the closing price of t+1 day.

3.1 Scaling of Input Data

The quality and accuracy of Input data are very important factors for the efficient performance of the model. After the selection of input variables and collection of data, the data was pre-scaled or data processed before applying to artificial neurons as input. In multilayer networks, sigmoid functions are mainly used. In general practice the input data is normalized or scaled before applying it to any model for de-trending of the data. In this way the network output falls into a normalized range. Both the input and target data are scaled before feeding them to the model and the output data is reverse transformed in the original form to find the generalized form of output data set.

In the research non-linear scaling method was used. Log scaling function was used to pre-process the input data. The equation below describes the log scaling function

$$ X_k = \ln(X_n) - \ln(X_{n-1}) $$

Where $X_n$ is the current day closing price and $X_{n-1}$ is the previous day closing price of the market index, in the research.

3.2 Neural Network Architecture and Design

Neural network, as we know, is processor composed of many simple elements called neurons, which are distributed parallel to store experiential knowledge and predict the outcome in case of similar event. These neurons are trained by providing the input data and target data. The network adjusts itself to present the output equal to the target vectors by taking the input vectors with optimal accuracy. The connection between the neurons determines the network function. The network adjusts itself by adjusting the weights of the connections between neurons.
When an input signal $X$ is applied to the neuron, it multiplies its strength to weight $W$ to form product $WX$ and adds a bias $A$ (see Fig 1), which then passes it through a transfer function to get the desired output. It then compares the calculated output to the target output and readjusts the weights and biases until the network output matches the target. Typically many such input/target pairs are needed to train a network. In this study we have taken 15% of the data for the network training.

The neural network works on feedback mechanism, i.e., the output generated from the input are compared with the target data fed. After a number of iteration the weights and biases are set for the network within an accepted range of accuracy of the outputs, and this model is used for predicting the output for a set of input vectors.

The network shown in the diagram (Fig 2) above is a Multi-Layer Neural Network. The network architecture is composed of input layer with number of input neurons, two hidden layers with hidden neurons in each of the hidden layers and one the output layer. The pre-processed input data is fed to hidden neurons and complex dynamic non-linear relation is established by the network between the inputs and output by back-propagation and feed forward network.

To establish a relationship between input and output transfer functions are used.MATLAB offers various transfer functions for scaling purpose. For this study tan sigmoid and Log sigmoid functions are considered for Trial and Error on both the hidden layer and output layers. Out of both the functions tan sigmoid function was selected as the results were found more accurate with tan sigmoid scaling function.

### 3.3 Neural Network Time Series Prediction

As the closing price of the stock depends on historical values of many input variables, we have used the time series prediction tool using neural networks in MATLAB i.e., ntstool. This tool creates a network to predict the future value of the output using historical values of the other input variables. There are 3 types of models that could be applied, viz,

1. Non Linear Auto-Regressive with External (Exogenous) Input Model
2. Non Linear Auto-Regressive (NAR)
3. Non-Linear Input-Output

We choose not to apply the NAR model as only one input variable can be applied, while we have multiple input variables to enhance the accuracy of prediction. The Non-Linear Input-Output model is also not selected since it does not support feedback mechanism of output as the input in the next input vector. As the closing prices of stocks are very much affected by the previous day closing prices this model could not prove on accuracy of results. NARX model however incorporates both multiple variable input facilities and the feedback mechanism which overcomes the drawback of the other models, thus providing better predictions than input-output model, because it uses the additional information contained in the previous values of predicted output.
3.4 Data Analysis & Interpretation

The standard NARX network is a two-layer feed forward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. This network also uses delay lines to store previous values of the x(t) and y(t) sequences. The output of the NARX network, y(t), is fed back to the input of the network (through delays), since y(t) is a function of y(t-1), y(t-2), ..., y(t-d).

In this type of time series, future values of a time series y(t) are predicted from past values of that time series and past values of a second time series x(t). The function can be mathematically represented as:

\[ y(t) = f(y(t-1), \ldots, y(t-d), x(t-1), \ldots, x(t-d)) \]  

(2)

of a stock or bond, based on such economic variables as unemployment rates, GDP, etc.

3.5 Performance Measurement

The study being a predictive modeling, the performance can be judged by the prediction accuracy of the model. Normalized Mean Square Error (NMSE) is used for measuring the prediction accuracy of the model. In the following equation, Pt is the actual value of the data series, Ot is the predicted value for the same day closing prices of the stock and Pt is the mean of the actual values.

\[ NMSE = \frac{\sum_{t=1}^{N}(P_t - O_t)^2}{\sum_{t=1}^{N}(P_t - \bar{P}_t)^2} \]  

(3)

The prediction of direction of the movement of the stock is also a valuable output of the prediction model. To test the performance of the model to measure the direction movement accuracy of the closing price of the next day’s closing; Sign Change Percentage (SCP) is used. The SCP is defined as:

\[ SCP = \frac{|\text{Sign}(P_t) - \text{Sign}(O_t)|}{N} \times 100 \]  

(4)

In equation, Pt is the actual value of the data series, O_t is the predicted value for the same days’ closing prices of the stock and N is the no of samples in the test data set. The SCP compares the movement of the predicted value with that of the actual value. SCP is a measure to calculate the predicting accuracy of the model to predict the direction of the movement (Rise or fall) of the stock Index from the previous day.

Also the correlation (r) between the predicted value and actual value of the closing price of the stock was recorded.

4 Results

Under the study, historical values of 5 input variables for the period 1st Jan 2009 to 24th Feb 2011 were analyzed, using which 14 neural networks with various combinations of number of input variables and number of neurons in the hidden layer were created. In each structure only one output node was created. For example in structure 10-45-1, last 10 days’ historical values of the input variables are provided as input (10 Input nodes), 45 neurons in the hidden layer and 1 output node which provide the 11th day’s predicted value of closing prices of the stock. The transfer function used for scaling is tan-sigmoid function.

For each structure Sign Change Percentage (SCP), Correlation between the actual value and predicted value of the closing price of stock, Normalized mean squared error and Mean squared error were recorded. Maximum 80% accuracy in predicting the direction of the closing prices for the next day i.e. SCP 80% with the structure 15-70-1 was recorded, While the maximum correlation was recorded for the structure 15-60-1. The minimum NMSE was recorded for the same structure 15-60-1.

Out of all the tested structures, 2 structures were determined as optimal structure i.e. 15-70-1 (See Fig 4) and 15-60-1 (See Fig 5).The structure 15-60-1 has he third highest SCP of 79.60% and the highest correlation and minimum NMSE. The second structure 15-70-1 has the highest SCP of 80% and the third highest correlation, third lowest NMSE and minimum MSE. Some of the graphs of the actual and predicted values of the closing prices of the Infosys for the month of February are as follows. The blue curve depicts the actual value of the stock and the red curve depicts the predicted value of the stock.
5 Conclusions & Discussion

As we can see from the empirical analysis and findings, with the help of neural networks, the direction of stock prices can be predicted. From
the data considered for this project work, we achieved an accuracy of 80%, with a high correlation and low mean squared error term. Neural networks can be helpful to an investor by giving an idea of the future direction of the market and the stock from the past values of the prices and other variables affecting market and the particular stock. A three layer feed-forward back propagation neural network with 15 input neurons, a hidden layer with 70 neurons and 1 output neuron with tan sigmoid and linear transfer function in the hidden and output layer respectively is considered as the optimal network structure. The input variables in the research are preprocessed historical values of the 5 input variables having strong positive or negative correlation with the stock prices (Infosys) to be predicted. In the volatile market conditions, the high performance and more accurate neural network models can be helpful for both investors and regulators. With this performance of the neural network, we can utilize this model as a promising tool for predicting future directions of the market and other economic factors.

5.1 Limitations & Future Scope
In this research study, we have considered only the Levenberg-Marquardt algorithm for training of the network. By incorporating various training algorithms available for the neural networks, and by including more input variables for analysis, more accuracy with reduced error in prediction can be achieved. Moreover, various transfer functions are available which can be analyzed for better performance of the model. In many places it has been seen that the model predicts the exact value as the actual but in some places difference in actual and predicted values has been seen. This difference is due to the news or events that affects the market and is not absorbed by the input variables selected. Whenever there is no extra event that is not covered by the selected input variables, the model predicts the exact value of the prices. By incorporating a higher number of variables which can cover more and more events and news affecting the market, more accurate results can be achieved. At present, the devised model predicts only the next day's direction of the stock. With extensive research, the model can be refined to predict the future values of the market for a prolonged period.

References:


