A comparison between Hybrid Approaches of ANN and ARIMA for Indian Stock Trend Forecasting

3 authors:

Nitin Merh
Jaipuria Institute of Management Indore

Vinod Prakash Saxena
Sagar Institute of Research and Technology

Kamal Raj Pardasani
Maulana Azad National Institute of Technology, Bhopal

A COMPARISON BETWEEN HYBRID APPROACHES OF ANN AND ARIMA FOR INDIAN STOCK TREND FORECASTING

Nitin Merh, Vinod P. Saxena and Kamal Raj Pardasani

Abstract

In this paper an attempt is made to develop hybrid models of three layer feed forward back propagation artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) for forecasting the future index value and trend of Indian stock market viz. SENSEX, BSE IT, BSE Oil & Gas, BSE 100 and S& P CNX Nifty. Simulations have been done using prices of daily open, high, low and close of SENSEX, BSE IT, BSE Oil & Gas, BSE 100 and S& P CNX Nifty. These are chosen as input data values and output is the forecasted closing price of SENSEX, BSE IT, BSE Oil & Gas, BSE 100 and S& P CNX Nifty for the next day and future trend. Simulation results of hybrid models are compared with results of ANN based models and ARIMA based models. Convergence and performance of models have been evaluated on the basis of the simulation results done on MATLAB®6.1 and SPSS®13.0.

Introduction

Many studies have established that non-linearity exits in financial data and artificial neural networks can be successfully used to model the relationship among this data. Neural network can mine valuable information from a mass of historical information and can be efficiently used in financial areas, so the applications of neural networks to financial forecasting have become very popular over the last few years [Zhang, et al, 2004, Windrow, et al, 1994, Refenes, 1995; Kate, et. al, 2000, J. T. Yao, et al, 1999; Abu- Mostafa, at el, 2001]. Artificial neural networks are a natural way for solving problems that involve learning and pattern recognition. It can detect patterns in data through learning and are much easier to program since they elicit general rules from exemplars [Afolabi, et al, 2007].

Though back-propagation (BP) neural networks are often used because of their better prediction ability than other models, they have certain limitations. Lawrence, Tsoi, and Giles (1996) pointed out that, when the training of a BP tends to be difficult due to the noise of data, then the networks fall into a naive solution such as always predicting the most common output. Miao et al. (2007) indicates that the solutions of the BP usually are forced to the local minimum due to the gradient descent algorithm used to get weights of connection (Yudong and Lenan, 2009).

New researches suggest that seeing the nature of stock market, combining two or more computational models synergistically is better approach for prediction problem. The basic idea behind this is that each models' unique capability can be used to model different patterns of data. The hybrid models can have similar components like two neural networks or can have dissimilar components like artificial neural networks (ANN) and auto regressive integrated moving average (ARIMA). Studies also show that using dissimilar models result in lower generalization error. Recent studies show that the techniques of ANN and ARIMA when combined, offer a competitive edge over each of the individual model. The benefits of such methods appear to be substantial especially when dealing with non-stationary series (Temizel and Ahemad, 2005). Results from study (Sallehudin, et al., 2008) show that GRANN_ARIMA(Grey Relational Artificial Neural Network_ Autoregressive Integrated Moving Averages) is better than the individual model and conventional hybrid model in terms of accuracy and robustness since it produces small forecasting errors and can work well in both small and large scale data. Studies also show (Tseng, et al., 2002) that SARIMABP (Seasonal Autoregressive Integrated Moving Averages Neural Network Back Propagation) is superior to the SARIMA (Seasonal Autoregressive Integrated Moving Averages) model, the BP (Neural Network Back Propagation) with depersonalized data, and the BP with differenced data for the test cases. A hybrid ARIMA–ANN model for time series prediction synergically combines the advantages of the easy-to-use and relatively easy-to-tune ARIMA models and the computational power of ANN.

The process of predicting as accurately as possible the stock price of the day is very difficult if not impossible (Afolabi and Olude, 2007; Brabazon, 2000). Technical analysis involves the analysis of the statistics generated by market activity like past prices and volume. In other words technical analysis uses only historical data like past prices, volume of trading, volatility etc., to determine the movements in the price of some financial assets. Technical approach is based on three premises that are (a) market action discounts everything

(b) prices move in trend (c) history repeats itself. The use of technical analysis (Malkiel, 1996; Skabar and Cloete, 2001) goes against the conservative academic opinion, which regards this behavior as irrational given the efficient markets hypothesis. According to Efficient Market Hypothesis (Samanta and Bordoloi, 2005; Fama, 1965 and Fama, 1970) all available market information are factored immediately into the formation of stock price and therefore, the best predictor of future stock price is the latest stock price is the latest available price. Technical analysis, (Murphy, 1999) refers to the various methods that aim to predict future price movements using past stock prices and volume information. It is (Mills, 1990; Priestley, 1988) possible that non-linear models are able to explain this residual variance and produce more reliable predictions of the stock price movements.

In the present study a hybrid models have been developed using techniques of both artificial neural network (ANN) and autoregressive integrated moving average (ARIMA) and their performances are compared to the ANN and ARIMA models. In the study MATLAB® 6.1.0 and SPSS® 13.0 for windows have been used for simulations.

Purpose of the Research

- Main objective of the study is to develop models for forecasting the next day's close value of SENSEX, BSE IT, BSE Oil & Gas, BSE 100 and S& P CNX Nifty. In the study hybrid model using techniques of feed forward back propagation supervised learning artificial neural network and autoregressive integrated moving average are developed and results are compared.
- Next objective is to compare the results and trends of actual and predicted values of above mentioned indices.
- To evaluate the performance of the models by calculating average absolute error (AAE), root mean square error (RMSE), mean absolute percentage error (MAPE), mean percent square error (MPSE) and compare their values.
- To compare the results and trends between the traditional hybrid and new hybrid approach for forecasting stock market in Indian context.
- To compare the trends between models based on single approach based of ANN/ ARIMA and hybrid approach using techniques of ANN and ARIMA.

Financial Time Series

A time series is a sequence of vectors $X_t = (x_{t-n}, \ldots\ldots, x_{t-i}, \ldots\ldots, x_{t-2}, x_{t-1})$ where x_{t-i} represents past value that varies with time. In stock prices also the data are typical example of time series. Models in the study predict the future trends on the basis of the past values. In the current study open, high, low and close of BSE 30 (SENSEX), BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty are used for the prediction of next stock value and future trend. The price of volume traded is not a significant predictor for the time series as during learning in the ANN model most of the prices against traded volume are zero as the data is not available with respect to volume traded. Thus traded volume is dropped as one of the variables or predictor.

Table 1. Input and Output Time Series for the Study.

Index Name	Input Variables				Output Variable
BSE 30 (SENSEX) from april 16,2004 to april 16,2009	Open	High	Low	Close	Next day close price
BSE IT from april 16,2004 to april 16,2009	Open	High	Low	Close	Next day close price
BSE Oil & Gas from august 23,2004 to april 16,2009	Open	High	Low	Close	Next day close price
BSE 100 from april 16,2004 to april 16,2009	Open	High	Low	Close	Next day close price
S&P CNX Nifty from april 16,2004 to april 16,2009	Open	High	Low	Close	Next day close price

Table 2. Control Parameters used in the Hybrid Model (Error Forecast through ANN to ARIMA) in the Study.

Model	Index	ANN	ARIMA (input for ARIMA will be residual from ANN) ARIMA(p.d.q.)=
ANN_ARIMA (Hybrid Model)	BSE30 SENSEX	ANN(4-4-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.01, Epochs=1000, Error Tolerance=0.001	ARIMA(1,0,0)
	BSEIT	ANN(4-4-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.01, Epochs=1000, Error Tolerance=0.001	ARIMA(0,0,1)
	BSE Oli & Gas	ANN(4-4-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.01, Epochs=1000, Error Tolerance=0.001	ARIMA(1,0,1)
	BSE 100	ANN(4-4-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.01, Epochs=1000, Error Tolerance=0.001	ARIMA(1,0,0)
	S&P CNX NIFTY	ANN(4-4-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.01, Epochs=1000, Error Tolerance=0.001	ARIMA(1,1,1)

Table 3. Control Parameters used in the Hybrid Model (Error Forecast through ARIMA to ANN) in the Study.

Model	Index	ARIMA	ANN (input for ANN will be residual from ARIMA)
ARIMA_ANN (Hybrid Model)	BSE30 (SENSEX)	ARIMA(p,d,q)= ARIMA(1,1,1)	ANN(5-5-1)Linear-log sigmoid-log sigmoid, α =0.5, η=0.001, Epochs=1000, Error Tolerance=0.01
	BSEIT	ARIMA(p,d,q)= ARIMA(0,1,2)	ANN(5-5-1)Linear-log sigmoid-log sigmoid, α =0.9, η=0.001, Epochs=1000, Error Tolerance=0.01
	BSE Oli & Gas	N/A	N/A
	BSE 100	ARIMA(p,d,q)= ARIMA(1,1,1)	ANN(5-5-1)Linear-log sigmoid-log sigmoid, α =0.5, η=0.001, Epochs=1000, Error Tolerance=0.01
	S&P CNX NIFTY	ARIMA(p,d,q)= ARIMA(1,1,1)	ANN(5-5-1)Linear-log sigmoid-log sigmoid, α =0.5, η=0.001, Epochs=1000, Error Tolerance=0.01

** In ANN 5-5-1 five input values are set containing samples of five values of residuals generated from ARIMA using suitable model.*

Use of ANN in Hybrid Modelling

Artificial neural network model has been used in two different ways in the first case where input data set comprising of open, high, low and close price time series of different indexes selected for the study is provided to the ANN model and in the second case the residual generated from suitable ARIMA model is provided to ANN model. Hence in the first case data set contains four different time series with respect to open, high, low and close price whereas in the second case data set contains time series generated as the residual from ARIMA model which are represented in the following diagrams (Figure 1 and Figure 2).

Input data for all Indices consist of data collected from http://www.cmie.com (CMIE Prowess database release 3.1, Date of Downloading: May 15, 2009).

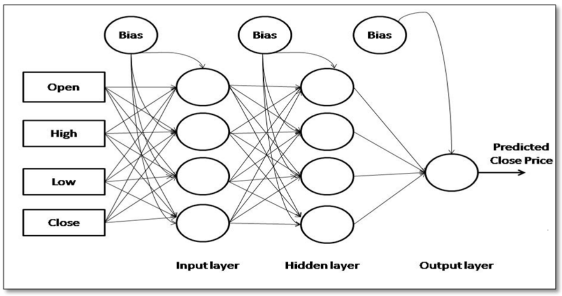


Figure 1. Three Layer Feed-forward Neural Network with One Hidden Layer for Hybrid Model (Error Forecast through ANN to ARIMA).

When data is loaded in the ANN (Artificial Neural Network), it must be preprocessed from its numeric range into the numeric range that the ANN can deal with efficiently. In this process, proper transformation of data simplifies the process of learning and may improve the generalizability of the learned results (Kim and Lee, 2004).

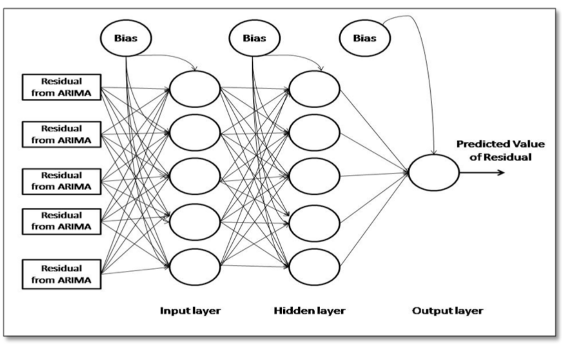


Figure 2. Three Layer Feed-forward Neural Network with One Hidden Layer for Hybrid Model (Error Forecast through ARIMA to ANN).

In artificial neural network model at the hidden and output layer log sigmoid transfer function has been used and its range lies in [0, 1]. So before providing the data to the network, the data is to be preprocessed or scaled. Once the input scaled data is processed through the artificial neural network and the output obtained, it is descaled. Methods for preprocessing and post processing data are as described as follows (Merh, Saxena and Pardasani, 2008):

[a] Procedure for Scaling of time series

(1)

$$F = \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}}$$

where $x_{\min}$ *to* $x_{\max}$ *are scaled from* $y_{\min}$ *to* $y_{\max}$

(2)

$$F = \frac{0.9 - 0.1}{x_{\max} - x_{\min}}$$

such that $y_{\max} = 0.9$ *and* $y_{\min} = 0.1$

(3)

$$Offset = (y_{\min} - (F) * x_{\min})$$

(4)

$$Offset = (0.1 - (F) * x_{\min})$$

The input data is preprocessed by using the following scaling function and the output data is again processed using reverse scale.

(5)

$$F * X + Offset$$

where X is the data which is to be processed

[b] Procedure for Descaling of time series

After processing the input value through the artificial neural network model proposed in the paper the data at the output layer is again scaled in the original numeric range using the following method:

(6)

$$y_{\max} = 0.9 \text{ and } y_{\min} = 0.1$$

(7)

$$F = \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}}$$

Where A is the scaled value which to be converted in the original numeric value.

(8)

$$val = \frac{A - y_{\min}}{F + x_{\min}}$$

Back propagation algorithm (supervised learning) is chosen for the training of the neural network, in which finite number of pairs of input patterns X_k and target output patterns Y_k in the form $(X_1, T_1), (X_2, T_2), \ldots$

..., (X_k, T_k) are initialized. Training dataset for artificial neural network module consists of approximately 1218 samples. The dataset used for prediction consist of 30 days closing price value. In model – I (Figure 1) for training open, high, low and close price has of about 1218 sample been provided to the model and next day close is provided as the target. While for the prediction next thirty days forecasted values have been generated from the model. In model-II (Figure 2) 1218 residuals generated from ARIMA are provided for training and residual of next day is the target for training. In the current research work control parameters learning rate, momentum, epochs, error tolerance, target and number of hidden layers are taken randomly.

Neural Network Architecture and Design

Network architecture consists of an input layer, one hidden layer and an output layer; index range for the above three layers are i=1,2,…….n, h=1,2,……..m and j=1,2,………p respectively and for input dataset index range is u=1,2,…n. In the model, selection of number of hidden layers is random. Linear transfer function is used at the input layer which receives the external real input and produces the same output, while for the output layer and the hidden layer log sigmoid transfer function is used as the transfer function. Results from the research (Merh, Saxena and Pardasani, 2008) show that in the case of only three layers an input layer, one hidden and an output layer results where linear function at the input layer, log sigmoid transfer function at the hidden & output layer are used and when linear function at the input layer, tan sigmoid transfer function at the hidden layer and output layer are used, generate similar patterns and errors. The selection of the sigmoid as a neuron signal function is often justified by the biological argument that it represents the behaviour of the firing rate of action potential, averaged over a population of neurons (Kumar, 2004; Koch, 1999).

More specifically, for the input layer the values are as follows:

$$x_i^k = \sum_{u=0}^{n} w_{ui}^k p_i^k, i = 1, 2, \dots\dots, n \tag{9}$$

$$f(x_i^k) = x_i^k, i = 1, 2, \dots\dots, n \tag{10}$$

where x_i^k is i^{th} component of the input vector X_k and k is the iteration index

Input for the hidden layer is as follows:

$$z_h^k = \sum_{i=0}^{n} w_{ih}^k x_i^k, h = 1, 2, \dots\dots, m \tag{11}$$

Output generated by the hidden layer will be as follows:

$$f(z_h^k) = 1/(1 + e^{-z_h^k}), h = 1, 2, \dots\dots, m \tag{12}$$

Input for the output layer of the neural network will be:

$$y_j^k = \sum_{h=0}^{m} w_{hj}^k f(z_h^k), j = 1, 2, \dots\dots, p \tag{13}$$

Output derived at the output layer is as follows:

$$f(y_j^k) = 1/(1 + e^{-y_j^k}), j = 1, 2, \dots\dots, p \tag{14}$$

In the above architecture biases are included with weights. The weights having index 0 will be treated as bias for the given layer and neuron i.e., w_{0i}^k, w_{0h}^k, *and* w_{0j}^k are the biases for the input layer, hidden layer and output layer respectively. The inputs datasets and output dataset is defined in Table 1.

After deciding the architecture of the neural network, next step is to evaluate the other control parameters as indicated in Table 2 and Table 3 which are important for the development of the overall network structure. These factors include learning rate, momentum factor, initial weight, error function, bias function and target output. For the estimation of above said parameters, various strategies are required rather it is an art to decide these parameters. The current work without emphasizing on these strategies assumes learning rate, momentum, epochs, error tolerance and target randomly.

Learning Process

Back propagation algorithm (supervised learning) is chosen for the training of the neural network, in which finite number of pairs of input patterns X_k and target output patterns T_k in the form $(X_1, T_1), (X_2, T_2), \ldots \ldots, (X_k, T_k)$ are initialized.

Instantaneous error E_k for k^{th} learning pattern will be:

$$E_k = T_K - f(y_k) \tag{15}$$

where

$$T_K = (e_1^k, e_2^k, \ldots\ldots, e_p^k)^T = (t_1^k - f(y_1^k), \ldots\ldots, t_p^k - f(y_p^k)) \tag{16}$$

$$M_k = \frac{1}{2}\sum_{j=1}^{p}(t_j^k - f(y_j^k))^2 \tag{17}$$

Mean Square Error on the entire training set T on a particular neural network will be as follows:

$$M = \frac{1}{Q}\sum_{k=1}^{Q}(M)_k \tag{18}$$

where M is the mean square error and k is the iteration index. We have a set of Q training vector pairs.

Multilayer Error Correction Learning

Processing element (PE) is a portion of the artificial neural network where all computing is performed. A continuously differentiable processing element function for the hidden layer PEs would allow chain rule of partial differentiation to be used to calculate weight changes for any weight in the network. This was realized while solving credit assignment problem (Barto, 1984; Minsky, 1961; Sinencio and Lau, 1992).

Multilayer learning (in this case three layer learning) for the Figure 3 can be explained as follows (Sinencio and Lau, 1992):

The output error across the entire F_Z processing element is found by the following cost function:

$$\Sigma = \tfrac{1}{2}\sum_{j=1}^{q}(b_{kj} - z_j)^2 \tag{19}$$

Following equation computes the output of F_Z processing element z_j:

$$z_j = \sum_{1=1}^{p} y_i W_{ij} \tag{20}$$

For the hidden layer F_Y PE, y_i is calculated by the following equation:

$$y_i = f\left(\sum_{h=1}^{n} a_{kh} v_{hi}\right) = f(r_i) \tag{21}$$

$$\text{where } r_i = \sum_{h=1}^{n} a_{kh} v_{hi} \tag{22}$$

The weights adjustment is performed by moving along the cost function in the opposite direction of the gradient to a minimum (minimum is the input /output mapping producing the least amount of the total error). The connection weights between the F_Y and F_Z processing elements are adjusted and for the two layers the error correction learning will be explained by the following equations:

$$\partial E/\partial w_{ij} = \partial/\partial w_{ij}\left[\tfrac{1}{2}\sum_{j=1}^{q}(b_{kj} - z_j)^2\right] \tag{23}$$

$$= (b_{kj} - z_j)\, y_i \tag{24}$$

$$= \delta_j y_i \tag{25}$$

A positive constant valued learning rate (η) has been added to adjust the amount of change with each move down the gradient.

Adjustments to the connection weights between the F_X and F_Y processing elements are derived using the chain rule (Sinencio and Lau, 1992)

$$\partial E/\partial v_{ih} = (\partial E/\partial y_i)(\partial y_i/\partial r_i)(\partial r_i/\partial x_h)(\partial x_h/\partial v_{hi}) \quad (26)$$

$$= \sum_{i=1}^{p}(b_{kl} - y_l)\, y_l\, w_{hl} f'(r_i)\, a_{kh} \quad (27)$$

A positive constant valued learning rate (β) has been added to adjust the amount of change with each move down the gradient.

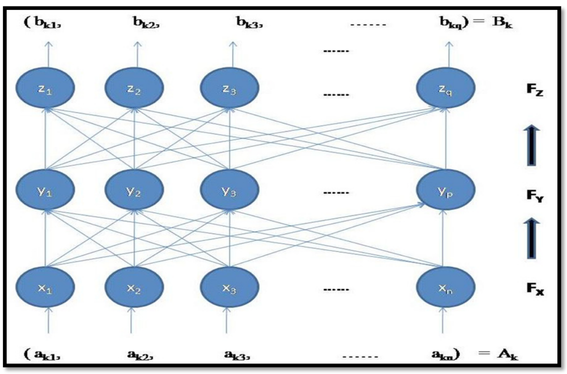


Figure 3. Three Layer Neural Network With set of Inputs and Outputs.

From equations 25 and 27 weight adjustment equation becomes

$$w_{ij}^{new} = w_{ij}^{old} - \eta(\partial E/\partial w_{ij}) \quad (28)$$

and

$$w_{hi}^{new} = w_{hi}^{old} - \beta(\partial E/\partial v_{hi}) \quad (29)$$

Large learning rate can lead to oscillations during learning. A refined method of increasing learning rate is to introduce a momentum term into weight update procedure, where (learning rate) > 0. The training process will continue until the error tolerance reaches some specific value let the range be 0.1 to .0001, depending upon the application.

Use of ARIMA in Hybrid Modelling

The ARIMA model is representative of linear models and has achieved great popularity since the publication of Box Jenkins' classic book: Time-Series Analysis: Forecasting and Control (Box and Jenkins, 1976; Zhang, 2001). Different stages of designing and implementing an ARIMA (p, d, q) based model are identification, estimation, diagnosis and forecasting. If seasonal component is also included then the model becomes ARIMA (p, d, q) $(P, D, Q)_s$ where s is number of periods per session and P, D, Q are seasonal components. Financial time series can be stationary or non-stationary. For converting a time series to a stationary series it is necessary to difference the time series d times. ARIMA (p, d, q) is applying ARMA (p, q) to a stationary series where p is the order of auto regression, q is the order of moving-average and d is the order of differencing. It is also known as Box – Jenkins (BJ) Method. Training dataset ARIMA consists of more than 1218 samples.

Hybrid Model of ANN and ARIMA (ANN_ARIMA)

It may be reasonable to consider a time series to be composed of a linear autocorrelation structure and a non-linear component. In the proposed model there are mainly two stages. In the first stage ANN is used to forecast the future value of the close price and then the residual generated is provided to the ARIMA which will forecast the error forecast. In the second stage the predicted close price by ANN is summed with the error forecast generated by ARIMA which produces the final forecasted value.

Let X_t be the time series for t = 1, 2, 3… which generates N_t as the forecasted series by implementing ANN and the residual (e_t) will be

$$e_t = X_t - N_t \text{ where } t = 1, 2, 3, 4... \quad (30)$$

Later ARIMA is used to model e_t (residual of the ANN forecasting), which will generate a series of forecast, let it be P_t . Studies (Zhang, 2003; Sallehuddin, et al., 2008) indicate that, this step can be said as a process of error generation of time series prediction and ARIMA is used as error correction of multivariate time series forecasting for ANN model (Sallehuddin, et al., 2008).

Final forecasting X_p will be as follows:

$$X_p = N_t + P_t \quad (31)$$

Framework of the proposed model is shown in figure 4.

As seen in the Figure 4 the first step is the input time series selection. In the present model five data sets BSE 30, BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty with four variables each has been used. In the next step the time series is used for training then prediction through ANN. The ANN forecasts the next thirty days closing price of SENSEX from February 26, 2009 to March 16, 2009. Residual generated by prediction through ANN is provided to the ARIMA which in turn generates its residual series. Finally the residual series generated by ARIMA is added to the predicted series generated by ANN, thus giving the final forecasted values.

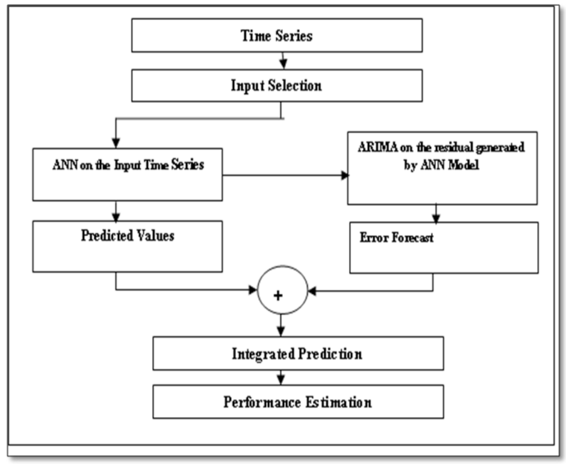


Figure 4. Proposed Hybrid Approach (ANN and ARIMA) (Flow Chart of Error Forecast through ANN to ARIMA).

Merged ARIMA: Model Identification, Estimation and Diagnosis

The input series is the residual or error generated from the difference of actual and predicted closing price of the indices from ANN model during the same time period. First stage is to verify that whether the series is stationary or not with the help of sequence graphs in SPSS 13.0®. It can be seen from the Figure 5 that series show almost constant mean and variance. Hence it is clear that the time series can be said a stationary series and no differencing is required. ACF and PACF on the input series will help to find out the tentative values of p, d and q three parameters of ARIMA (p, d, q)

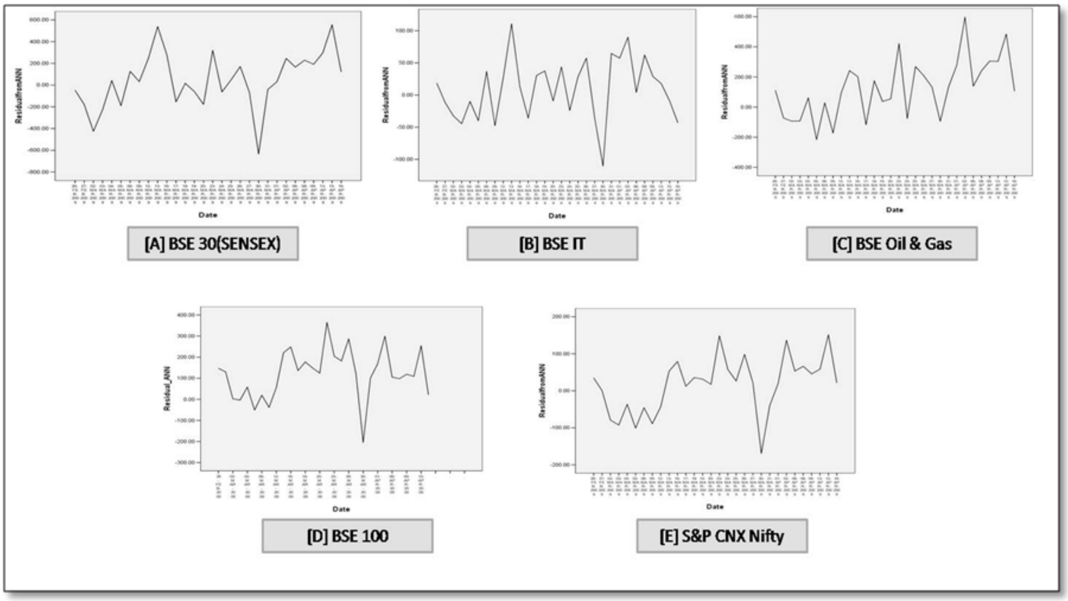


Figure 5. Sequence Graph of Residual Generated from ANN Model Indicates that for Indices BSE 30, BSE IT, BSE Oil& Gas, BSE 100 and S&P CNX Nifty Input Time Series are Stationary.

Values of ACF, PACF, P-Value (significance) and LB Statistics (Q*) for the maximum lag 10 also indicates that the residual series is stationary do not needs any differencing or transformation using natural log (Table 4). It is also found that in case of S&P CNX Nifty after taking first order differencing results are better.

Table 4. ACF, PACF, Significance Value and LB -Statistics Q Value for all the Indices Taken in The Study at Ten Degrees of Freedom.*

Index	ACF	PACF	Significance (P-Value)	LB - Statistic (Q*)
BSE 30(SENSEX)	-0.073	-0.230	0.190	13.641
BSEIT	-0.080	-0.178	0.345	11.166
BSE Oil & Gas	-0.160	-0.199	0.242	12.684
BSE 100	-0.240	-0.285	0.708	7.188
S&P CNX Nifty	-0.087	-0.161	0.311	11.625

After analyzing ACF, PACF and performing various iterations by taking different p and q parameter values with d as 0 and best fit was found having minimum Akaike Information criteria (AIC) value. The best fit was selected after applying various p and q values the model. Table 5 shows the best fit model for five cases. To check the adequateness of the model the ACF and PACF plots of the error between the input series and its fit were generated.

Table 5. Best Fit ARIMA Model for the Given Data Set.

Index	p	d	q	ARIMA (p, d, q)
BSE 30(SENSEX)	1	0	0	ARIMA(1,0,0)
BSEIT	0	0	1	ARIMA(0,0,1)
BSE Oil & Gas	1	0	1	ARIMA(1,0,1)
BSE 100	1	0	0	ARIMA(1,0,0)
S&P CNX Nifty	1	1	1	ARIMA(1,1,1)

In the process of diagnosis it is verified whether the residual or error generated if white noise or not. The autocorrelations specifying Q*-value, P-value (significance) at different degrees of freedom having a lag of maximum 10 are shown in the table (Table 6). To determine whether the time series is white noise or not, the Box-Ljung Q* statistic is compared with the chi-square distribution with (h-m) degrees of freedom. Here h is the number of lags and m is the number of parameters. Table 6 clearly demonstrates that the Q* statistic has same distribution as chi-square with (h-m) degrees of freedom. The plots and the autocorrelations generated indicate that the model fits well.

Table 6. LB Statistics and P-Value of the Error or Residual Generated from ARIMA Model for the Verification of White Noise.

Index	(h-m) Degrees of Freedom	LB - Statistics Q*	Significance (P-Value)	Is Series White Noise
BSE 30(SENSEX)	(10-1)=9	10.209	0.334	Yes
BSEIT	(10-1)=9	10.438	0.316	Yes
BSE Oil & Gas	(10-2)=8	5.234	0.732	Yes
BSE 100	(10-1)=9	6.240	0.716	Yes
S&P CNX Nifty	(10-2)=8	2.930	0.939	Yes

Forecasting Using Hybrid ANN and ARIMA

Figure 4 outlines the forecasting procedure. The above mentioned ARIMA models are not used for forecasting purpose. The fit of the input (residual of ANN) series generated by the ARIMA models is summed with the predicted close values generated by ANN. This results in the final predicted value.

Hybrid Model of ARIMA and ANN (ARIMA_ANN)

The model has same components as the ANN_ARIMA model. The difference is that here ARIMA is used to forecast the values and ANN is used to model the residuals summation of the ARIMA forecasted stock price and the error forecast generated by ANN produces the final forecasted value.

Let X_t be the time series for t = 1, 2, 3… which generates N_t as the forecasted series by implementing ARIMA and the residual (e_t) will be

$$e_t = X_t - N_t \text{ where } t = 1, 2, 3, 4... \quad (32)$$

Later ANN is used to model e_t (residual of the ANN forecasting), which will generate a series of forecast, let it be P_t. Studies (Zhang, 2003; Sallehuddin, et al., 2008) indicate that, this step can be said as a process of error generation of time series prediction and ANN is used as error correction of multivariate time series forecasting for ARIMA model (Sallehuddin, et al., 2008).

Final forecasting X_p will be as follows:

$$X_p = N_t + P_t \quad (33)$$

Framework of the proposed model is shown in figure 6.

As seen in the Figure 6 the first step is the input time series selection. In the present model five data sets BSE 30, BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty each has been used. In the next step the time series is used for model identification and the 30 samples of hold out data for forecasting through ARIMA. This forecasts the next thirty days closing price of SENSEX from February 26, 2009 to April 16, 2009. Residual generated by ARIMA is provided to ANN for residual analysis. Finally the residual predicted is added to the ARIMA forecasts thus giving the final forecasted values.

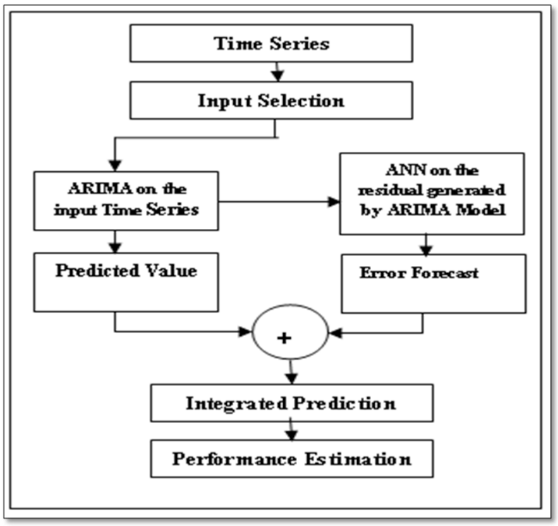


Figure 6. Conventional Hybrid Approach (ARIMA and ANN) (Flow Chart of Error Forecast through ARIMA to ANN).

It was found that without transforming the series using natural logarithm and differencing it was non-stationary (Figure 7 (a), Figure 7 (b), Figure 7 (c), Figure 7 (d)). If the series is transformed only by taking differencing or natural logarithm then also series is not becoming stationary. After taking the natural logarithm and differencing of first order all input series (BSE 30, BSE IT, BSE 100 , S&P CNX Nifty) became stationary (Figure 8) which is evident from the values of ACF, PACF, LB statistics (Q*) and P-value (Table 7). In case of BSE Oil & Gas even after transformation through natural logarithm and differencing of order two the series does not become stationary this is because of heteroscedasticity in the data; hence it shows that ARIMA modelling is not suitable technique for BSE Oil & Gas.

Table 7 ACF, PACF, Significance Value and LB -Statistics Q Value for all the Indices Taken in the Study at Three Hundred (Maximum) Degrees of Freedom.*

Index	ACF	PACF	Significance (P-Value)	LB - Statistic (Q*)
BSE 30(SENSEX)	0.018	-0.007	0.191	321.255
BSEIT	0.005	-0.011	0.353	308.625
BSE 100	0.022	-0.004	0.114	329.811
S&P CNX Nifty	0.021	-0.011	0.222	318.425

†BSE Oil & Gas is not fitting in ARIMA modelling.

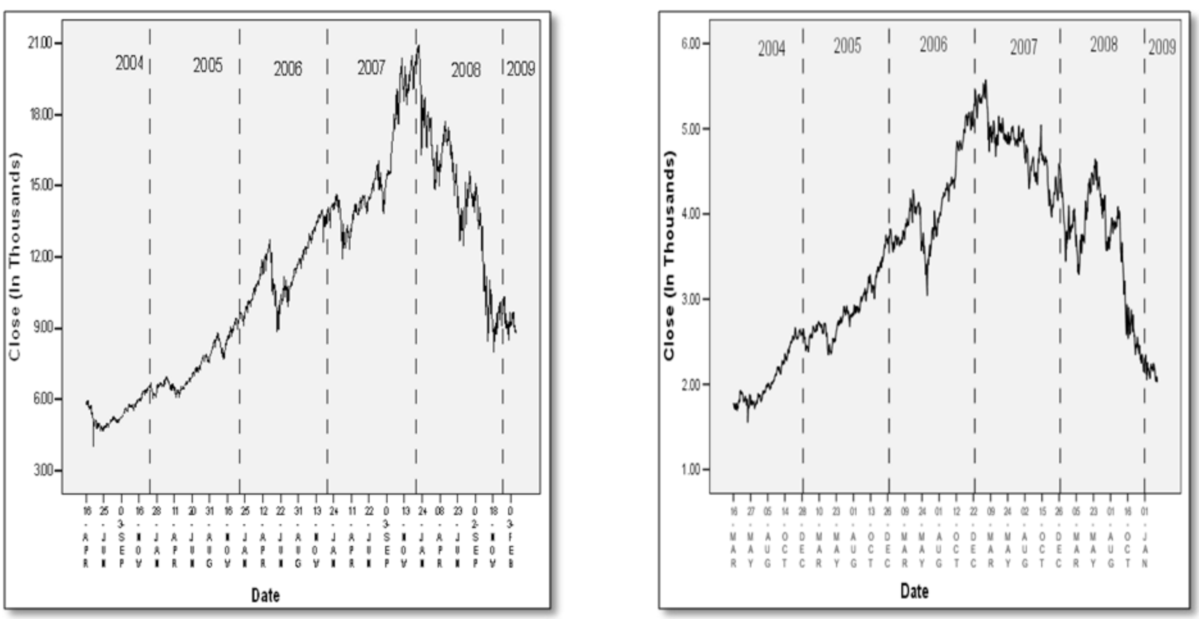
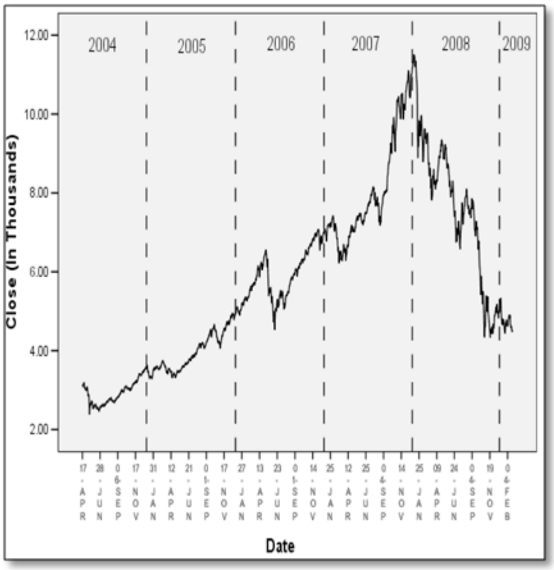


Figure 7. (a) Sequence Graph of BSE 30 Close Price from April 16, 2004 to April 16, 2009. Figure 7 (b) Sequence Graph of BSE IT Close Price from April 16, 2004 to April 16, 2009.



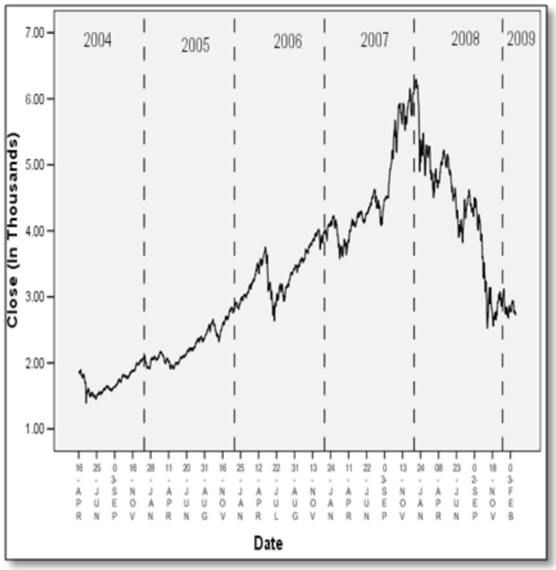


Figure 7. (c) Sequence Graph of BSE 100 Close Price from April 16, 2004 to April 16, 2009. Figure 7 (d) Sequence Graph of S&P CNX Nifty Close Price from April 16, 2004 to April 16, 2009.

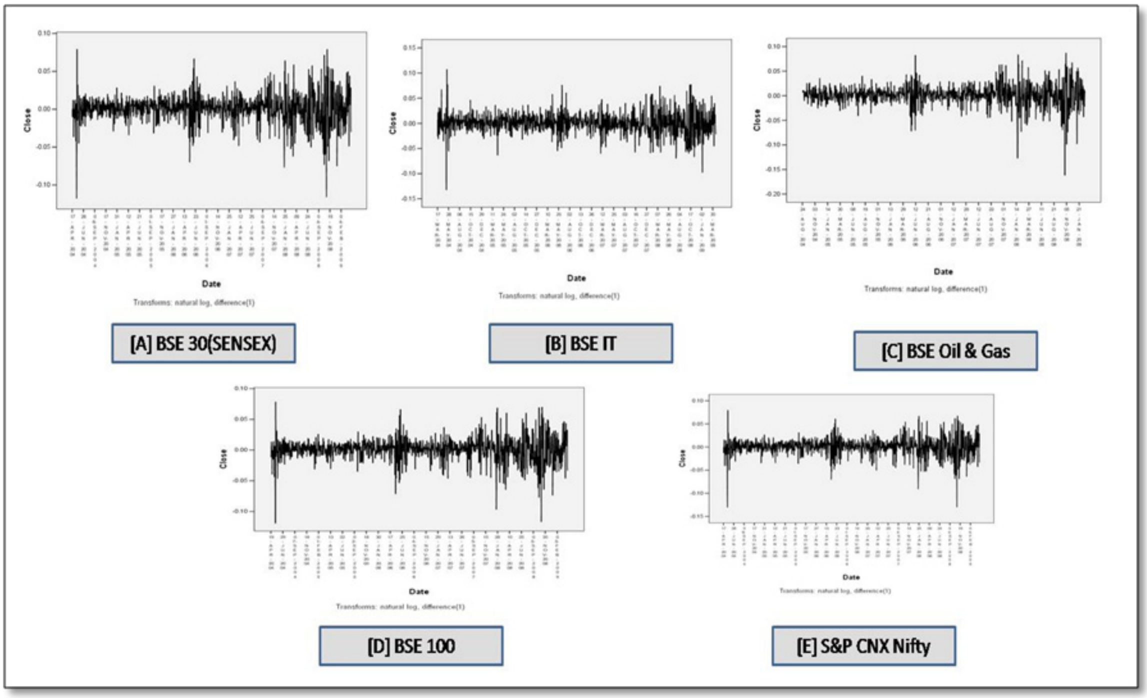


Figure 8. Sequence Graph of Transformed Input Time Series for Indices BSE 30, BSE IT, BSE Oil& Gas, BSE 100 and S&P CNX Nifty. Series is Transformed Using Natural Logarithm and Differencing of First Order.

After analyzing ACF, PACF and performing various iterations by taking different p and q parameter values with d as 0 and best fit was found having minimum AIC value. The best fit was selected after applying various p and q values the model. Table 8 shows the best fit model for five cases. To check the adequateness of the model the ACF and PACF plots of the error between the input series and its fit were generated.

Table 8. Best Fit ARIMA Model for the Given Data Set.

Index	p	d	q	ARIMA (p, d, q)
BSE 30(SENSEX)	1	1	1	ARIMA(1,1,1)
BSEIT	0	1	2	ARIMA(0,1,2)
BSE 100	1	1	1	ARIMA(1,1,1)
S&P CNX Nifty	1	1	1	ARIMA(1,1,1)

†BSE Oil & Gas is not fitting in ARIMA modelling.

In the process of diagnosis it is verified whether the residual or error generated if white noise or not. The autocorrelations specifying Q*-value, P-value (significance) at different degrees of freedom having a lag of maximum 300 are shown in the table (Table 9). To determine whether the time series is white noise or not, the Box-Ljung Q* statistic is compared with the chi-square distribution with (h-m) degrees of freedom. Here h is the number of lags and m is the number of parameters. Table 10 clearly demonstrates that the Q* statistic has same distribution as chi-square with (h-m) degrees of freedom. Plots and the autocorrelations generated indicated that the model fits well.

Table 9. LB Statistics and P-Value of the Error or Residual Generated from ARIMA Model for the Verification of White Noise.

Index	(h-m) Degrees of freedom	LB-Statistics Q*	Significance (P-Value)	Is Series White Noise
BSE 30 (SENSEX)	(300-2)=298	289.956	0.620	Yes
BSEIT	(300-2)=298	279.365	0.774	Yes
BSE 100	(300-2)=298	280.680	0.608	Yes
S&P CNX Nifty	(300-2)=298	295.677	0.527	Yes

†BSE Oil & Gas is not fitting in ARIMA modelling.

Residual derived from suitable ARIMA model is supplied to ANN 5-5-1 model for residual analysis and residual is forecasted as shown in Figure 2. Control parameters for models are explained in Table 3. Various errors derived are explained in table 10.

Results and Simulation Comparisons

Comparisons and performance of above models have been evaluated by calculating errors between the actual close price and the predicted close price generated by all the models (Hybrid ARIMA_ANN and Hybrid ANN_ARIMA). In the study average absolute error (AAE), root mean square error (RMSE), mean absolute percentage error (MAPE)and mean square percent error MSPE) have been calculated. These measures help in calculating and comparing accuracy of different techniques. It will not only help in measuring a particular technique's reliability, usefulness but also help in searching the optimal technique (Hanke and Wichern, 2007).Figure 9 displays the comparison between the actual close of BSE 30 (SENSEX), BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty forecasted close price generated by the models discussed above during Feb 26, 2009 to April 16, 2009.

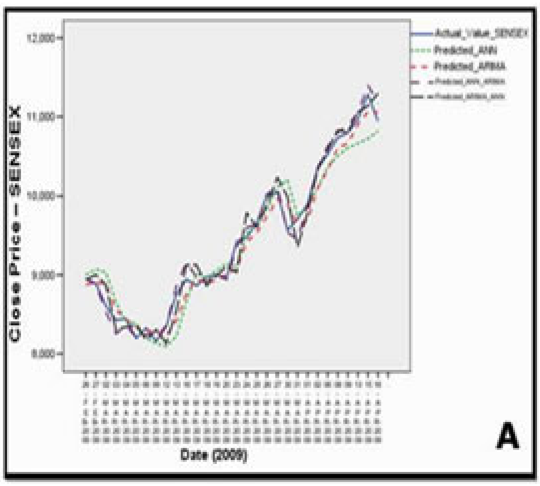
Table 10. Comparison between the Errors of Hybrid Model of ANN_ARIMA and Hybrid ARIMA_ANN Models.

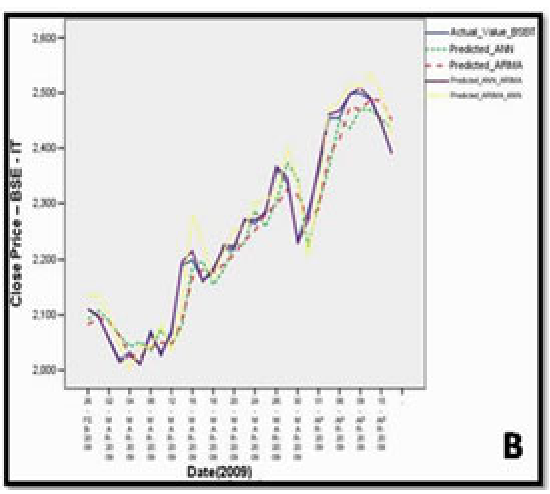
Model	Index	Errors			
		AAE	RMSE	MAPE	MSPE
ANN_ARIMA	BSE30 (SENSEX)	77.8306	102.1076	0.8200	0.011496
	BSEIT	5.8395	7.1693	0.2591	0.001001
	BSE Oil & Gas	91.8356	117.5227	1.2688	0.023937
	BSE 100	90.9771	108.8592	1.8602	0.04916
	S&P CNX Nifty	36.4592	44.1206	1.2430	0.022582

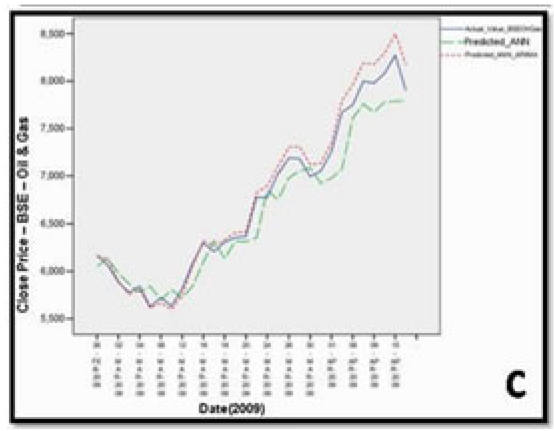
Model	Index	Errors			
		AAE	RMSE	MAPE	MSPE
ARIMA_ANN	BSE30 (SENSEX)	164.0544	198..9514	1.7619	0.0451531
	BSEIT	40.4625	46.74907	1.8194	0.44264
	BSE 100	87.1768	104.6332	1.8152	0.0469203
	S&P CNX Nifty	43.8553	56.8648	1.4678	0.034431

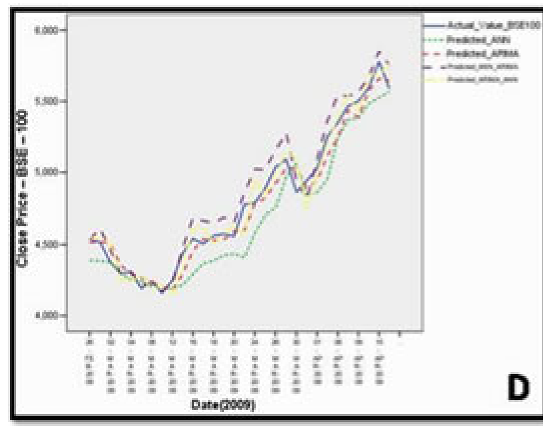
f†BSE Oil & Gas is not fitting in ARIMA modelling.

Snapshot of the Output that Compares Actual and Predicted Close Prices of SENSEX, BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty.









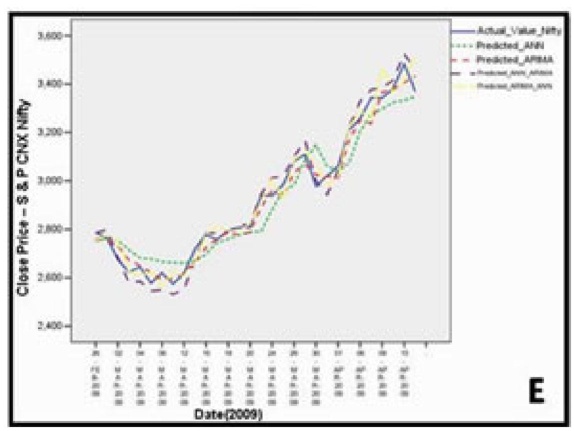


Figure 9. Snapshots of the Output that Compares Actual and Predicted Close Prices of SENSEX (A), BSE IT (B), BSE Oil & Gas (C), BSE 100 (D) and S&P CNX Nifty (E) Generated by ANN, ARIMA, ANN_ARIMA and ARIMA_ANN for the year 2009.

Conclusion

In the current research an attempt was made to study whether a hybrid model achieves better results than an individual mode? Two hybrid models were developed using techniques of ANN and ARIMA comparison is also made between the conventional hybrid approach using ARIMA, ANN and proposed hybrid approach using ANN, ARIMA. A hybrid model has been developed using techniques of back propagation artificial neural networks (ANN) and ARIMA for analyzing BSE 30 (SENSEX), BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty for forecasting the next day's closing price and predicting the stock trend. In the current study, BSE 30 (SENSEX), BSE IT, BSE Oil & Gas, BSE 100 and S&P CNX Nifty open, high, low, and close prices have been taken as the input data set and the output generated is the predicted close price of the same.

Results show that prediction of hybrid ANN_ARIMA model for BSE 30 (SENSEX), BSE IT and S&P CNX Nifty are better than hybrid ARIMA_ANN whereas in the case of BSE 100 results of ARIMA_ANN are marginally better then hybrid ANN_ARIMA. Hybrid Model ARIMA_ANN(Error Forecast through ARIMA to ANN) exhibited over-fitting problem. Various trials were done by modifying the learning rate and momentum. Finally the learning rate was reduced and momentum changed accordingly to resolve this problem. Results also demonstrate that for BSE Oil & Gas ANN and hybrid ANN_ARIMA are able to manage the input data set and predict the future closing price, whereas ARIMA and ARIMA_ANN fails to forecast future values.

References

Abu- Mostafa, Y. S., et al. (2001). Neural Networks in Financial Engineering, IEEE Transactions on Neural Networks, 12(4): 653-656.

Afolabi, Mark O. and Olude, Olatoyosi (2007). Predicting Stock Prices Using a Hybrid Kohonen Self Organizing Map (SOM), Proceeding of 40th International Conference on Systems Sciences, 1560-1605.

Barto, A., (1984). Simulation Experiments with Goal-Seeking Adaptive Elements, Air Force Wright Aeronautical Laboratory, Technical Report, AFWAL-TR-84-1022.

Box, G.E.P. and Jenkins, G.M., (1976). Time Series Analysis: Forecasting and Control. San Francisco, CA: Holden-Day.

Brabazon, T. (2000). A Connectivist Approach to Index Modelling in Financing Markets, Proceedings of Coil / EvoNet Summer School. University of Limerick, Briys, E., Bellalah, M., Mai, H. M. and Varenne, F. de (1998). Options, Futures and Exotic Derivatives. England: John Wiley & Sons Ltd.

Fama, E.F. (1965). The Behaviour of Stock Market Prices, Journal of Business, 38: 34-105.

Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance, 25: 1465-1468.

Hanke, John E. and Wichern, Dean W. (2007). Business Forecasting. New Delhi: Pearson Education, Inc, 79-80.

http//www.cmie.com (CMIE Prowess database release 3.1, Date of Downloading: May 15, 2009).

Kate A., Simth, Jatinder, N. D, Gupta (2000). Neural Networks in Business: Techniques and Applications for the Operations Researcher, Computer and Operations Research, 27: 1023-1044.

Kim, Kyoung-jae and Lee, Won Boo (2004). Stock Market prediction using Artificial Neural Networks with Optimal Feature Transformation, Neural Computing and Application, 13: 255–260.

Koch, C., (1999), "Biophysics of Computation: Information Processing in Single Neuron", Oxford university Press, New York.

Kumar, Satish (2004), Neural Networks: A Classroom Approach, Tata McGraw-Hill Publishing Company Limited, New Delhi.

Lawrence, S., Tsoi, A. C., and Giles, C. L. (1996). Noisy Time Series Prediction Using Symbolic Representation and Recurrent Neural Network Grammatical Inference. Technical report UMIACS-TR-96-27 and CS-TR-3625. Institute for Advanced Computer Studies, University of Maryland.

Malkiel, B.G. (1996). A Random Walk Down Wall Street. New York: W. W. Norton, 6th Edition.

Miao, K., Chen, F., and Zhao, Z. G. (2007). Stock Price Forecast Based on Bacterial Colony RBF Neural Network, Journal of QingDao University, 20 (2): 50–54 (in Chinese).

Mills, T. C. (1990). Non-linear Time Series Models in Economics, Journal of Economic Surveys, 5: 215–241.

Minsky, M. (1961). Steps Towards AI, Proceeding of the IRE, 49: 5-30.

Merh, Nitin, V.P. Saxena and Kamal Raj Pardasani (2008), Artificial Neural Network for Stock Market Forecasting, Nirma University Journal of Business and Management Studies (NUJBMS), 2(3 & 4): 3-19.

Murphy, J. J. (1999). Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York: Institute of Finance, 4-5.

Priestley, M. B. (1988). Non-linear and Non-stationary Time Series Analysis. London: Academic Press.

Refenes, A. P., (Ed) (1995). Neural Network in the Capital Markets. England:John Wiley & Sons Ltd, 149-161.

Sallehuddin, Roselina., Shamsuddin, Siti Mariyam. and Hashim, Siti Zaiton Mohd (2008). Hybridization Model of Linear and Non-linear Time Series Data for Forecasting. Second Asia International Conference on Modelling & Simulation, IEEE Conference Proceedings, DOI 10.1109/AMS.2008.142, 597 – 602.

Samanta G.P. and Bordoloi, Sanjib (2005). Predicting Stock Market- An Application of Artificial Neural Network Technique Through Genetic Algorithm, Finance India, 19(1): 173-188.

Sinencio, Edgar Sánchez and Lau, Clifford (1992). Artificial Neural Networks:

Paradigms, Applications and Hardware Implementations. New York: IEEE Press, 16-17.

Skabar, Andrew and Cloete, Ian (2001). Neural Networks, Financial Trading and The Efficient Markets Hypothesis, Australian Computer Society, Inc. , Twenty-Fifth Australasian Computer Science Conference (ACSC2002), Melbourne, Australia, Conferences in Research and Practice in Information Technology, 4: 241 – 249.

Termizel, Taskaya T., and Casey, M.C. (2005). A Comparative Study of Autoregressive Neural Network Hybrids, Neural Networks, 18(5-6): 781-789.

Tseng, Fang-Mei., Yu, Hsiao-Cheng and Tzeng, Gwo-Hsiung (2002). Combining Neural Network Model with Seasonal Time Series ARIMA model, Technological Forecasting & Social Change, 69(1): 71–87.

Windrow, B, D Rumelhart, E. and Lehr, M.A (1994). Neural Networks Applications in Industry, Business and Science, Communications of the ACM, 37(3): 93-105.

Yao, J.T., and Poh, H. L. (1996). Equity Forecasting: A Case Study on the KLSE Index. Neural Networks in Financial Engineering, Proceedings of the Third International Conference on Neural Networks in the Capital Markets. World Scientific, Singapore, 341–353.

Yudong, Zhang and Lenan, Wu (2009). Stock Market Prediction of S&P 500 via Combination of Improved BCO Approach and BP Neural Network, Expert Systems with Applications, 36: 8849–8854.

Zhang, Defu., Jiang, Qingshan and li, Xin (2004). Application of Neural Networks in Financial Data Mining, International Journal of Computational Intelligence, 1(2): 116-119.

Zhang, G. Peter (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model, Neurocomputing, 50: 159 – 175.

Zhang, Guoqiang Peter (2001). An Investigation of Neural Networks for Linear Time-Series Forecasting, Computers and Operations Research, 28: 1183-1202.

Annexure

Table 11. Comparison between the Actual and Predicted Values Generated by ANN, ARIMA, Hybrid ANN_ARIMA and Hybrid ARIMA_ANN Models for BSE 30 (SENSEX).

Date	Actual Values of BSE 30 (SENSEX)	Predicted Close Value of BSE 30 Using Following Model			
		ANN	ARIMA	Hybrid ANN_ ARIMA	Hybrid ARIMA_ ANN
26-FEB-2009	8954.90	9001.30	8875.57	8954.90	8940.26
27-FEB-2009	8891.60	9070.80	8908.93	8872.95	9003.59
02-MAR-2009	8607.10	9033.20	8875.15	8535.06	8873.20
03-MAR-2009	8427.30	8653.80	8558.21	8256.00	8305.54
04-MAR-2009	8446.50	8403.70	8441.90	8355.44	8342.52
05-MAR-2009	8197.90	8388.60	8338.91	8215.11	8358.90
06-MAR-2009	8325.80	8199.10	8304.31	8249.14	8178.69
09-MAR-2009	8160.40	8128.70	8270.04	8211.33	8306.92
12-MAR-2009	8343.80	8095.30	8228.51	8356.54	8134.24
13-MAR-2009	8756.60	8217.80	8415.51	8856.50	8546.12
16-MAR-2009	8943.50	8663.90	8769.14	9160.10	9125.61
17-MAR-2009	8863.80	9019.10	8938.47	8976.20	9128.24
18-MAR-2009	8976.70	8960.20	8922.82	8914.27	8863.54
19-MAR-2009	9001.70	9060.60	8926.35	9008.33	8995.96
20-MAR-2009	8966.70	9145.20	9018.28	8943.02	9109.10
23-MAR-2009	9424.00	9103.50	9070.86	9352.24	9034.63
24-MAR-2009	9471.00	9534.10	9420.80	9599.84	9789.33
25-MAR-2009	9667.90	9619.70	9536.24	9642.53	9601.84
26-MAR-2009	10003.10	9832.50	9742.85	10022.48	9889.88
27-MAR-2009	10048.50	10116.90	9963.24	10117.08	10238.85
30-MAR-2009	9568.10	10202.60	9891.09	9540.60	9991.71
31-MAR-2009	9708.50	9748.20	9663.57	9453.42	9356.06
01-APR-2009	9902.00	9870.90	9686.73	9886.04	9791.43
02-APR-2009	10348.80	10103.00	10102.59	10361.30	10333.22
06-APR-2009	10534.90	10369.70	10329.97	10633.71	10591.58
08-APR-2009	10742.30	10513.10	10614.78	10808.71	10835.05
09-APR-2009	10803.90	10612.20	10668.07	10896.04	10811.00
13-APR-2009	10967.20	10669.70	10907.09	11044.27	11058.25
15-APR-2009	11284.70	10728.30	11063.38	11404.30	11138.88
16-APR-2009	10947.40	10825.60	11058.48	11171.08	11295.20

Table 12. Comparison between the Actual and Predicted Values Generated by ANN, ARIMA, Hybrid ANN_ARIMA and Hybrid ARIMA_ANN Models for BSE IT.

Date	Actual Values of BSE IT	Predicted Close Value of BSE IT Using Following Models			
		ANN	ARIMA	Hybrid ANN_ ARIMA	Hybrid ARIMA_ ANN
26-FEB-2009	2110.70	2092.10	2082.62	2110.70	2132.37
27-FEB-2009	2096.20	2107.70	2094.06	2098.99	2136.73
02-MAR-2009	2057.60	2090.00	2088.89	2055.41	2099.85
03-MAR-2009	2019.20	2063.80	2062.57	2014.56	2046.14
04-MAR-2009	2032.80	2042.70	2028.77	2026.67	2000.29
05-MAR-2009	2011.00	2051.10	2023.89	2010.42	2042.54
06-MAR-2009	2072.30	2035.40	2039.78	2066.23	2041.72
09-MAR-2009	2025.30	2073.00	2050.24	2031.89	2084.09
12-MAR-2009	2073.50	2046.40	2045.77	2065.17	2035.71
13-MAR-2009	2190.30	2079.40	2090.56	2195.74	2122.25
16-MAR-2009	2199.20	2186.20	2166.94	2215.38	2281.13
17-MAR-2009	2161.30	2197.30	2181.95	2160.81	2228.77
18-MAR-2009	2184.80	2154.40	2176.88	2179.35	2165.62
19-MAR-2009	2219.20	2181.80	2191.26	2224.70	2213.85
20-MAR-2009	2217.40	2226.60	2209.45	2222.30	2252.05
23-MAR-2009	2271.60	2227.70	2232.91	2269.44	2255.47
24-MAR-2009	2263.00	2287.00	2250.89	2270.07	2300.15
25-MAR-2009	2286.00	2257.30	2280.31	2281.23	2307.30
26-MAR-2009	2361.20	2303.70	2297.32	2366.34	2317.39
27-MAR-2009	2337.80	2373.80	2325.16	2345.84	2403.54
30-MAR-2009	2234.30	2344.70	2314.72	2227.54	2342.29
31-MAR-2009	2285.70	2221.10	2270.98	2269.79	2205.26
01-APR-2009	2358.10	2300.70	2290.72	2370.45	2319.84
02-APR-2009	2454.70	2364.60	2389.35	2461.61	2470.73
06-APR-2009	2454.70	2450.60	2416.18	2467.47	2482.51
08-APR-2009	2498.40	2436.00	2474.98	2497.07	2516.52
09-APR-2009	2499.30	2470.40	2469.76	2509.08	2508.05
13-APR-2009	2487.00	2469.20	2488.09	2489.93	2532.51
15-APR-2009	2444.60	2453.50	2486.54	2446.88	2500.39
16-APR-2009	2391.80	2435.00	2450.64	2390.08	2423.57

Table 13. Comparison between the Actual and Predicted Values Generated by ANN and Hybrid ANN_ARIMA Models for BSE Oil & Gas.

Date	Actual Values of BSE Oil & Gas	Predicted Close Value of BSE Oil & Gas Using Following Models	
		ANN	Hybrid ANN_ARIMA
26-FEB-2009	6165.30	6053.20	6165.30
27-FEB-2009	6064.10	6137.10	6123.74
02-MAR-2009	5886.00	5979.70	5896.71
03-MAR-2009	5775.60	5867.90	5756.51
04-MAR-2009	5842.20	5781.00	5805.64
05-MAR-2009	5625.30	5842.50	5611.64
06-MAR-2009	5724.10	5694.90	5667.44
09-MAR-2009	5633.90	5806.60	5596.39
12-MAR-2009	5814.60	5719.40	5750.73
13-MAR-2009	6088.50	5846.80	6058.02
16-MAR-2009	6298.30	6096.20	6322.54
17-MAR-2009	6206.60	6323.90	6265.31
18-MAR-2009	6308.50	6133.80	6331.08
19-MAR-2009	6355.20	6318.10	6407.08
20-MAR-2009	6368.00	6313.40	6415.65
23-MAR-2009	6775.40	6353.10	6823.19
24-MAR-2009	6772.80	6848.90	6892.79
25-MAR-2009	7020.90	6751.70	7099.39
26-MAR-2009	7190.70	6982.90	7304.55
27-MAR-2009	7181.80	7049.10	7311.15
30-MAR-2009	6992.40	7087.20	7119.09
31-MAR-2009	7053.00	6919.90	7133.05
01-APR-2009	7256.00	6977.40	7344.40
02-APR-2009	7671.80	7074.20	7795.20
06-APR-2009	7747.60	7609.40	7960.75
08-APR-2009	7998.80	7757.70	8191.81
09-APR-2009	7978.10	7672.40	8175.60
13-APR-2009	8083.20	7780.20	8296.84
15-APR-2009	8273.40	7787.80	8499.08
16-APR-2009	7901.40	7795.90	8172.23

Table 14. Comparison between the Actual and Predicted Values Generated by ANN, ARIMA, Hybrid ANN_ARIMA and Hybrid ARIMA_ANN Models for BSE 100.

Date	Actual Values of BSE 100	Predicted Close Value of BSE 100 Using Following Models			
		ANN	ARIMA	Hybrid ANN_ ARIMA	Hybrid ARIMA_ ANN
26-FEB-2009	4537.10	4390.20	4513.22	4537.10	4547.03
27-FEB-2009	4516.40	4385.90	4519.83	4612.21	4545.90
02-MAR-2009	4377.20	4374.40	4506.72	4462.31	4505.59
03-MAR-2009	4293.50	4297.20	4355.14	4295.33	4244.34
04-MAR-2009	4311.00	4252.50	4290.44	4308.59	4243.66
05-MAR-2009	4193.50	4244.80	4265.99	4231.65	4289.49
06-MAR-2009	4242.20	4222.50	4240.53	4208.74	4179.18
09-MAR-2009	4160.40	4198.80	4209.41	4173.25	4212.41
12-MAR-2009	4246.80	4191.40	4182.86	4221.76	4150.97
13-MAR-2009	4437.00	4215.80	4289.33	4473.13	4354.95
16-MAR-2009	4541.40	4292.60	4455.78	4685.67	4605.23
17-MAR-2009	4504.60	4368.60	4539.99	4666.87	4626.68
18-MAR-2009	4562.40	4385.00	4529.72	4651.10	4505.22
19-MAR-2009	4575.40	4426.40	4535.12	4691.10	4570.97
20-MAR-2009	4558.60	4434.60	4594.27	4655.78	4637.39
23-MAR-2009	4775.60	4410.30	4600.63	4856.47	4569.78
24-MAR-2009	4784.80	4580.40	4779.09	5023.05	4955.80
25-MAR-2009	4886.00	4703.90	4818.57	5019.31	4830.35
26-MAR-2009	5036.40	4749.40	4921.18	5155.17	4989.71
27-MAR-2009	5091.60	4969.10	5030.01	5278.78	5147.70
30-MAR-2009	4861.00	5065.70	5030.27	4940.90	5094.32
31-MAR-2009	4942.50	4841.20	4891.13	4808.99	4740.67
01-APR-2009	5028.40	4859.90	4942.23	5094.47	4996.14
02-APR-2009	5253.10	4953.00	5117.91	5363.00	5206.57
06-APR-2009	5348.20	5243.00	5243.84	5543.93	5379.26
08-APR-2009	5467.30	5369.50	5436.97	5535.91	5542.77
09-APR-2009	5500.80	5381.90	5385.92	5564.59	5418.97
13-APR-2009	5593.10	5484.30	5575.56	5670.65	5692.91
15-APR-2009	5780.00	5525.50	5658.77	5850.96	5681.28
16-APR-2009	5588.10	5566.30	5644.33	5754.09	5767.80

Table 15. Comparison between the Actual and Predicted Values Generated by ANN, ARIMA, Hybrid ANN_ARIMA and Hybrid ARIMA_ANN Models for S&P CNX Nifty.

Date	Actual Values of S&P CNX Nifty	Predicted Close Value of S&P CNX Nifty Using Following Models			
		ANN	ARIMA	Hybrid ANN_ ARIMA	Hybrid ARIMA_ ANN
26-FEB-2009	2785.70	2750.90	2763.00	2785.70	2755.55
27-FEB-2009	2763.70	2763.60	2750.75	2798.50	2775.04
02-MAR-2009	2674.60	2753.90	2732.20	2684.40	2753.26
03-MAR-2009	2622.40	2715.00	2672.34	2577.57	2624.76
04-MAR-2009	2645.20	2681.50	2649.49	2585.52	2608.59
05-MAR-2009	2576.70	2677.90	2619.97	2545.20	2623.93
06-MAR-2009	2620.20	2665.70	2591.73	2550.42	2558.31
09-MAR-2009	2573.20	2662.30	2595.33	2530.72	2624.30
12-MAR-2009	2617.40	2660.30	2619.95	2550.34	2607.91
13-MAR-2009	2719.30	2666.10	2663.82	2674.98	2664.47
16-MAR-2009	2777.30	2698.20	2724.20	2785.15	2781.09
17-MAR-2009	2757.50	2745.60	2759.13	2783.97	2812.85
18-MAR-2009	2794.70	2759.20	2787.89	2790.69	2795.22
19-MAR-2009	2807.20	2775.90	2776.86	2817.22	2792.45
20-MAR-2009	2807.10	2789.70	2787.23	2817.32	2823.17
23-MAR-2009	2939.90	2791.10	2890.55	2944.78	2914.73
24-MAR-2009	2938.70	2881.10	2959.90	3013.36	3010.07
25-MAR-2009	2984.40	2958.10	2932.74	3018.16	2921.73
26-MAR-2009	3082.30	2984.10	3036.67	3101.88	3088.70
27-MAR-2009	3108.60	3087.60	3067.77	3166.06	3119.52
30-MAR-2009	2978.20	3147.50	3024.81	2999.19	3066.33
31-MAR-2009	3021.00	3061.20	3018.87	2943.59	2982.25
01-APR-2009	3060.40	3041.10	3011.76	3041.89	3020.31
02-APR-2009	3211.10	3074.30	3170.05	3221.46	3221.02
06-APR-2009	3256.60	3203.60	3249.85	3328.29	3291.46
08-APR-2009	3342.90	3276.90	3237.04	3376.84	3252.16
09-APR-2009	3342.10	3296.40	3365.83	3384.32	3472.01
13-APR-2009	3382.60	3324.20	3378.66	3416.34	3365.20
15-APR-2009	3484.20	3332.80	3406.00	3525.46	3412.40
16-APR-2009	3369.50	3348.10	3432.62	3460.00	3512.79