DEVELOPMENT OF ANN FOR FORECASTING OF BSE INDEX

[1]Udit Narayan Bera, [2]Bhavesh B Kunbi

[1,2]*Mechatronics, IIITDM Jabalpur, MP, (India)*

ABSTRACT

Forecasting stock price index is one of the major challenges in the trade market for investors. Time series data for prediction are difficult to manipulate, but can be focused as segments to discover interesting patterns. In this paper we have used several functional link artificial neural networks (ANN) to get such patterns for predicting stock indices. During the past few decades, a large number of neural network models have been proposed to solve the problem of financial data and to obtain accurate prediction result. We have used "K-nearest neighbor method" based on Correlation and Regression. The forecasting accuracy is analyzed and measured with reference to an Indian stock market index such as Bombay Stock Exchange (BSE). In this work we have taken 1242 number of BSE index (closed price) starting from 15/04/2009 to 17/04/2014 for training and forecasting. Result shows more the number of training data, better would be the forecasted output.

Keywords: ANN, BSE forecasting, K-Nearest method

I. INTRODUCTION

Stock markets are highly correlated in the era of globalization. The world has become a small global village. Established in 1875, BSE Ltd. (formerly known as Bombay Stock Exchange Ltd.), is Asia's first Stock Exchange group [5]. More than 5000 companies are listed on BSE making it world's No. 1 exchange in terms of listed members. BSE also provides a host of other services to capital market participants including risk management, clearing, settlement, market data services and education. The companies listed on BSE Ltd command a total market capitalization of USD 1.32 Trillion as of January 2013 [5]. A stock market index should capture the behavior of the overall equity market. stock markets in virtually every developed and most developing economies, with the world's largest markets being in the United States, United Kingdom, Japan, India, China, Canada, Germany, France, South korea and Netherlands. In this study, the BSE and NIFTY, MIDCAP50 are considered for stock market index evaluation and prediction. Historically, from 1979 until 2012, Indian Stock Market (SENSEX) averaged 5462 Index points reaching an all time high of 21005 Index points in November of 2010 and a record low of 113 Index points in December of 1979 [5].

BSE index is one of the important financial parameter. A lot of matters are related to it. More precisely the "Profit & Loss" profile of any organization or any person directly or indirectly depends on this. A huge amount of money invests here. So it should not be entirely depends on "luck" rather there may be some kind of information about it. To answer the question "what will happen tomorrow?" we need to predict.

A neural network is a processing device, either an algorithm or an actual hardware. The computing world has a lot to gain from neural networks, also known as artificial neural network or neural net. ANNs have three great advantages over traditional methods [1-3]. At first, they have universal approximation capabilities second, they can recognize ''on their own'' implicit dependencies and relationships in data third, they can ''learn'' to adapt their behavior viz., prediction, to changed conditions quickly and without complication. According to Dimitri Pissarenko, [8] the most neural network business application studies utilize multilayered feed forward neural networks with the back propagation learning rule. W K Wong et al [7] pointed that, time series forecasting is used to forecast the future based on historical observations. Traditional methods, such as time-series regression, exponential smoothing and Auto Regressive Integrated Moving Average (ARIMA) are based on linear models. All these methods assume linear relationships among the past values of the forecast variable and therefore non-linear patterns cannot be captured by these models. Although ANNs have the advantages of accurate forecasting [4,6], their performance in some specific situation is inconsistent. Based on the study of, Lean Yu et al. noted that ANNs are a kind of unstable learning methods, i.e., small changes in the training set and/or parameter selection can produce large changes in the prediction.

II. TECHNICAL APPROACH

We follow the steps like create the network, configure the network, initialize the weights and biases, and train the network. To forecast the BSE index there number of artificial neural network method available. Such as Back propagation method, ARIMA based method, Modular neural network, Radial Basis Function, Multilayer Perceptron and Recurrent Network, Branch Neural Network, K-Nearest Neighbor Method, Functional Link ANN, ANN & Genetic Algorithm, Machine Learning Algorithm, Combinatorial Algorithm. From the above mentioning methods we have selected K-Nearest Neighbor Method. The reason behind choosing this methods are a) it can handle huge amount of data b) it is stationery and non volatile c) simplest machine learning algorithm d)it depends on the “Objects” those are “Near”. It is based on Correlation and Regression. Correlation is the technique concern with predicting some variable knowing other and Regression describes strength of linear relation between two variables. So for predicting some value we need to study previous data available with us.

Figure 1: BSE INDEX (1991-2013)

Figure 1 indicates the total variation of the data from 1991 to 2013.but if we concentrate on these data closely for last three year it can be concluded that the BSE sensex has been almost same since 2010.

III. K-NEAREST NEIGHBOR METHOD

The nearest neighbor method is defined as a non-parametric class of regression. Its main idea is that the series copies its own behavior along the time. In other words, past pieces of information on the series have symmetry with the last information available before the observation on t+1. Such way of capturing the pattern on the times series behavior is the main argument for the similarity between NN algorithm and the graphical part of technical analysis, charting.

The way the NN works is very different than the popular ARIMA model. The ARIMA modeling philosophy is to capture a statistical pattern between the locations of the observations in time. For the NN, such location is not important, since the objective of the algorithm is to locate similar pieces of information, independently of their location in time. Behind all the mathematical formality, the main idea of the NN approach is to capture a nonlinear dynamic of self-similarity on the series, which is similar to the fractal dynamic of a chaotic time series.

IV. STEPS OF K-NN METHOD (METHOD-CORRELATION)

1. The first step is to define a starting training period and divide such period on different vectors (pieces) y_m^T of size m, where t = m,...,T . The value of T is the number of observation on the training period. The term m is also defined as the embedding dimension of the time series. For notation purposes, the last vector available before the observation to be forecasted will be called y_T^m and the other pieces will be addressed as y_i^m.
2. The second step is to select k pieces most similar to y_m^T .For the method of correlation, in a formal notation, it is searched the k pieces with the highest value of |ρ| , which represents the absolute (Euclidian) correlation between y_i^m and y_m^T . The only difference between the univariate and the multivariate case is on this step: the way that is going to be searched for the k pieces with highest symmetry withy_m^T.
3. With the k pieces on hand, each one with m observations, is necessary to understand in which way the k vectors can be used to construct the forecast on t+1. Several ways can be employed here, including the use of an average or of a tricube function, Fernández-Rodríguez et al [2]. The method chosen for this case of the function is the one used on. The method chosen for this of the function which consists on calculation of the following equation.

$$\hat{y}_{T+1} = \hat{\alpha}_0 + \hat{\alpha}_1 y_{T-1} + \hat{\alpha}_2 y_{T-2} + \ldots + \hat{\alpha}_m y_{T-m} \quad (1)$$

The coefficients in Equation (1), $\widehat{\alpha_0}$, $\widehat{\alpha_1}$, $\widehat{\alpha_2}$, are the ones derived from the estimation of a linear model with the dependent variable as y_{i_r+1} and the explanatory variables as $y_{i_r}^m=(y_{i_r}, y_{i_r-1}, \ldots \ldots \ldots, y_{i_r-m+1})$

where *r* goes from 1 (one) to k. In order to facilitate the understanding of such regression, Equation (1) is presented on a matricial form on nextexpression, Equation (2).

$$\begin{bmatrix} y_{i_1+1} \\ y_{i_2+1} \\ y_{i_3+1} \\ \vdots \\ y_{i_k+1} \end{bmatrix} = \hat{\alpha}_0 + \hat{\alpha}_1 \begin{bmatrix} y_{i_1} \\ y_{i_2} \\ y_{i_3} \\ \vdots \\ y_{i_k} \end{bmatrix} + \hat{\alpha}_2 \begin{bmatrix} y_{i_1-1} \\ y_{i_2-1} \\ y_{i_3-1} \\ \vdots \\ y_{i_k-1} \end{bmatrix} + \ldots + \hat{\alpha}_{m-1} \begin{bmatrix} y_{i_1-m+1} \\ y_{i_2-m+1} \\ y_{i_3-m+1} \\ \vdots \\ y_{i_k-m+1} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_k \end{bmatrix} \quad (2)$$

For a clarified view of Equation (2), is necessary to comprehend that the NN algorithm is non temporal. The values of y_{i_k+1} are the observations one period ahead of the pieces chosen by the correlation criteria defined earlier.

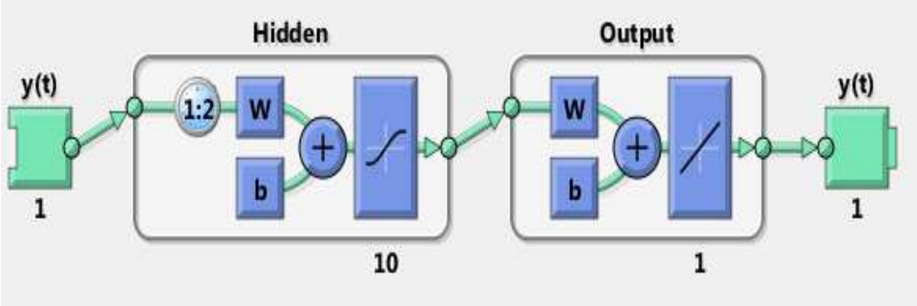

Figure 2: Neural Network Diagram (10 Hidden Layers and 1 delay);w=weight, b=bias

The term y_{i_k-m} indicates the first values of the k chosen pieces, while the term y_{i_k} represents the last terms of each piece chosen. It's easy to see that the number of explanatory series on [2] is m, and that each one of those will have k observations. The term $\widehat{\alpha_1}$, Equation [2], is the coefficient aggregated to the last observation of the chosen series and $\widehat{\alpha_2}$ is the coefficient for all the second last observations of all k series. This logic for the coefficients continues until it reaches the first observation of all k chosen series, $\widehat{\alpha_{m-1}}$. The values of the coefficients on Equation [2] are estimated with the minimization of the sum of the quadratic error $\sum_{i=1}^{k}\varepsilon_k^2$ The steps 1-3 are executed in a loop until the point that all forecasts on (t+1) are created.

V. RESULT

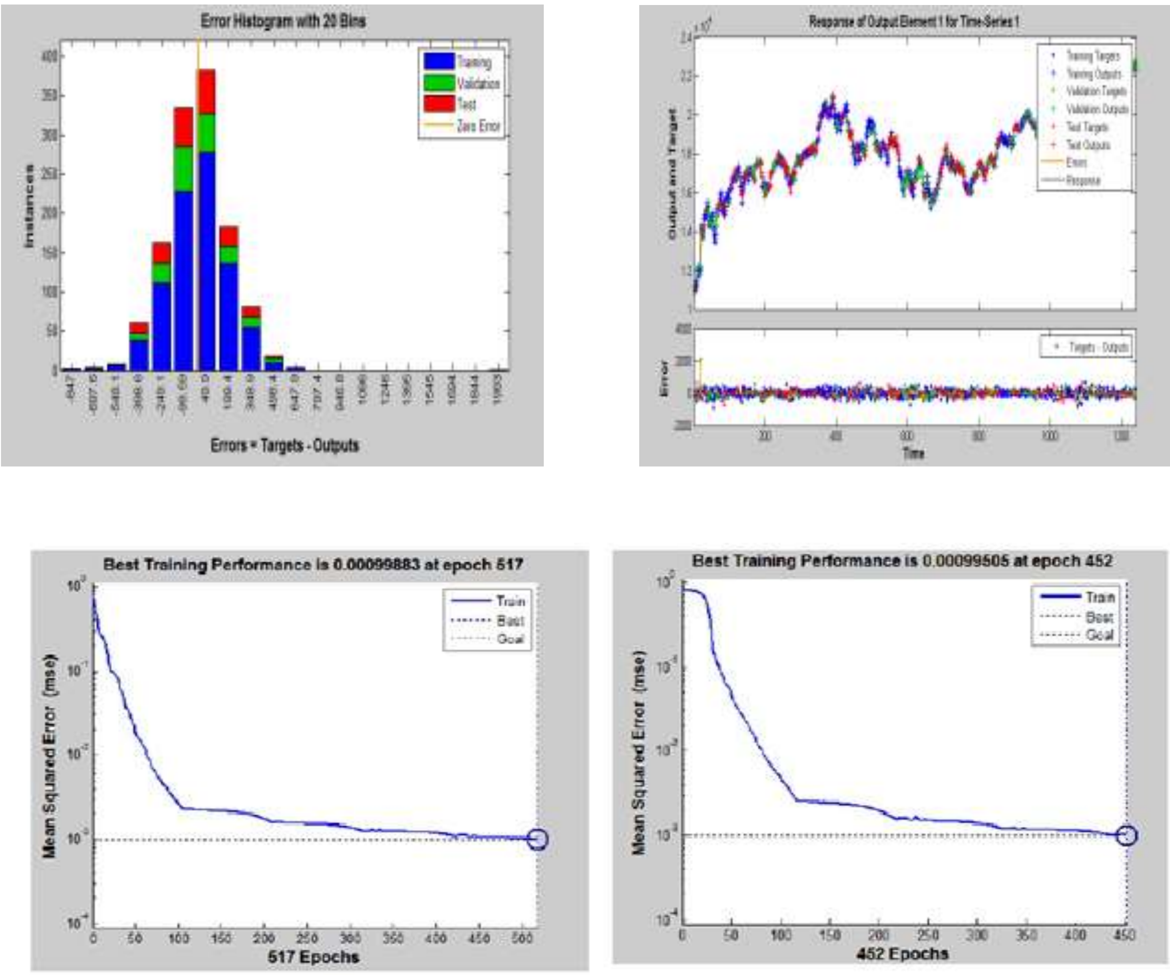

Figure 3: Errors

Figure 3 shows when forecasting BSE index in one day, the network can come to the desired error at epoch 517, a fast convergence speed. And network can be convergent with a range of initial weights. By Comparison of these images, we can know that a faster speed of network training also accelerates the convergence rate of the error. Hence the network here is working well in this case.

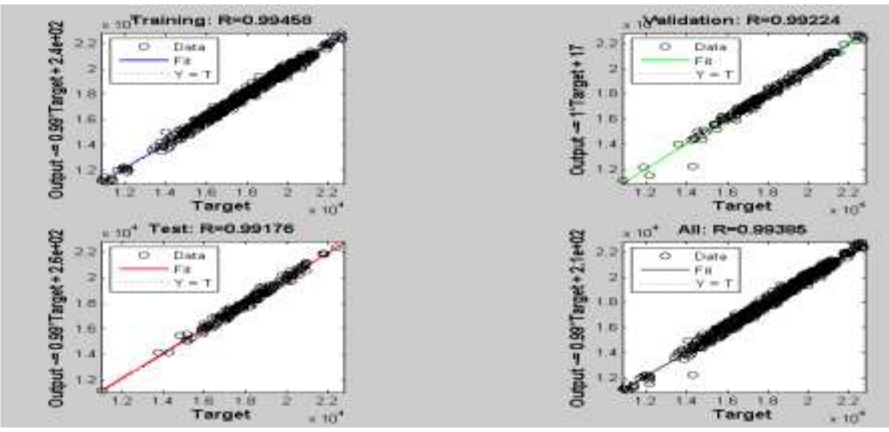

Figure 4: Target Vs Output

Figure 4 shows that regression R Values measure the correlation between outputs and targets. All the plots have a high value better than 0.99. The outputs of the training network are quite close to the targets. So the network model has a good training performance. Fig-5(k=20, d=870), Fig-6(k=5, d=870), Fig-7(k=5, d=1000), Fig-8(k=20, d=1000), Fig-9(k=5, d=1200),Fig-10(k=20, d=1200) shows the comparison between real data and forecasted data. In the Fig-5 to Fig-10, k indicates number of nearest neighbor and d is for number of training data. As the predicted output entirely depends on the number of sample data used for network training, we have shown that the more the value of parameter "d" which is the number of input samples to be trained, better will be the forecasted output. The predicted BSE index depends too on the number of nearest neighbor taken, which is parameter "k". Error is less if the value of "k" is more.

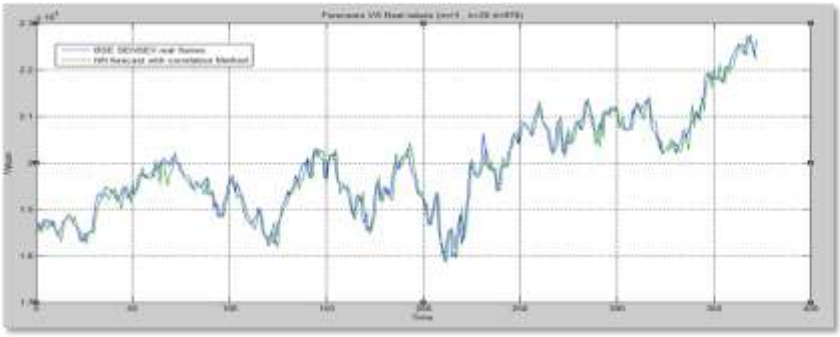

Figure 5: k=20, d=870

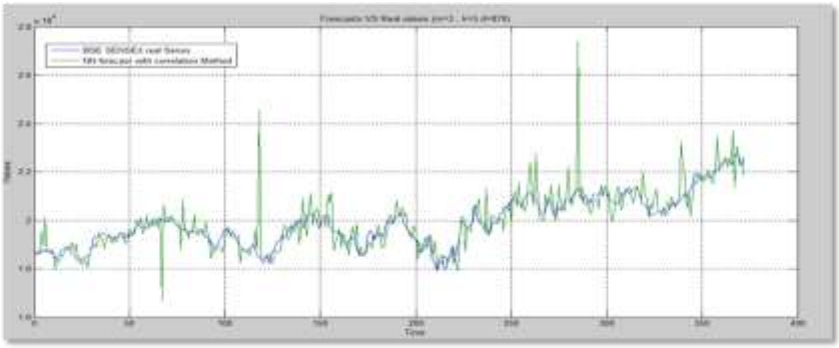

Figure 6: k=5, d=870

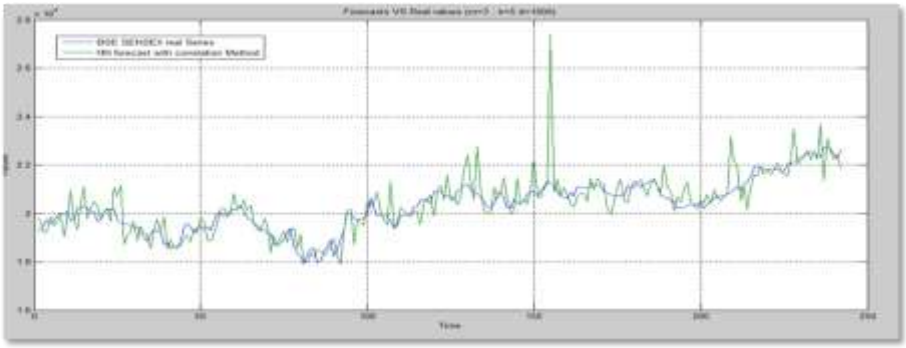

Figure 7: k=5, d=1000

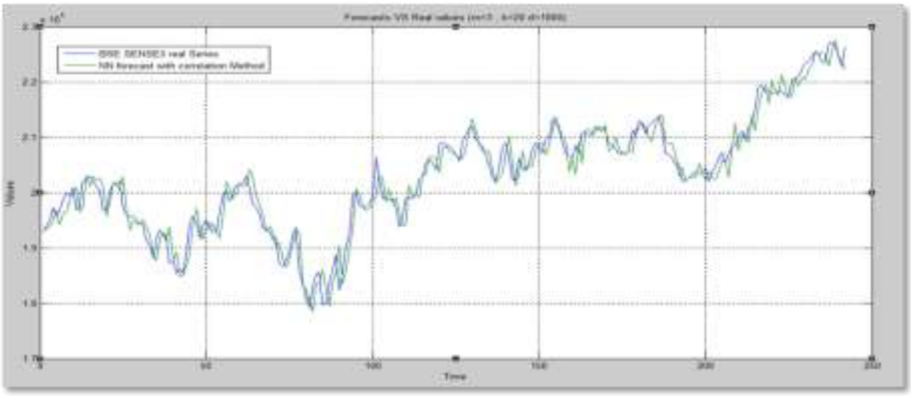

Figure 8: k=20, d=1000

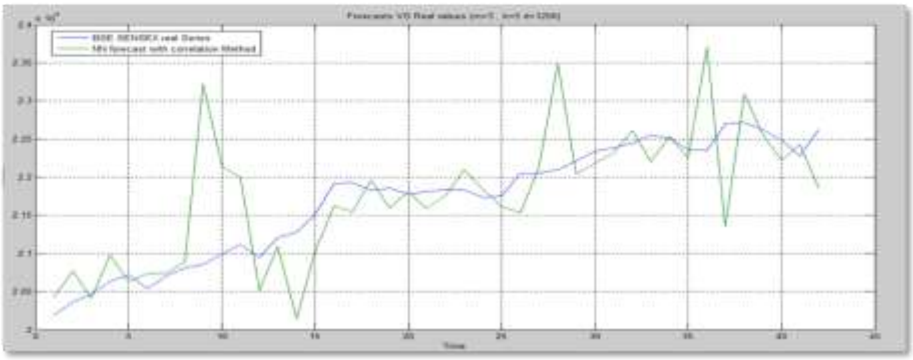

Figure 9: k=5, d=1200

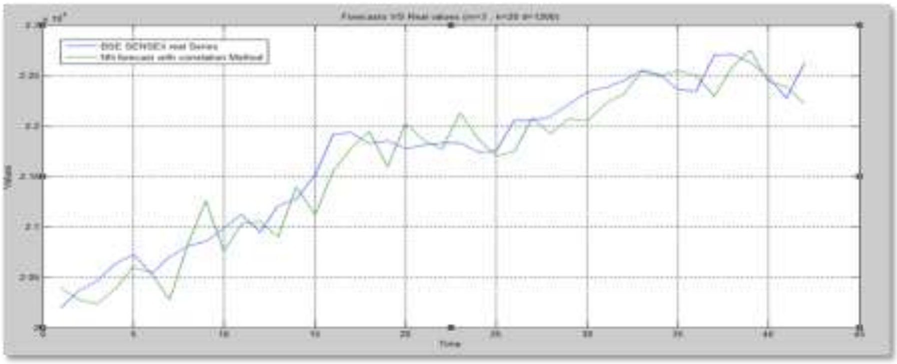

Figure 10: k=20, d=1200

VI. CONCLUSION

The predicted BSE index is very close to original data. Hence the error has been reduced. Also it has been found out that the method K-Nearest Neighbor is faster and simpler related to other. By analyzing all the results we come to some conclusions like following,

- Large value of nearest neighbor (k), less predicted error and better forecasting error.
- More neurons used, longer time and of higher complexity the training network can be.
- MSE (mean square error) shows low correlation with number of hidden neurons. To some degree, to step over a specified order of magnitude.
- To increase the performance we have to increase the number of neurons in hidden layer, train it again and get larger training data set.

REFERENCES

[1] FARMER, D., SIDOROWICH, J. Predicting chaotic time series. Physical Review Letters, v. 59, p. 845-848, 1987.

[2] FERNÁNDEZ-RODRÍGUEZ, F., RIVERO, S. S., FELIX, J. A. Nearest Neighbor Predictions in Foreign Exchange Markets. Working Paper, n. 05, FEDEA, 2002.

[3] FERNÁNDEZ-RODRÍGUEZ, F., SOSVILLA-RIVERO, S., GARCÍA-ARTILES, M. Using nearest neighbor predictors to forecast the Spanish Stock Market. Investigaciones Económicas, v. 21, p. 75-91, 1997.

[4] FERNÁNDEZ-RODRÍGUEZ, F., SOSVILLA-RIVERO, S., GARCÍA-ARTILES, M. Anempirical evaluation of non-linear trading rules. Working paper, n. 16, FEDEA, 2001

[5] http://www.bseindia.com

[6] Chen, N., Roll, R., & Ross, S. (1986). Economic forces and the stock market. Journal of Business 59, 383–403.

[7] Wong WK, Hurwitz EL. Predictors and outcomes of scleroderma renal crisis: the high-dose versus low-dose D-penicillamine in early diffuse systemic sclerosis trial. Arthritis Rheum. 2002 Nov;46(11):2983-9.

[8] Dimitri Pissarenko. Neural networks for financial time series prediction: Overview over recent research 2002.