Abstract—This paper presents a scheme using Differential Evolution based Functional Link Artificial Neural Network (FLANN) to predict the Indian Stock Market Indices. The Model uses Back-Propagation (BP) algorithm and Differential Evolution (DE) algorithm respectively for predicting the Stock Price Indices for one day, one week, two weeks and one month in advance. The Indian stock prices i.e. BSE (Bombay Stock Exchange), NSE, INFY etc. with few technical indicators are considered as input for the experimental data. In all the cases, DE outperforms the BP algorithm. The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are calculated for performance evaluation. The MAPE and RMSE in case of DE are found to be very less in comparison to BP method. The simulation study has been done using Java-6 and NetBeans.

Keywords—Stock Market Prediction; Functional Link Neural (FLANN); Differential evolution (DE); Back Propagation (BP) Algorithm; Least Mean Square (LMS) method.

I. INTRODUCTION

Due to the recent financial crisis, the world economy has gone down drastically. When the major contributor of the global economy i.e. US economy is suffering from the debt crisis, Indian stock market is also showing clear signs of slow down. Financial time series[1][2][3][4] is highly noisy, irregular, random, non-linear, non-seasonal and chaotic in nature. So it has always remained as a challenge for the common investors, stock buyers/sellers, policy makers, market researchers and capital market role players to gain knowledge about the daily stock market price values. Making money and gaining high profit is the dream of every investor, but it requires proper financial knowledge, analytical capability and ability for discovering the non-linear pattern hidden[5] within the particular stock market data. As a lot of risk is involved in the stock market, the investors become highly insecure to invest their amount. Though a lot of research has already been done on US Stock market, European Stock market, China and Japan Stock markets but less attention is given for Indian Stock market in comparison to them. Due to rapid growth in IT, Telecom sectors in the last decade Indian Stock market is setting its upcoming platform not only in Asia as well as globally. Asia’s oldest stock exchange i.e. Bombay Stock Exchange (BSE) along with NSE is contributing a lot to the global economy. It has been found that not only the economic factors but also the non-economic factors like political scenario, continuous terrorist attacks and the moods of typical individual investors have become the major role players in contributing the uncertainty in Indian stock market. Researchers are highly motivated to develop more and more efficient and advanced models for making huge profits.

Stock market analysts have adopted many statistical techniques likes autoregressive moving average(ARMA) [6][7], autoregressive integrated moving average(ARIMA) [7], autoregressive conditional heteroscedasticity(ARCH) [6], generalized autoregressive conditional heteroscedasticity (GARCH)[8], ARMA-EGARCH [6], Box and Jenkins approach[4] along with various soft computing and evolutionary computing methods[9][10]. Artificial Neural Network(ANN), fuzzy set theory, Support Vector Machine(SVM) etc. are considered under soft computing techniques where as the various evolutionary learning algorithms include Genetic algorithm(GA) [9], Ant Colony Optimization(ACO), Particle Swarm Optimization (PSO) [11][12][13], Differential Evolution (DE)[14][15], Bacterial Foraging Optimization (BFO)[16] etc. It has been found that the Soft computing and Evolutionary computing methods outperforms over statistical techniques in Stock market prediction. As Statistical method can handle only linear data, they become unable to follow the non-linear pattern hidden within the stock data. The Literature survey reveals that different types of ANN’s like Radial Basis Function(RBF) [5], Multi Layer Perceptron(MLP) [17], Recurrent Neural Network(RNN)[18], Time delay Neural Network (TDNN) [18], Machine learning techniques [1], Functional Link Artificial Neural Network(FLANN) [2][17][19], Local Linear Wavelet Neural Network (LLWNN) [20], Evolutionary
Neurofuzzy NN[7][10], and various Neurofuzzy hybrid models have been used for stock prediction. In this paper, a simple FLANN network is proposed whose randomly chosen weights are optimized with BP training algorithm and DE algorithm respectively. Initially the FLANN weights are updated using LMS method. Whereas later on they are updated with DE and results are compared. The FLANN model is designed to improve the information processing capability and uncertainty modeling.

Our paper is organized as follows. Section II deals with the basic principle of FLANN and BP [12] training algorithm. The basic principle of Differential Evolution (DE) is dealt in section III. The application of DE based training to the FLANN model is proposed in section IV. The simulation study is carried out in section V. The training and testing results of the proposed model is discussed in section VI. Finally conclusions are given in section VII.

II. BASIC PRINCIPLE OF FUNCTIONAL LINK ARTIFICIAL NEURAL NETWORK AND BACK PROPAGATION

A FLANN network provides large reduction in computational requirement in comparison to Multi Layer Perceptron (MLP). It possesses high convergence speed. The multilayer perceptron has one or more hidden layers. Sometimes due to the large number of nodes in the hidden layer, it suffers from slow convergence rate and high computational complexity. So to reduce the computational burden on the network, the single layer ANN known as FLANN was originally proposed by Pao [22]. The FLANN is a single layer, single neuron architecture which has the capability to form complex decision regions by creating non-linear decision boundaries. An artificial neural network (ANN) can approximate a continuous multivariable function f(x).

In this model, each component of the input vector is subjected to a functional expansion to yield the enhanced representation of the original pattern. The function used in this may be a subset of orthonormal basis function spanning over an n-dimensional representation space like { cos(πx), sin(πx), cos(2πx), sin(2πx) …..} and so on. Besides this trigonometric function other orthogonal function such as Legendre, Chebyshev can be used, but the trigonometric function represents the Fourier series expansion of the time series. So the trigonometric function given by \{1, \cos(\pi x), \sin(\pi x), \cos(2\pi x), \sin(2\pi x) ….., \cos(N\pi x), \sin(N\pi x) \} provide a compact representation of the function in the mean square sense. When suitable trigonometric polynomials are used after training, the FLANN weights represent a multi-dimensional Fourier series decomposition of a periodic of the desired response function. Different application of trigonometric expansion can be found by Giles and Maxwell[21].

In our proposed FLANN model the final output “Y” is represented as the sum of Y1 and Y2. At first a set of six inputs including moving average input as the technical indicator are getting functionally expanded through \{ \cos(\pi x_1), \sin(\pi x_1), \cos(2\pi x_1), \sin(2\pi x_1) \} respectively. Then they get multiplied with their corresponding randomly chosen weights and the sum is represented as “Y1”.

\[ Y_1 = [w_1 \cos(\pi x_1) + w_2 \sin(\pi x_1) + \ldots + w_{14} \cos(\pi x_6) + w_{15} \sin(\pi x_6)] \]  \hspace{1cm} (1)

Now the set of original six inputs are directly multiplied with their weights and summed to generate “Y2”.

\[ Y_2 = [w_{16} x_1 + w_{17} x_2 + \ldots + w_{18} x_6] \]  \hspace{1cm} (2)

Finally, the FLANN output

\[ Y = \tanh(Y_1 + Y_2) \]  \hspace{1cm} (3)

Now the error “e” is calculated by calculating the difference between output “Y” and the desired output “D”.

Then the weights \[ w_1, w_2, w_3, w_4, w_5, w_6 \] are updated using BP learning algorithm.

\[ W_j(\text{new}) = W_j(\text{old}) + \mu \cdot e(i) \cdot \frac{\partial(Y)}{\partial W_j} \]  \hspace{1cm} (4)

Where \( \mu \) represents the learning rate , \( e(i) \) represents the error during \( i \)-th iteration, and \( \frac{\partial(Y)}{\partial W_j} \) represents the change in weight.

The structure of the proposed FLANN model is shown in figure 1(a) and figure 1(b).

Figure 1.(a) FLANN model with LMS
III. BASICS OF DIFFERENTIAL EVOLUTION ALGORITHM

Differential Evolution is introduced by Kenneth Price and Rainer Storn in 1995[14] is a population based stochastic search which behaves as an efficient global optimizer in the continuous search domain. DE has been successfully applied in global numerical optimization, pattern recognition, image pixel clustering, text document clustering, large scale power dispatch problem etc. DE algorithm is like genetic algorithm which uses similar operators: crossover, mutation and selection. DE can find the true global minimum regardless of the initial parameter values. It has other advantages such as fast convergence and requires few control parameters. The main difference in constructing better solution between GA and DE is that GA relies on crossover while DE relies on mutation operation. The mutation operation is based on the differences of randomly sampled pairs of solution in the population. The algorithm uses mutation operation as a search mechanism and the selection operation directs the search towards the prospective region in the search space. The DE algorithm also uses a non-uniform crossover which takes child vector parameters from one parent more often than it takes from others. When the components of the existing population members are used to construct the trial vectors, the recombination (crossover) operator efficiently shuffles information about successful combinations, enabling the search for a better solution space.

A. INITIALIZATION

DE generates an initial population size of NP, D-dimensional parameter vectors called individual which encode the candidate solution.

\[ X_{i,G} = [x_{1,i,G}, \ldots, x_{D,i,G}] \]  

Where \( i, j = 1, \ldots, \) NP represents the total population. Each parameter values is selected randomly and uniformly between \( X_{\min} \) and \( X_{\max} \). The lower and upper parameters bound are \( X_{\min} = [x_{1,\min}, \ldots, x_{D,\min}] \) and \( X_{\max} = [x_{1,\max}, \ldots, x_{D,\max}] \) respectively. The initial value of \( j^{th} \) parameter in case of \( i^{th} \) individual at generation \( G=0 \) is generated by

\[ x_{j,i,0} = x_{j,\min} + \text{rand}(0,1) \times (x_{j,\max} - x_{j,\min}) \]  

where, \( \text{rand}(0,1) \) represents uniformly distributed random number lying between 0 and 1, \( j=1,2,\ldots,D \).

B. MUTATION OPERATION

For each target vector, one mutant vector \( V_{i,G} \) has to be produced by using the mutation operation. For each target vector \( X_{i,G} \) at the generation \( G \), its associated mutant vector is

\[ V_{i,G} = [v_{1,i,G}, \ldots, v_{D,i,G}] \]  

The mutant vector can be generated using any one of the following mutation strategies. Most frequently used five mutation strategies are given below
"DE/rand/1"
\[ V_{i,G} = X_{i,G} + F \times [X_{i,G} - X_{r_1,G}] \]  (8)

"DE/best/1"
\[ V_{i,G} = X_{best,G} + F \times [X_{i,G} - X_{r_1,G}] \]  (9)

"DE/rand to best/1"
\[ V_{i,G} = X_{i,G} + F \times [X_{best,G} - X_{i,G}] + F \times [X_{r_1,G} - X_{r_2,G}] \]  (10)

"DE/rand/2"
\[ V_{i,G} = X_{i,G} + F \times [X_{r_1,G} - X_{r_2,G}] + F \times [X_{r_3,G} - X_{r_4,G}] \]  (11)

"DE/best/2"
\[ V_{i,G} = X_{i,G} + F \times [X_{r_1,G} - X_{r_2,G}] + F \times [X_{r_3,G} - X_{r_4,G}] \]  (12)

The indices \( r_1, r_2, r_3, r_4 \) are mutually exclusive integers randomly generated within the range 1 to NP. These indices are randomly generated once for each mutant vector. "F" represents the scaling factor.

C. CROSSOVER OPERATION

In this phase a trial vector \( U_{i,G} \) is generated from each pair of the target vector \( X_{i,G} \) and its corresponding mutant vector \( V_{i,G} \) after going through the crossover operation.

\[ U_{i,G} = [u_{1,1,G}^{\text{1st}}, u_{2,2,G}^{\text{2nd}}, \ldots, u_{D,D,G}^{\text{Dth}}] \]  (13)

Generally DE uses binomial crossover defined below

\[
u_{i,G}^{1,1,G}, \text{ if } (\text{rand}[0,1] \leq CR) \text{ or } (j=j_{\text{rand}})
\]
\[ u_{i,G}^{j,G} = x_{i,G}^{j,G}, \text{ otherwise} \]

With \( j = 1, 2, 3, \ldots, D \) and \( CR \) is crossover rate. \( j_{\text{rand}} \) is a randomly chosen integer. The crossover operator the jth parameter of \( V_{i,G} \) to the corresponding element in the trial vector \( U_{i,G} \) if \( (\text{rand}[0,1] \leq CR) \) or \( (j=j_{\text{rand}}) \) otherwise it is copied from the corresponding target vector \( X_{i,G} \).

D. SELECTION OPERATION

In selection operation phase DE checks the upper and lower bounds of each parameter of the newly created trial vector. When they exceed, again they have to be re-initialized within pre specified range. The objective function values of all trial vectors are evaluated and selection operation is carried out. The objective function values of all trial vectors \( F(U_{i,G}) \) is compared to that of its corresponding target vector \( F(X_{i,G}) \) in the current population. If the vector has less or equal objective function values than corresponding target vector, the trial vector will replace the target vector and enter the population of next generation. Otherwise the target vector will remain in the population for the next generation. It can be defined as

\[
X_{i,G+1} = \begin{cases} U_{i,G+1}, & \text{if } F(U_{i,G+1}) \leq F(X_{i,G}) \\ X_{i,G} & \text{otherwise} \end{cases} \]  (15)

The above steps are repeated until some specific criteria are satisfied.

IV. APPLICATION OF DE BASED TRAINING TO THE (FLANN) MODEL

(1) The coefficients or weights of the proposed FLANN model are chosen from a target population of ‘NP’ number of individuals. Each member of the population constitutes ‘D’ number of parameters. Here we have considered \( NP = 20, D = 18 \).

(2) ‘P’ number of patterns each containing six features including moving average (technical indicator) is obtained from the stock market price data.

(3) For each of the ‘P’ no. of patterns, \( X_p \) are passed through functional expansion unit multiplied with the weights to generate partial sum

\[
Y_{i1} = \{w_1 \cdot \cos(\pi x_1) + w_2 \cdot \sin(\pi x_1) + \ldots \} + w_3 \cdot \cos(\pi x_2) + w_4 \cdot \sin(\pi x_2)
\]

and the same “P” number of patterns ‘\( X_p \)’ is passed through the linear combiner and multiplied with the weight to generate the partial sum \( Y_{i2} = \sum_{n=1}^{P} w_n \cdot x_n \) (17)

(4) Finally the resultant sum is passed through hyperbolic tangent activation function to generate

\[
Y_i = \tanh(Y_{i1} + Y_{i2}) \]  (18)

(5) Now the Final calculated output ‘\( Y_i \)’ is compared with the actual output ‘\( Act_i \)’ to calculate the error ‘\( e_i \)’. At the completion of all the training sample inputs, (I) no of errors are produced. The mean square error (MSE) for a set of parameters using the relation

\[
MSE(I) = \frac{1}{I} \sum_{i=1}^{I} e_i^2 \]  (19)

This is repeated for I number of times.

(6) The perturb vector is then calculated using equation (8) and the mutant population is obtained.

(7) Then the crossover operation is carried out using equation (13) and the trial population is created.

(8) The selection of members of the target population for the next generation is performed using Equation (15) and the whole process is repeated for some generations.

(9) In each generation the MSE is determined. MSE is plotted over number of generations. This is used to show the convergence speed of the DE model specifying its learning characteristics.

V. SIMULATION STUDY

A. Experimental Data for training and testing

The daily closing price of various Indian Stock market data i.e. BSE, NIFTY and NSE are considered here as the experimental data. All these data are obtained from www.finance.yahoo.com and all the models are predicting for 1-day, 1-week, 2-weeks, and 1-month ahead. All the inputs are...
normalized within a range of [0, 1] using the following formula.

\[ X_{\text{norm}} = \frac{X_{\text{orig}} - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \]  \hspace{1cm} (20)

Where \( X_{\text{norm}} \) = normalized value, \( X_{\text{orig}} \) = current daily closing price, \( X_{\text{min}} \) and \( X_{\text{max}} \) are the daily minimum and maximum prices of the stock data respectively.

Details of datasets and Technical Indicators are given in the following tables.

<table>
<thead>
<tr>
<th>TABLE I. STOCK DATA SETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>STOCK DATA SETS</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>NSE NIFTY</td>
</tr>
<tr>
<td>INFY</td>
</tr>
<tr>
<td>BSE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II. FORMULAS USED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical indicators</td>
</tr>
<tr>
<td>Simple moving average(SMA)</td>
</tr>
<tr>
<td>Exponential Moving Average(EMA)</td>
</tr>
</tbody>
</table>

B. Training and testing of the forecasting model

Training of the FLANN model is carried out using the DE algorithm given in Section III and the optimum weights are obtained. Then using the trained model, the forecasting performance is tested using test patterns for 1-day, 1-week, 2-weeks, and 1-month ahead. The Mean Absolute Percentage Error (MAPE) defined in (21) is computed to compare the performance of various models.

\[ \text{MAPE} = \frac{1}{K} \sum_{k=1}^{K} \left| \frac{d_k - y_k}{d_k} \right| \]  \hspace{1cm} (21)

where \( K \) is the number of test patterns.
A. Performance of FLANN with BP for one day ahead prediction during training

B. Performance of FLANN with BP for one day ahead during Testing:

C. Performance of FLANN with DE for one day ahead prediction during training
VII. CONCLUSION

Accurate stock prediction is always a very challenging task. The proposed FLANN model trained with back propagation is giving good result as per the recorded RMSE, and MAPE values during testing for one day, one week, two weeks and one month ahead respectively. The DE optimized FLANN is proving it’s superiority as far as RMSE and MAPE are concerned. Further the prediction performance of DE is to be compared with PSO.

REFERENCES


