A Neuro-Genetic Technique for Index Prediction

S. C. Nayak
Silicon Institute of Technology
E-mail: sarat_silicon@yahoo.co.in

B. B. Mishra
Silicon Institute of Technology
misrabijan@gmail.com

Abstract: Artificial Neural Network (ANN) has preeminent learning ability, but often exhibit inconsistent and unpredictable performance for noisy data. In addition, it may not be possible to train ANN or the training task cannot be effectively carried out without data reduction when the amount of data is so large. In this paper, we have used a Neuro-genetic model to predict the index value for Stock Price Index of Bombay Stock Exchange. Here we used the genetic algorithm (GA) to optimize simultaneously the connection weights between layers and a selection task for relevant parameters such as Bias to the hidden as well as output layer. The globally evolved weights mitigate the well-known limitations of gradient descent algorithm. In addition, genetically selected weights and parameters shorten the learning time and enhance prediction performance. Experimental results show that the Neuro-genetic approach is a promising method for Index Prediction than ANN model.

Index Terms: Genetic algorithms; artificial neural networks; Index Prediction, Bombay Stock Exchange.

INTRODUCTION

Recently forecasting stock market return is gaining more attention, maybe because of the fact that if the direction of the market is successfully predicted the investors may be better guided and also monetary rewards will be substantial. If any system which can consistently predict the trends of the dynamic stock market be developed, would make the owner of the system wealthy. In recent years, the rapid development of computer and artificial intelligence technology provide many new technology methods for the modeling and forecast of the stock market [1]. Another motivation for research in this field is that it possesses many theoretical and experimental challenges. The most important of these is the efficient market hypothesis which proposes that profit from price movement is very difficult and unlikely. In an efficient market, stock market prices fully reflect available information about the market and its constituents and thus any opportunity of earning excess profit ceases to exist any longer. So it is ascertain that no system is expected to outperform the market predictably and consistently. The stock market is a very complicated nonlinear dynamic system, it has both the high income and high risk properties. So the forecast of stock market trend has been always paid attention to by stockholders and

Invest organization. However, because of the high nonlinearity of the stock market, it is difficult to reveal the inside law by the traditional forecast methods, so we are not satisfied with many of the applied effect for forecast analysis method. In recent years, the rapid development of computer and artificial intelligence technology provide many new technology methods for the modeling and forecast of the stock market [2]. The neural network gains wide application in the aspect of nonlinear forecast due to its broad adaptive and learning ability. There is many neural networks applied in forecasting stock price, at present, the most widely used neural network is BP NN, but BPNN has many shortcomings such as the slow learning rate, large calculate amount, easy to get into local minimum and bigger randomicity and so on. This affects the predicted results of the stock price. RBF neural networks is also a very popular method to predict stock market, this network has better calculation and spreading abilities, it also has stronger nonlinear mapped ability [3]. But the stock market is not only with nonlinearity but also chaos, and it is a dynamic system related to time (time is an independent variable). Therefore the network for predicting itself is a dynamic system. In researcher have adopted a neural network model trained with weights and other parameters optimized by GA. The Neural Network-GA model to forecasting the index value has got wide acceptance. In this paper work we suggest a ANN and Neuro-genetic Model for Index Prediction implemented on closing index price of BSE.

The section 2 describes some research backgrounds on applications of ANN and GA to the Stock Market. Section 3 provides the architecture of the proposed Neuro-Genetic model. The experimental results are given in section 4 which is followed by a conclusion with brief references.

Related Works on Index Prediction:

In the last two decades forecasting of stock returns has become an important field of research. In most of the cases the researchers had attempted to establish a
linear relationship between the input macroeconomic variables and the stock returns.

One of the earliest studies, Kimoto [5], Asakawa , Yoda & Takeeka(1990) used several learning algorithms and prediction methods for developing the Tokyo stock exchange price index(TOPIX) prediction systems. They used the modular neural networks to learn the relationship among various market factors.

Recent research tends to hybridize several AI techniques. Hiemstra [6] suggested that ANN and fuzzy logic could capture the complexities of functional mapping because they do not require the specification of the function to approximate.

Researchers have tested the accuracy of ANN in predicting the stock market index return of most developed economics across the globe. Literature are available for forecasting index returns of U.S markets like NYSE, FTSE, DJIA, S&P 500. Few papers are also available in context to Asian stock markets like Hang Seng stock Exchange, Korea store exchange, Tokyo store Exchange & Taiwan store Exchange.

A literature by Panda, C. and Narasimhan, V [8]. in Indian context used the artificial neural network to forecast the daily returns of BSE sensitive Index. They compared the performance of the neural network with performances of random walk & linear autoregressive models. They reported that the neural network out performs liner autoregressive & random walk models.

After a brief literature survey we motivated to incorporate the application of both ANN and GA to the Index value prediction in Indian context.

3. Architecture of Neuro-Genetic Model:

In Fig.1, X1 to Xn represents the set of input data. V11 to Vmn represents the weight value between the input and hidden layer. W1 to Wm represents the weight value between the hidden layer and output layer. Yi is the calculated value produced at the output layer neuron for the ith input data set. The bias value (b) is given to the hidden layer neuron as well as to the neuron of the output layer.

Genetic Algorithms work on a population of potential solutions in the form of chromosomes, attempting to locate the best solution through the process of artificial evolution.GA are based on biological evolutionary theory and used to solve optimization problems [9] which works with encoding parameter instead of parameter itself. It consists of the following repeated artificial genetic operations: evaluation, selection, crossover, and mutation. The weights and other parameters are optimized by GA and used to train the network.

In general the genetic evolution process consist the following basic steps:
1. Initialization of the search node randomly.
2. Evaluation of fitness of individuals.
3. Application of selection operator.
5. Application of mutation operator.
6. Repetition of the above steps until convergence.

The fitness of the best and average individual in each generation increases towards a global optimum.

The Neuro-genetic model is a hybrid model which exhibits the characteristics of both ANN and GA. It can be used as the tool for decision making in order to solve the complex nonlinear problems. In this method first we define a network structure with a fixed number of inputs, hidden nodes and outputs. Second we employed the GA in the learning phase of the network, as it is capable to search in a large search space. The hybridization of ANN and GA is able to select the optimal weight sets as well as the bias value for prediction. The following table represents the chromosome structure in the above Neuro-genetic model.

<table>
<thead>
<tr>
<th>Table 1: Chromosome Representation for Neuro-genetic Model.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights between Input and Hidden Layer</td>
</tr>
<tr>
<td>V11, V12, ..., Vmn</td>
</tr>
</tbody>
</table>

![Fig.1: Architecture of Artificial Neural Network.](image)
Here \( m \) is the number of inputs, \( n \) is the number of hidden nodes and \( B \) is the bias value. Each cell of \( V_{11} \) to \( V_{mn} \), \( W_1 \) to \( W_n \) and \( B \) constitute of 17 binary bits. For calculation the decimal equivalent of the binary chromosome is considered. A randomly initialized population with 40 genotypes is considered. GA was run for maximum 200 generations with the same population size. Parents are selected from the population by elitism method in which first 10% are the best parents and the rest are selected by binary tournament selection method. Then a new offspring is generated from these parents using uniform crossover followed by mutation operator. In this experiment the crossover probability is taken as 0.7 and mutation probability is taken as 0.03. In this way the new population generated replaces the current population and the process continues until convergence occurs.

The fitness of the best and average individuals in each generation increases towards a global optimum. The uniformity of the individuals increases gradually leading to convergence. A gene is considered to have converged when more than 95% of the population shares the same value.

### 4. Experimental Result:

The network is employed to use the input data as the daily closing index price (S&P100 and S&P500) which are collected from the historical data available on the website of BSE. The index values are collected during the financial year starting from 2005 to 2010. The data are normalized into the range \([0, 1]\). The normalized data are then used to form a training bed for the network model. The above genetic operators are applied on the network and the output is calculated at the output layer neuron. Table 1 and table 2 represents the simulated parameters of ANN and Neuro-genetic Model respectively.

The Average Percentage Error is considered for performance evaluation. We present the error in output node \( j \) in the \( n \)th data point by:

\[
e_{j}^{(n)} = d_{j}^{(n)} - y_{j}^{(n)}
\]

Where \( e_{j}^{(n)} \) is the error function, \( d_{j}^{(n)} \) is the desired output and \( y_{j}^{(n)} \) is the calculated output for the \( n \)th data set.

The following tables show the Average Percentage Error for the data set taken during the period 2005 to 2010. The analysis is done by taking the parameters as shown in table 2 and table 3. After the network is trained, for testing purpose the record immediate to the training set is used. This process has been conducted for the total data of one financial year.

#### Table 4: Simulation result Showing Average Percentage of Errors of BSE-S&P100 Data

<table>
<thead>
<tr>
<th>Financial Year</th>
<th>ANN-model</th>
<th>Neuro-Genetic-Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>0.0300</td>
<td>0.0200</td>
</tr>
<tr>
<td>2006</td>
<td>0.0500</td>
<td>0.0300</td>
</tr>
<tr>
<td>2007</td>
<td>0.0520</td>
<td>0.0320</td>
</tr>
<tr>
<td>2008</td>
<td>0.0800</td>
<td>0.0780</td>
</tr>
<tr>
<td>2009</td>
<td>0.1000</td>
<td>0.0910</td>
</tr>
<tr>
<td>2010</td>
<td>0.1500</td>
<td>0.1410</td>
</tr>
</tbody>
</table>

#### Table 5: Simulation result Showing Average Percentage of Errors of BSE-S&P500 Data

<table>
<thead>
<tr>
<th>Financial Year</th>
<th>ANN-model</th>
<th>Neuro-Genetic-Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>0.0300</td>
<td>0.0200</td>
</tr>
<tr>
<td>2006</td>
<td>0.0500</td>
<td>0.0300</td>
</tr>
<tr>
<td>2007</td>
<td>0.0520</td>
<td>0.0320</td>
</tr>
<tr>
<td>2008</td>
<td>0.0800</td>
<td>0.0780</td>
</tr>
<tr>
<td>2009</td>
<td>0.1000</td>
<td>0.0910</td>
</tr>
<tr>
<td>2010</td>
<td>0.1500</td>
<td>0.1410</td>
</tr>
</tbody>
</table>

From the above table it is concluded that during the financial year 2005 and 2007 the Neuro-genetic model gives much more better performance in average percentage of errors. It was found that in all cases the Neuro-genetic model outperforms the ANN model trained with feed forward back propagation learning algorithm. The following Fig 2-5 represents the graph plot by simulation result performed by taking the data set from financial year 2005-2010. The Target value is represented along X-axis and the Estimated value is represented along Y-axis.
5. Conclusion:

In this paper we have suggested a Neuro-genetic model for stock index prediction. This paper compares the forecasting accuracy of this model with the ANN model trained with back propagation. We apply the models to predict BSE index values. The stock market in our country is far from effective market and the information of the market can not entirely reflected the stock price on time. So there is some error on price predicted, but in a short time the stock price can be predicted, and these models are comparatively exact in the aspect of predicting the trend of stock price. For spacious stockholders, the most important thing is not the stock price but the trend of the stock market, therefore the model of neural network trained by Genetic Algorithm is feasible. The objective of the proposed research work is to forecasting the index value on financial data by using both ANN and Neuro-genetic model and it is concluded that the later model gives better performance than the former one.

Our future work will focus on the optimal number of neurons in each layer and rest of the genetic parameters dynamically by GA.

Acknowledgements:

The authors would like to thank Prof(Dr.) S . Padhy for his valuable guidance and suggestions.

REFERENCES