Issues in Business Management and Economics Vol.2 (6), pp. 094-102, June 2014
Available online at http://www.journalissues.org/IBME/
© 2014 Journal Issues
ISSN 2350-157X

Original Research Article

Hybrid of neuro-fuzzy inference system and quantum genetic algorithm for prediction in stock market

Accepted 9 June, 2014

**Osman Hegazy,
Omar S. Soliman
and
Ahmed A. Toony***

Faculty of Computers and
Information – Cairo University, Egypt

*Corresponding Author
Email: ahmed.a.eltoony@gmail.com
Tel.:+201223334656

Many classical soft computing approaches have successfully applied in the prediction of stock price and showed a good performance. This paper investigates the power of Quantum Genetic Algorithm in a neuro-fuzzy system composed of an Adaptive Neuro Fuzzy Inference System (ANFIS) controller used in prediction of stock market, identified using an optimization technique based on a double chains quantum genetic algorithm. In this paper the ANFIS controller and the stock market process model inputs are chosen based on a comparative study of fifty different combinations of past stock prices to determine the stock market process model inputs that return the best stock trend prediction for the next day in terms of the minimum Mean Square Error (MSE). The proposed model are tested with actual financial data and show the weak form of compared approach by demonstrating much improved and better predictions by finding optimal value for optimization variable in ANFIS using a double chains quantum genetic algorithm.

Key words: Fuzzy logic, neural network, stock market prediction, neuro-fuzzy, quantum genetic algorithm

INTRODUCTION

Stock market forecasting is the trying to determine the future value of a company stock or other financial instrument traded on a financial exchange. The successful forecasting of a stock's future price could lead to huge profit (Liang, 2013). Some believe that the changes in the stock price are based on the random assumptions and therefore unpredictable. Others disagree and those with this viewpoint have a plentiful number of methods and technologies which allow them to predict future price information. They assumed that the behaviors of the stock market in the future could be predicted using historical information; this is the idea behind Technical Analysis of Investment trading. Therefore, there exists a function

$$P(t) = f(p(t-n), x, y) \quad (1)$$

where $P(t)$ is the stock price for the day, $p(t-n)$ is the stock price for the days before the day while x and y are the other influence factors like daily lowest price, daily highest price, open price and volume of trading, ... etc . Thus, the work of predicting the future price of the stock turns to the problem of regression data in the sense of computer science (Hemanth et al., 2012).

Many effective methods are implemented to solve this problem in computer science and operation research. Among them the artificial neural networks (ANN) are well suited for getting accurate solutions in an environment described by noisy, complex, partial or irrelevant information**.** Furthermore, fuzzy logic is an effective rule-based modeling system in artificial intelligence not only supports unfaithful information, but also makes a scope of approximate reasoning. But the fuzzy logic lacks the ability to self-learning. So, the concept of combining fuzzy logic techniques with a neural network model becomes more prevalent (Qinghual et al., 2010).

(Chiroma et al., 2013) says that "among the fuzzy models that exist, Takagi-Sugeno (TS) fuzzy model is the most

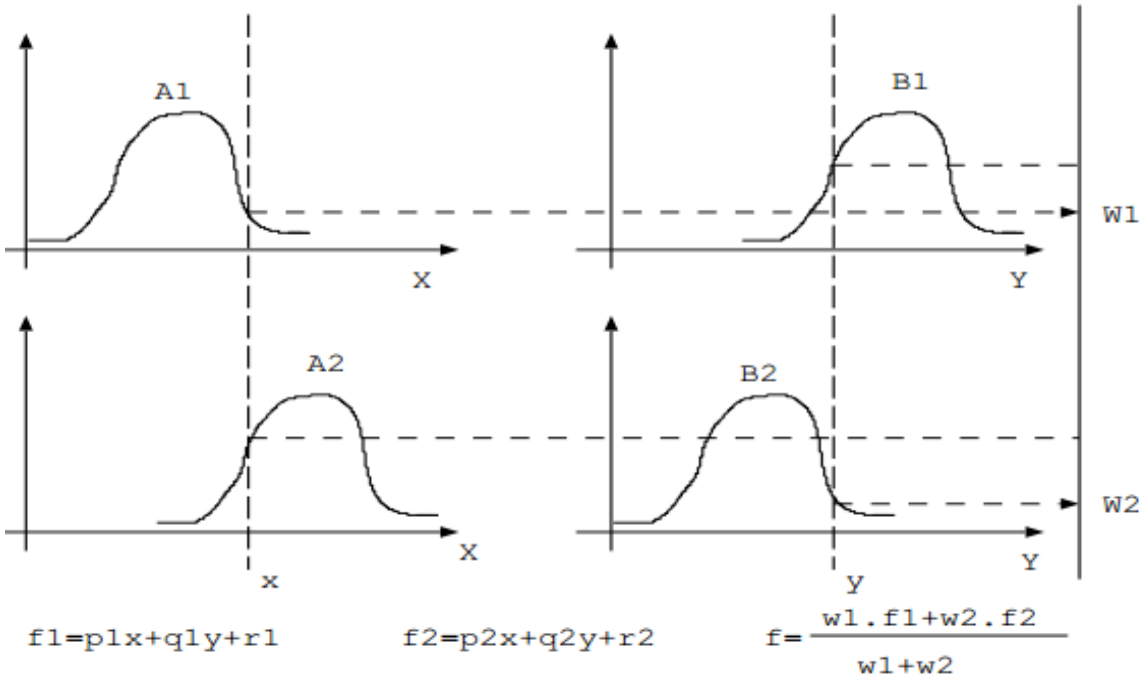

Figure 1: Sugeno fuzzy model (George et al., 2011)

preferable model for multiple independent variables and single dependent variable". on other hand prediction using Artificial Neural Networks (ANNs) model outperform all methods of prediction (Bello and Chiroma, 2011).

The proposed model refines the forecasting accuracy of the proposed model by Double Chains Quantum Genetic Algorithm method, in which the value of radii in subtractive clustering fuzzy technique used in ANFIS model is optimized using Quantum Genetic Algorithm (QGA) and follow the framework for selecting data mining techniques presented by Chiroma et al.,(2014).

Based on the concept above, this paper proposes a novel model to forecast the Bombay Stock Exchange (BSE) Sensex 500 index. This paper optimizes the fuzzy inference system parameters by an adaptive network, which can overcome the limitations of statistical methods. Finally, the optimized value model in which value of radii used in fuzzy subtractive clustering method is optimized using QGA is applied to the proposed model for refining forecasting accuracy. The experiment results indicate that creating a considerate optimization model makes the performance of the proposed model is better than the compared model.

The rest of the paper is organized as follows: Section 2 describes the background and related works describes the Sugeno type fuzzy rule based system, fuzzy subtractive, ANFIS Structure, Genetic Algorithms (GAs) Quantum Genetic Algorithm (QGA). Section 3 introduces details of proposed model. Section 4 describes implementation and Section 5 demonstrates experimental results; where the last section is devoted to conclusions.

Takagi Sugeno Type fuzzy rule based system

For a first order Takagi –Sugeno and Kang fuzzy model, a common rule is represented as follows:

- When the fuzzy inference system has two inputs x and y and one output f
- A first-order Sugeno fuzzy model shown in Figure 1 has rules as the following:
- Rule 1: If x is A_1 and y is B_1 , then
  $Z_1 = p_1x + q_1y + r_1$ (2)
- Rule 2: If x is A2 and y is B2 , then
  $Z_2 = p2x + q2y + r2$ (3)

where x and y are linguistic variables and A_1, A_2, B_1 and B_2 are corresponding fuzzy sets and p, q, r are linear parameters called consequent parameters (Liang, 2013).

There are many variables that the stock prices might be affected by so many factors. And many papers have dealt with selection of input to mapping financial indexes and stocks (Akbar and Werya, 2010). Proposed model has dealt with stock price variation in each day, it takes the open price as input 1, highest price as input 2, lowest price as input 3 and close price as output**.** The performance measurements have been used for appraising performance and error appraisal of results of model. These measures are

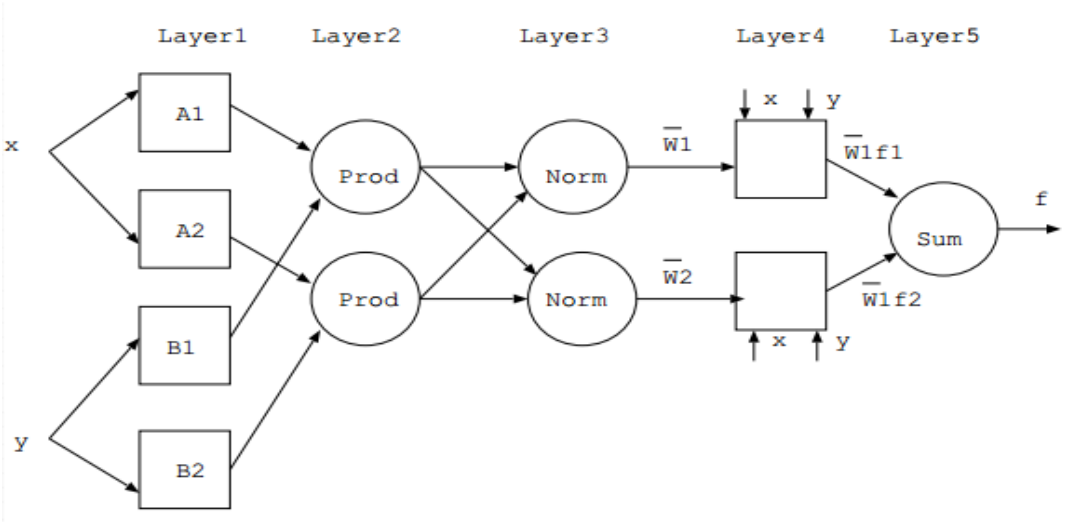

Figure 2: Anfis structure

used widely for assess the gap between real output and predicted results produced by model (Akbar and Werya, 2010). MSE is used to evaluate proposed model in the experiment.

$$MSE = (1/N) * \sum error^2 \quad (4)$$

The fuzzy subtractive clustering

If there not had a clear idea how many clusters there should be for a given set of data, Subtractive clustering (Stephen, 1994), is a fast, one-pass algorithm for estimating the number of clusters and the cluster centers in a set of data.

ANFIS Structure

Neural networks and fuzzy logic are complementary tools in building intelligent systems. While neural networks are low-level computational structures that perform well when dealing with raw data, fuzzy logic deals with logic at a higher level, using linguistic information which is given from system analyst (Toha and Tokhi, 2010). However, fuzzy systems don't have the ability to learn and cannot adjust themselves to a new environment. On the other hand, although neural networks can learn, they are unclear to the user (Loung and Suet, 2006).

A neuro-fuzzy system is a neural network which works like a fuzzy inference model. It can be trained to develop IF-THEN fuzzy rules and determine membership functions for input and output variables of the system. Expert knowledge can be added to the structure of the neuro-fuzzy system (Avagaddi et al., 2013).

ANFIS is a class of adaptive networks that are functionally equivalent to fuzzy inference systems. it represents Sugeno & Tsukamoto fuzzy models and uses a hybrid learning algorithm (Rajpal et al., 2012). The main structure of ANFIS model contains 5 layers as shown in Figure 2 (Erik et al., 2009).

Genetic Algorithm

Genetic Algorithms (GAs) are adaptive algorithms for finding the global optimum solution for an optimization based on the evolutionary ideas of natural selection and genetics (Y and Q, 1999). As such they represent an intelligent exploitation of a random search used to solve optimization problems. Although randomized, GAs are random at all, instead they use historical information to force the search into the region of better performance within the search space. It is better than conventional AI in that it is more robust. Unlike older AI systems, they do not break easily even if the inputs changed slightly, or in the presence of reasonable noise. Also, in searching a large state-space, multi-modal state-space, or n-dimensional surface, a genetic algorithm may offer significant benefits over more typical search of optimization techniques (Dhubkarya et al., 2010). GAs simulate the survival of the fittest among individuals over consecutive generation for solving a problem. Each generation consists of a population of character strings that are analogous to the chromosome that seen in Deoxyribonucleic acid (DNA). Each individual represents a point in a search space and a possible solution. The individuals in the population are then made to go through a process of evolution (Shalini and Ejaz, 2013).

Gas are based on an analogy with the genetic structure and behaviour of chromosomes within a population of individuals using the following foundations:

Individuals in a population compete for resources and mates. Those individuals most successful in each 'competition' will produce more offspring than those individuals that perform poorly.

Genes from `good' individuals propagate throughout the population so that two good parents will sometimes produce offspring that are better than either parent. Thus each successive generation will become more suited to their environment.

Quantum Genetic Algorithm (QGA)

QGA is qubit based encoding chromosomes, such that each chromosome can be found in superposition of many states at quantum population and it represents only one state in classic population by applying measurement. According to the information of optimal individual in a population, quantum gates can lead chromosomes to update themselves (Omar et al., 2012).

Qubit encode

Qubit is the smallest information unit in quantum computing, it represents a two - state quantum system.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (12)$$

Compared with classical-bit, the state of quantum-bit can be any linear superposition state between "0" and "1" state, namely:

$$|\alpha|^2 + |\beta|^2 = 1 \quad (13)$$

The system will give "0" state when the whole probability is $|\alpha|^2$ and will give "1" state when the whole probability is $|\beta|^2$, also normalization condition must be satisfied such that

$$|0\rangle \text{ and } |1\rangle \text{ (Dirac's notation)}$$

In this model the state is a superposition of two basis states, “0” and “1” or

$$|\alpha|^2 \Rightarrow \textit{The probability of observing the outcome } 0$$

This state is unknown before we make a measurement, After we perform a measurement the system is in one of the two basis states:

$$|\beta|^2 \Rightarrow \textit{The probability of observing the outcome } 1$$

METHODOLOGY OF THE PROPOSED MODEL

Proposed model works to determine the best value for radii of the genfis2 method which help in the training phase for stock price data using ANFIS model in Matlab using double chain quantum genetic algorithm.

A double chains quantum genetic algorithm (DCQGA) based on the probability amplitudes of quantum bits. In this method, the probability amplitudes of each qubit are regarded as two genes, each chromosome contains two gene chains, and each of gene chains represents an optimization solution. The number of genes is determined by the number of optimization parameters. Taking each qubit in the optimal chromosome as the goal, individuals are updated by quantum rotation gates, and mutated by quantum non-gates to increase the diversity of population. This mean that the mutation process implemented by quantum non-gates and the crossover and selection operations were done by quantum rotation gates. Where the changing trend of fitness function in search point is applied to design of the rotation angle computing function. If the changing ratio of fitness function in a certain search point is greater than that in other points, the rotation angle is properly decreased. otherwise, the rotation angle is properly increased. Such method can make each chromosome stride in plainness of search process to accelerate convergence and rush in sweep of the search process to avoid missing the global optimum solution (Maojun and Fuhua, 2010).

The main steps of the proposed model are described in Algorithm 1.

Algorithm 1: Proposed model

Step 1: Data Pre-processing (Phase one)
Step 2: Generate random population
Step 3: Calculate the value of Radii
Step 4: Initialize Anfis model
Step 5: Perform optimization algorithm using DCQGA (Phase two).
Step 6: Finding best accuracy and performance.
Step 7: Stop

The conceptual design of the proposed model is shown in Figure 3.

Implementation

The implementation of the proposed model consists of two main phases.

Phase one: Data Pre-processing

Data Pre-processing is a process that converts the raw inputs and outputs (target values) into a form understandable or acceptable before the training process. Often, this is used to reduce the dimensionality of input data and to optimise the generalization performance (Christopher, 1995).

The stock price is quite different in value may affect the performance of the prediction algorithm. So the original

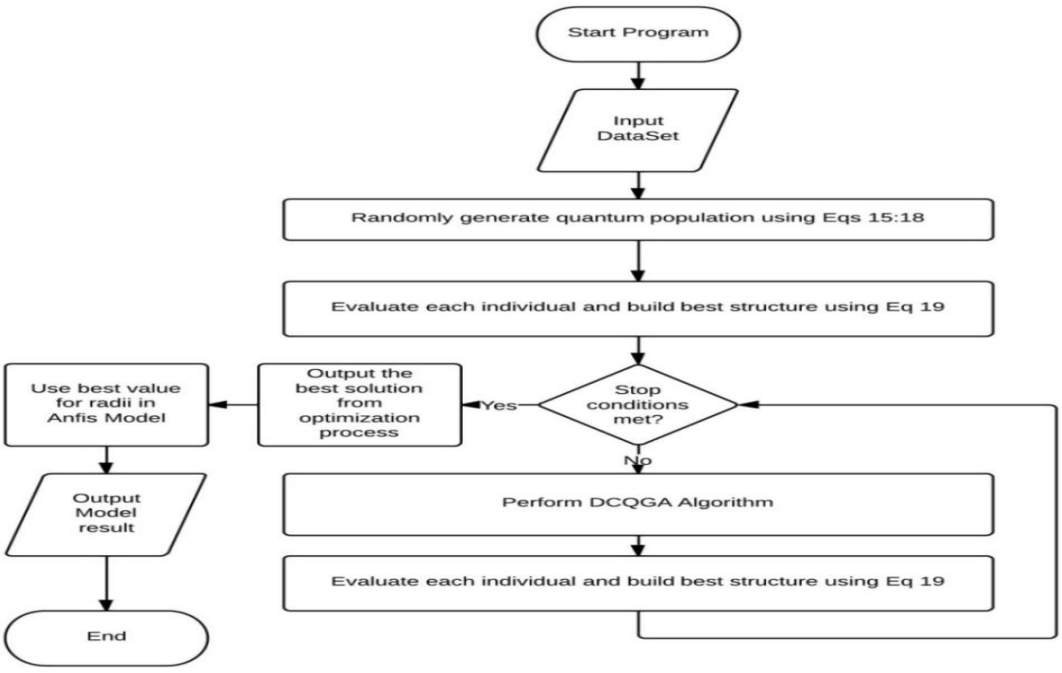

Figure 3: Flow Chart for the proposed model

data is mapped to [0,1] by min-max normalization:

$$X(i) = \frac{x(i) - m}{M - m} \qquad (14)$$

for the time series data x, m=min{x}, M=max(x). In the experiment, four kinds of time series, e.g. the open price, close price, highest price and lowest price are normalized independently (Qinghual et al., 2010).

Phase two: Optimization algorithm

The used algorithm in optimization phase based on double chain quantum genetic algorithm (Maojun and Fuhua, 2010) . it is described as follows:

1. Generate initial angle to produce double chain from it

$P_{i,1}=(\cos(y_{i,1}),\cos(y_{i,2}),\ldots,\cos(y_{i,n}))$ (15)

$P_{i,2}=(\sin(y_{i,1}),\sin(y_{i,2}),\ldots,\sin(y_{i,n}))$ (16)

where $y_{i,n}$ is a random number between 0 and 2π, $P_{i,1}$ called cosine solution and $P_{i,2}$ called sine solution.

2. Perform solution space transform

$X_{(i,j)c} = 0.5 * [b_i(1+\alpha_{i,j})+a_i(1-\alpha_{i,j})]$ (17)

$X_{(i,j)s} = 0.5 * [b_i(1+\beta_{i,j})+a_i(1-\beta_{i,j})]$ (18)

where $i = 1: m$, $j = 1: n$, m the number of qubits and n the population size.

3. Calculate fitness value which equal to

$$\frac{1}{1+MSE} \qquad (19)$$

4. Update best structure.

$$\Delta\theta_{i,j} = -\mathrm{sgn}(A)\,\Delta\theta_o \exp\left(-\frac{|\nabla f(X_{i,j})| - \nabla f_i \min}{\nabla f_i \max - \nabla f_i \min}\right) \qquad (20)$$

5. Perform quantum rotation gates to update corresponding qubits on the current chromosome for each chromosome.

Where $\Delta\theta_0$ is the initial value of rotation angle,

$$A = \begin{vmatrix} \alpha_0 & \alpha_1 \\ \beta_0 & \beta_1 \end{vmatrix} \qquad (21)$$

Then the direction of rotation angle Δθ can be determined by such rules as follows: if $A \neq 0$, then sgn (Δθ) =- sgn(A), else if A = 0, the direction of Δθ is arbitrary,

$$\nabla f(X_{i,j}) = \frac{f(X_{i,p}) - f(X_{i,c})}{(X_{i,j})^p - (X_{i,j})^c}; \qquad (22)$$

$$\nabla f_i \max = \max\left(\left|\frac{f(X_{1,p})-f(X_{1,c})}{(X_{1,j})^p-(X_{1,j})^c}\right|, \left|\frac{f(X_{2,p})-f(X_{2,c})}{(X_{2,j})^p-(X_{2,j})^c}\right|, \ldots, \left|\frac{f(X_{n,p})-f(X_{n,c})}{(X_{n,j})^p-(X_{n,j})^c}\right|\right) \qquad (23)$$

$$\nabla f_i \min = \min\left(\left|\frac{f(X_{1,p})-f(X_{1,c})}{(X_{1,j})^p-(X_{1,j})^c}\right|, \left|\frac{f(X_{2,p})-f(X_{2,c})}{(X_{2,j})^p-(X_{2,j})^c}\right|, \ldots, \left|\frac{f(X_{n,p})-f(X_{n,c})}{(X_{n,j})^p-(X_{n,j})^c}\right|\right) \qquad (24)$$

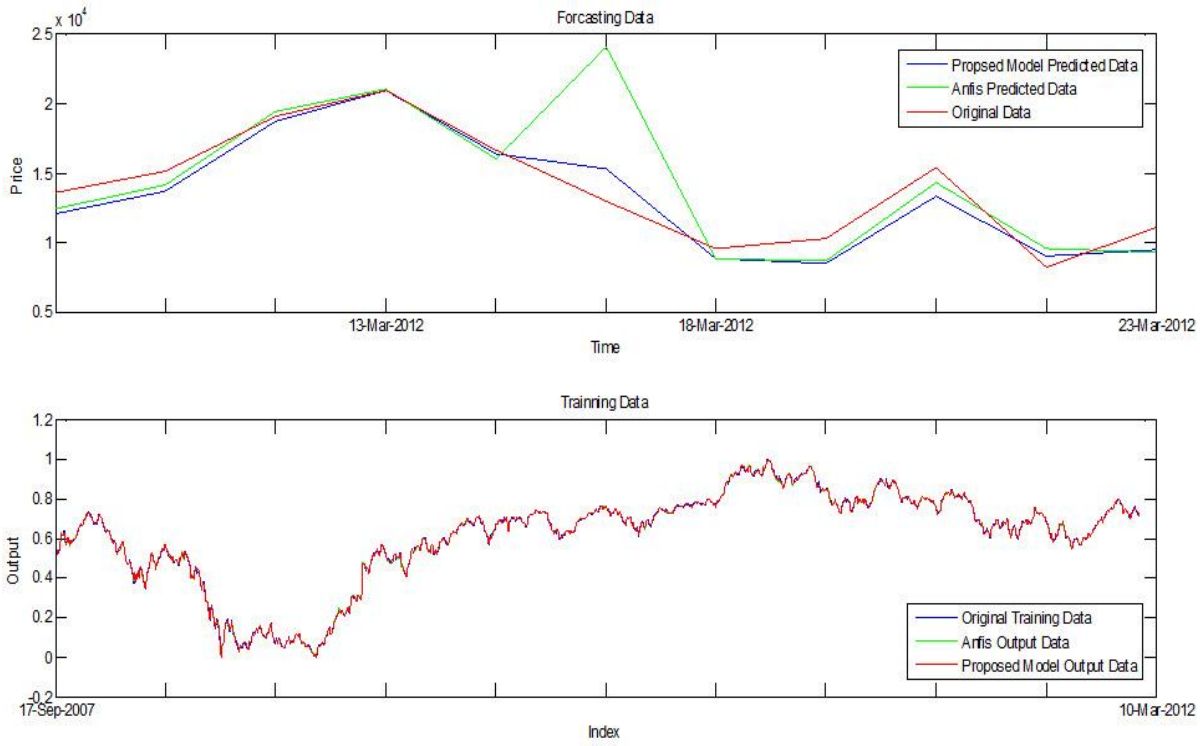

Figure 4: The proposed model result

where $X_{i,p}$ and $X_{i,c}$ represents ith vector in solution space corresponding to the parent colony and child colony, and $(X_{i,j})^p$ and $(X_{i,j})^c$ represents the jth variable of vector $X_{i,p}$ and $X_{i,c}$, respectively.

6. Mutate qubits according to mutation probability (Pm = 0.1) which equal to 1 over the population size (pop size = 10) and these values comes from experimental for reasonable speed and performance for the model .

The procedures mentioned above was used to find the best value for the optimization parameter using double chain quantum genetic algorithm and calculate fitness function according to ANFIS model using subtractive clustering to generate initial FIS for ANFIS system.

RESULTS AND DISCUSSION

A set of stock market data collected by yahoo finance from BSE Sensex index for a period from 17th September 2007 to 12 march 2012 which contains the open price, highest price, lowest price and close price for each working day (Hemanth et al., 2012). BSE Limited is the oldest stock exchange in Asia it is now popularly known as the BSE was established as "The Native Share & Stock Brokers' Association" in 1875. Over the past 135 years, BSE has facilitated the growth of the Indian corporate sector by providing it with an efficient capital raising platform. Today, BSE is the world's number 1 exchange in the world in terms of the number of listed companies (over 4900). It is the world's 5th most active in terms of number of transactions handled through its electronic trading system. And it is in the top ten of global exchanges in terms of the market capitalization of its listed companies. BSE is the first exchange in India and the second in the world to obtain an ISO 9001:2000 certification (Yusuf and Asif, 2012).

To show the performance of the proposed model, testing was done on the BSE Sensex index and compare its result with compared model result in Hemanth et al., (2012), obtained result of the proposed model as shown in Figure 4 and Figure 5 display the results of the compared model.

The first part in Figure 4 illustrate future ten days forecasting using proposed model with blue colour , using ANFIS model with green colour and the original corresponding data in red colour, the second part represent the original data in blue colour, Anfis data in green colour and proposed model data in red colour. As shown above the matching between input data and predicted data in training data is clear. After running the proposed model for 30 times the average of MSEs is calculated and equal to 0.012187.

As reported in the Table 1 below the differences between the values of MSE is very small, as seen clearly the range of

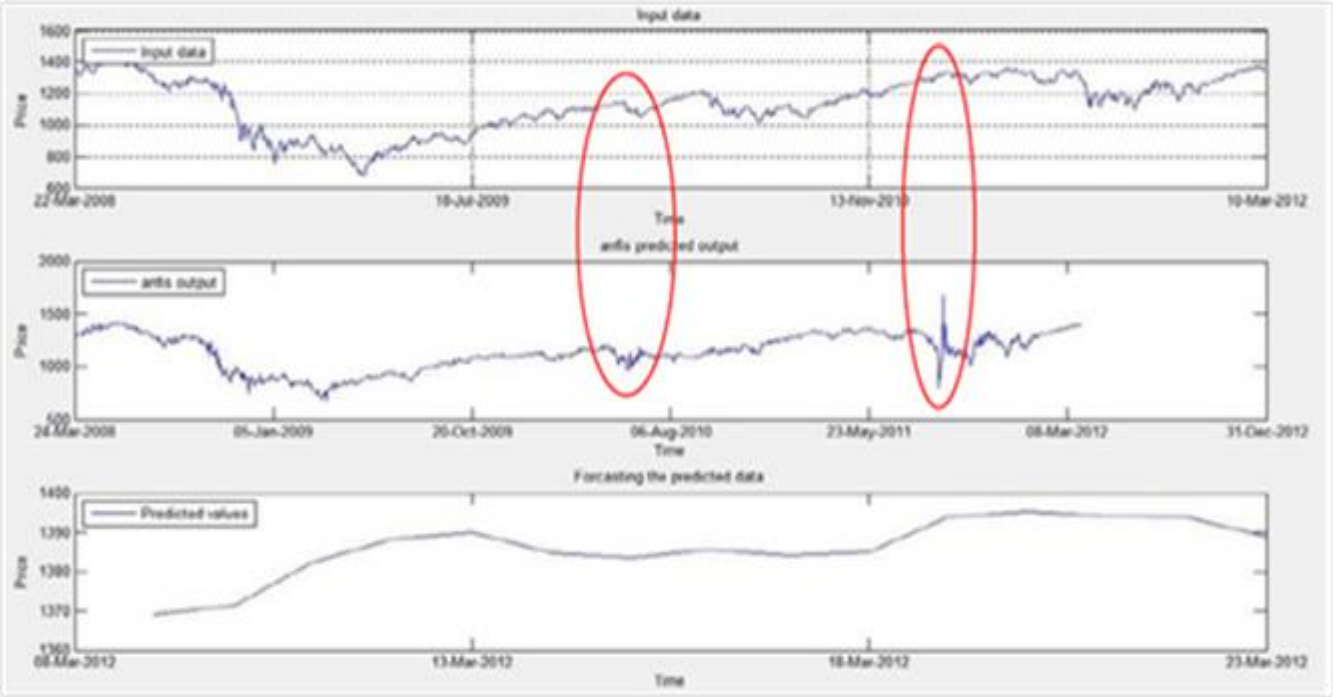

Figure 5: Result mentioned in Hemanth et al., (2012)

Table 1: Results of running 30 times

No.	Best MSE	radii	Best Fitness	Compared model MSE
1	0.012	0.5002	0.9881	0.2066
2	0.0121	0.5371	0.9881	0.2066
3	0.013	0.7818	0.9872	0.2066
4	0.0121	0.5391	0.9881	0.2066
5	0.0122	0.1184	0.9879	0.2066
6	0.013	0.1537	0.9872	0.2066
7	0.013	0.7577	0.9872	0.2066
8	0.012	0.1019	0.9881	0.2066
9	0.012	0.102	0.9881	0.2066
10	0.012	0.499	0.9881	0.2066
11	0.012	0.6026	0.9882	0.2066
12	0.012	0.5761	0.9881	0.2066
13	0.0119	0.6603	0.9883	0.2066
14	0.012	0.1034	0.9882	0.2066
15	0.012	0.5008	0.9881	0.2066
16	0.0119	0.4329	0.9882	0.2066
17	0.0127	0.7327	0.9875	0.2066
18	0.0121	0.1001	0.9881	0.2066
19	0.0119	0.4304	0.9882	0.2066
20	0.0121	0.5327	0.9881	0.2066
21	0.012	0.5787	0.9881	0.2066
22	0.0128	0.1423	0.9873	0.2066
23	0.012	0.4994	0.9881	0.2066
24	0.0121	0.1	0.9881	0.2066
25	0.0122	0.1166	0.9879	0.2066
26	0.0127	0.3758	0.9875	0.2066
27	0.0119	0.1131	0.9882	0.2066
28	0.012	0.1029	0.9881	0.2066
29	0.012	0.1041	0.9882	0.2066
30	0.0119	0.6603	0.9883	0.2066

differences about 0.0001 except 6 from 30 value have a difference about 0.0006 to 0.0009 and the proposed model has higher accuracy so it has in average mse 0.012187 while the comparable model have an average of 0.2066

The compared model introduces that neural predictors are used to expect the general behavior of financial indexes in the context of stocks and options trading. It presents the prospect of prediction is usually a few days and trading strategies are studied by historical data. The model investigates the neuro fuzzy approaches in prediction of financial time series and it shows well performance in the context of various trading strategies involving stocks. First step is data collection, the stock market closing data is extracted from the BSE stock market, then data is normalized to prepare the data for training by applying Adaptive Neural Network based techniques, then predicting the future data by fuzzy technique. Data collected is same as data used in the proposed model and the same period chosen.

In the Figure 5 the differences between original data and data predicted in training phase is shown, this means that the system is not efficient.

It can be seen clearly that the system mentioned in the compared paper has two fully observed differences but in the proposed model the matching between original data and predicted data is obviously shown.

CONCLUSION

Many classical soft computing approaches have successfully applied in the prediction of stock price and showed a good performance. In this study, the objective is to investigate the power of QGA by using it to optimize the value of radii used in subtractive clustering method applied to generate initial fuzzy inference systems and use it in neuro-fuzzy model "ANFIS" to train the system by historical data of BSE Sensex Index and predict the stock price. The proposed model was implemented and evaluated using BSE Sensex index and compared with ANFIS model. The obtained result showed the performance of the proposed model is better than ANFIS model. The soft computing approaches employed in this paper are the ANFIS model and the double chain quantum genetic algorithm model.

REFERENCES

Akbar E, Werya A (2010). Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis. Expert Syst. Appl. 37(7): 4742-4748.

Avagaddi P, Durga PB, Ravindra BM, Manohar K, Siva MRN (2013). A Wavelet-Based Adaptive Neuro Fuzzy Inference System Approach To Fault Classification For High Speed Protective Relaying. 1(1):2347-2812.

Bello MY, Chiroma H (2011). Utilizing artificial neural network for prediction in the Nigerian stock market price index. GESJ,30(1)1512-1232.

Cao YJ, Wu QH (1999). Teaching Genetic Algorithm Using Matlab. Int. J. Elect. Enging. Educ. Vol. 36:139–153. Manchester.

Chiroma H, Abdulkareem S, Abubakar A, (2014). A Framework for Selecting the Optimal Technique Suitable for Application in a Data Mining Task. In *FIT* (pp. 163-169). Lecture Notes in Electrical Engineering .Springer Berlin Heidelberg Germany.

Chiroma H, Abdulkareem S, Abubakar A, Zeki A, Gital AYU and Usman MJ (2013). Co—Active neuro-fuzzy inference systems model for predicting crude oil price based on OECD inventories. *(ICRIIS). 2013* International Conference *on* (pp. 232-235). IEEE.

Christopher MB (1995). Neural Networks for Pattern Recognition. Oxford University Press, New York, NY, USA.

Dhubkarya DC, Deepak N, Jay K (2010). Function Optimization Using Genetic Algorithm by VHDL. GJCST, Vol 10, No 9

Erik C, Daniel Z, Raul R (2009). Neurofuzzy prediction for gaze control. CJECE , October 2009.

George SA, Constantinos DZ (2011). Elliott Wave Theory and neuro-fuzzy systems, in stock market prediction. The WASP system. Expert Syst. Appl. 38(8): 9196-9206.

Hemanth KP, Prashanth KB, Nirmala TV, Basavaraj PS (2012). Neuro Fuzzy based Techniques for Predicting Stock Trends. Int. J. Comp. Sci. Issues, 9(4)385-391.

Liang-Ying W (2013). A GA-weighted ANFIS model based on multiple stock market volatility causality for TAIEX forecasting. APPL SOFT COMPUT (13)2 :911–920.

Luong TT, Suet PY (2006). A Fuzzy Neural Based Data Classification System. DMIN : 383-388.

Maojun C, Fuhua S (2010). Double Chains Quantum Genetic Algorithm with Application in Training of Process Neural Networks. ETCS. Second International Workshop on

Omar SS, Roba B, Amr A (2012). Associative Classification using a Bio-Inspired Algorithm. AusDM. 134: 119-125.

Qinghua W, Zehong Y, Yixu S (2010). Hybrid Approaches for Stock Price Prediction. NISS. 4th International Conference.

Rajpal SB, Ramvir S, Pradeep Kumar S (2012). Adaptive Neuro-Fuzzy Inference System for Prediction of Effective Thermal Conductivity of Polymer-Matrix Composites. MNSMS, 2(3): 43-50.

Shalini S, Ejaz AL (2013). Study of Variation in TSP using Genetic Algorithm and Its Operator Comparison. Int. J. Sci. Comput. Eng., 3(2): 2231-2307.

Stephen LC (1994). Fuzzy Model Identification Based on Cluster Estimation. J. Intell. Fuzzy Syst, 2(3):267-278.

Toha SF, Tokhi MO (2010). ANFIS modelling of a twin rotor system using particle swarm optimisation and RLS. CIS, IEEE 9th International Conference, pp.1-6.

Yusuf P, Asif P (2012). Prediction of the Bombay Stock Exchange (BSE) Market Returns Using Artificial Neural

Network and Genetic Algorithm. J. Intell. Learn. Syst. Appl.,22(4): 108-119.