2019 5th International Conference for Convergence in Technology (I2CT)
Pune, India. Mar 29-31, 2019

Prediction of Stock Market Using Artificial Intelligence

Hemil N. Shah
Computer Engineering
Mukesh Patel School of Technology Management & Engineering
Mumbai, India
shahhemil1999@gmail.com

***Abstract*—All over the world there are enormous amount investments in stock markets. The financial system all over the world is interconnected to the stock markets. Nowadays stock markets particularly as trading has become like money making medium.**

Artificial intelligence is used to identify the unknown samples and analyze them and taking these values as input they are used to predict the stock market. Predicting stock market is a very ambiguous tasks it has many methods which make it possible to predict such artificial neural network (ANN), adaptive neuro fuzzy inference (ANFIS), Swarm Intelligence etc. There are future methods like Levenberg Marquardt which is a part of neural which help not only to predict accurately but also being more efficient than ANFIS in many ways such as time taken, memory allocation, accuracy etc.

Keywords—Artificial Neural Network (ANN); Adaptive Neuro Fuzzy Iinference (ANFIS); Swarm Intelligence (SI); Stock Market;other methods for prediction

I. BRIEF DESCRIPTION

The growth of stock market has been the major strength and predicting it is the major concern as the values of the stock keep on changing making it ambiguous task. Many people invest in stock market to expand their capital. There are many methods which are used for prediction of a stock exchange for e.g.:-

ANFIS is used to predict the Istanbul stock exchange [2] ANFIS consist of three stages in it where there is wavelet transform and recurrent method which is combined as in artificial bee algorithm.

But in Dhaka stock exchange the neural network worked very little but ANFIS showed more accuracy. [2]

So the methods used depend from country to country.

The methods for prediction are as follows:-

- Artificial neural network
- Artificial neuro fuzzy inference
- Swarm intelligence

Artificial neural network (ANN):-

The ANN basically is like an artificial human brain in which every artificial neuron is connected to each other. [1]This artificial neuron has a memory bank where all knowledge is stored. In the network the signals are broadcast from input to output and are called the weight between both input and output nodes.[1]

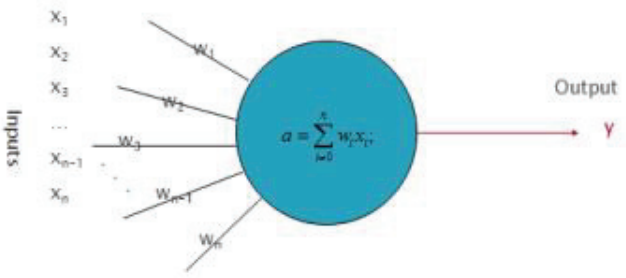

Fig.1 Graphical representation of artificial neuron

In the figure above:

X_i = neuron inputs

W_i= neuron weights

The input for this is taken by a which contains the values of X_i and W_i . The output of this equation is within (0,1) as shown below:

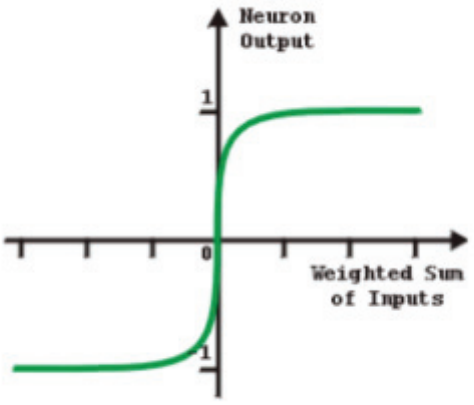

Fig.2 The threshold function

There are three layers in the artificial neuron:

Input layer, output layer and hidden layer. The role of the hidden layer is that whenever there is input it moves towards output node but the hidden executes the input before it reaches the output.

Disadvantages:

-It is difficult for ANN to learn patterns

- It might be volatile when complicated finance data is applied.

Feedforward network:-

There two types of feedforward network:-

- Single layered
- Multilayered

Single layered have two layers input and output.

Input layer takes values as input and output layer gives output.

But there are hidden layers in this neural network which executes the input values. Inputs are like input neurons mulplied by the weight of the neurons. Then this input is executed by hidden layer and output is given in the form like multiply input neurons and weight of the neurons.

Multilayered are similar to single layered but there are range of hidden layers.

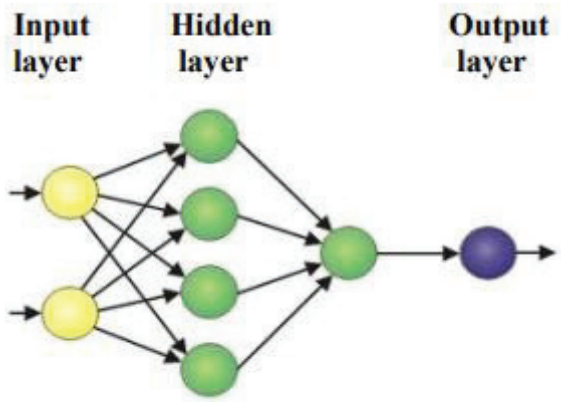

Fig.2 Feedforward Neural Network

Single layered feedforward network is most used for DELTA rule while multi layered feedforward network is used for generalization of delta rule. [1]

Advantages:-

-As they are good in DELTA rule so they try to get the correct output by overcoming errors using backpropogation algorithm.

Disadvantages:-

- Sometimes there are complicated task so it might be possible that while solving errors and going back and forth in layers they might they a long time to get the work done.

Prediction methods:-

As the stock prediction is a really complex task it requires a lot of calculation analysis so this method has various categories through which approximate predicion of stock market is possible.The categories are as follows:-

- Technical Analysis
- Fundamental Analysis
- Time series Prediction
- Machine learning

Technical analysis is basically used for prediction of statistics like cost, as they change rapidly. The present behaviour of stocks is seen and then prediction of future prices of stocks if given.[1]

Fundamental analysis provide additional security approach to the financial system by looking into the different issues.[1]

In Time Series prediction the values of stocks are taken for particular period of time and the time series variables forecast the values for the futures as the time series prediction are data series whih are placed in successive manner.[1]

Machine learning method is baed on leaning from training and practice. ANN use machine learning to improve their capabilities.[1]

Disadvantages:

- Technical analyis are subjective and our personal biases are reflected
- Fundamental analysis is time consuming .

Neural Network Learning:-

The error can cause change in the expected output so neural network has backward algorithm in which the error moves in backward direction from output layer to the hidden layer and this process continues until the error is solved or minimized.[1]

Disadvantage:-

It is the time consuming process as it has to go back and forth again and again until there are no errors or exact value.

Backpropogation algorithm:-

In the backpropoation algorithm the errors are given from output layer to the hidden layer so that the eerors resolved making the ANN trained. So now it can be used in fresh data.[1]

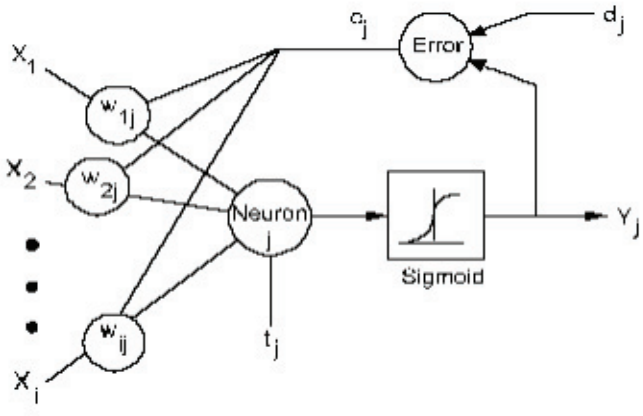

Fig 5. Adjustment of Neuron Weights

Disadvantages:

- It is a very slow process as it takes lot of time to get the desired output.
- Someimes we do not get the desired output.
- Not sure how many inputs can be given together.
- It requires different activation function for all neurons.

The methods used above are efficient but they have a disadvantage that many are time consuming and they have difficulty in learning patterns so there are improved neural network algorithm which are efficient they are as follows:-

- Gauss Newton algorithm.
- Levenberg Marquardt algorithm

Gauss Newton algorithm:-

Gauss Newton method improves the slow convergence. In this method it takes the approximation for prediction. The Gauss Newton method is used to solve non- linear least square problems where the random values are approximated and then they are solved to get the correct output. It works efficiently in the quadratic surface.

Disadvantages:-

If there is no reasonable approximation of error function then the algorithm is mostly divergent.

Levenberg Marquardt algorithm:-

As it combines steepest descent method and Newton method,it has an update rule which is:-

$$\Delta w = (J^T J + \mu I)^{-1} J^T e$$

Where w = weight vector

J= Jacobian's matrix

I= identity matrix

e= error vector

$$J = \begin{vmatrix} \frac{\partial e_{11}}{\partial w_1} & \frac{\partial e_{11}}{\partial w_2} & \cdots & \frac{\partial e_{11}}{\partial w_N} \\ \frac{\partial e_{12}}{\partial w_1} & \frac{\partial e_{12}}{\partial w_2} & \cdots & \frac{\partial e_{12}}{\partial w_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{1M}}{\partial w_1} & \frac{\partial e_{1M}}{\partial w_2} & \cdots & \frac{\partial e_{1M}}{\partial w_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{P1}}{\partial w_1} & \frac{\partial e_{P1}}{\partial w_2} & \cdots & \frac{\partial e_{P1}}{\partial w_N} \\ \frac{\partial e_{P2}}{\partial w_1} & \frac{\partial e_{P2}}{\partial w_2} & \cdots & \frac{\partial e_{P2}}{\partial w_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{PM}}{\partial w_1} & \frac{\partial e_{PM}}{\partial w_2} & \cdots & \frac{\partial e_{PM}}{\partial w_N} \end{vmatrix}$$ [2]

Where p= no. of training patterns

m= no. of outputs

n= no. of weights.

Error vector is calculated by

$$e_{pm} = d_{pm} - O_{pm}$$

Where d_{pm} and O_{pm} are desired output and actual output respectively.

The algorithm adjusts whenever the error is increasing or decreasing:-

When we update the values we take the weight vector and see if it is increasing or decreasing. If it is increasing then we have to take the previous values and then try increasing the value of μ and if the error decreases by this we take the we take the new value of weight vector which has been found and decrease the value of μ.

Disadvantage:

- We cannot solve the large jacobian's matrix which would acquire large amount of memory and time.

Improved Levenberg Marquardt:-

To reduce the cost of traditional Levenberg algorithm there is a performance index that is optimized by:-

$$F(w) = \sum_{p=1}^{p} \left[\sum_{m=1}^{m} e_{pm}^2 \right]$$ [2]

Now this computation is generalized into more improved computation to reduce the cost and time which is given by:-

$$F(w) = \sum_{p=1}^{p} \left[\sum_{m=1}^{m} e_{pm} \right]^2$$

The more optimized way was made which helped to solve the larger and more complex problems in which if μ value was decreased in the factor of 5 and increased in factor of 1.5 to get better result. It was updated by:-

$$\Delta w = (Q + \mu I)^{-1} g$$

Where Q= Quasi-Hessian

g= gradient vector

In this Quasi Hessian matrix is equal to sum of sub matrices which is q_{pm}.

$$q_{pm} = j_{pm}^T j_{pm}$$

$$j_{pm} = \begin{bmatrix} \frac{\partial e_{pm}}{\partial w_1} & \frac{\partial e_{pm}}{\partial w_2} & \cdots & \frac{\partial e_{pm}}{\partial w_N} \end{bmatrix}$$

Gradient vector is g is equal subvector η_{pm}

$$\eta_{pm} = j_{pm} e_{pm}$$

By improving the computation when quasi matrix and gradient are combined there is no need of jacobian matrix> It will be stored for temporary basis and memory used is very less.

Algorithm 1 Improved Levenberg Marquardt Algorithm

Step 1: Set Q = 0, g = 0 and Performance Index,

$$F(w) = \sum_{p=1}^{p} \left[\sum_{m=1}^{m} e_{pm}^2 \right]^2 ;$$

Step 2: According to performance Index F(w), compute j_{pm};
Step 3: Compute q_{pm};
Step 4: Compute η_{pm};
Step 5: Compute Δ w according to

$$\Delta w = (Q + \mu I)^{-1} g;$$

Step 6: Evaluate the error at the new weigh vector;
Step 7: If error <previous error

$$w = w + \Delta w$$

$$\mu = \mu \times 1.5$$

$$gotostep2;$$

else

$$\mu = \frac{\mu}{5}$$

$$gotostep4;$$

Result Analysis:-

The result analysis is done by RMSE (root mean square error) which is used to measure the accuracy and performance and coefficient of multiple determinations is used which is R^2 which is used to compare predicted value.

The RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{n=1}^{n} (y_{\text{predicted}} - y_{\text{actual}})^2}{n}}$$

The R^2 is defined as:

$$R^2 = 1 - \frac{\sum_{n=1}^{n} (y_{\text{predicted}} - y_{\text{actual}})^2}{\sum_{n=1}^{n} (y_{\text{actual}} - y_{\text{mean}})}$$

n= total no. of data samples.

$y_{predicted}$ = the predicted value of data sample.

y_{mean} = it is the average of output value.

In this case if the value of RMSE is near to 0 then there are less errors and if R^2 value is near to 1 that means high correlation.

TABLE I
ANN (IMPROVED LM ALGORITHM), ANFIS AND ANN (TRADITIONAL LM) TEST RESULT

	ANN (improved LM)	ANFIS Result	ANN (Traditional)
RMSE	0.87	1.86	1.881
R^2	0.99797	0.98116	0.98

In this case the improved LM has less errors and more correlation then any of the other methods so improved LM is more efficient than any other methods.

Modular neural network:-

Another method is of modular neural network which predicts the neural network based on index values. The neural network predicts by the combine linear and logistic regression models.[3]

Mizuno and his co- researchers use neural networks which has the precision of 63% of genetic algorithm. The have combined the genetic algorithm with indexed stock market changes. [3]

Swarm Intelligence (SI):-

Swarm is the term used for interacting agents. SI is the behavior of insects with set of rules. SI's main property is self-organization, it does not need help of external entity and in SI, agents interact with each other to achieve goal.[4]

Ant colony algorithm (ACO):-

ACO is like how the ants work the go in the random direction in search of food and after they get the food they come back to their colony leaving behind the substance called pheromone on the ground to make way for other ants. The no. of times the path is visited the pheromone gets concentrated and with time pheromone gets evaporated. By considering these conditions the shortest path is calculated.[4]

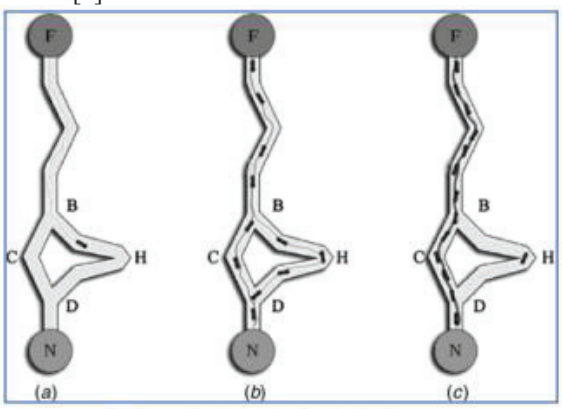

Fig.1. Shortest path selection by Ants

Artificial bee colony (ABC):-

The ABC works like a honey bee. There are 3 types of bees:- Employee bees (EB), on lookers bees (OB) and scouts:-
There is only one employee bee for one food source so the no. of food source is equal to no. of employee bees.[4]

Algorithm:-

i. Create various sources of food for all employed bees
ii. Repeat
iii. Each EB tray to reach to the food source in her memory
iv. Identify the neighbor source of food
v. Evaluates nectar amount and dances in the hive
vi. Each OB keep watch on the activity of employed bees
vii. OB selects one of the sources based on the activity and reach to the source.
viii. After selecting a neighbor around that, ant evaluates its nectar amount.

ix. Food sources are identified and replaced with the new food sources predicted by scouts.

x. The best available food source is registered.

xi. UNTIL (requirements are achieved)

Particle Swarm Optimization (PSO):-

PSO has many particles all over the space they are used to get better candidate solution. First they perform their functions locally so that they get a better solution and then they are used across the space. [4]

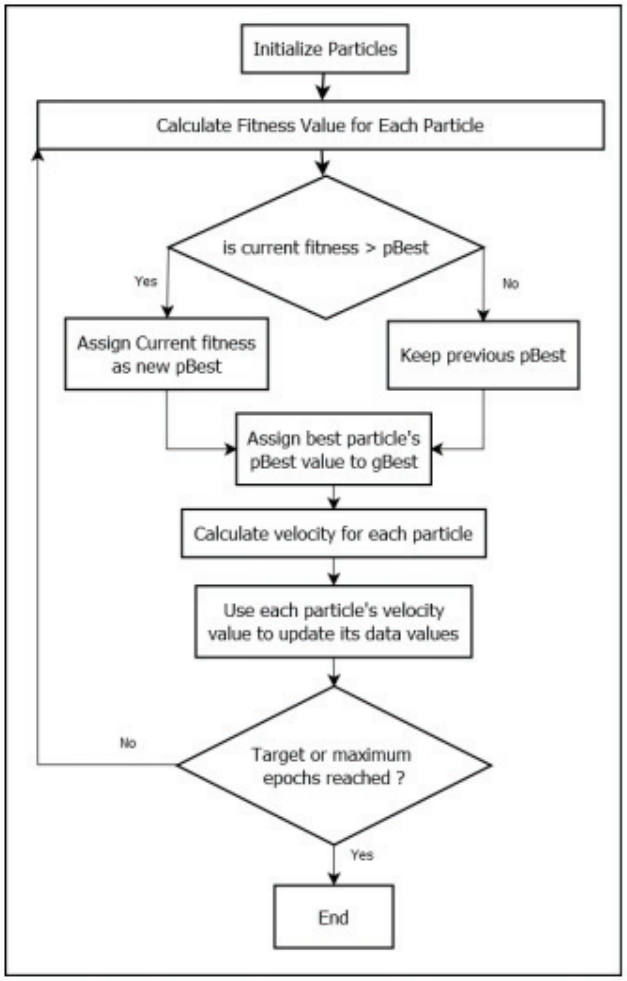

Fig.2. PSO Flowchart [6]

Cuckoo Search (CS):-

CS works like the bird cuckoo which keeps their eggs in another birds nest. Now see the nest egg as solution and cuckoo egg as better solution.

- Cuckoo stores in a nest at a time by replacing another egg
- The nest with high quality egg is carry forward from generation to generation
- The number of available nests is fixed and the cuckoo egg keeping is estimated by the host bird with a function of probability. [4]

Firefly Algorithm:-

In firefly algorithm the fireflies make a flashing activity. The less bright firefly is attracted to firefly with more brightness. Brightness tells the distance between the two fireflies. If there is no brightness then it moves randomly. [4]

SI for stock market prediction:-

SI is combined with ANN to predict the stock market. So first the Bat algorithm is combined with ANN to get the accurate results. For performance evaluation it takes seven year old data of private bank. There were three combinations considered:-

Back propagation with ANN, Partial Swarm optimization with ANN and Bat with ANN. The Bat algorithm with ANN was most accurate among three. Bat algorithm uses micro bats to fly randomly with different velocities at different positions with different frequency, wavelength and loudness and these things were intensified and got the results.

Bat – ANN algorithm:-

i. Begin

ii. BAT is Initialized then passes its first population to ANN as weight's value

iii. Load data

iv. ANN starts training and computes the accuracy of the model

v. Bat finds the initial best solution by means of the ANN's results

vi. While I < Max number of iterations

vii. Bat population performs local search to find new solution and passes it to ANN

viii. Bat population finds the best solution by means of ANN's results

ix. Until network convergence, Bats keep on finding the best weight at each epoch

x. End While

xi. End

Another method to predict accurate result was combining cuckoo search (CS) with ANN or SVM. SVM is support vector Machine which is used in machine learning. CS with SVM proved more accurate than CS with ANN.[4]

Cuckoo Search Algorithm:-

i. Start

ii. Initialize with n number of nest

iii. Select a cuckoo randomly by using equation $x_i^{t+1} = x_i^t + \propto \oplus Levy(\lambda)$

iv. Calculate fitness function (F_c)

F_c = Error rate = 100 - Accuracy

Where, Accuracy= (Number of Right Prediction/Total Predictions)*100

v. Randomly select the nest

vi. Compute fitness function (Fn)

vii. If (Fc < Fn) is true then replace the nest with the cuckoo

viii. A fraction Pa of the nest is replaced by new nests

ix. Keep the best nests as optimal fitness value

x. Check stopping criteria

xi. If it is 'no' then go to step 3

xii. Stop

INFERENCES:-

In terms of accuracy, memory and time Improved LM is much better than ANFIS and ANN.

TABLE I
ANN (IMPROVED LM ALGORITHM), ANFIS AND ANN (TRADITIONAL LM) TEST RESULT

	ANN (improved LM)	ANFIS Result	ANN (Traditional)
RMSE	0.87	1.86	1.881
R^2	0.99797	0.98116	0.98

TABLE II
MEMORY AND TIME COMPARISON OF PROPOSED LM AND TRADITIONAL LM

	Traditional LM	ANFIS	Improved LM
Time needed(s)	4.23	5.65	2.948
Memory allocated(kb)	5,674.00	6,453.00	2,592.00

REFERENCES

[1] NEELIMA BUDHANI, Dr. C. K. JHA, SANDEEP K. BUDHANI, "Prediction of Stock Market Using Artificial Neural Network", 2014 International Conference of Soft Computing Techniques for Engineering and Technology (ICSCTET), Aug. 2014

[2] Mustain Billah, Sajjad Waheed, Abu Hanifa " Stock Market Prediction Using an Improved Training Algorithm of Neural Network", 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Dec. 2016

[3] N.Srinivasan, C.Lakshmi " Stock Prediction and Analysis Using Intermittent training data with artificial neural networks", 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

[4] Mr. Pankaj K. Bharne, Dr. Sameer S. Prabhune "Survey on Combined Swarm Intelligence and ANN for Optimized Daily Stock Market Price", 2017 International Conference on Soft Computing and its Engineering Applications (icSoftComp), Dec. 2017