

# An Integrated Approach to Predictive Analytics for Stock Indexes and Commodity Trading Using Computational Intelligence

Chandra J, Dr Nachamai.M

Associate Professor,  
Department of Computer Science, Christ University,  
Bangalore, Karnataka, India.

Dr Anitha S Pillai

Professor and Head,  
Department of MCA, Hindustan University,  
Chennai, Tamil Nadu, India

---

**Abstract** — Investment prediction is a method to decide the future values of stock indexes and commodity exchanges or trading of financial services. The aim of the model is to perform optimized prediction on commodities and stock market indexes. The investment prediction is an important task for an investor to maximize his or her return on investment. The purpose of the paper is to propose an optimized model using computational intelligence and it is a step by step method that follows an integrated approach which can solve several complex problems in predictive analytics. The integrated approach in this paper uses genetic algorithm, pearson's correlation coefficient and multilayer perceptron adaline feed forward neural network to predict the next business day high values of stock indexes and commodities trading. As an integrated method, the model uses genetic algorithm as first step to check the data optimization, since the data is considered as an important element in data analytics. The optimized data is extracted using correlation coefficient and the classifier prediction is done with multilayer perceptron adaline feed forward neural network for making the prediction. The proposed model was implemented continuously on three months data to evaluate the performance and to check the accuracy on the NSEindia. The predicted values were checked against the next business day of original values, the predicted result is very close to the original values. The model is evaluated with the statistical parameter MRE, MMRE and the accuracy rate. In comparison with other existing methods, the current method outperforms other testing patterns.

**Keywords-** Predictive Analytics (PA); Computational Intelligence (CI); Genetic Algorithm (GA); Pearson's Correlation Coefficient(PCC); Multi Layer Perceptron (MLP); Adaptive Linear Element (ADALINE); Neural Network (NN).

---

## I. INTRODUCTION

Investment prediction in predictive analytics is challenging. PA is generally used to generate better decisions with greater consistency at lower cost. Investment prediction is a very essential and a difficult task in financial services. It is very much essential to develop an integrated method to analyze stock indexes and commodity exchanges. So it is necessary to develop a robust and efficient Computational Intelligence technique to predict future values. Generally, a Computational Intelligence methodology gives optimized solutions to many real world problems but the traditional methods are ineffective and infeasible to solve many real world problems. The predictive analytics is a set of business intelligent methodologies which is looking forward using past actions to anticipate the future. The applications of the

predictive analysis are cross ceiling, budgeting and forecasting. PA is an interesting method in the process of decision making on data analytics and it is considered as one of the most efficient and effective methods in the field of data analytics. The PA uses different method available in statistics, machine learning, NN, robotics and computational mathematics and artificial intelligence to discover new knowledge from any data set. CI mainly uses ANN, GA and evolutionary algorithm. The intelligence reflects the characteristics of human which is directly connected. The CI is an emerging area of research which facilitates powerful tools for modeling and analyzing complex systems. The CI is mainly focused on discovering the structure in data and recognition of some patterns. Many times, the CI is used in predictive accuracy, many researchers build integrated model for prediction based on accuracy. The CI is also used in hybrid

prediction for solar power, the performance is measured with the help of predictive accuracy in advance to ensure constant solar power in the smart grid operation [2].

The focus of this research is to find the prediction reliability of the hybrid method to improve the performance of the prediction [3]. CI embraces biology inspired paradigms like evolutionary computation and swarm intelligence that have been established in market survey, medical research pattern matching and manufacturing domain [4].

## II. RELATED WORKS

Niall O'Connor et al., [5] has developed NN methods to predict stock index movements using external futures which estimate the efficiency of movements using external indicators such as currency trading, commodity exchanges for predicting activities in Dow Jones industrial average index and involves the use of trading simulations to assess the experimental value of predictive model. The neural network model gives the strategy for the investor to take the direction as buy or hold.

M.Mozaffari Legha et al., [6] has proposed a new intelligent method based on NN to predict stock prices of Tehran stock market indexes which contains two steps. The first step is feature selection for best input data and the second step is to do improved hybrid NN classification algorithm to predict the next business day's high values of stock indexes.

Osman Hegazy et al., [7] have proposed a machine learning model for daily stock index prediction. The author has proposed integrated model for predicting daily stock prices. The proposed model integrates particle swarm optimization and Least Square Support Vector Machine. This model uses five different technical indicators to do the calculation. Due to integrated approach the proposed model behaves better than existing methods.

Meysam et al., [8] proposed group method of data handling for stock market prediction of petrochemical industry in Iran's Tahrn stock exchange. The author has compared the predicted values against experimental values and the result of NN is considered excellent in stock price prediction. Md

Rafial Hassan et al., [9] developed integrated model to predict financial market behavior using Hidden Markov Model, ANN and GA and the purpose of the model is to do detailed investigation of stock indexes using ANN for daily stock prices.

Aiken and Bsat [10] have developed a predictive model on market share price using NN with the given input variables of the share market is also incorporated for analysis. ANN can remember data for any number of years, which can be used for training network and thus predicting the future based on past data and the proposed method makes use of FF architecture for prediction. The authors used FF NN trained by a genetic algorithm to estimate three-month U.S. Treasury Bill rates and the author concluded that the NN can be used to accurately predict these rates.

Birgul Egeli [11] has proposed for predicting the Istanbul stock exchange index value using ANN. The inputs to the system includes previous days index value, previous days TL/US exchange rate, previous day's overnight interest rate

and five dummy variables each representing the working day of the week. The network architecture includes MLP and Generalized FF NN. Training and Testing is performed with these two network architectures. Results were compared with moving averages where ANN proved to be better in performance. From the related work, it was found that researchers have worked with integrated method for forecasting stock market index. The current work has also adopted an integrated method for maximizing the value of ROI. The computational intelligence approaches for stock price forecasting using NN and neuro-fuzzy approach to find the accurate prediction on individual investor decision based on Taiwan Stock Exchange and the experimental result showed that the hybrid CI approaches are recommended for stock price prediction with the lowest test error of standard statistical parameter [12].

## III. METHODOLOGY

The Figure1 shows the different steps involved in predictive analytics which performs all the procedures stepwise. The proposed model was trained for top twenty products of stocks and commodities. The data sets used for this research was segmented into 60% and 40% for training and testing the model.

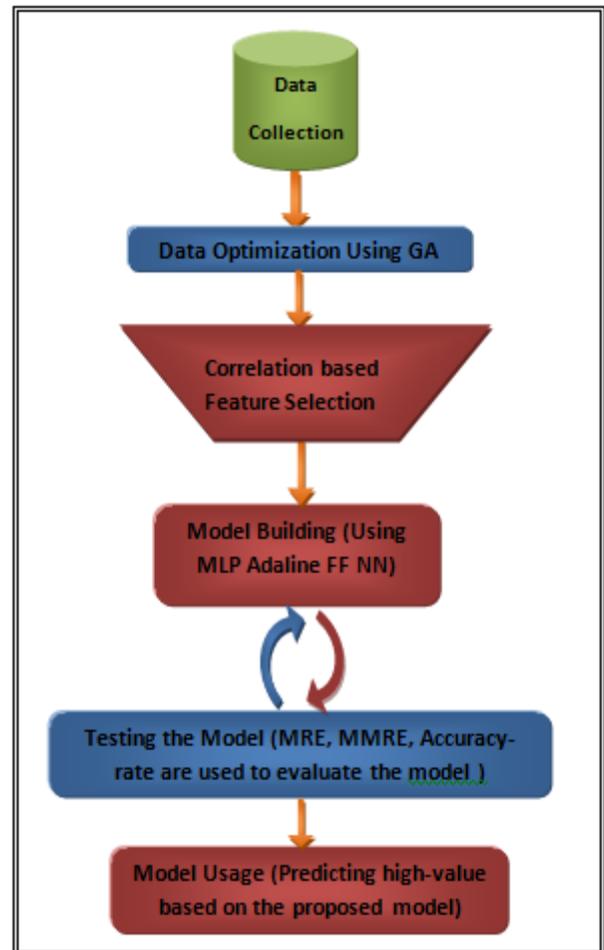


Figure 1. Proposed frame work for predictive modeling

The integrated method uses more than one computational method, as the first step, the data was captured from Asia’s fastest exchange market, the popular Indian websites MCX and NSE [12, 13]. The NSE and MCX are popular stock market exchanges and famous commodity exchanges in India. The MCX offers futures trading in bullion, ferrous metals, energy and a number of agricultural commodities like cardamom, potatoes, palm oil, metal, fiber and others. The NSE index contains many stock prices like IT, CG TECH, FMCG, PSU, METAL, OILGAS and BANKEX etc. are used for the experiment. The data set contains more than thirteen attributes such as item-no, item-name, expiry-month(dd-mm-yy), open(Rs),high(Rs),low(Rs), ltp(Rs), pcp(Rs), change(%), buyqty, sellqty, sellprice(Rs).The historical data of NSE for stock market indexes and MCX for commodities are used to build the predictor model with higher accuracy.

Once the entire data is selected, it uses the GA to check whether the selected data is suitable for further analysis. The purpose of GA is to verify the collected data is optimized or not. In this research work, out of all GA operations, the chromosome is generated randomly at each iteration with fitness function values. The selection criteria involved in this experiment is based on fitness function values ranges from -1 to +1 then the data is considered for further processing otherwise the given data is rejected for data analysis. Once the data is verified for optimization using GA, as the second step of the proposed framework is feature selection using PCC. To reduce the computational complexity and to increase the predictor accuracy, the highly correlated variables are used for building the MLP Adaline FF NN classifier. Once the NN model is built, the hidden values are used to calculate the final result. The 80% of the data is used for building MLP Adaline NN classification and 20% are considered to check the accuracy of the model. The model is trained continuously for three months data to evaluate the patterns. The accuracy-rate is considered as one of the major parameters for checking prediction.

A. An Integrated Approach to PA

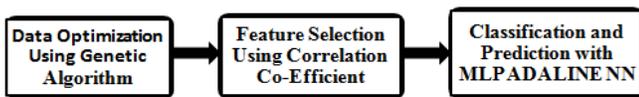


Figure 2. Overview of Integrated Approach

Many researchers use one specific method to evaluate the performance with other method. The current research works deploy GA, PCC and MLP ADALINE FF NN is integrated to optimize predictive analytics. The ADALINE NN uses adaptive linear regression to form the adaptive neuron. The future values of commodity trading and stock indexes of financial services are predicted using the proposed model. The purpose of the research is to find a good predictive model for any product under financial services. Due to the prevalence of high volatility, it is not easy to do the prediction on stock indexes and commodities. To avoid the complexity and risk, the integrated approach is used to investigate with MCX and NSE India data of every day closing prices of high values and

it is checked with the next business day closing price of high values of commodities and stock indexes.

B. Data Optimization Using GA

The data set used for data analysis in prediction needs to be preprocessed because the data is considered as a main entity. The main aim of GA is to perform the optimality check for the data set which is to be used in predictive analysis. In each iteration, the value of fitness and functional aspects for each entity is calculated. Generally, the GA stops once it reaches the highest amount of generations. There are different genetic operations that are to be used for checking the optimality of the data set used for this experiment. The different operations involved in GA are initialization, chromosome representation, mutation, crossover, termination and fitness function. Initialization is done at the beginning; all individual results are randomly generated to form initial population. The volume of the population depends on the nature of the problem. Usually, the population is generated randomly with all possible combinations.

$$f(x) = \log \left( 1 - \frac{N_p}{N_c} \right), k = 1,2,3, \dots, m \text{ ----- (1)}$$

In equation (1), were  $N_c$  is the measure that counts that occurs and  $N_p$  is the sum of individual values obtained. The fitness function  $f(x)$  is passed to the next level. The new chromosomes are organized on the basis of fitness function. The initial population is created randomly, which represents the genetic structure of an individual. The crossover is as in biological system, applicant solutions join to generate children in each iteration is called a generation. The fittest survive to become applicant results in the subsequent formation. For mutation, the newly created chromosomes are given from the crossover process. The purpose of doing mutation is to generate the new chromosome. The mutation is to do the arbitrary changes in single, extra digits in the gnome of representative object. The chromosome productions are repeated till its reaches stopping criteria. In general, the reproduction is stopped till it does satisfy the least criteria. The aim of the fitness function is to review the values of chromosome. To measure the given designed pattern, the fitness value is used to measure the given specified patterns [13]. The fitness function guarantees that the progress is towards optimization by measuring the robust assessment [14]. Among these operations of GA, the fitness function can be considered as one of the important functions in evaluating the data set. The fitness value always depends on the nature of the problem. The data optimization is done with the help of the tool “GA optimization for Excel” is used for this implementation [15] [16]. It is a free tool developed by Alexscherey for research purpose which allows the user to use only excel data. The constraint for selecting the data for data analysis depends on fitness function. If the fitness function lies between -1 to +1, then the selected data can be considered as an optimal data for predictive analytics. Once the data is selected from the MCXindia and NSEindia , the entire data is

incorporated into numerical values for creating the model. The data transformation is done with the help of excel and then loaded into GA tool to check whether the data is optimized or not. After loading the entire data into the 'GA Optimization for Excel tool', the results of default values for populations' parameter are shown in the Figure 3. The focus is to check the data file using fitness function

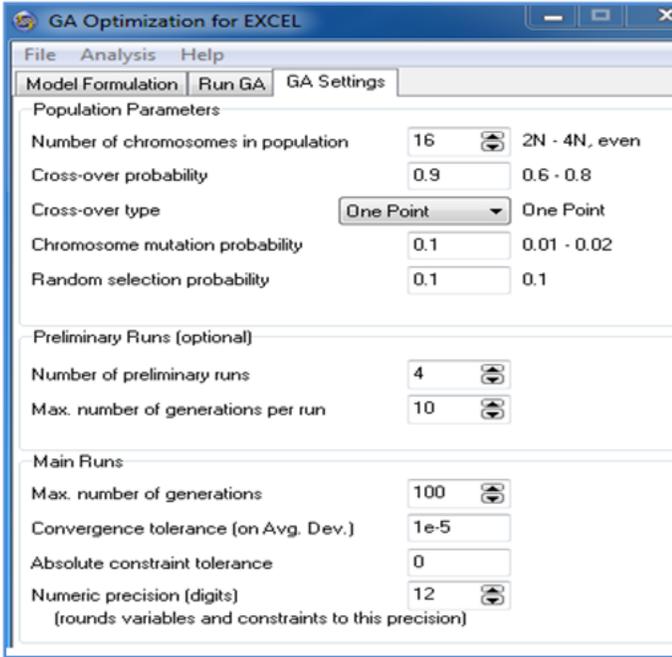


Figure 3 Data optimization using GA with default parameter



Figure 4. Experimental result on Data optimization Using GA(MCX data set)

The Figure 4 describes the result of GA for checking the data optimization. Based on the fitness values generated from various chromosomes, it is decided that the selected data is considered as an optimal data for predictive analytics. The same experiment is repeated for all data sets. Finally, selection of the data set is considered for further analysis only if it satisfies necessary conditions. The identified data is used for further analysis when the value of  $f(x)$  lies between 1 and -1. By this experimental result, it is found that the GA can be considered as an optimization tool for identifying right data for data analysis.

*C.Feature Selection Using Pearson's Correlation Coefficient*

The feature selection is a most challenging problem faced by real life applications for computational intelligence. Good predictor algorithms (like neural network, data mining and machine learning) need tools for feature selection algorithm that integrates learning system which suffers from high computational complexity. There are multiple benefits while applying this method. It improves its general efficiency, accuracy, speeding up the training process and in reducing the computational complexity of the data model making it more comprehensible. Real world problems may have several irrelevant features and such feature selection provides high reduction rate preserving important information in the full data set. To face these requirements, the PCC is used to reject irrelevant features. Here, the work is mainly concentrated on effective removals and it is implemented based on PCC which works better than other correlation coefficient and is verified with the help of P-value measures in statistics. The PCC is a statistical measure of the strength of linear association between paired data. To categorize the variable as positively correlated, negatively correlated and no correlation, the following are the constraints used [17]

- i) The values of one variable increases the values of another variable when the variables are positively correlated,
- ii) If it is negatively correlated, then increasing the values of one variable decreases the values of another variable and
- iii) If there is no change in two variables values such as increase or decrease of one variable will not affect the values of another variable.

TABLE 1 EXPERIMENTAL RESULT ON CORRELATION SUMMARIES (PEARSON'S)

PCC	BuyPrices	BuyQty	Change	High.Rs.	Low.Rs.	LTP.Rs.	Open.Rs.	PCP.Rs.	SellPrice	sellQty
BuyPrice.Rs.	1	-0.06	-0.13	1	1	1	1	1	1	-0.02
BuyQty	-0.06	1	0.24	-0.06	-0.06	-0.06	-0.06	-0.06	-0.07	0.32
Change...	-0.13	0.24	-0.13	0.24	1	-0.13	-0.13	-0.13	-0.14	-0.16
High.Rs.	1	-0.06	-0.13	1	1	1	1	1	1	-0.02
Low.Rs.	1	-0.06	-0.13	1	1	1	1	1	1	-0.02
LTP.Rs.	1	-0.06	-0.13	1	1	1	1	1	1	-0.02
Open.Rs.	1	-0.06	-0.13	1	1	1	1	1	1	-0.02
PCP.Rs.	1	-0.07	-0.13	1	1	1	1	1	1	-0.02
SellPrice.Rs.	1	0.32	-0.14	1	1	1	1	1	1	-0.02
SellQty	-0.02	0.32	-0.16	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	1

Table 1 describes Pearson’s correlation matrix for all the attributes in the given data set. From table 1 it is understood that ‘High,Low,LTP,Open, PCP,Sellprice’ are considered as highly correlated variables for building MLP classifier. The same data set is checked with other correlation methods like partial correlation, spearman rank order are used for feature selection but the PCC with MLP gives higher accuracy than other methods. The variables in the optimized data set are open, close, high, low and previous values are considered as highly correlated variables for building NN Adaline FF model for the prediction. The PCC is used for extracting the feature of all data sets for the MLP ADALINE FF NN classifier.

**D.MLP ADALINE FF NN**

Rosen blatt (1958) proposed the perceptron models where the weights are adjustable by the learning perceptron rule. Widrow and Hoft (1969) proposed an ADALINE model for computing elements and LMS (Least Mean Square) learning algorithm to adjust the weights of a learning algorithm to adjust the weights of an ADALINE model. Hopfield (1982) has found the energy analysis of feedback network provided the network has symmetrical weights [19].

Rumelhart et al., (1986) described that it is possible to adjust the weights of MLP ADALINE FF in a systematic way to implement mapping in correlation input and output patterns. The main characteristics of the ANN are to exhibit the mapping capabilities to map the input patterns to their associated output pattern [20].

The ADALINE FF NN is considered as a most accurate computational model that consist of a number of input unit that communicate by sending signals to one another over large number of weighted connections. These NN originally developed from the inspiration of human brain, consists of processing units called artificial neurons and the connections are called as weights between them. The important feature of ADALINE network uses learning-by-example as an adaptive nature. The MLP ADALINE FF NN consists of input, hidden and output layers of neurons. The aim of the ADALINE FF NN is to build the classifier model for stock market and commodity trading.

The FF NN calculates the error rate as  $d_i - y_i$  where  $d_i$  is the measured output and  $y_i$  is the real output produced by the MLP adaline FF NN based result on input  $x_i$ . The CI has been used when complete domain or incomplete information of the problem prevails, which can be solved where the training model is available readily. The NN model is based on the biological neuron or cognitive science. The ADALINE FF NN method is a powerful tool for pattern recognition and classification based on learning from experience. The MLP ADALINE FF NN is trained, since learning by example is used to make an inference which further finds unknown instances. This ADALINE FF NN is capable of predicting new results or show outcomes from past events.

**IV. US EXPERIMENTAL RESULT**

The current experimental work focuses on finding accurate prediction. The accuracy is measured with MRE, MMRE, and accuracy-rate. The experiment predicted high values which are also checked with the next business day’s high values. The

experimental work resulted in an average accuracy of 95% and above. The predicted values were checked with real time data of next day high values. The Table 2 describes the comparative result on MCX data set with few commodities like alumni, aluminum, three types of cardamom products, two types of copper and copper products are considered for building the MLP NN model. The hidden layers and the input variables, low, high, open, current, previous values of the date 23-Feb-15 are used for the calculation. The predicted values are compared with the next business day (ie 24-Feb-2015) original values which are already shown in the Table 3. The performance evaluation is done with the evaluation parameter MRE, error and accuracy rate with the actual difference. From Table 1, it is found that the predicted high values of different commodities are almost equal to the original high values of the next business day. The measured accuracy rate for the given experiment is 98.337 and all MRE values are also equal to zero. The expected accuracy rate for proposed model is greater than 90%. The model is tested continuously for three months. The same model is implemented with NSE data set for stock market indexes. So it is observed that all experimental result for commodities and stock indexes gives an average accuracy of 95%. The predicted values of commodities and stock market indexes are very close to original values of the next business day.

TABLE 2 COMPARATIVE RESULT FOR MCX DATA

Item No	High(Rs)-original	Predict High(Rs)
1	114.45	116.3146207
2	114.05	116.1646161
3	113.5	116.2662867
4	112.55	116.5486838
5	113.55	115.3864472
6	113.5	116.4852521
7	112.6	115.3439898
8	1100.1	1101.203716
9	1120	1126.109981
10	1129	1128.319954
11	1126.9	1135.386375
12	1146.4	1158.645407
13	374	374.9793843
14	369.95	371.4053639
15	366.85	00366.87984

(As on 23-02-2015 and predicted value for 24-02-2015 with accuracy rate = 98.337)

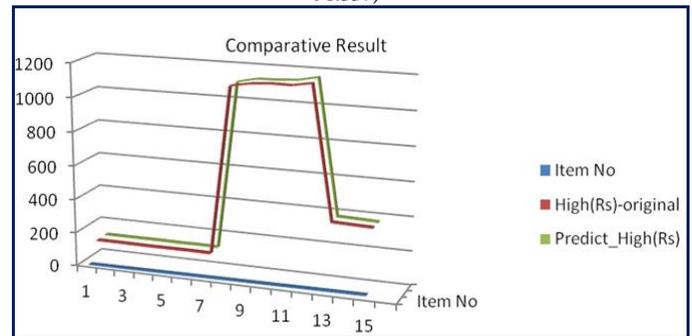


Figure 5 Performance Graph on MCX data (As on 24-02-15)

The Figure 5 shows the performance graph on MCX data. The comparative result shows the difference between actual values and experimental values. The comparison values are nearly equal to the original values. The Table 2 describes Serial number of various stock indexes, original values and predicted values of NSE stocks as on 07-03-2015 and the predicted value as stocks are compared with original high values of the next business day ( as on 09-03-2015). The measured accuracy rate for the given experiment is 99.898 and MMRE value is equal to 0.1015967.

TABLE 3 COMPARATIVE RESULT FOR NSE DATA

SL No	Original_High_values	Predicted_High_values
1	8,848.20	8956.544
2	19,751.84	19950.652
3	4,021.61	4140.134894
4	17.92195219	16.93193612
5	8,842.72	8945.8625
6	7,225.27	7292.554
7	13,179.87	13307.42
8	3,443.28	3500.075
9	5,731.30	5772.948
10	4,573.51	4616.3825

(As on 07-03-2015 and predicted value for 09-03-2015 with accuracy rate = 99.898)

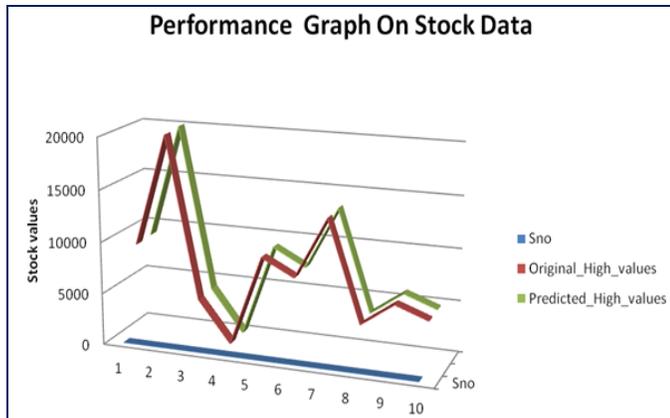


Figure 6. Performance Graph on NSE data (As on 07-03-15)

The Figure 6 shows the performance graph on NSE stocks. The comparative result shows the different between the actual values and the experimental values. The comparison values of stock indexes are very near to original values.

V. CONCLUSION

Given the predictions of the integrated model, an experimental investigation was conducted to compare the values of original values with the predicted high values for the next business day. The result of the implemented model gives close agreement between experimental values and the original

values of stock prices and commodity indexes. Financial services are considered as most challenging problems in the investment field. The purpose of the research is to avoid the investment risk. This paper, proposed a computational intelligence model that integrates GA, PCC and MLP ADALINE FF NN for stock price prediction and commodity trading. The strength of the model is to predict accurate prediction on stock market indexes and commodity exchanges. Based on this integrated model, the decision support system tools can be developed to perform all kinds of future prediction where all the inputs are considered as numeric variable. The same model can also be implemented to diagnose the diseases in health care. As a future work, the experimental work can be extended to predict other kinds of financial series.

REFERENCES

- [1] Md Rahat Hossain, Amanullah Maung Than Oo, A. B. M. Shawkat Ali ,Hybrid Prediction Method for Solar Power Using Different Computational Intelligence Algorithms, Smart Grid and Renewable Energy, 4, 76-87, February - 2013,
- [2] Olivier C., Blaise, Pascal University; Neural network modeling for stock movement prediction, *state of art*. 2007.
- [3] Abhishek, Kumar, Anshul Khairwa, Tej Pratap, and Surya Prakash. A stock market prediction model using Artificial Neural Network. *In Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference on IEEE*, PP: 1-5, 2012.
- [4] Niall O'Connor and Michael G. Madden, A neural network approach to predicting stock exchange movements using external factors. *Knowledge-Based Systems*, No: 19, Issue No-5, 371–378, September 2006.
- [5] M.Mozaffari Legha, F.Keynia, M.Mozaffari Legha, A new intelligent method based on NN for stock price index prediction, *International Journal on Technical and Physical Problems of Engineering(IJTPE)*, Iussue-19,Volume-6,No-2,PP:24-30,June-2014.
- [6] Abdul Salam, A machine leaning model for Stock Market Prediction, *International journal of computer science and Telecommunication[IJCST]*, Volume-4,issue 12, December-2013.
- [7] Meysam Shaverdi, Saeed Falahi, Vahhab Bashiri Prediction Stock price of Iranian Petro Chemical Industry Using GMDH-Type Neural Network and Genetic algorithm, *Applied mathematical sciences*, Vol-6,no-7,PP: 319-332,2012.
- [8] Md.Rafial Hussan, Bai Kunth Nath, Michael Kirley, A fusion model of HMM, ANN and GA for Stock market forecasting, *Journal Systems with Applications: An international Journal*, Volume -33, issue-1,pages 171-180,July,2007.
- [9] Aiken and Bsat, Forecasting market trends with neural networks. *Information Systems Management*, Volume No-16, Issue No-4, 1999.
- [10] Birgul Egeli, "Stock Market Prediction Using Artificial Neural Networks."*Decision Support Systems* 22: 171-185.
- [11] Jui-Yu Wu,Chi-Jie Lu ,Computational Intelligence Approaches for Stock Price Forecasting, *IEEE International Symposium on Computer, Consumer and Control (IS3C)*, pp. 52 - 55, 2012, DOI: 10.1109/IS3C.2012.23.

- [12] Azzini.A, A New Genetic Approach for Neural Network Design. *Ph.D. Thesis*, 2007.
- [13] Azzini.A., Tettamanzi.A, A neural evolutionary approach to financial modeling. *proceeding of soft the Genetic and Evolutionary Computation Conference, GECCO- 2006*, volume No. 2, pp. 1605–1612.
- [14] <http://www.alexschreyer.net/projects/xloptim>.
- [15] [http://www.nseindia.com/live\\_market/dynaContent](http://www.nseindia.com/live_market/dynaContent) and <http://www.mcxindia.com>.
- [16] Mark A. Hall, *PhD Thesis*, submitted to Waikato University, New Zealand, Correlation-based Feature Selection for Machine Learning, April 1999.
- [17] Rosenblatt, Frank. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, No. 6, PP. 386–408.
- [18] Mehrotra K., Mohan C.K., Ranka S. Elements of artificial neural networks, *MIT Press*, 1997, ISBN: 0262133288.
- [19] Rumelhart,D.E.,Hinton G.E., & Williams, R.J, Learning representations by back propagation errors, *Nature*,vol No-323. PP No: 533-536.

AUTHORS PROFILE



**Chandra.J** is an Associate Professor in Computer Science Department, Christ University holds a Masters in Computer Applications from Bharathidasan University. She has MPhil in Computer Science from Vinayaka Mission University and doing PhD in Hindustan University. Her research interests include Artificial Neural Network, Data Mining, Genetic algorithm, big data analytics and Predictive analytics.



**Dr Nachamai M.** is an Associate Professor in Computer Science Department, Christ University holds a Masters in Computer Applications from the University of Madras. She has both an MPhil and PhD in Computer Science from Bharathidasan University and Mother Teresa Women’s University respectively. Her research interests include Perception, Computer Vision, Image Processing, Digital Signal Processing and Artificial Intelligence.



**Dr Anitha S Pillai** is working as a Professor and Head in department of MCA, Hindustan University, Chennai, Tamil Nadu, India. Her research interests include Software Engineering, Data Mining, Natural Language Processing, Soft Computing and Artificial Intelligence.