International Journal of Engineering & Technology, 7 (4.28) (2018) 710 -716

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Application of Genetic Algorithm to Predict Stock Price Index Using Artificial Neural Networks

M.Tawarish[1] Dr. K.Satyanarayana[2]

[1]Research Scholar, Department of Computer Applications (Ph.D.),
Bharath Institute of Higher Education and Research (BIHER),
(Declared as Deemed-to-be-University under Section 3 of UGC Act 1956)
 # 173 Agharam Road, Selaiyur, Chennai - 600 073, Tamil Nadu, India.
[2]Research Supervisor / Principal, Sindhi College of Arts and Science,
Chennai – 600 077, Tamil Nadu, India

Abstract

Investment is the vital part of our country economy. The prediction of price index is very essential for stockholders to earn the highest profit from their investment. The changes of stock market rage by several factors such as political, economic and public factors, using classic strategies for stock market prediction lead to precise results. Since, the intelligent techniques have this capability that contemplate the complicated effects of the factors within the analysis,thus we cameto use them in stock index prediction. Therefore, using genetic algorithm for predicting and optimizing the input variable in the neural network. The study of this research in artificial neural network and genetic algorithm for predicting the stock price for National Stock Exchange. For this cause, the economic data such as open market price of Forex, CPCL stock price and the price as inputs to compound model and total index that compared with outputs in compound model.The prediction may giveenhanced result.

Keywords: *Artificial Neural Networks, Genetic Algorithm, Price Index, Stock Market, Compound Model as Hybrid-ARIMA.*

1.Introduction

The Prediction of Price index on stock market will be an interesting part of research making researchers in the area field always desiring to improve existing predictive models. The reason is that organizations and individuals are empowered to make investment decisions and talent to plan and develop effective strategy about their daily and future events. Stock price prediction is regarded as one of most difficult work to accomplish in economic forecasting due to complex nature of stock market. The need of many investors is to put hold of any forecasting method that could assure easy profiting and minimize investment risk from the stock market. This remains an inspiring factor for researchers to change and evolve new predictive models.

The artificial neural networks (ANNs) model by approaching genetic algorithm, which are very popular due to its ability to learn patterns from data and assume solution from unknown data. Hardly, many related works that engaged ANNs model to stock price prediction. Currently, hybrid model [1] approaches has also been engaged to improve stock price predictive models by exploiting the unique strength of each of them.

The stock price prediction by hybrid models are used by correlation and RMSE. ARIMA models are from statistical models perspectives. ARIMA [2] models are known to be tough and efficient in financial predictions especially short-term prediction than even the most popular ANNs techniques and GeneticAlgorithm.

It has been broadly used in field of economics and finance. Other statistics models are regression method, Bayesian method and Chi-Square test,related works that has engaged ARIMA model for price prediction.Some impact factors that affect stock price are company news and act, industry performance, investor sentiment and economic factors. Hence, we have introduced the above models to predict an efficient result by the application of genetic algorithm by using the artificialneural networks.

2. Artificial Neural Networks

Artificial neural networks [3] is a wide range of existing practical observations and commonly used heuristics. Artificial Neural Networks are viewed here as parallel computational model. With varying degrees of complexity, comprised if closely interconnected adaptive processing units. These networks are very well grained parallel implementations of nonlinear static or dynamic systems. A very significant feature of these networks is their adaptive nature, where “learning by example” replaces traditional” programming in solving problems. This feature makes such computational models very appealing in application domains where one has little or incomplete understanding of the problem to be solved but where training data is readily available. Another key feature is the basic parallelism that allows for fast computations of solutions when these networks are implemented on parallel digital computers or ultimately, when implemented in modified hardware.

Artificial neural networks are possible computational models for a wide selection of problems including pattern classifications, speech fusion and detection, adaptive interfaces between humans and complex physical systems, purposeestimate image data compression, associative memory, clustering, forecasting and prediction, combinatorial optimization, nonlinear system modeling and control.

In fact the most of the network models covered in this volume are more closely related to usualarithmetical and or statistical models such as perfect algorithms, nonparametric pattern classifiers, clustering algorithms linear and nonlinear filters and statistical regression models than they are to neurobiological nodes.
The theories and techniques of artificial neural networks workedthere by fairly mathematical although the level of mathematical force is relatively low. In my exposition I have used mathematical to provide imminent and understanding rather than a establish exact mathematical basics.

3. Genetic algorithm

A genetic algorithm is a heuristic seektechnique used in artificial intelligence and computing. It is used for finding perfect solutions to search problems based on the theory of natural selection [4].
Genetic algorithms are excellent for searching through huge and difficult data sets. They are considered capable of finding realistic solutions to difficult issues as they are highly capable of solving free and forced optimization issues.
Genetic algorithms are widely used in many fields such as robotics, self-propelled design, optimized broadcastings routing, engineering design and computer-aided molecular design.
Genetic algorithms are machine learning search techniques inspired by Darwinian Relevance models. The benefit of genetic algorithm over factor analytic and other such statistical models is that Genetic Algorithm models can address problems for which there is no human expertise or where the problem pursuing a solution is too difficult for expertise based approaches. Genetic Algorithm can be applied to challenges which can be formulated as function optimization problems.
This makes genetic algorithm ideal for application to discrete combinatorial problems and mixed problems. Thus the genetic algorithm approach is appropriate for finding solutions that require efficient searching of a subset of features to find combinations that are near optimal for solving high-structural classification problems, particularly when the search space is huge, difficult or poorly understood.
Impacts' of Genetic Algorithm [4]:

- An effective hybrid model genetic algorithm was developed inwhich we predict solution forindex into different prices and each be the stock issues, this assure the range of the algorithm.
- Plan the operators so as to search only the feasible space; thus, we save figuring time by avoiding infeasible space.
- Once perfectly tuning the algorithm, it is compared with different techniques used in the literature.

4. Literature Survey

A lot of researchers have been worked on stock market prediction. All researches have different feature of selecting the models using different computing techniques. Some of the models are in similar kind has been discussed.Most of the elegant computing techniques [5] are developed over the last two decades. This component describes briefly some of the work that has previously been done in the turf of stock price prediction. The stock market prediction is a very complicated vital system.
The using of Statistical models have been taken to prediction of price index. There are difficulty in the modeling, problems of traditional statistical economics methods are rigidity in dynamic circumstances and change of stock price is feasibly a substandard economic factor.
The predicting of stock price has been one of the main challenge to the artificial intelligence area due to complex nature of stock markets according to Naeini et al., (2010). Stock markets are draw by Anish and Majhi (2015) are a difficult, evolutionary and nonlinear dynamic system whose prediction is considered a challenging task. Chai et al., (2015) is of the view that predicting the stock markets is amongst the most complicated and difficult tasks owing its movement to being affected by a multiplicity of factors such as government policy, investor's expectations, global economic situations and correlations with other markets. [6]
Prove of such challenges to stock price prediction, Taran et al., (2015) propose various methods that have been applied to prediction of stock markets ranging from time series forecasting, numerical analysis, basic analysis and technical analysis to technological analysis. Hard work to come up with prediction model has been current with movements form statistical approaches such as Autoregressive Integrated Moving Averages (ARIMA) models.
The day2kgreat researcher K. J. Kim et al estimated new hybrid model of Artificial Neural Network and Genetic Algorithm for attribute discretization. Attribute discretization concepts to convert continuous data into discrete data using certain threshold. Properly discretized data makes of the issue ijpam.eu 357 learning simple and may be develop the quality of the learned results. [7]
Debadrita Banerjee's[8] Stock Prediction Model is based on the statistical approach that utilizes Autoregressive Integrated Moving Average (ARIMA). The author has analyzed month-wise data that has been obtained from the S&P Bombay Stock Exchange Sensitive Index. The most suitable ARIMA model is shortlisted by trying out different combinations of the number of auto-regressive terms, non-seasonal differences and lagged predict errors in the prediction equation.
By Artificial Neural network approach [9]O. Abdalla, A. Elfaki, and Y. Murtadha have introduce a new line for the optimization of ANN parameters [2]. Their method introduces a process of training the ANN which is efficient and also less humandependent. Based on their survey of the related work in this field, they have processed ANN designing problems with the help of a Genetic Algorithm. 5.

5.Methodology Used

The Methodology used for this proposed system is prediction of stock price index is shown in Figure1. So, we have proposed that AGPSISA [10]. It consists of three main elements: Extraction, Hybrid Model-Correlation (covariance)-RMSE and ARIMA Models.

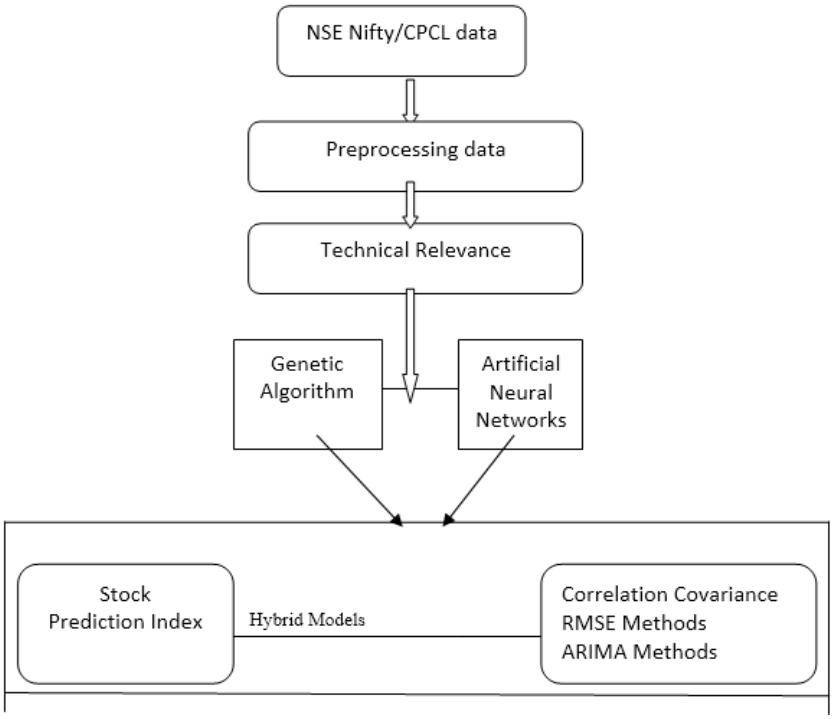


Figure 1: Framework of the proposed prediction of stock price index

6.Extraction

This extraction is the first phase of the Stock price index. The Data Set is collected from the National Stock Exchange, CPCL. There are resources that can be used to share information acquisition. This paper describes the application of Artificial Neural Network (ANN) for the prediction of Stock Market using Genetic Algorithms. A new model is proposes of ANN for feature Extraction and selection to get more accurate prediction of stock market. Therefore, the extraction techniques to be able to create a more accurate prediction model to build detailed features.

National Stock Exchange (India): Data Sets for NIFTY 50 and CPCL stock price.
Historical index data for NIFTY 50 for the period October [11].
The Price Index Changes in vitality.

Daily	Price	Open	High	Low	Vol:	Change
01-Oct-2018	10930.90	11035.65	10821.55	11008.30	398987274	20391.72
03-Oct-2018	10982.70	10989.05	10843.75	10858.25	398756507	21225.59
04-Oct-2018	10754.70	10754.70	10547.25	10599.25	438202008	23711.57
05-Oct-2018	10514.10	10540.65	10261.90	10316.45	625153832	25254.21
08-Oct-2018	10310.15	10398.35	10198.40	10348.05	470279031	22130.94
09-Oct-2018	10390.30	10397.60	10279.35	10301.05	443795275	18285.41
10-Oct-2018	10331.85	10482.35	10318.25	10460.10	373844130	19592.59
11-Oct-2018	10169.80	10335.95	10138.60	10234.65	498509417	21907.13
12-Oct-2018	10331.55	10492.45	10322.15	10472.50	354056762	18606.04
15-Oct-2018	10524.20	10526.30	10410.15	10512.50	267517728	14372.15
16-Oct-2018	10550.15	10604.90	10525.30	10584.75	273444095	15453.03
17-Oct-2018	10688.70	10710.15	10436.45	10453.05	293586908	17173.14
19-Oct-2018	10339.70	10380.10	10249.60	10303.55	368552204	20505.66
22-Oct-2018	10405.85	10408.55	10224.00	10245.25	306472776	17241.94
23-Oct-2018	10152.60	10222.10	10102.35	10146.80	312042090	18114.71
24-Oct-2018	10278.15	10290.65	10126.70	10224.75	362272428	19413.24
25-Oct-2018	10135.05	10166.60	10079.30	10124.90	522421914	23503.76

CPCL historical data (data set) for stock price index for the moth October 2018[12]

Date	Open	High	Low	Close
01/10/18	279	283	270.3	281.65
03/10/18	279	301	274	290.2
04/10/18	286	288	272	278.8
05/10/18	266.8	270.95	253.8	256.2
08/10/18	255.2	257.95	222.2	232.2
09/10/18	223	233.8	220.65	225.95
10/10/18	227	244.4	225.65	238.2
11/10/18	228	255	225	246.95
12/10/18	252.2	262.55	250.5	258.35
15/10/18	264	264	246.7	256.9
16/10/18	259	268.1	255.2	266
17/10/18	265.55	271.15	253.5	256.1
19/10/18	247	268	246.3	265.25
22/10/18	264	269.4	256.1	257.6
23/10/18	257	269	246.3	264.95
24/10/18	272.7	274.9	258.1	265.5
25/10/18	263.7	265.65	256	258
26/10/18	259.9	262	251.1	257.2
29/10/18	257.5	267.3	255.35	262.9
30/10/18	253	261.9	245.15	252.1
31/10/18	252.1	259.4	248.5	256.1

7. Technical Models

This technical model is delivered as the Second part of this Stock price index. In this model, the Intensities are calculated for the correlation, RMSE as Hybrid and ARIMA models. The intensity is calculated drawn as graphical by the two Models as Hybrid model namely.

* Correlation and Covariance methods.
* Root Mean Sum of Squares (RMSE) both of them as Hybrid Technique andARIMA technique

Correlation and Covariance methods

A evaluate used to represent how strongly two data sets are related known as correlation. Covariance is nothing but a measure of correlation. The reverse, correlation refers to the scaled form of covariance.

Calculation methods:

Initiate by selecting a time p over which you will calculate the correlation between the two stocks. Calculate Mean and Deviation. Calculate the average price for each stocks by totaling up daily prices and isolating the sum by the number of days. To estimate the Coefficient.

The result for covariance of a stock starts with finding a list of prior prices. It is shown as "historical data" on most quote pages. Typically, the closing price (stock) for each day is used to find list (Opening Stock) to begin the calculations [13].

$$correl(x,y) = \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}} * \sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n}}}$$

$$COV(X,Y) = \frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{n-1}$$

Where X and y are the opening stock values and closing stock values.

RMSE:

The Root Mean Square Error (**RMSE**) also known as, RMSD) is a frequently used measureof the difference between values predicted by a model and the values actually observed from the environment that is being modeled. The Prediction of stock based prices with opening and closing stocks. [14]

The RMSE of a model prediction with esteem to the estimated variable X_{model} is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

WhereX_{obs} is observed values and X_{model} is modeled values at of (time/place)i.
The observed as opening stock price data and modeled as closing stock price data to get the trend value.

7.Arima ModeL

Autoregressive Integrated Moving Average is the expansion of ARIMA model. **ARIMA** is a predicting technique that projects the future values of a series based entirely on its own indolence. Its main purpose is in the area of modest term forecasting (predict) need at least 40 historical data points.
This model consists of two parts, anautoregressive (AR) part and a Moving Average (MA) part. The model is usually referred to as the ARMA(p,q) model where p is the order of the autoregressive part and q is the order of the moving average part [15]

AR Model is predicting the price for the stock index.
(t) = A(1) * X(t-1) + E(t)
where X(t) = time series under prediction; A(1) = the autoregressive parameter of order 1;
X(t-1) = the time series lagged 1 period to predict, E(t) = the error term of the model.
Finally the subject of the ARIMA is
x(t) = A(1) * X(t-n) + A(2) * X(t-n) + E(t)
MA model is randomly predicts the prices of the index. With the help of AR Model. As given:
X(t) = -B(1) * E(t-1) + E(t)
B (1) is called the Moving Average of the I st Predicted indeed.
The use of ARIMA methods is to fit a stochastic model to a given set of time series data, such that the model can closely approximate the prediction that is truly generating the data. So by extracting the CHI –Square test [16], we can predict the approximation (Table A).

	ARIMA for chi-square		NSE	
	Opening Inde	Price	Closing Index	Price
01-Oct-18	10930.9	0.109309	11008.3	0.110083
03-Oct-18	10982.7	0.109827	10858.25	0.108583
04-Oct-18	10754.7	0.107547	10599.25	0.105993
05-Oct-18	10514.1	0.105141	10316.45	0.103165
08-Oct-18	10310.15	0.103102	10348.05	0.103481
09-Oct-18	10390.3	0.103903	10301.05	0.103011
10-Oct-18	10331.85	0.103319	10460.1	0.104601
11-Oct-18	10169.8	0.101698	10234.65	0.102347
12-Oct-18	10331.55	0.103316	10472.5	0.104725
15-Oct-18	10524.2	0.105242	10512.5	0.105125
16-Oct-18	10550.15	0.105502	10584.75	0.105848
17-Oct-18	10688.7	0.106887	10453.05	0.104531
19-Oct-18	10339.7	0.103397	10303.55	0.103036
22-Oct-18	10405.85	0.104059	10245.25	0.102453
23-Oct-18	10152.6	0.101526	10146.8	0.101468
24-Oct-18	10278.15	0.102782	10224.75	0.102248
25-Oct-18	10135.05	0.101351	10124.9	0.101249

(Table A)

After applying (extracting) by Chi test by ARIMA Table (B) as shown below, gives approximate predicting index data. [16]

	ARIMA for chi-square		NSE		
	Opening Inde	Price	Closing Index	Price	chi
01-Oct-18	10930.9	0.109309	11008.3	0.110083	0.544
03-Oct-18	10982.7	0.109827	10858.25	0.108583	1.426
04-Oct-18	10754.7	0.107547	10599.25	0.105993	2.28
05-Oct-18	10514.1	0.105141	10316.45	0.103165	3.787
08-Oct-18	10310.15	0.103102	10348.05	0.103481	0.139
09-Oct-18	10390.3	0.103903	10301.05	0.103011	0.773
10-Oct-18	10331.85	0.103319	10460.1	0.104601	1.572
11-Oct-18	10169.8	0.101698	10234.65	0.102347	0.411
12-Oct-18	10331.55	0.103316	10472.5	0.104725	1.897
15-Oct-18	10524.2	0.105242	10512.5	0.105125	0.013
16-Oct-18	10550.15	0.105502	10584.75	0.105848	0.113
17-Oct-18	10688.7	0.106887	10453.05	0.104531	5.312
19-Oct-18	10339.7	0.103397	10303.55	0.103036	0.127
22-Oct-18	10405.85	0.104059	10245.25	0.102453	2.517
23-Oct-18	10152.6	0.101526	10146.8	0.101468	0.003
24-Oct-18	10278.15	0.102782	10224.75	0.102248	0.279
25-Oct-18	10135.05	0.101351	10124.9	0.101249	0.01

(Table B)

8. Results

The fig (2) represents the results of data set taken from historical data on NSE from opening and closing Stock. The Correlations Covariance for the given data hybrid model is Correlation: 0.887646 and Covariance is 56196.06
Stock Price Index variation for NSE (Nifty) data set for the month of October 2018

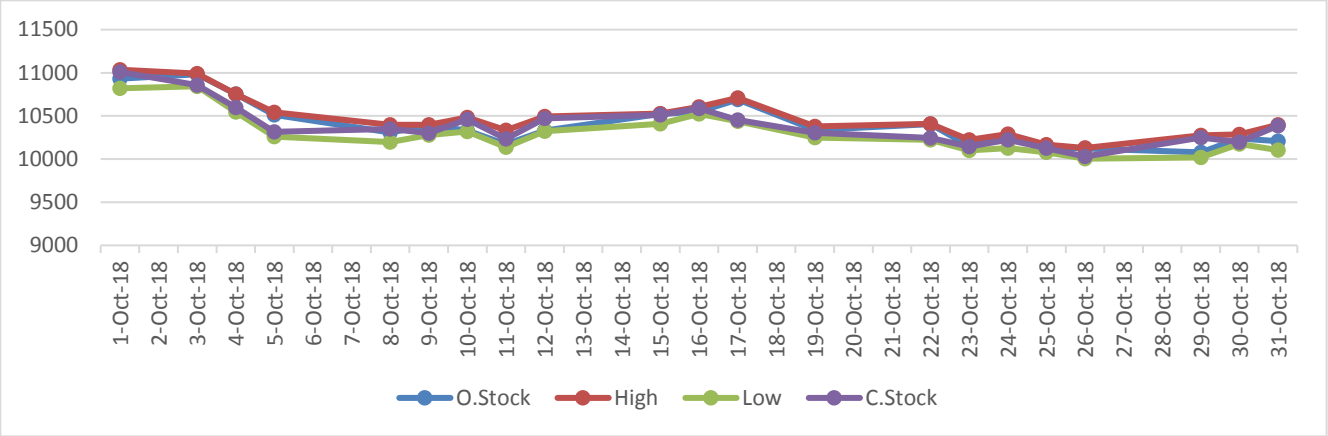


Fig (2)

The fig (3) represents the results of data set taken from historical data on CPCL from opening and closing Stock. The Correlations Covariance for the given data hybrid model is Correlation: 0.791517 and Covariance is 186.8746

The RMSE for the predicting index is 107.3749284 from the Opening Stock and Closing stock of NSE (NIFTY).

Stock Price Index variation for CPCL data set for the month of October 2018

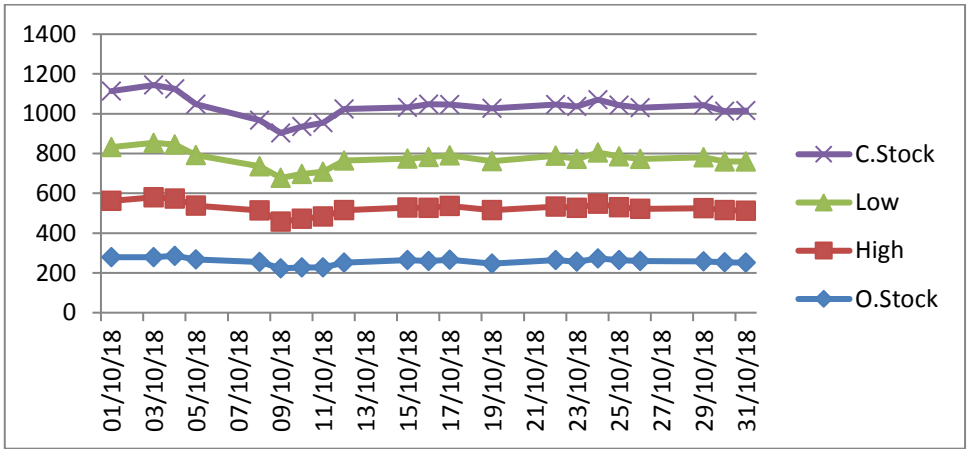


Fig (3)

The RMSE has be calculated for the prediction stock index by the data of NSE for the period October as shown in the below chart (Fig 4).

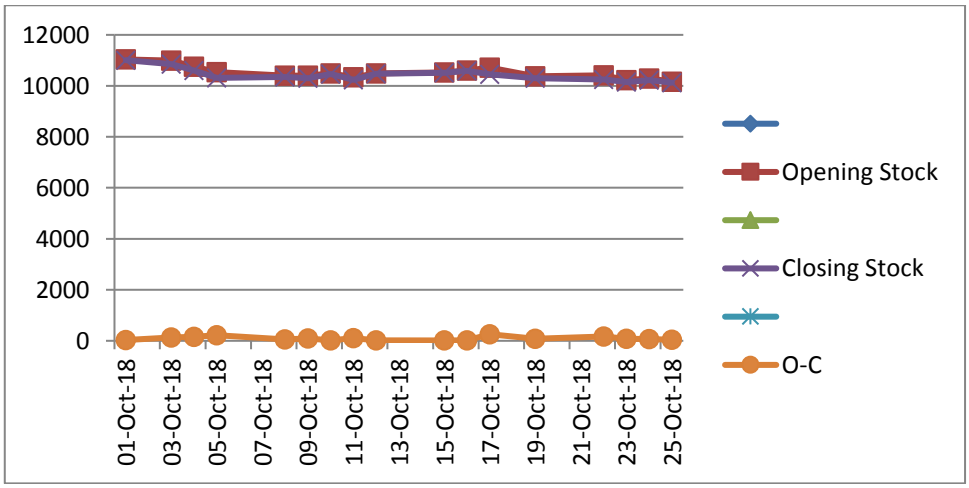
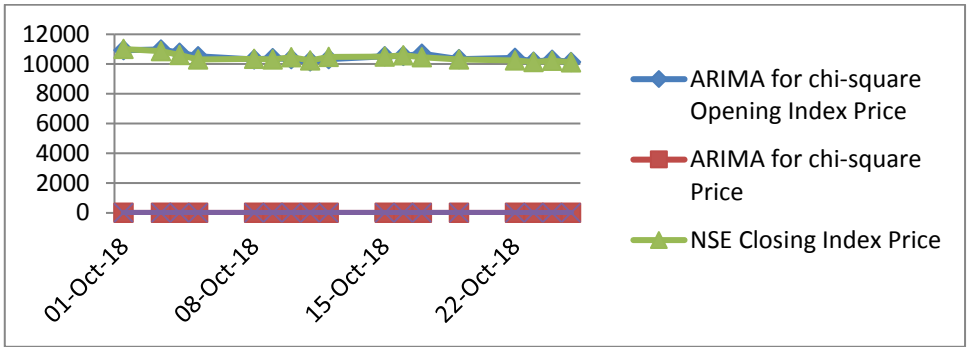


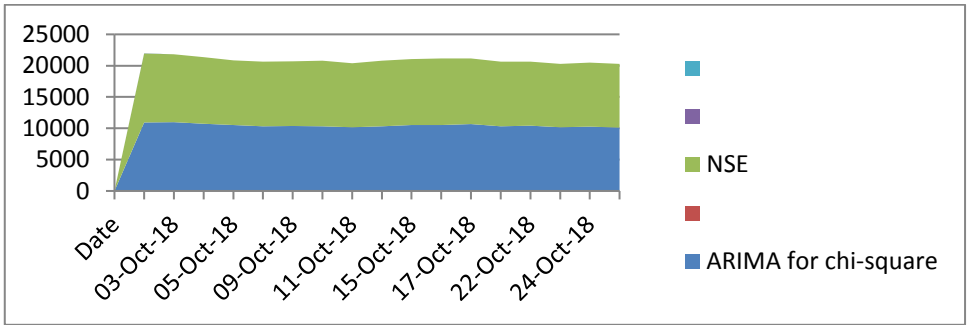
Fig (4)

The ARIMA Model data set(Fig 5) taken on by the moving average and auto regressive is by fluctuating the appropriate predicting results of the price index casted before applying the chi square.



(Fig 5)

The Result of the Line diagram showing (Fig 6) the prediction of approximate index after applying the chi-square extract model.



(Fig 6)

9.Conclusion

This research paper vitally studies the major determinants of stock market price index inIndia which has been in an unpredictable inclination in the current years.Stock markets seem to be verydifficult and unpredictable. Sowe have studied different methods for stock market prediction using Artificial Neural Networks and genetic Algorithm. Because of effectiveness and needs from the people, opinion predicting techniques became adynamic research area.As the size of the following facts increases, analyzing and shortening opinionated data is becoming more crucial. To satisfy these needs, various kinds of techniques are proposed. The approaches of compound models, statistics techniques ofCorrelation Co-variance,Root Mean Sum of square Error (RMSE) analysis, finally by Compound Models (CHI) - Hybrid-ARIMA Methods we have been predicted. But, we have to conclude that stock prediction is very complex task and various influence should be considered for predicating the market more perfectly and efficiently. In imminent work, we plan to determine the critical impacts of particular basic analysis variables on quality of stock price index prediction.

References

[1] J.J. Wang, J.Z. Wang, Z.G. Zhang and S.P. Goo, "Stock index forecasting based on a hybrid model", Omega vol.40 pp.758-766, 2012.

[2] P. Pay and C. Lin, "A hybrid ARIMA and support vector machines model in stock price prediction", Omega vol.33 pp. 497-505, 2005

[3] Fundamentals of Artificial Neural Networks By Mohamad H. Hassoun, Assistant Professor of Computer Engineering Mohamad H Hassoun

[4] https://www.techopedia.com/definition/17137/genetic-algorithm. Stock Trend Forecasting", Journal of Business Intelligence, vol. 3, no.2, pp. 23-43, 2M. Chisel, M. Bihari, and G.A.R Argali, "Improvement of AutoRegressive Integrated Moving Average models using Fuzzy logic and pp. 956-967, 2009. 010.

[5] 3.N. Mesh, V.P. Sabena, and K.R. Par Dasani, "A Comparison Between Hybrid Approaches of ANN and ARIMA For Indian

[6] A Review of Artificial Neural Networks Application to Stock Market Predictions Dennis Murekachiro1, 2 1.WITS Business School, University of Witwatersrand, No.2 St. David's Place,2193 Parkton, Johannesburg, South Africa 2. Bindura University of Science Education, P Bag 1020, Bindura, Zimbabwe

[7] A HYBRID FORECASTING MODEL FOR STOCK VALUE PREDICTION USING SOFT COMPUTING TECHNIQUE Mohammed Siddique 1, DebdulalPanda 2, Sumanjit Das 3, Sudhir Kumar Mohapatra 4

[8] D. Banerjee, "Forecasting of Indian Stock Market Using Time-Series ARIMA Model", In: Proceedings of the 2nd International Conference on Business and Information Management, IEEE, 2014, pp.131–135.

[9] Survey: Stock Market Prediction Using Statistical Computational Methodologies and Artificial Neural Networks A. Rao1, S. Hule2, H. Shaikh3, E. Nirwan4, Prof. P. M. Daflapurkar5 .

[10] Hybrid Model for Stock Trading System M. J. Soumiya* and N. Sairam Indian Journal of Science and Technology

[11] https://www.nseindia.com/products/content/equities/indices/historical_index_data.html.

[12] https://in.investing.com/equities/chennai-petroleum-corporation-historical-data

[13] https://www.investopedia.com/articles/financial-theory/11/calculating-covariance.asp Calculating covariance for stock By Peter Cherewyk | Updated April 24, 2018 — 1:00 PM

[14] International Journal of Computer Applications (0975 - 8887) Volume 129 - No.11, November 2015 Predicting Closing Stock Price using Artificial Neural Network andAdaptive Neuro Fuzzy Inference System (ANFIS): The Case of the Dhaka Stock Exchange.

[15] https://www.quantinsti.com/blog/forecasting-stock-returns-using-arima-model. Forecasting Stock Returns Using ARIMA Model

[16] http://faculty.smu.edu/tfomby/eco6375/BJ%20Notes/Stats.pdf