

Stock Price Prediction Using Time Series Analysis and Business Intelligence

Gaurav Priyadarshi^{1*}, Avneet Ranjan², Sharath Kumar³, Bipul Mohanta⁴

^{1,2,3,4}School of computing and Information Technology, REVA University, Bangalore, India

DOI: <https://doi.org/10.26438/ijcse/v7si14.2831> | Available online at: www.ijcseonline.org

Abstract— Stock price prediction has always been a curious, interesting and complex topic in business studies. Stock market is very unreliable for forecasting since there are no major rules or algorithms to estimate or predict share price in the stock market. Several methods like Random Forest analysis, neural networks, time series analysis algorithms like ARIMA, statistical analysis, SVM and many more have been used to predict the stock price of shares in the stock market but not all of these implementations have been correctly identified as a consistent acceptable prediction tool. This paper presents a comparative study of time series analysis using Autoregressive Moving Average i.e. (ARIMA MODEL) and Tableau (a powerful business intelligence tool) to predict the closing index of Google Inc. This paper also presents a process to build stock price prediction model with the help of time series analysis i.e. (ARIMA). The model has been built with the help of R Programming and Tableau. With the upcoming of machine learning and neural networks many researchers are trying to predict the stock price of companies and the trend that it will follow in the near future because it affects the investors as well as the competitors that are present for that company in the market. The prediction of stock price can also be done using neural networks, SVM etc. But here time series analysis has been used because it is easy to implement and it gives better results for short term predictions. The results obtained by ARIMA model shows that it is one of the best methods for the analysis of time series data.

Keywords—: Stock market, forecasting, ARIMA Model, Business intelligence.

I. INTRODUCTION

Prediction continues to be an interesting area in almost all the domains. In case of business studies predicting the future stock prices has always been an interesting and challenging area for the researchers. The researchers have been trying to improve the existing models, the reason behind it is that, it will help the investors to make an intelligent decision based on the trend forecasted by the model in the near future.

Several methods have been used in past to predict the future stock prices, one of the famous methods is neural networks which uses backpropagation method i.e. the data is fed to the neural network and the results as well as the error term is calculated, again the result is fed to the neural network as input. This process continues until we get an optimal solution but the problem with neural networks is that it takes a lot of time in computing and sometimes it might end up in an infinite loop resulting in wastage of resources.

So here we are predicting the stock price closing index using the ARIMA Model, ARIMA Model are best suited for

Financial and economic studies for short term predictions. ARIMA also known as autoregressive integrated moving average uses the past data points and calculates the error between the past and the future data points using various

tests like ADF and DFC tests. ARIMA model are classified using three terms (p,d,q), where p = number of autoregressive terms, d = how many non-seasonal trends are required to achieve stationarity and q = number of lagged forecasted errors. So to make the process easy, effective and less computationally expensive, We have used Auto ARIMA model which is present under the forecast function in R. Auto ARIMA works on stationary data and it tests data with all p, d and q values provides the best results that can provide best forecasted indices[1].

Generally, it has been found in the survey that future prediction of stock price might be done using two stances: statistical analysis methods which involves statistical measures like mean, median, variance, different statistical tests like T test and neural networks methods [2]. ARIMA models works with stationary data sets and the data set should be univariate i.e. it works to predict only one target variable in time series analysis mainly for short-term predictions. We have used ARIMA model with the confidence of 95% and we are predicting the stock of GOOGLE for the next two years i.e. 2020 and 2021. ARIMA models are considered one of the best methods for predicting statistical time series and it has given better results. Other models such as regression and neural network also gives good results but they are more computationally expensive and require more training data. In this paper we have presented a method of predicting stock price using ARIMA

models for short-term stock price prediction. The real time data has been downloaded from yahoo finance for two fiscal years. ARIMA models helps the investors to provide short-term prediction.

It should include important findings discussed briefly. Wherever necessary, elaborate on the tables and figures without repeating their contents. Interpret the findings in view of the results obtained in this and in past studies on this topic. State the conclusions in a few sentences at the end of the paper. However, valid colored photographs can also be published.

I ARIMA MODEL

As we all know that ARIMA model is applied on time series data so first we need to make sure that the data set is time series. The data points are collected over a period of time and the Autoregressive term is calculated based on the relation between the current data points and previous data points, the integrated part is used to identify the rate of growth or decline in data points, finally the noise is reduced using moving average part and the trend is predicted. ARIMA Model is known to give better results with less time complexity and resources. There are three components of time series data

1. **TREND:** - when there is either increase or decrease in data over a period of time it is called trend.
2. **SEASONAL:** - when trend is affected by seasonal factors.
3. **CYCLIC:** - when trend rise or fall are not fixed.

ARIMA Model is a combination of two models (AR+MA). AR stands for autoregressive and MA stands for moving average, what binds together is the integration part indicated by I. ARIMA Model is identified by three parameters (p,q,d). P = autoregressive lags. Q= moving average. D= order of integration.

To predict the value of p we plot PSEF graph i.e. partial autocorrelation graph. Similarly to predict the value of q we plot ACF graph i.e. autocorrelation graph and finally to the value of d depends upon the order of integration.

II. TABLEAU

TABLEAU is a powerful data visualization tool, it is used for making interactive graphs, dashboards and many more. The reason why we have selected Tableau here because it has gained much popularity over a period of time. It is considered one of the best data visualization tool by GARTNER.

We have used tableau to understand the data better by building lots of graphs and charts. We have also used the forecast function of tableau to predict the future stock prices of google in order to verify that the predictions made by

ARIMA and tableau are identical and the results that we have found are quite satisfactory.

II. RELATED WORK

Fahad and Alshatri (2014) have prepared and examined about 24 clustering calculations theoretically and observationally by considering big data properties like volume, variety and speed [10]. At last, they have contingent that intentions like DENCLUE and OptiGrid demonstrate some better execution in contrast to others [4,5,6]. Li and Xie (2014) endeavoured to foresee stock value utilizing sentiments analysis in addition to classification strategies [5,12]. They have utilized some semi-automated and manual word references from Harvard, Loughran and McDonald, cambria and so on. They reasoned that this strategy beat bag of words display in approval and autonomous testing using data sets [7, 8]. Sul and Dennis (2014) attempted to evaluate total possessed valence with normal every day stock exchange returns of firms [6,15]. They presumed that clients with numerous supporters have vast effect on intraday returns conversely with client with couple of adherents have extensive effect on stock return in future couple of days.

III. METHODOLOGY

The methodology involves the steps from data collection i.e. downloading the data from YAHOO FIANACE to final forecasting of stocks. This methodology consists of following steps:

- 1) Downloading the data set from yahoo finance.
- 2) Storing the data into local system.
- 3) Importing the data into R Studio using read.csv.
- 4) Visualizing the fetched data using R studio and TABLEAU.
- 5) Forecasting the trend of stock index using time series model (ARIMA MODEL)

STEPS INVOLVED IN ANALYSIS

1. Importing the data to R Studio.

#	Date	Open	High	Low	Close	Adj.Close	Volume
1	2016-12-01	null	null	null	null	null	null
2	2017-01-01	778.809998	841.950012	775.799988	796.789978	796.789978	33191600
3	2017-02-01	799.679993	833.450012	791.190002	823.210022	823.210022	25682500
4	2017-03-01	828.849976	853.500000	803.369995	829.559998	829.559998	33835100
5	2017-04-01	829.219971	916.849976	817.020020	905.960022	905.960022	25161900
6	2017-05-01	901.940002	979.270020	901.450012	964.859985	964.859985	32180100
7	2017-06-01	968.950012	988.250000	908.309988	908.729980	908.729980	43577200
8	2017-07-01	912.179993	986.200012	894.789978	930.500000	930.500000	36679200
9	2017-08-01	932.380005	941.979980	903.400024	939.330017	939.330017	27958000
10	2017-09-01	941.130005	959.786011	909.700012	959.109985	959.109985	28053800
11	2017-10-01	959.979980	1048.390015	947.840027	1016.640015	1016.640015	31779700
12	2017-11-01	1017.210022	1062.375000	1013.010010	1021.409973	1021.409973	24265500
13	2017-12-01	1015.799988	1078.489990	988.280029	1046.400024	1046.400024	28163700
14	2018-01-01	1048.339966	1186.890015	1045.229980	1169.939941	1169.939941	28734500
15	2018-02-01	1162.809985	1174.000000	992.559998	1104.729980	1104.729980	42384100
16	2018-03-01	1107.869995	1177.050049	980.640015	1031.790039	1031.790039	45353300
17	2018-04-01	1022.820007	1094.165039	990.369995	1017.330017	1017.330017	41715900
18	2018-05-01	1013.659973	1110.750000	1006.289978	1084.989990	1084.989990	31818700
19	2018-06-01	1099.349976	1186.286011	1096.010010	1115.650024	1115.650024	32096000
20	2018-07-01	1099.000000	1273.890015	1093.800049	1217.260010	1217.260010	31940100
21	2018-08-01	1228.000000	1256.500000	1186.239990	1218.189941	1218.189941	28762400
22	2018-09-01	1204.270020	1212.989990	1146.910034	1193.469971	1193.469971	28847000

Fig.1 Data set after importing to R studio

The data has been downloaded from yahoo finance and imported to R studio. The data has been taken from 1/1/2017 to 1/12/2018.

- To import the data set read.csv function has been used in R studio.
- To analyse the data set view function has been used.

2. Data pre-processing

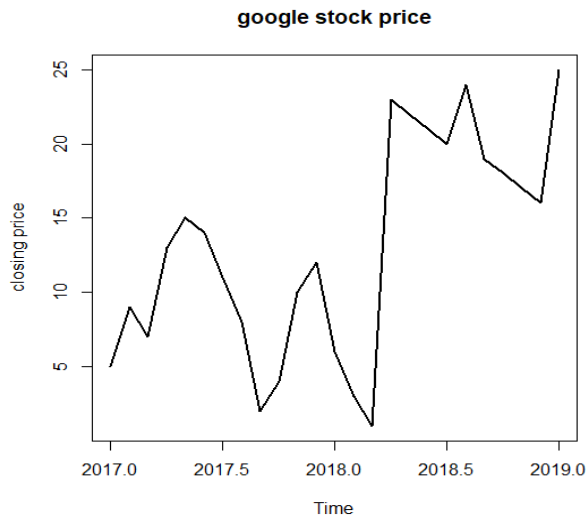


Fig.2 Graph shows the closing price of stock vs time

- Plot function has been used to plot the graph.
- X-axis represents time in years.
- Y-axis represents the google stock closing price.
- From the graph we can observe that there is no regular trend.

3 Data Analysis

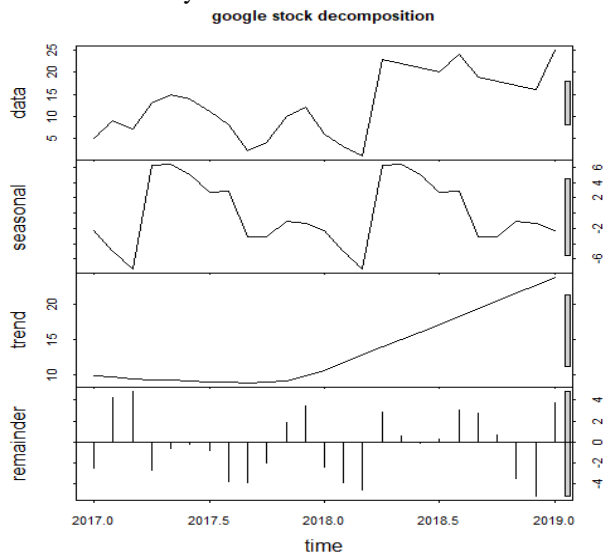


Fig.3 Google stock decomposition vs time

- The above graph represents different components of time series data.
- Seasonal trend can be observed from the graph.

4 Predicted forecast of ARIMA model
Forecasts from STL + ARIMA(0,1,0)

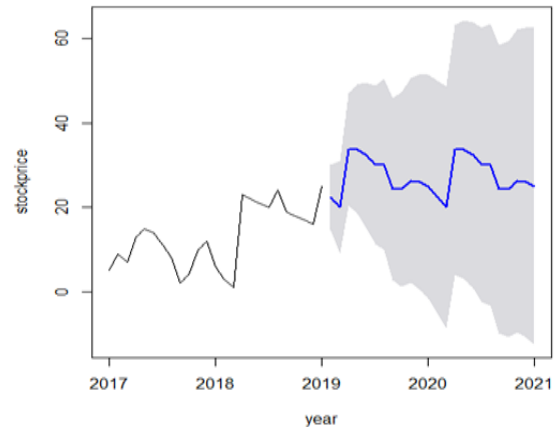


Fig.4 forecast of arima model

- We have used forecast function to predict the trend.
- Under forecast function we have used ARIMA.
- The confidence level is 95%.
- The trend is predicted for next two years.

5 Forecast in tabular format

	Point	Forecast	Lo 95	Hi 95
Feb 2019	22.41555	14.7662595	30.06484	
Mar 2019	20.01370	9.1959676	30.83142	
Apr 2019	33.63279	20.3838294	46.88174	
May 2019	33.75188	18.4532994	49.05045	
Jun 2019	32.39535	15.2910245	49.49968	
Jul 2019	30.03883	11.3019770	48.77568	
Aug 2019	30.16550	9.9273885	50.40362	
Sep 2019	24.29218	2.6567214	45.92763	
Oct 2019	24.29485	1.3469881	47.24272	
Nov 2019	26.29753	2.1083566	50.48670	
Dec 2019	25.97015	0.6003283	51.33997	
Jan 2020	25.00000	-1.4979127	51.49791	
Feb 2020	22.41555	-5.1643539	49.99545	
Mar 2020	20.01370	-8.6073217	48.63471	
Apr 2020	33.63279	4.0072187	63.25835	
May 2020	33.75188	3.1547223	64.34903	
Jun 2020	32.39535	0.8565291	63.93418	
Jul 2020	30.03883	-2.4143519	62.49201	
Aug 2020	30.16550	-3.1769720	63.50798	
Sep 2020	24.29218	-9.9164817	58.50083	
Oct 2020	24.29485	-10.7585900	59.34830	
Nov 2020	26.29753	-9.5808126	62.17587	
Dec 2020	25.97015	-10.7145508	62.65485	
Jan 2021	25.00000	-12.4737075	62.47371	

Fig.5 Forecast in tabular format

- After the forecasting the trend is obtained in tabular format.
- We have table function to put the data into a tabular format.

ANALYSIS OBTAINED BY TABLEAU

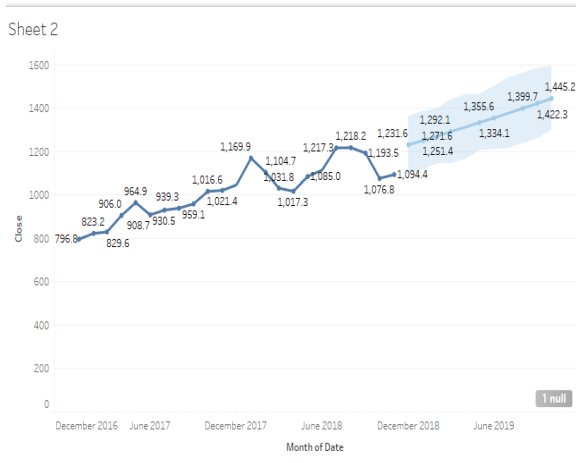


Fig.6 Month vs closing value

- Forecast function gives final prediction in tableau.
- Forecast function is present under the analysis part of tableau.
- X-axis represents time.

Y-axis represents closing stock price

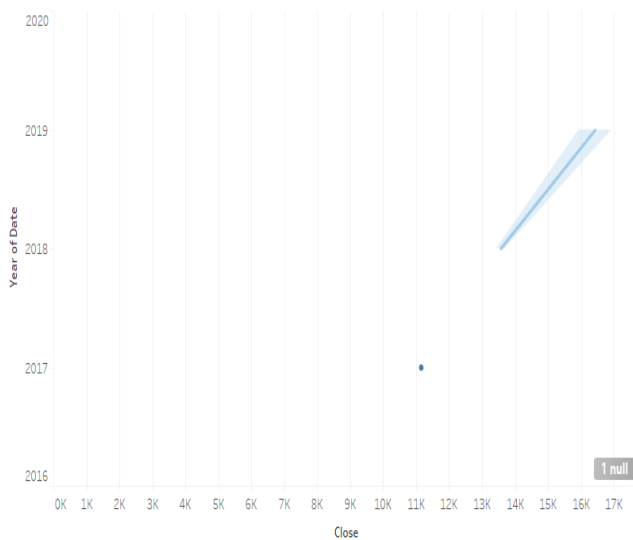


Fig.7 closing value vs year

- The figure represents closing value of stock in yearly basis.

IV. RESULTS AND DISCUSSION

From the graph shown below and the data visualized using Tableau we can conclude that the result is satisfactory and ARIMA model works better with the data set for future prediction of stock index. The results have been calculated with the confidence of 95%

V. CONCLUSION

The paper presents a concise and effective approach for building ARIMA model for the prediction of stock prices of google. The model build can be applied to any company’s data set and the stock price can be predicted effectively. Although any method cannot predict the prices effectively but by building this model we have tried to make an effort towards a big problem in financial studies i.e. stock prediction and with the results obtained we can say that ARIMA model works well with the data and can be applied to problems of these kind.

REFERENCES

- [1] N. Chatterjee, S. Mohan, “Extraction-Based Single-Document Summarization Using Random Indexing”, Proc. 19th IEEE International Conference on Tools with Artificial Intelligence, Greece, pp. 448-455, 2007.
- [2] Sumona Mukhopadhyay, Santo Banerjee, “Global optimization of an optical chaotic system by Chaotic Multi Swarm Particle Swarm Optimization”, Expert Systems with Applications, Vol.39, Issue.1, pp. 917-924. 2012.
- [3] Yang Shi, Hongcheng Liu, Liang Gao, Guohui Zhang, “Cellular particle swarm optimization”, Information Sciences, Vol.181, Issue.20, pp.4460-4493, 2011.
- [4] H. Huang, XS. Chen, H. Lei, ZQ. Li, WG. Li, “The Modified Temperature Field of Ceramic Roller Kiln Based on DEPSO Algorithm”, Advanced Materials Research, Vols. 219, Issue.220, pp. 1423-1426, 2011
- [5] Gitali Rakshak and Amit Pimpalkar, “A Review on an optimized path finding on road network using Ant colony algorithm”, International Journal of Compurt Science and Engineering, Vol.2, Issue.10, pp.26-29, 2014
- [6] GJO. Osório, JCO. Matías, JPS. Catalão, “Hybrid evolutionary-adaptive approach to predict electricity prices and wind power in the short-term”, 2014 Power Systems Computation Conference, Wroclaw, pp. 1-7, 2014.
- [7] R. Majhi ,G. Panda ,G. Sahoo , “Efficient prediction of exchange rates with low complexity artificial neural network models”, Expert Systems with Applications, Vol.36, Issue.1, pp.181-189, 2009.
- [8] JC. Hung, “Adaptive Fuzzy-GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization”, Information Sciences, Vol.181, Issue.20, pp.4673-4683, 2011.
- [9] MT. Leung, H. Daouk, AS. Chen, “Forecasting stock indices: A comparison of classification and level estimation models”, International Journal of Forecasting, Vol.16, Issue.2, pp.173190, 2000.
- [10] P. Chang, D. Wang, C. Zhou, “A novel model by evolving partially connected neural network for stock price trend forecasting”, Expert Systems with Applications, Vol.39, Issue.1, pp.611-620, 2012.
- [11] HH. Chu, TL. Chen, CH. Cheng, CC. Huang, “Fuzzy dual factor time series for stock index forecasting”, Expert