ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 12 , December 2019

Apple Stock Price Prediction using ARIMA Model

URANG Awajionyi S. , KABARI Ledisi G.

Department of Computer Science, Ignatius Ajuru University of Education, Port Harcourt, Rivers State, Nigeria.
Department of Computer Science, Ignatius Ajuru University of Education, Port Harcourt, Rivers State, Nigeria.

ABSTRACT: Time series data analysis of monthly Apple Stock Price (APPL) from the New York Stock Exchange (NYSE) was conducted. The time plot reveals a positive upward trend. Stockpile price forecast is constantly attracting attentions owning to its direct monetary advantage plus the connected difficulty. In this paper, we have conducted the prediction of Apple stock price using the Autoregressive Integrated Moving Average (ARIMA) Model. The ARIMA model was selected because of its wide suitability of the model. The parameters selected for the ARIMA model have been done using Akaike information criterion (AIC) and Akaike information criterion corrected (AICc). The results of the analysis using r software, shows it yielded 97.20% accuracy.

KEY WORDS: ARIMA Model, Apple Stock, Prediction, AIC.

I. INTRODUCTION

The data observed over time give rise to temporary (temporal) series. The analysis of time series data makes it possible to understand the behavior of data over time. The reason of time series study procedures is to identify trends, investigate possible cases, and often make predictions for the unknown time. The temporal information of a datum plays an important role in the pattern identification step. If a temporary reason is identified, the possible causes of the late reason will be examined. Then, the response variable can be modeled as a function of time or as a function of the direct cause of the prediction. The importance of time in a dataset is not limited to those cases in which we intend to do a time series analysis. Time is often a vital element of the explore dataset. This is partly because it can often be adjusted for underlying variables and the effects of unmeasured or misunderstood factors. Time series data can be decomposed into trend, seasonality, cycle and random fluctuation components. The trend component of a time series is the long-term trend or the change in the regularity of the data. The existence of a trend in the data of a time series can be identified simply by comparing the averages of the data of a time series at different intervals or simply by means of a regression analysis of the data for the duration of the time. Time, in years, months, days or hours, is a device that links a phenomenon to a set of common and stable reference points. The concept of time series is essentially based on historical observations. This involves explaining past observations to try to predict, those of the future.
The major problem in stock prediction is the accuracy of the prediction. The better the accuracy the more likely that the future price prediction will hold and investors can use the prediction.
The aim of the research is to foretell the upcoming price of the apple stock; it is ahead of the year 2020 and the objectives are

1. Get the best AMIRA model
2. Get predictions are based on absolute adjusted close price of the historical data.
3. Check the prediction accuracy since it counts.
4. The prediction is for a period of twelve months or one year

A).ARIMA Model

The Autoregressive integrated moving average ARIMA model can be regarded as an expansion of the ARMA model. The progression X_t is said to be an autoregressive integrated moving average process ARIMA (p, d, q) if $W_t = \nabla^d X_t = (1-B)^d X_t$ is a stationary ARMA (p, q) process.

In general, we can write the model as an extension of the ARMA model. The process X_t is said to be an autoregressive integrated moving average process, ARIMA (p, d, q) if $W_t = \nabla^d X_t = (1-B)^d X_t$ is a stationary ARMA (p, q) process.

In general, we can write the model as $\emptyset(B)(1-B)^d X_t = \theta(B)\varepsilon_t$ ………………… (1.1)

Where $\nabla = B - 1$ is the differentiation operator. We can usually take $d = 1$ or at most 2 we can write this kind of model as $\Phi(B)(1-B) X_t = \varepsilon_t$

Where $\Phi(B)$ is a fixed autoregressive operator $\Theta(B)$ is a stationary moving operator, and ε_t is white noise. If the processes contain no autoregressive conditions, we describe it an integrated moving average and denoted by ARIMA (p, d, q). If no moving average conditions are there, we denoted the model as ARIMA (p, d, q).

II. LITERATURE SURVEY

Stock prediction has been the subject of numerous research and reviews details of diverse statistical techniques. Currently, most of the study is based on forecasting of stock price trends using neural networks based on ARIMA. ARIMA is used as PACAP - Database of China, formed as a result of the peaceful -Basin Capital Markets (PACAP) Research Center of the University of Rhode Island (USA) and the SINOFIN Information Service Inc., associated with the China Economic Research Center (CCER) of the University of Beijing (China). ARIMA has been applied to solve real-world problems in Nifty Midcap-50 using MATLAB alongside through the performance measure.

Combining the diffuse regression model with The ARIMA model, diffuse ARIMA model (FARIMA) that was formed for the intention of predicting the financial exchange rate from NT dollars to US dollars. An additional reason in support of the ARIMA model is its usefulness in predicting or forecasting the next day, mainly in the study along with experiments a specific stock market associated toward stock price prediction.

However, the problem here is not on the relative performance and superiority of the ARIMA model over the ANNs model, in prediction but how realizable and accuracy is the model when applied to real live situation, is more important. This work therefore seeks to further clarify the accurateness of the ARIMA model along with its predicting ability to stock prices.

III. METHODOLOGY

A) Data

The data for this work is four hundred and seventy one (471) Apple Stock weekly prices observations from 01 January 2010 to 1 November, 2019, obtainable from the www. Yahoo.com/finance website. The data will be divided into two parts train and test data respectively. The train dataset is for model estimation and the test is for testing our predicted model. The train dataset is 450 and test dataset data set is 21 observations respectively.

The data structure for our analysis is shows that our data type is data.frame with 471 observations of two variables. The variables are weekly date as a factor variable and Adj.close price of the stock as a numeric variable.

Table 1. The data structure of our stock.
'data.frame': 471 obs. of 2 variables: $ Date : Factor w/ 471 levels "2010-11-01","2010-11-08",..: 1 2 3 $ Adj.Close: num 39.5 38.4 38.2 39.3 39.6.

B) Methodology

Getting a suitable time series model is always a difficult task. Box-Jenkins (1976) procedure that contains four major steps, which is now well established, to build an ARIMA model, each of which may be used several times.

1. Model specification (or identification).
2. Model fitting (or estimation of parameters).
3. Model diagnostics (or checking)
4. Forecasting (or prediction).

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 12 , December 2019

IV. EXPERIMENTAL RESULTS

Table 2. Descriptive statistics of Apple Stock Price

Statistics	Apple Stock
Minimum	38.24
1st Quartile	66.70
Median	98.80
Mean	108.39
3rd Quartile	149.22
Maximum	255.82

From table 2, it gives us a quick summary of the Apple stock price, showing it minimum price at $38.24 and it maximum price at $255.82.

We present in figure 1 and 2, we present Apple stock price plot over time, ACF and PACF of Apple stock price for proper analysis

Figure 1. Apple stock price plot over time.

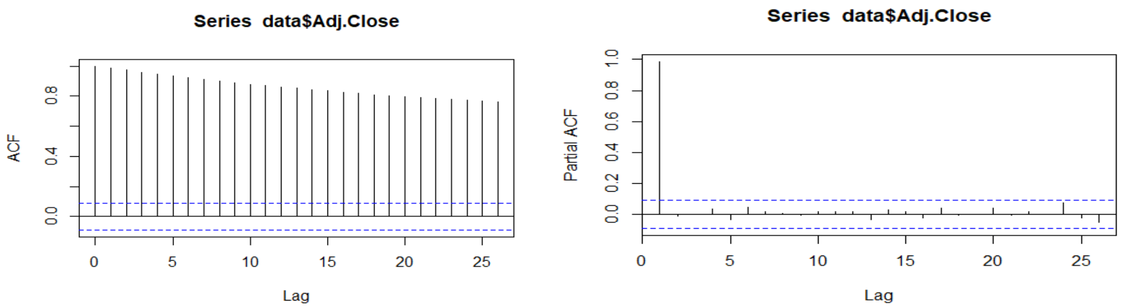

Figure 2. The ACF and PACF of our Apple stock price.

From figure 1, we can observed that the variance is changing over time, showing that there is no constant variance, hence indicating that our Apple stock price is not stationary. We can further test for the stationary of stock price using the Augmented Dickey Fuller (ADF) test.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 12 , December 2019

Table 3. First ADF test to check stationaity of data.

```
Augmented Dickey-Fuller Test

data:  data$Adj.Close
Dickey-Fuller = -1.8895, Lag order = 7, p-value = 0.6248
alternative hypothesis: stationary
```

From table 3, the p-value is 0.6428 above 0.05, hence the sequence series is not stationary. The data needed to differentiated by 1 and re-conduct the ADF test again.
In figure 3, we present the first difference of Apple stock price, including its ACF and PACF.

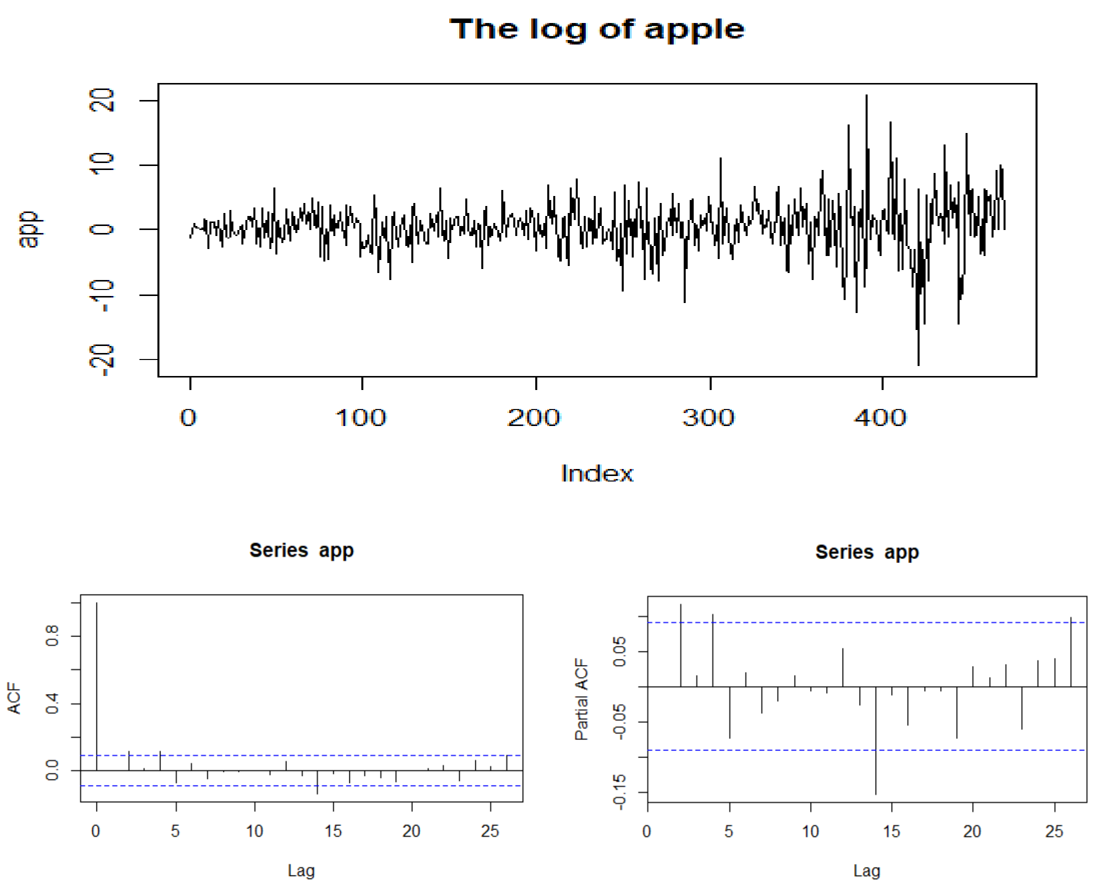

Figure 3. The first differencing of Apple stock price and ACF and PACF after diff.

Table 4. Second ADF test, after first differences

```
Augmented Dickey-Fuller Test

data:  app
Dickey-Fuller = -7.3343, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

From table 4, the p-value is now 0.01 below the 0.05, hence the series is now stationary. It means no more differencing is needed for the data.
From table 5, based on the auto.arima command applied we have the following parameters, p=1,d=1 and q=1, hence our ARIMA model is ARIMA(1, 1,1).

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 12 , December 2019

Table 6. Arima Parameters Estimates

```
Series: data$Adj.Close
ARIMA(1,1,1) with drift

Coefficients:
          ar1     ma1   drift
      -0.8859  0.8333  0.3395
s.e.   0.0923  0.1093  0.1874

sigma^2 estimated as 16.79:  log likelihood=-1268.89
AIC=2545.79   AICc=2545.88   BIC=2562.22

Training set error measures:
                       ME     RMSE      MAE        MPE    MAPE      MASE
 ACF1
Training set 0.000263278 4.079401 2.791818 -0.1273876 2.80134 0.9723825 0.054
35183
```

From the Table 6, after applying the ARIMA (1, 1, 1) model to our train data, our estimated parameters are

Coefficients:	ar1	ma1
	-1.0923	1.1389

and using the equation (1) the fitted model in this case is (1+1.0923B)(1-B)$\hat{X}_t$ = (1-1.1389B) $\hat{\varepsilon}_t$ …………...(4.1)
with estimated variance, $\hat{\sigma}^2 t$ = 16.79 and log likelihood = −1268.89.

Table 6. Arima prediction result for 21weeks

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
451	192.3984	187.1470	197.6497	184.3671	200.4296
452	192.6949	185.4609	199.9290	181.6314	203.7584
453	193.0724	184.1517	201.9931	179.4294	206.7154
454	193.3782	183.1503	203.6062	177.7359	209.0205
455	193.7475	182.2758	205.2192	176.2030	211.2920
456	194.0606	181.5365	206.5846	174.9066	213.2145
457	194.4234	180.8720	207.9749	173.6983	215.1486
458	194.7422	180.2830	209.2013	172.6288	216.8555
459	195.1000	179.7476	210.4525	171.6205	218.5796
460	195.4232	179.2597	211.5867	170.7032	220.1432
461	195.7771	178.8134	212.7408	169.8333	221.7208
462	196.1038	178.3997	213.8079	169.0277	223.1799
463	196.4546	178.0196	214.8895	168.2608	224.6484
464	196.7840	177.6633	215.9047	167.5415	226.0266
465	197.1324	177.3350	216.9298	166.8549	227.4099
466	197.4640	177.0249	217.9030	166.2051	228.7228
467	197.8104	176.7384	218.8825	165.5835	230.0374
468	198.1437	176.4665	219.8210	164.9912	231.2962
469	198.4887	176.2146	220.7627	164.4234	232.5539
470	198.8233	175.9750	221.6716	163.8798	233.7668
471	199.1670	175.7525	222.5815	163.3576	234.9764

Table 7. Arima Model accuracy

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.000263278	4.079401	2.791818	-0.1273876	2.80134	0.9723825	0.05435183

From table 7, the model accuracy is 97.20%, because the mean absolute percentage error (MAPE) is 2.80134, so our accuracy of the model is 100% - 2.80134 = 97.20%.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 6, Issue 12 , December 2019

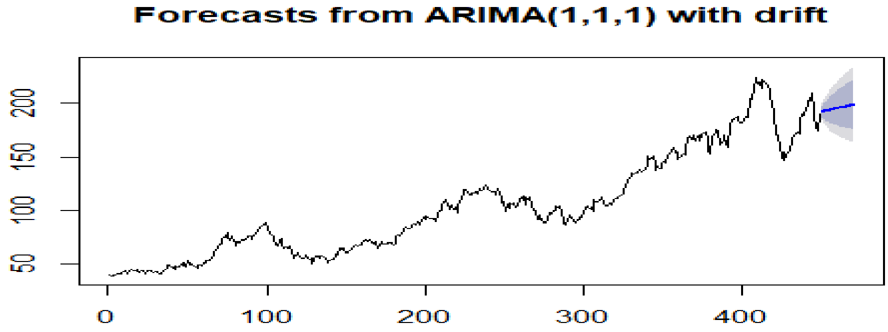

Figure 4. The ARIMA (1, 1, 1) prediction plot.

The figure 4 shows the Apple stock price prediction, the indication of the blue line in the figure shows an uptrend for the next 21 weeks. Therefore we can wrap up, by saying the Apple price stock trend will continue upward for the period of 21 weeks from

V. CONCLUSION AND FUTURE WORK

This study has highlighted ARIMA model to predict the Apple stock price in the short term. After collecting enough real data to create stock market data, an ARIMA model is implemented in the data set used to improve the short term forecast. The application of the model in the case of bank action data has allowed verifying its accuracy and demonstrating its presentation capabilities. About 471 observations were collected to implement its predictions and the best ARIMA model was selected based on the best known criteria, the AIC. Another important observation is that the predictive accurateness of the ARIMA model is 97.20% and it gradually decreasing at this stage of the growth process from one period to the next. This model can be applied and it is suitable for high-tech market cases, especially banks, as it provides a significant indicator for the future. The method was limited to a short-term forecast and is not useful in the long term. Future research on this topic includes other prospective market data, such as industrial data and any data that is measured over time.

When applying to real life data to trade with, the outcome news should be considered greatly since it plays major role in the price actions of most stocks in the world.

REFERENCES

[1] Abu-Mostafa Y. S., and Atiya A. F., "Introduction to financial forecasting," Applied Intelligence, Vol.6, 3, pp.205–213, 1996.
[2] Akcora, C. G., Dey, A. K, Gel, Y.R., & Kantarcioglu, M.," Forecasting Bitcoin price with graph chainlets",Journal of Lecture Notes in Computer Science, 10939. Springer, Cham.
[3] Box, G., and Jenkins, G.," Time series analysis forecasting and control", Holden day, San Francisco, 1976.
[4] Catania, L., and Grassi, S.," Modelling cryptocurrencies financial times series", SSRN Electronic Journal, 2017.
[5] Catania, L., Grassi, S., Ravazzolo, F., (2018). Forecasting crypto-currencies finance time series CAMP working paper series, Vol 6, 2018. [6] Chu J., Chan S., Nadarajah S., and Osterrieder J., "GARCH modelling of cryptocurrencies", Journal of Risk and Financial Management, 2017.
[7] Hall J. W., "Adaptive selection of U.S. stocks with neural nets," in Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Nancial Markets, Deboeck, G. J.,Edition, John Wiley & Sons, New York, NY, USA.1994
[8] Engle, R. (1982), the use of ARCH/GARCH models in applied econometrics; Journal of Economic Perspectives Vol.15, 4, pp.157-168, 1982.
[9] Engle, R. F.," Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation econometrics", 50, pp.987-1007, 1982.
[10] Engle, R. F., and Bollerslev, T., "Model the persistence of conditional variances", Econometric Reviews Vol.5, pp1-50, 1986.
[11] Franses, P. H., and Dijk, D.," Forecasting stock market volatility using (Non-linear) GARCH models", Journals of Forecasting, Vol.15, pp.229-235, 1996.
[12] Gangwal S., and Longin, F., "Extreme movements in Bitcoin prices: a study base on extreme value theory", 2018.
[13] Jarrett, Jeffrey E., and Kyper E., (2011). ARIMA modeling with intervention to forecast and analyze Chinese stock prices. International Journal of Engineering Business Management, Vol 3, 3, pp.53-58, 2011.
[14] Ljung, G. M., and Box, G. E. P., " On a measure of lack of fit in time series models", Biometrika Vol.65, pp.297-303,1978.
[15] Nayak, S. C., and Mishra B.B., "A Neuro-Genetic Technique for Index Prediction" Silicon Institute of Technology Bhubaneswar, the capital of Odisha, India**, ND**.
[16] Osterrieder, J., The statistics of bitcoin and crypto-currencies. Zurich University of Applied Sciences, School of Engineering, Switzerland, 2016.
[17] Philip, A., Chan, J. S. K., and Peiris S.," A new look at cryptocurrencies", Economics Letters Vol.163, 2017.
[18] Uma Devi B., Sundar D., and Alli P.," An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50", International Journal of Data Mining & Knowledge Management Process (IJDKP), Vol.3, 3, 2013.