Time-Series Forecasting: ARIMA Model Stock Price Prediction

Benjamin Tilden
Western University
London, Canada
250959344
btilden@uwo.ca

Ian Borwick
Western University
London, Canada
250950449
iworick@uwo.ca

Lucas Fraulin
Western University
London, Ontario
250963527
lfraulin@uwo.ca

***Abstract*—Stock price prediction is an interesting yet challenging topic in investing. The results of correctly predicting stock prices have caused researchers to improve and develop predictive models continuously. The autoregressive integrated moving average (ARIMA) models have been explored in literature and are perhaps the best type of model to use on time series data, particularly stock price data. This paper presents an extensive process of building stock price predictive models using the ARIMA model. Published stock data is obtained from the Nasdaq Composite are used with stock price predictive models developed. Results obtained revealed that the ARIMA model has a strong potential for short-term prediction and can compete favourably with existing techniques for stock price prediction [1].**

***Index Terms*—ARIMA, models, stock price prediction**

I. Introduction

Predictions and forecasting will continue to be a crucial aspect of company analysis in investment decisions. The ability to plan and develop effective strategies to reduce investment risk or increase potential success remains a task that researchers will continuously try to improve. Financial statement projection is typically the most common form of forecasting in investment management [1]. Projecting financial statements incorporate current trends, qualitative and quantitative analysis to arrive at an economic point that an investor believes the company could attain in the future. Although projected financial statements are perhaps the most essential element of company forecasting, some researchers and investors have challenged themselves with predicting the actual company stock price.

"Stock price prediction is regarded as one of the most difficult tasks to accomplish in financial forecasting due to the complex nature of the stock market [1, 2, 3]." The possibility of developing a model that could guarantee easy profit remains a motivating factor for researchers to improve and develop predictive models. Historically, there have been several models and different techniques designed to predict stock prices [4]. Artificial neural networks (ANNs) were prevalent due to their pattern learning ability and inference results from anonymous data. Several different methodologies are used as the backbone of ANNs, with each having a resulting strength. Thus, recently researchers have been exploiting each of those models' strengths and have started building 'hybrid' models to improve stock price prediction [5, 6].

Generally, models can be categorized as a statistical or artificial intelligence model [2]. ARIMA models are statistical models "known to be robust and efficient in financial time series forecasting, especially with short-term predictions [1]." Other statistical models include regression models, exponential smoothing, and generalized autoregressive conditional heteroskedasticity (GARCH) models [7,8,9]. Our paper describes the extensive process of developing an ARIMA model, the reasoning behind our decisions, and the results obtained from real-life stock data to demonstrate the potential strength of ARIMA models.

II. ARIMA Model

An ARIMA model is a statistical model designed for analyzing and forecasting time series data. The model "caters to a suite of standard structures in time-series data, and as such provides a simple yet powerful method of making skillful time series forecast." An ARIMA model can be further described by breaking down its acronym [10].

- AR - Autoregression: A model that uses the dependent relationship between an observation and some number of lagged observations. Our report will examine the relationship between a stock price at some date and the stock price at some other previous lagged date [10].
- I - Integrated: The use of the difference of raw observations (e.g. subtracting an observation from observation at the previous time step) to make the time series stationary [10].
- MA - Moving Average: A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations [10].

ARIMA models have demonstrated efficient capability to generate short-term forecasts and have constantly outperformed complex structural models [1]. Proper model selection is key to ensuring accurate results and the starting point of the overall project. Model comparison of the Apple (AAPL) stock data (dating from 1984 to 2017) was made between various forecasting models [11]. Models included; linear regression (LR), lasso regression (LS), elastic net regression (EN), k-nearest neighbours (KNN), random forest regression (RF), and an ARIMA model. Each model was first put through hyper-parameter tuning (a process of finding the most ideal model parameters for best results) before evaluation [12]. From this custom, metrics were computed to allow for proper comparison. The models compared are those classically used for predicting stock data, therefore the natural choice to evaluate ARIMA against.

Accuracy was compared between all of the models to determine the best model. Accuracy represents the percentage that each model correctly predicted the stock price going up or down in the future. The KNN and RF provided null results as, without K-fold cross-validation, the maximum train value is the predicted values for all data exceeding the maximum train date. As such, performing the K-fold validation proved to over-fit the models; therefore KNN and RF were removed from consideration.

The LR, EN, LS, and ARIMA models all had comparable results for accuracy; thus, another metric was used to determine the best overall model. Y true change (yTrueChange) is the average magnitude of change in the actual test data, while the y predicted change (yPredChange) is the average magnitude of change in the predicted values. These metrics assist in demonstrating how accurately the model is predicting. ARIMA proved to provide the best results and, as such, was selected for the model used. See Table 1 for the raw data and see Figure 1 for the predicted values by each model.

TABLE I
MODEL COMPARISON METRICS

Model	*Accuracy*	*yTrueChange*	*yPredChange*
LR Test Scores	52.821128	4.745959	3.050159
LS Test Scores	52.821128	4.745959	3.050159
EN Test Scores	52.821128	4.745959	2.296856
KNN Test Scores	NaN	4.745959	4.49079
RF Test Scores	NaN	4.745959	4.488776
ARIMA Test Scores	52.821128	4.745959	4.782466

To ensure the ARIMA model would perform over time, the model must prove to have substantial results on a year-by-year basis. Table 2 depicts the raw data where the results can be seen. 2017 provides the most robust results with a 56.88% directional accuracy and a 0.02 difference between the predicted and real values. 2015 provides the weakest results with only a 48.41% directional accuracy and a 0.12 difference between predicted and real values.

"Each of these components is explicitly specified in the model as a parameter. A standard notation of ARIMA (p,d,q)

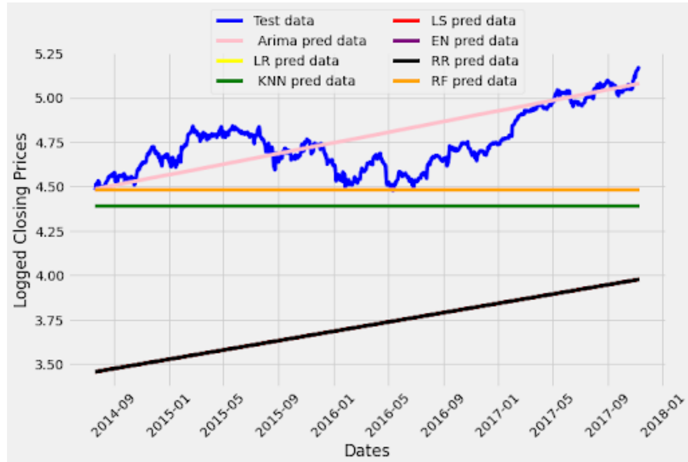

Fig. 1. Model Comparison on AAPL Close Prediction.

TABLE II
ARIMA YEAR BY YEAR METRICS

Year	*Accuracy*	*yTrueChange*	*yPredChange*
2014	53.982301	4.525321	4.588490
2015	48.412698	4.656029	4.739679
2016	53.200000	4.835533	4.621567
2017	56.880734	5.002928	4.980080

where the parameters are substituted with integer values to indicate the specific ARIMA model being used quickly [10]."

The parameters of the ARIMA model are defined as follows:

- p: The number of lag observations included in the model, also called the lag order [10, 12].
- d: The number of times the raw observations are differenced, also called the degree of difference [10, 12].
- q: The size of the moving average window, also called the order of moving average [10, 12].

A value of 0 can be used for a parameter, which indicates not to use that element of the model. This way, the ARIMA model can be configured to perform an ARIMA model and even a simple AR, I, or MA model [10].

A linear regression model is constructed, including the specified number and type of terms. The data is prepared by a degree of differencing to make it stationary, i.e. to remove trend and seasonal structures that negatively affect the regression model [10].

III. RELATED WORK

The ARIMA model is directly related to other time series forecasting models. These include Autoregression (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Seasonal Autoregressive Integrated Moving

Average (SARIMA) models. [13].

AR modelling assumes the time series's current value is a linear combination of past values with a random error. They try to explain the momentum and mean reversion effects in trading markets. An AR model of order P is described as:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t [13] \quad (1)$$

The benefit of AR models is the wide range of different time series patterns they can represent. In the definition of the AR model of order P, ε_t is representative of white noise, y_t are the predictors, and ϕ_p are the parameters. When the parameters are changed, the result is a different time series pattern. Adjusting the white noise only affects the scale of the time series [13].

MA modelling uses a regression-like model using past forecast errors instead of variable values. MA models try to show the effects of unexpected events that affect markets through white noise terms. A MA model of order q can be described as:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} [14] \quad (2)$$

Similar to the AR model, ε_t represents white noise and will only affect the scale of the series and not the pattern when changed. θ_q describes the parameters and, like the AR models, will mean a different time series pattern. θ_q represents the weighted moving average of the past forecast errors. Unlike usual regression models, we don't observe the values of white noise. MA models are similar to AR models to the extent that it is possible to write any stationary AR(p) model as an MA(∞). Also, all invertible MA(q) processes can be represented as AR(∞) [14].

ARIMA models combine AR and MA models. They attempt to capture the shock effects from unexpected events through MA and explain market participant effects through AR. An ARMA(p, d, q) model, where p is the order of the AR polynomial and q is the order of the MA polynomial, can be described as:

$$[15,16] X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \Theta_i \varepsilon_{t-i} \quad (3)$$

In this equation φ_i is the AR models parameters, Θ_i is the MA models parameters, c is a constant, and like in the AR and MA definitions, ε_t is the white noise. In an ARMA model the MA part predicts using the mean and previous errors while the AR predicts using previous values of the dependent variable. The difference between an ARMA and ARIMA model is that an ARMA model is stationary. If no differencing is involved in an ARIMA model then it becomes an ARMA model [15,16].

SARIMA models are simply arima models with an extra seasonal part. This seasonal part consists of three additional numbers P, D, Q.

- P represents the number of seasonal autoregressive terms.
- D represents the number of seasonal differences.
- Q represents the number of seasonal moving-average terms [17, 18].

SARIMA models generalize the regression approach using seasonal lags and differences to fit a seasonal pattern. The complete SARIMA model can be written as:

$$ARIMA(p, d, q) \times (P, D, Q)[17, 18] \quad (4)$$

SARIMA models often do well when compared with other seasonal models and potentially result in more accurate short-term forecasts [17, 18].

IV. Methods

A. Data Collection and Interpretation

Apple stock data used in this study covers the period from September 7th, 1984, to November 10th, 2017. This data set includes the date, open, high, low, close, volume, and openInt [12]. However, the close price is the primary concern as that is what will be the predicted value. Figure 2 depicts the original pattern of the series to show whether the time series is stationary or not. However, observations alone are not substantial enough to determine if the data is stationary or not.

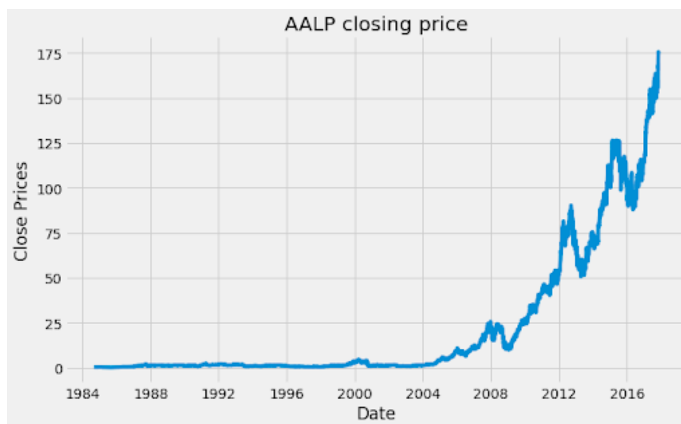

Fig. 2. Apple Stock Closing Price [11, 18].

To determine if the data is stationary, an Augmented Dickey-Fuller (ADF) test must first be performed. The ADF test is one of the most popular statistical tests used to determine the presence of a unit root in a series of data and helps to demonstrate if the series is stationary [12, 19]. There are two hypotheses within this:

- Null Hypothesis: the series data has a unit root [12, 19].
- Alternative Hypothesis: the series data has no unit root [12, 19].

To determine which hypothesis the data falls under, both mean and standard deviation lines must be flat, (must have a constant mean and constant variance) for the series to be considered stationary. See Figure 3 for the rolling mean and standard deviation and Table 2 for the dickey fuller test results [12, 19].

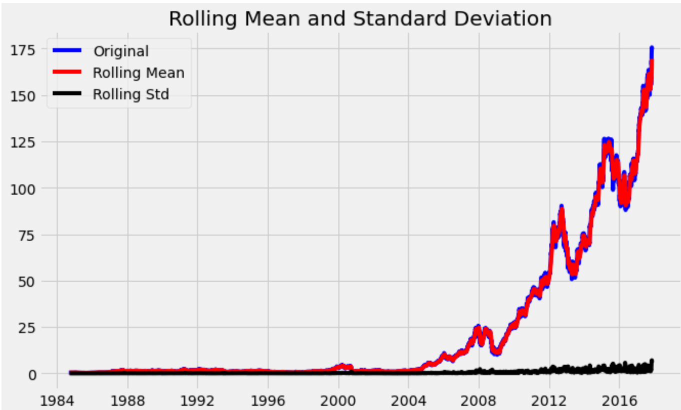

Fig. 3. Rolling Mean and Standard Deviation [12, 19].

Test Statistics	4.373924
p-value	1.000000
No. of lags used	37.000000
Number of observations used	8326.000000
critical value (1%)	-3.431136
critical value (5%)	-2.861887
critical value (10%)	-2.566955

Table 2

Figure 3 shows that the mean and standard deviation is increasing. Thus, the series is not stationary. Additionally, the observed p-value is greater than 0.05 (37.0); therefore, the null hypothesis is rejected. Furthermore, the data is non-stationary due to the test statistics being greater than the critical value [19].

Seasonality and trend factors must be separated from the series data to perform a time series analysis [19]. The resultant series will become stationary through this process, see Figure 4.

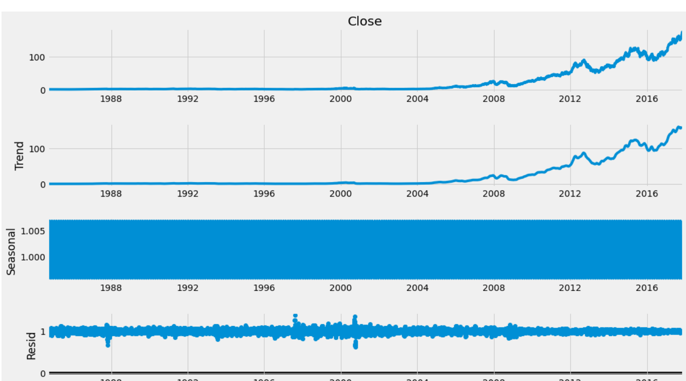

Fig. 4. Close Trend, Seasonality, and Resid.

Firstly, the series data is converted into a logarithmic scale to reduce the magnitude of the value and reduce the rising trend in the series. Then, the rolling average of the series data is calculated. A rolling average is calculated by taking input for the past 12 months and giving a mean consumption value at every point further ahead in the series [19]. See Figure 5.

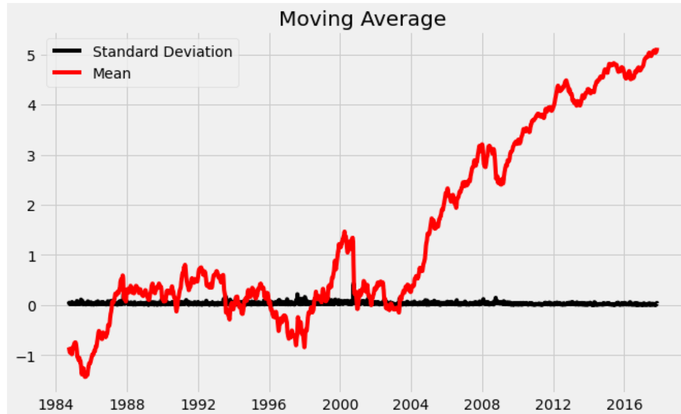

Fig. 5. Moving Average of Standard Deviation and Mean.

B. Dataset Splitting and Hyper Parameter Tuning

The model is trained on the closing price of the stock data. Figure 6 depicts the split between the training and test data. The test data makes up the past 10% of data, while the training set used to train the model consists of the first 90% [19].

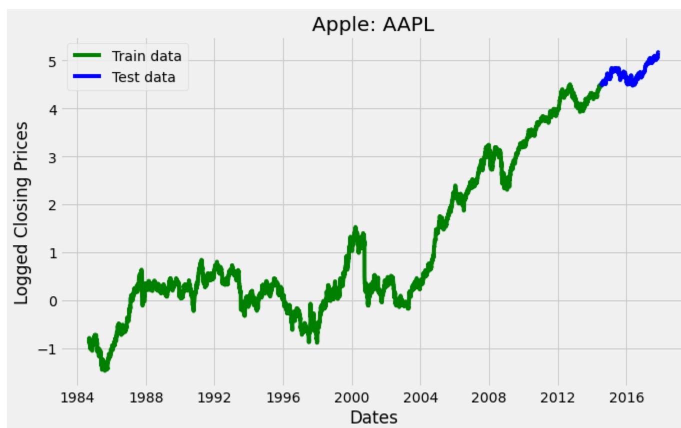

Fig. 6. Train and Test Split for AAPL Close Dataset.

The model must also be tuned for the most ideal parameters of p, q, and d. In order to do this a Auto ARIMA must be used [19].

Auto ARIMA: automatically discover the optimal order for an ARIMA model. The auto ARIMA function seeks to identify the optimal parameters for an ARIMA model and returns a fitted ARIMA model [19].

The ARIMA model functions by conducting various tests, including the Kwiatkowski-Phillips-Schmidt-Shin test, Augmented Dickey-Fuller test or the Phillips-Perron test. These tests determine the order of difference, d, and then fits models within ranges of defined start p, max p, start q, and max q ranges. If the seasonal optional is enabled, the function will also seek to identify the optimal P and Q through hyper-parameter tuning after conducting the Canova-Hansen test to determine the optimal order of seasonal difference, D [19].

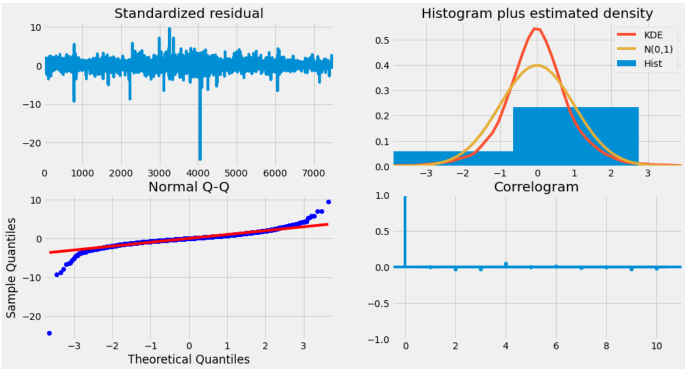

Fig. 7. ARIMA Specific Metrics for Fitting.

Interpreting Figure 7 metrics:

- The **standardized residual** errors fluctuate around a zero mean and have a uniform variance [19].
- The **histogram plus estimated density** plot suggests a mean zero with normal distribution [19].
- The **normal Q-Q** dots should fall along the red line, any significant deviations would imply the distribution is skewed [19].
- The **correlogram** or the ACF plots demonstrates the residual errors are not autocorrelated. Any amount of autocorrelation would imply that there is a pattern in the residual errors which are not explained in the model [19].

These metrics demonstrate an above average fit of the model when applied to this dataset.

```
                             ARIMA Model Results
==============================================================================
Dep. Variable:                D.Close   No. Observations:                 7523
Model:                 ARIMA(1, 1, 2)   Log Likelihood               15693.022
Method:                       css-mle   S.D. of innovations              0.030
Date:                Wed, 10 Mar 2021   AIC                         -31376.043
Time:                        17:01:43   BIC                         -31341.415
Sample:                             1   HQIC                        -31364.154

=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const             0.0007      0.000      2.109      0.035    5.04e-05       0.001
ar.L1.D.Close     0.2385      0.206      1.158      0.247      -0.165       0.642
ma.L1.D.Close    -0.2340      0.206     -1.138      0.255      -0.637       0.169
ma.L2.D.Close    -0.0236      0.011     -2.152      0.031      -0.045      -0.002
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            4.1937           +0.0000j            4.1937            0.0000
MA.1            3.2250           +0.0000j            3.2250            0.0000
MA.2          -13.1514           +0.0000j           13.1514            0.5000
-----------------------------------------------------------------------------
```

Fig. 8. ARIMA Model Results.

Figure 8 applies to the model creation and hyper-parameter tuning used in the auto ARIMA, giving the ideal parameters and best metrics [10. 19].

C. Model Forecasting

Upon completion of the preprocessing and methods, forecasting is now able to occur. The previous information is all regarding the apple stock data. To determine if this model works well with other stock data, the model was trained and tested with the same size datasets for other stocks. This demonstrates the overall effectiveness of ARIMA model when regarding stock data. In addition analysing multiple stock predictions demonstrates general trends in the models faults.

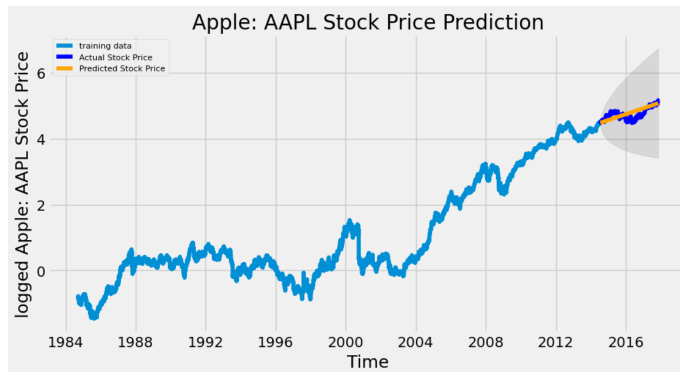

Fig. 9. AAPL Stock Price Prediction.

Figure 9 depicts the forecasted close price of Apple (AAPL) stock data fitted with optimal parameters p, d, and q. The test dataset used kept a 95% confidence level.

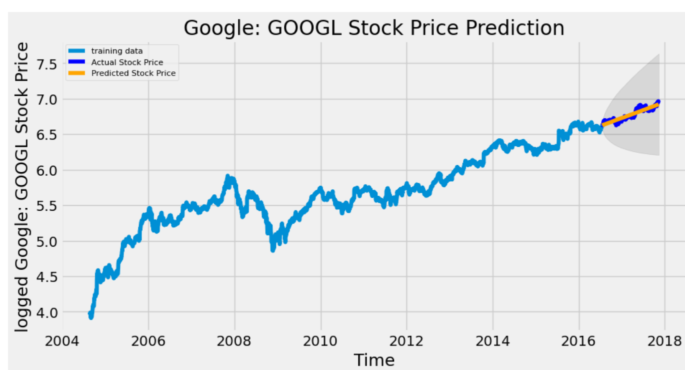

Fig. 10. Google Stock Price Prediction.

Figure 10 depicts the forecasted close price for Google (GOOGL) stock data, this model was fitted with optimal parameters p, d, and q. The test dataset used kept a 95% confidence level.

Fig. 11. Amazon Stock Price Prediction.

Figure 11 depicts the forecasted close price for Amazon (AMZN) stock data fitted with optimal parameters p, d, and q. The test dataset used kept a 95% confidence level.

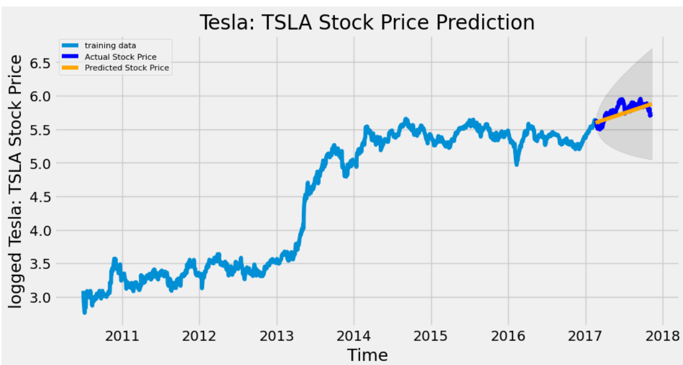

Fig. 12. Tesla Stock Price Prediction.

Figure 12 depicts the forecasted close price for Tesla (TSLA) stock data, this model was fitted with optimal parameters p, d, and q. The test dataset used kept a 95% confidence level.

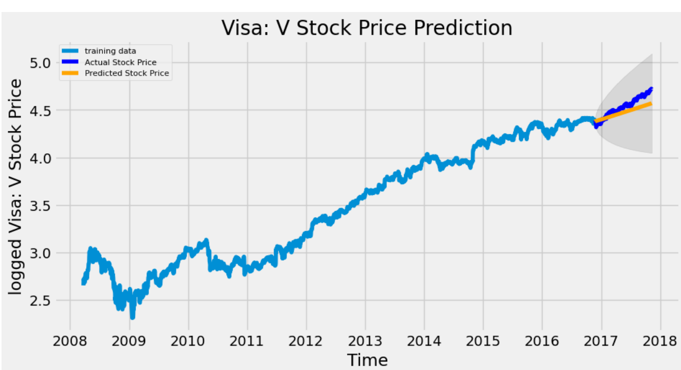

Fig. 13. Visa Stock Price Prediction.

Figure 13 depicts the forecasted close price for Visa (V) stock data fitted with optimal parameters p, d, and q. The test dataset used kept a 95% confidence level.

V. Experimental Results

In this experiment seven measurements are compared to determine the resulting prediction accuracy for the AAPL stock. Each metric addresses prediction accuracy through a different approach.

Mean Squared Error (MSE) is the average of the squared forecast error. Larger errors have more weight on the overall score. MSE is commonly used to evaluate and find ideal models for data. If MSE is similar between models, other measures such as mean absolute error (MAE) are also considered. MSE is extremely sensitive to outliers since they grow quadratically. In the experiment the MSE is roughly 0.021265. MSE validity is not absolute, but a way to interpret the value is a lower value corresponding with the data range is more accurate. For example, for a datum with range 0 to 1000 an MSE of 0.7 is small, but for a datum with range 0 to 1, 0.7 is not considered small anymore. The data range of the logged AAPL stock is from 0 to 8 so an MSE of 0.02 is relatively small which means the prediction forecast in the experiment is quite accurate [20, 21].

MAE is a metric that represents the average of the absolute values of the deviation, which makes it worthwhile to measure prediction errors in the same units as the original series data. It is a very robust measurement which means it is resistant to outliers making it useful when training data is corrupted with outliers that are not likely to be recreated in the future. It is one of the most straightforward regression error metrics to understand, so it is frequently used in data science: the lower the MAE, the more accurate the regression model. The MAE calculated from the ARIMA model on AAPL stock data is 0.116, a very low metric. Thus our prediction model is relatively accurate when measured in terms of MAE [20, 21].

Root Mean Squared Error (RMSE) is directly related to the MSE. It is the square root of the MSE value obtained from the data. As a result of squaring error values to get the MSE, using RMSE instead can make it simpler to interpret the data because it relates to the datum range with a one to one correlation. The RMSE calculated in the experiment is 0.146, which is small when compared to our datum range of 0 to 8, meaning a relatively accurate prediction forecast[20, 21].

Mean Absolute Percentage Error (MAPE) simplifies error measures and makes results easier to understand. It calculates a percentage error which allows a clear idea of the relative error. It is also good to use when comparing different forecasts that may have different scales of data since the percentage error will not be affected by the scale. The experiment resulted in a prediction forecast with a MAPE of 2.48%. This is a relatively low value considering that it represents that every prediction in the resulting experimental forecast is on average 2.48% off of the true value. Therefore this MAPE indicates an accurate prediction forecast [20, 21].

Median Absolute Error (MedAE) is closely related to MAE. It is calculated by taking the absolute differences of the data and then finding the median value. An advantage MedAE has over MAE is that the score allows for missing values. MedAE trims out outliers and results in a reducing the bias that would be in favour of a low forecast. The prediction forecast calculated in the experiment has a MedAE of 0.1091, which is a low value when compared to our datum range of 0 to 8 meaning the prediction forecast is relatively accurate [20, 21].

R-Squared (RSQ) is a statistic in the form of a percentage representing the amount the dependent variable can explain the variance in the independent variable. "In scholarly research that focuses on marketing issues, RSQ values of 0.75, 0.50, or 0.25 can, as a rough rule of thumb, be respectively described as substantial, moderate, or weak." [22].A high RSQ it indicates that the variance of the true values is highly correlated with the variance of the predicted values. A weak value means the two values are not significantly correlated. For the AAPL stock price prediction, our RSQ value is 0.248, a recognizably weak RSQ value indicating no significant correlation between the dependent and independent variables [20, 21, 22].

Explained Variance (EV) is a metric used to describe and measure the difference between actual data and predicted data. EV is the variance in a model that is not the result of error variance and can be explained by factors. Higher percentages of EV indicate higher association strength and better predictions. The EV calculated in the experiment for the prediction model is 29.5%, which is a relatively low explained variance score since this means about 70% of the variance results from error variance and only 30% is a result of actual factors in the true data[23].

TABLE III
RESULTING METRICS

Metric	***Score***
MSE	0.02126501088734349
MAE	0.11625444332131539
RMSE	0.14582527520064378
MAPE	0.02483169448098529
MedAE	0.10909132476230443
RSQ	0.24759461122539284
Explained Varience	0.2947530589675551

VI. CONCLUSION

This paper presents time-series forecasting using ARIMA modelling for stock price prediction. The experimental results achieved within this report demonstrates the potential of ARIMA models when used to predict stock prices on a short term. This report can act as a guide when considering profitable investment decisions. All of the resulting data obtained from the experiment was derived from historic AAPL stock price data. The metrics obtained from the ARIMA model demonstrate an above-average fit of the dataset, indicating the methods used are a reasonable indicator of future stock prices.

REFERENCES

[1] A.A. Adebiyi, A. O. Aderemi, and C. K. Ayo, "(PDF) Stock price prediction using the ARIMA model," ResearchGate. [Online]. Available: https://www.researchgate.net/publication/261179224_Stock_price_prediction_using_the_ARIMA_model. [Accessed: 02-Apr-2021].

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

[3] L.Y. Wei, "A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX", Economic Modelling vol. 33 pp. 893-899, 2013.

[4] G.S. Atsalakis, E.M Dimitrakakis. and C.D. Zopounidis, "Elliot Wave Theory and neuro-fuzzy systems, stock market prediction: The WASP system", Expert Systems with Applications, vol. 38, pp.9196–9206, 2011.

[5] S. K.Mitra, "Optimal Combination of Trading Rules Using Neural Networks", International Business Research, vol. 2, no. 1, pp. 86-99, 2009.

[6] G.S. Atsalakis and P.V. Kimon, "Forecasting stock market short-term trends using a neuro-fuzzy methodology", Expert Systems with Applications, vol. 36, no. 7, pp.10696–10707, 2009

[7] C. Javier, E. Rosario, J.N. Francisco and J.C. Antonio, "ARIMA Models to Predict Next Electricity Price", IEEE Transactions on Power Systems vol. 18 no.3, pp.1014-1020, 2003.

[8] N. Rangan and N. Titida, "ARIMA Model for Forecasting Oil Palm Price", Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, Universiti Sains Malaysia, 2006.

[9] M. Khasel, M. Bijari, and G.A.R Ardali, , "Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and pp. 956-967, 2009.

[10] J. Brownlee, "How to Create an ARIMA Model for Time Series Forecasting in Python," Machine Learning Mastery, 09-Dec-2020. [Online]. Available: https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/. [Accessed: 02-Apr-2021].

[11] B. Marjanovic, "Huge Stock Market Dataset," Kaggle, 16-Nov-2017. [Online]. Available: https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs. [Accessed: 02-Apr-2021].

[12] Workneh, "Stock price prediction using ARIMA Model," Medium, 06-Jun-2020. [Online]. Available: https://medium.com/@derejeabera/stock-price-prediction-using-arima-model-251ddb4ee52a. [Accessed: 02-Apr-2021].

[13] J. Jordan, "Hyperparameter tuning for machine learning models.," Jeremy Jordan, 05-Dec-2018. [Online]. Available: https://www.jeremyjordan.me/hyperparameter-tuning/. [Accessed: 02-Apr-2021].

[14] "Forecasting: Principlesnbsp;andnbsp;Practice (2nd ed)," 8.3 Autoregressive models. [Online]. Available: https://otexts.com/fpp2/AR.html. [Accessed: 02-Apr-2021].

[15] "Forecasting: Principlesnbsp;andnbsp;Practice (2nd ed)," 8.4 Moving average models. [Online]. Available: https://otexts.com/fpp2/MA.html. [Accessed: 02-Apr-2021].

[16] Stephanie, "ARMA model," Statistics How To, 14-Jan-2021. [Online]. Available: https://www.statisticshowto.com/arma-model/. [Accessed: 02-Apr-2021].

[17] Auquan, "Time Series Analysis for Financial Data IV- ARMA Models," Medium, 19-Dec-2017. [Online]. Available: https://medium.com/auquan/time-series-analysis-for-finance-arma-models-21695e14c999. [Accessed: 02-Apr-2021].

[18] R. Nau, in Lecture notes on forecasting.

[19] Nageshsingh, "Stock market forecasting/ARIMA," Kaggle, 02-Jan-2021. [Online]. Available: https://www.kaggle.com/nageshsingh/stock-market-forecasting-arima. [Accessed: 02-Apr-2021].

[20] E. Woschnagg and J. Cipan, rep. https://homepage.univie.ac.at/robert.kunst/procip.pdf
[21] J. Bhattacharjee, “Common metrics for Time Series Analysis,” Medium, 29-Dec-2019. [Online]. Available: https://joydeep31415.medium.com/common-metrics-for-time-series-analysis-f3ca4b29fe42. [Accessed: 02-Apr-2021].
[22] E. Mooi, M. Sarstedt, and R. Gorsuch, “(PDF) A concise guide to market research: The process, data, and methods using IBM SPSS Statistics. New York: Springer,” ResearchGate, 14-Oct-2010. [Online]. Available: https://www.researchgate.net/publication/260192478_A_concise_guide_to_market_research_The_process_data_and_methods_using_IBM_SPSS_Statistics_New_York_Springer. [Accessed: 03-Apr-2021]
[23] S. Glen, “Explained Variance / Variation,” Statistics How To, 08-Jun-2020. [Online]. Available: https://www.statisticshowto.com/explained-variance-variation/. [Accessed: 02-Apr-2021].