New Energy Vehicles Industry Stock Price Prediction Based on ARIMA: Tesla, NIO and BAIC BluePark

Runxin Li *

Department of Finance, Capital University of Economics and Business, Beijing, China

*Corresponding author: lirunxin@cueb.edn.cn

Abstract. Stock price prediction plays a main role and serve as a key for quantitative analysis and investment in financial field. With the boosted of various state-of-art statistical approaches as well as computing strategies, various forecasting scenarios have been proposed. Among numerous prediction schemes, ARIMA is one of widely implemented and explored models. On this basis, this study presents extensive process of predicting stock price based on the ARIMA model. The stock prices of Tesla, NIO and BAIC BluePark in new energy vehicles industry are collected from investing.com and money.com and used for prediction. The parameters of ARIMA model which is p, q and d ranging from 0 to 4. According to the analysis, it is feasible for short-term forecasting in new energy vehicles industry and is competitive with other forecasting models in helping both institutions and individuals' investment. To be specific, the predicted stock price is close to the actual data and institutions and individuals can make profitable decisions based on the results. Overall, these results shed light on guiding further exploration of price forecasting of underlying assets.

Keywords: Forecasting; ARIMA; new energy vehicles.

1. Introduction

Forecasting of the price for certain underlying assets remain a key issue for quantitative finance, which have been discussed in a large number of studies. It is considered as one of the hardest problems [1, 2] owing to the stochastically and non-stationary features [3, 4]. This motives many researches to proposed new scenarios, such as artificial neural network model (ANNs). Thereinto, ARIMA is a widely adopted and useful models even nowadays [5]. It is from statistical view and is famous on account of the advantages from the statistical view. Thus, ARIMA is chosen to predict stock price of Tesla, NIO and BAIC BluePark in this paper. The new energy vehicles industry is booming recent years. The new energy market will have a strong prospect in the future due to the advantage of eco-friendly. Taking China as an example, the government has introduced several encourage policies [6]. There will be a great chance for investors to make profit in new energy vehicles stock market. Thus, predicting the stock price of new energy vehicle companies is meaningful and significant.

There three well-known corporations in the new energy vehicle industry (i.e., Tesla, Nio and BAIC), which are the investigation target of this study. Tesla Inc. is an electric vehicle and energy company, which has been a global giant in new energy vehicles industry nowadays. NIO is a smart electric vehicles brand, registered and established in Shanghai, China on November 25, 2014. BAIC BluePark New Energy Technology Co., Ltd. is a high-tech listed company controlled by BAIC Motor Group Co., Ltd. The company is a representative of Chinese traditional car company which transform to new energetic vehicles area and is also one of the leading enterprises in China. Its main business includes design, manufacturing, and sales services of BEV (Battery Electric Vehicle). The company is currently making efforts to build its two major brands: ARCFOX and BEIJING.

Ariyo et al. revealed that ARIMA can compete favorably with existing methods for daily stock price prediction [7]. Dong et al. present extensive process of forecasting daily stock price based on the ARIMA method and stock data of Apple Inc. The results shows that the ARIMA model has a strong potential for short-run forecast and has some advantages compared with other prediction methods [8]. Wadi et al. utilize ARIMA model to forecast closed price, which has good results for short-term prediction [9]. Huang applies the ARIMA model to analyze and predict stock price of China Merchants Bank, based on R Studio, which provides some suggestions for investors [10]. Liu and Zhang forecast future stock price of Southeast Securities with Python. Results revealed that that

the difference of predicted and real closed stock price is no larger than 0.04, showing that ARIMA is a good way to predict future stock price in short-term and is helpful for investors while making decisions [11].

This paper will utilize ARIMA to predict the selected 3 companies. The results based on the data from Tesla, NIO and BAIC BluePark stock price demonstrated the availability for prediction of ARIMA. This paper is organized as follows: Section 2 presents data and methodology used while section 3 discusses the results obtained. Limitations and future outlooks are presented in section 4. Finally, in section 5 is the conclusion.

2. Data and Methodology

2.1 Data

Daily stock closing price is chosen to build ARIMA model and prediction. The stock price time series data of Tesla is obtained from investing.com and the data of NIO and BAIC BluePark is from money.com. The train set and test set is divided by 8-2 for all the three companies. For Tesla, stock price data of train set starts from 3rd, September, 2018 to 12th, January, 2022, 840 in total, and the test set is from 13th, January, 2022 to 11th, November, 2022, 210 in total. For NIO, stock price data of train set starts from 13th, September, 2018 to 12th, January, 2022, 840 in total, and the test set is from 13th, January, 2022 to 11th, November, 2022, 210 in total. For BAIC BluePark, stock price data of train set starts from 27th, September, 2018 to 18th, January, 2022, 803 in total, and the test set is from 19th, January, 2022 to 18th, November, 2022, 201 in total. The daily stock price data is shown in Fig. 1, Fig. 2 and Fig. 3, respectively for Tesla, NIO and BAIC BluePark.

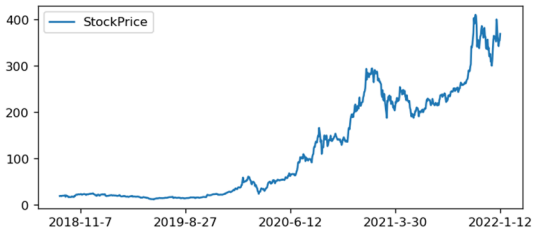

Fig. 1 Stock Price Data of Tesla.

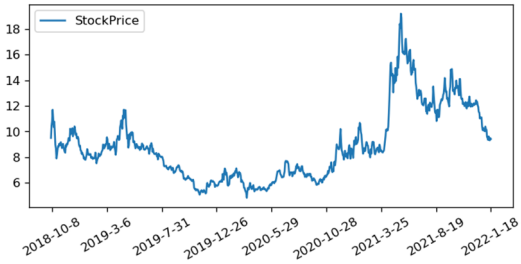

Fig. 2 Stock Price Data of NIO.

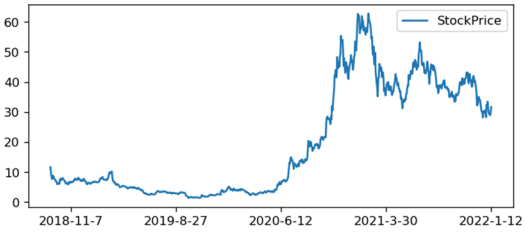

Fig. 3 Stock Price Data of BAIC BluePark.

2.2 The ARIMA Model

The ARIMA is originated from the ARMA, the difference between the two is: ARMA is only available to be used to analyze stationary time series, while ARIMA model can be used to cope with non-stationary time series sequences. The ARIMA model is established after the differential process

of the time series sequence. In practice, time series are mostly non-stationary sequences, therefore, the ARIMA model is more widely used. In the ARIMA(p, d, q) model, AR represents autoregression and p is the order of autoregression; MA represents the moving average and q is the number of moving average terms; d is the number of differences when the data become stationary time series sequences:

$$\Delta d y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (1)$$

Here, y_t is the actual stock price time series sequence and ε_t is the random error at t, representing white noise process, φ_i and θ_j are the coefficients, p and q are orders of the model which need to be determined, Δdy represents the stationary time series sequence derived by d times of differential transformation of the initial non-stationary sequence y_t.

2.3 Evaluations

To evaluate the ARIMA model, R-square and MSE are chosen as the metrics. R-square, the coefficient of determination, measures the proportion of the dependent variable which can be explained by the independent variable through the fitness. R-square is a number ranging from 0 to 1, the closer to 1, the better the model is. It is often regarded as an effective indicator to estimate the effectiveness of the model. Besides, MSE, the standard error of estimate, measures the spread of real-world observations around the predicted line. The lower of MSE, the less difference between the predicted value and actual value. Through these two indicators, the ARIMA model can be appraised reasonably and the efficiency can be tested.

3. Results and Discussion

3.1 Tesla

3.1.1 ADF test

In section 2.1, the original stock price data of Tesla is visualized through Python. By observing the figure, the author find that the stock price time series data of Tesla presents non-stable characteristics, but it is far from enough for just observing the data. Therefore, a unit root test on the original data is performed and concludes that the t-statistic of the ADF unit root test is 0.614, which is less than the critical values of -3.438, -2.865 and -2.569 (1%, 5%, and 10% SL), respectively. It can be obtained that the ADF test results of the original data fall within the acceptance null hypothesis interval, that is, the stock price of Tesla is non-stationary. The non-stationary of the time series data can be solved by the difference method, so the first-order differentiation on the original data is conducted and the ADF test results is -6.271, which is greater than the critical values of -3.438, -2.865 and -2.569 (1%, 5%, and 10% SL), respectively. Besides, the p-value is 0.000, largely less than 0.05. Therefore, both t-statistic and p-value suggest that after the first-order differential processing of the original data, the stock price time series data of Tesla is no longer non-stationary and can be forecasted based on the ARIMA model.

3.1.2 Parameters estimation

Parameter estimation is selecting the proper value of parameters to select the best ARIMA model. In the last section, after the first-order differential processing, the stock price is stationary, so d, the number of differences, is 1. To determine the value of p and q, ACF and PACF, as well as BIC matrix has been introduced. According to Fig. 4 ACF and PACF of Tesla, the parameters of p and q can be used 1 and 1 respectively. From Fig. 5, the BIC value is the lowest when (p, q) is (1,0) or (2,0). Since the predicted result is better for ARIMA(1,1,1), it is chosen as the fitted model to forecast stock price.

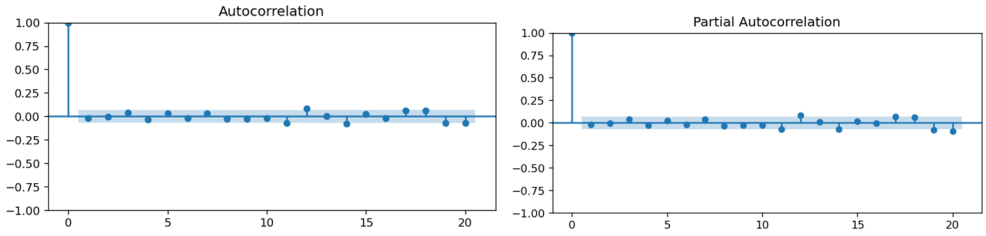

Fig. 4 ACF and PACF of Tesla

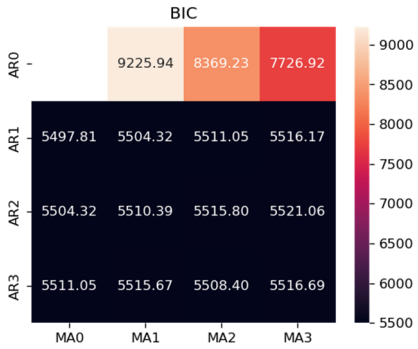

Fig. 5 BIC matrix of first-order differentiation stock price of Tesla.

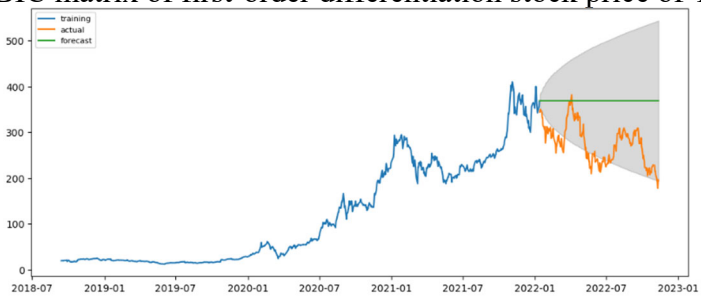

Fig. 6 Predicted and Actual Stock Price of Tesla

3.1.3 Prediction results

Based on the ARIMA(1,1,1), the author derive the forecasting result, as it is shown in Fig. 6. It suggests that the results of stock price prediction indicates almost stationary in the following days, but the actual result is a slight decrease, so the predicted result isn't very good. This may be attributed to the relatively large change and volatility in stock price of Tesla. But in general, the ARIMA (1,1,1) model is able to make stock price prediction.

3.2 NIO

3.2.1 ADF test

Similar to Tesla, the original stock price data of NIO is also visualized through Python in section 2.1, and the observation result is the same that the stock price time series data of NIO presents non-stable characteristics. A unit root test on the original data is performed and concludes that the t-statistic of the ADF test is -1.068554, which is less than the critical values of -3.438, -2.865 and -2.569 (1%, 5%, and 10% SL), respectively. It means that the stock price time series data of NIO is non-stationary. After the first-order differentiation on the original data, the ADF test results is -5.886, larger than the critical values of -3.438, -2.865 and -2.569 (1%, 5%, and 10% SL). Besides, the p-value is 0.000, largely less than 0.05. Therefore, both t-statistic and p-value suggest that after the first-order differential processing of the original data, the stock price time series data of NIO is no longer non-stationary and can be forecasted based on the ARIMA model.

3.2.2 Parameters estimation

For stock price time series data of NIO, the second-order differentiation is better for predicted result, so d is chosen as 2. To determine the value of p and q, similar to previous processing, ACF and PACF, as well as BIC matrix has been introduced. According to Fig. 7 ACF and PACF of NIO, and Fig. 8 BIC matric, combing with the predicted result, the ARIMA(2,2,2) is the best model for forecasting.

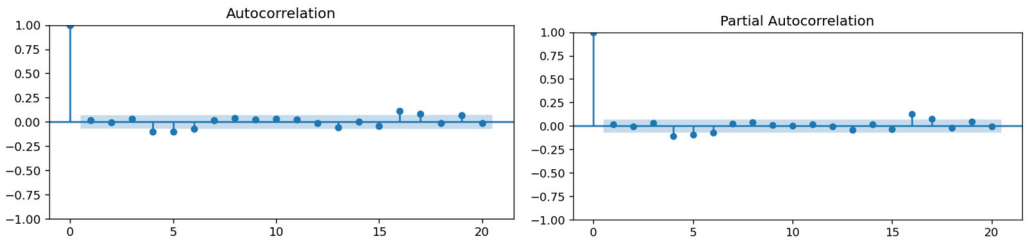

Fig. 7 ACF and PACF of NIO

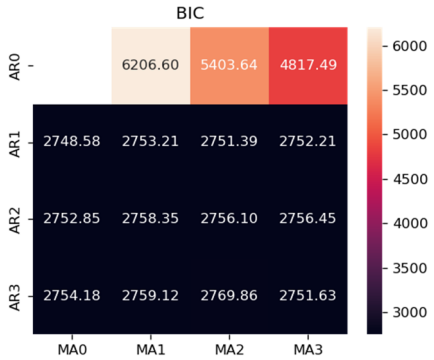

Fig. 8 BIC matrix of first-order differentiation stock price of NIO

3.2.3 Prediction results

Based on the ARIMA(2,2,2), the predicted result is acquired and visualized in Fig. 9, which shows that the predicted stock price is slowly decease, whose trend is same to the actual stock price data. Therefore, the forecasting result is quite good, which is a proper way to make prediction and can provide useful suggestions for investors.

Fig. 9 Predicted and Actual Stock Price of NIO

3.3 BAIC BluePark

3.3.1 ADF test

The same processing method is used on BAIC BluePark stock price time series data. A unit root test on the original data is performed and concludes that the t-statistic of the ADF test is -1.930, which is less than the critical values of -3.438, -2.865 and -2.569 (1%, 5%, and 10% SL), respectively. It means that the stock price time series data of BAIC is non-stationary. After the first-order differentiation on the original data, the ADF test results is -5.810, which is greater than the critical values of -3.439, -2.865 and -2.569 (1%, 5%, and 10% SL), respectively. Besides, the p-value is 0.000, largely less than 0.05. Therefore, both t-statistic and p-value suggest that after the first-order

differential processing of the original data, the stock price time series data of BAIC BluePark is no longer non-stationary and can be forecasted based on the ARIMA model.

3.3.2 Parameters estimation

In the last section, after the first-order differential processing, the stock price is stationary, so d, the number of differences, is 1. Then, ACF and PACF, as well as BIC matrix has been introduced to select the value of parameters p and q. According to Fig. 10 ACF and PACF of BAIC BluePark, the p and q can be chosen as 1 and 1 respectively. From Fig. 11, the BIC value is the lowest when (p, q) is (1,0). Taking all into account, ARIMA(1,1,0) is chosen as the fitted model to forecast stock price.

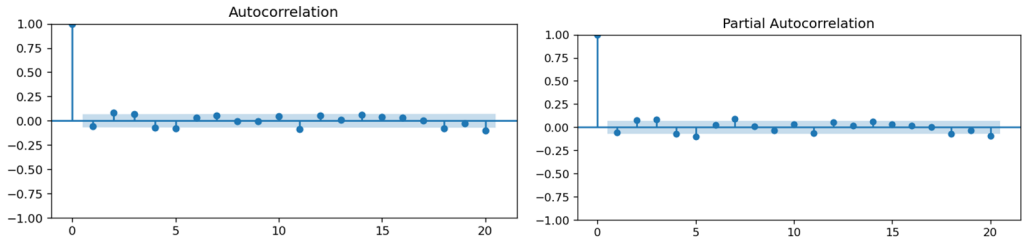

Fig. 10 ACF and PACF of BAIC BluePark

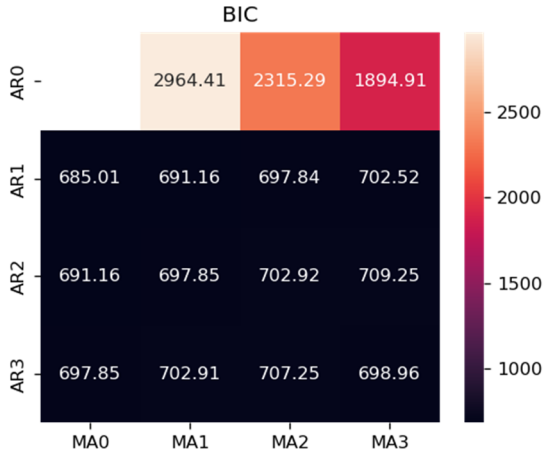

Fig. 11 BIC matrix of first-order differentiation stock price of BAIC BluePark

3.3.3 Prediction results

Based on the ARIMA(1,1,0), the forecasting result is demonstrated in Fig. 12, which shows that the overall predicted value is slightly higher than actual stock price value. But in general, the predicted result is good and is helpful for investment decisions.

Fig. 12 Predicted and Actual Stock Price of BAIC BluePark

3.4 Comparison and Suggestion

Based on the predicted results of the three companies, a comparison can be made and some suggestions can be provided to investors. The ARIMA model can be used to predict daily stock price

for all the three companies in new energy vehicles industry. Among the three, NIO has the best forecasting results, the predicted value is close the actual stock price, which is quite accurate for prediction. While for Tesla, the result isn't very well, this can be ascribed to the relatively large change in stock price and limited data sample. But overall, the predicted results are good and meaningful for investment decisions. Therefore, the author can conclude that it is available to predict daily stock price in new energy vehicles industry and can provide useful suggestions for investors to make profitable investment decisions.

4. Limitations and Future Outlooks

Nevertheless, this paper has some shortcomings and drawbacks. First, the data sample used is still limited and the time frame maybe not long enough. It may affect the accuracy of the stock price forecasting based on ARIMA model. Second, the ARIMA model selected maybe isn't the best. In this paper, only ACF and PACF method are chosen to determine p and q, but other approaches maybe can derive better parameters estimation results. For further study, a large data sample can be employed. By choosing a longer time frame, more numbers of data can be collected and therefore can increase the accuracy of prediction results. Moreover, the determination of p and q can be further discussed. However, many methods can be used to choose the best ARIMA model (e.g., AIC and BIC). The forecasting results may be better than this paper based on some other approaches.

5. Conclusion

In summary, this paper investigates new energy vehicles industry stock price prediction based on ARIMA model. To be specific, the stock price of Tesla, NIO and BAIC BluePark has been predicted based on the daily stock price from 2018 to 2022. According to the evaluation, it has a strong potential for short-run forecasting especially for relatively stable stock price time series, for example, NIO. In addition, the predicted result maybe not very good for new energy vehicles companies whose stock price fluctuation is relatively large. Nevertheless, this paper still has some shortcomings and drawbacks that the data sample used is limited and the way to select ARIMA model maybe isn't the best. In the future, a larger data sample can be employed and several approaches can be chosen and compared to determine the parameters of ARIMA model. Overall, these results will offer a good guideline for earning profit in financial markets of a specific industry.

References

[1] Pai P F, Lin C S. A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 2005, 33(6): 497-505.

[2] Wang J J, Wang J Z, Zhang Z G, et al. Stock index forecasting based on a hybrid model. Omega, 2012, 40(6): 758-766.

[3] Tabachnick B G, Fidell L S. Using multivariate statistics, 4th ed., Person Education Company, USA, 2001.

[4] Ahmed R A, Shabri A B. Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and support vector machines. American Journal of Applied Sciences, 2014, 11(3): 425.

[5] Kyungjoo L C, Sehwan Y, John J. Neural network model vs. SARIMA model in Forecasting Korean Stock Price Index (KOSPI). Issues in Information System, 2007, 8(2): 372-378.

[6] Xiang Z, Jia W, Jianzhong Y, et al. Prospects of new energy vehicles for China market. Hybrid & Eco Friendly Vehicles Conference 2008 (HEVC 2008), 2008.

[7] Ariyo A A, Adewumi A O, Ayo C K. Stock price prediction using the ARIMA model. UKSim-AMSS 16th international conference on computer modelling and simulation. IEEE, 2014: 106-112.

[8] Dong Y, Li S, Gong X. Time series analysis: An application of ARIMA model in stock price forecasting. 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017). Atlantis Press, 2017: 703-710.

[9] Wadi S A, Mohammad A, Ahmed A A. Predicting Closed Price Time Series Data Using ARIMA Model. Modern Applied Science, 2018, 12: 11.

[10] Huang S. Stock price analysis and prediction based on the ARIMA model: Taking China Merchants Bank Inc. as an example. Management & Technology of SME, 2022, 11:184-187.

[11] Liu S, Zhang S. An empirical study of stock price prediction using the ARIMA model. Economic Research Guide. 2021, 25: 76-78.