Forecasting EV Stock Trends Based on ARIMA Model Represented by Tesla and BYD

Kaiyue Yang*
Department of Mathematics and Department of Economics, Brandeis University, Waltham, MA, 02453, USA
*Corresponding author: kaiyueyang@brandeis.edu

Abstract. In recent years, an increasing number of people have turned their attention to renewable energy, and sales of electric vehicles (EVs) have skyrocketed, making EVs a very popular topic. To better understand the development of the electric vehicle market and provide some insights for investors, this study forecasts EV stock trends of two leading companies in the Electric Vehicle industry, Tesla and BYD, by the Autoregressive Integrated Moving Average (ARIMA) model. The data in this research includes a daily time series of stock prices. ARIMA model is chosen in this research for its effectiveness in predicting stationary time series data, particularly suitable given the limited data available in the recently emerged EV industry. This research demonstrates that both Tesla and BYD's stock closing prices are not stationary data. After differencing, for Tesla, the predicted result suggests a slightly increasing trend, while BYD's prices are predicted to remain at the same level and relatively stable in future 100 days. The study provides insights into the EV stock market, indicating a potential upward yet not pronounced trend in the future. But such a stable price might be a good choice for the investor low-risk takers.

Keywords: renewable energy; electric vehicles; stationary time series data; EV stock market; ARIMA model.

1. Introduction

As global warming becomes increasingly severe, the issue of carbon dioxide emissions has garnered heightened attention. Roughly thirty percent of the nation's greenhouse gas (GHG) emissions come from transportation. Significant modifications to the transportation fleet have been contemplated as a means of lowering greenhouse gas emissions, and electric vehicles are one such solution [1]. Unlike traditional cars, electric vehicles run on gas or fuel. As a result, greenhouse gases are not produced by electric vehicles. Because they are so fuel-efficient, electric cars assist drivers in using less fuel [2]. For many consumers, electric vehicles also prove to be more cost-effective. Making the switch to an electric vehicle will save $300 a year for every $1 increase in gasoline prices [3]. In 2022, global expenditure on electric vehicles surpassed USD 425 billion, a 50% increase from 2021. Consumers accounted for the remaining 90% of spending, with government support accounting for just 10% of the total [4]. For investors, a flourishing industry is often mirrored in the stock market. The stock price is decided by the public market. Therefore, especially for the industry that is flourishing with consumers, the stock price implies the public consuming tendency. Among all EV brands, Tesla leads the US market and is a dominant player in the global BEV market. Among all automakers, Tesla had the biggest market share for BEVs in 2018, which takes 11% of global BEV sales [5]. Besides the US market, China's EV market is flourishing rapidly. The Chinese government has implemented several incentive programs since 2005 to encourage the growth of the electric vehicle (EV) market [6]. In 2018, China sold 1.2 million plug-in electric vehicles (PEVs), accounting for 56% of PEV sales worldwide [7]. In 2023, Tesla began slashing pricing in January and has continued to do so ever since. BYD and other Chinese EV manufacturers dropped prices in reaction to Tesla and a general production decrease [8]. These actions further promote sales and concurrently impact the stock market. Therefore, many investors turn their attention to Tesla and BYD and their stock prices. Beyond these macroeconomic backgrounds, investors are more interested in the future trajectory of stocks, as this relates directly to their investment returns. This research is focused on
Tesla and BYD, the most representative brand of electric vehicles in China, contributing to the understanding of EV market dynamics and offering a valuable perspective for investors.

2. Data and Method

2.1. Data Selection

There are two groups of time series data: Tesla stock price and BYD stock price, which are collected from Yahoo Finance. These two companies are leading companies in the EV industry. Tesla is widely regarded as the leader in the EV market around the world, especially in the United States and Europe. Tesla is at the forefront of EV technology and innovation. As a newly emerging Chinese company, BYD has gained significant popularity in recent years. Its diversity and rapid advancement in technology enable it to emerge as a formidable challenger to Tesla and, concurrently, a prominent representative in the electric vehicle sector. According to the statistics of Best-selling plug-in electric vehicle models worldwide in 2022, the x-axis represents the models and the y-axis represents the sales in thousand units. As shown in Fig. 1, 8 models are Tesla and BYD models among the top 10 best-selling models [9]. Therefore, Tesla and BYD can be seen as the representative of the EV market. By analyzing and forecasting the performance of these two stocks, a general projection and outlook can be provided for the entire electric vehicle market.

![Fig. 1 Best-selling plug-in electric vehicle models worldwide in 2022 [9]](image)

2.2. Tesla

The time range of Tesla is from Jun 28, 2010, to Oct 31, 2023, and the frequency is daily. Opening price, highest price, lowest price, closing price, adj close price, and volume for each day are included in the dataset. Since the closing price is the final traded price of the asset for the day, which reflects the market sentiment, the closing price is plotted and analyzed. The horizontal axis represents time, ranging from the year 2014 to just beyond 2022. As shown in Fig. 2, the vertical axis represents the closing price over time. Starting from 2020, the stock price dramatically increased until 2022. And then there’s a significant drop from 2022 to 2023 and a rebound after 2023. By Oct 31, 2023, the price stayed around 200 dollars per share.
2.3. BYD

The time range of BYD is from Feb 12, 2009, to Oct 31, 2023, and it’s also daily data. Similar to Tesla, opening price, highest price, lowest price, closing price, adj close price, and volume for each day are also included in the dataset, and the focus is on the closing price. Based on Fig.3, around the same time as Tesla, in 2020, BYD's stock price experienced rapid growth. However, BYD exhibits greater volatility compared to Tesla and displays a certain degree of cyclicality. The price of the stock is significantly lower than Tesla's, and by Oct 31, 2023, the closing price was around 30 dollars per share.

2.4. Method

Both the Tesla and BYD stock price data are time series. Time series always exhibit seasonal variations and they are time-dependent. Over the period, there’s a trend displayed by time series data. Forecasting on time series data helps identify causal relationships and future decision-making. One important statistical method used for predicting time series data is the Autoregressive Integrated Moving Average (ARIMA) model. It’s effective in predicting stationary time series data, and it can predict based on a relatively small amount of data. Due to the recent emergence of the electric vehicle industry in the past few years, there is limited data available. Therefore, in comparison to machine learning models that require extensive training and testing data, the ARIMA model is a relatively better choice.
3. Result

3.1. Tesla

3.1.1 Stationarity

Before building the ARIMA model to complete the prediction, one important step is to test the data stationary. As shown in Fig. 4, Tesla’s mean and standard deviation exhibit an upward trend. However, the overall plot does not consistently demonstrate an upward trend. Before 2020, the mean and standard deviation remained stable. Yet, from around 2020 to 2022, the curve began to rise sharply. Following 2022, the trajectory initially declined, before rebounding with an upward surge. On the other hand, the results of the Dickey-Fuller test revealed a p-value of 0.697, which is an indicator reflecting the data stationarity.

Fig. 4 Tesla rolling mean and standard deviation [original]

3.1.2 ARIMA model

In the stock price forecast plot, the price remains at a low value and is stable from 2010 to about 2019. As shown in Fig.5, the red line represents the historical closing price, and the blue line in the shaded area exhibits the prediction trend on Tesla’s closing price in the next 100 days. And the shaded area represents the confidence interval, which is the uncertainty, of the forecast.

Fig. 5 Tesla price forecast in future 100 days [original]

Starting from 2019, the price dramatically increase until 2022. The highest price reaches approximately 400 dollars per share. From 2022 to Oct 31, In 2023, the price dropped to approximately 100 dollars per share and rebounded to about 200 dollars per share. The blue straight line is slightly increasing with shaded. According to the ARIMA Model, it predicts that in the next 100 days, the closing price of Tesla might slightly increase.
3.2. BYD

3.2.1 Stationary and differencing

As shown in Fig. 6, similar to Tesla, BYD's mean and standard deviation also exhibit an upward trend. BYD exhibits a similar stable trend as Tesla before 2020. But after 2020, the mean and standard deviation of BYD is more fluctuated, compared to Tesla. Especially in the span from 2021 to October 31, 2023, BYD's mean and standard deviation exhibited considerable oscillations. The numerical indicator, p-value, resulting from the Dickey-Fuller test, is 0.811.

Fig. 6 BYD rolling mean and standard deviation [original]

3.2.1 ARIMA model

Based on the red line of Fig. 7, there is a noticeable upward trend starting around the beginning of 2020, where the values sharply increase. Before this point, the values fluctuate but do not display a clear long-term trend. Compared to Tesla, there are more ups and downs. However, the magnitude of price fluctuations for BYD was relatively modest, primarily oscillating around 30 dollars per share. In the future 100 days, the blue line, which is the predicting line, is horizontal. It indicates that the price might remain the same.

Fig. 7 BYD price forecast in the future 100 days [original]

4. Discussion

4.1. Tesla

Based on the stationary analysis, the increasing mean and standard deviation indicate that this group of data is not stationary. On the other hand, while testing, the null hypothesis was set to be this group of data is stationary. However, the high p-value, 0.697, is greater than 0.05. So the null hypothesis is rejected. Numerically, the data is non-stationary. To apply the ARIMA model for forecasting, it is imperative to ensure data stationarity. Therefore, for Tesla's stock data, differencing is both necessary and crucial. The ARIMA model is described as (2, 2, 2) for Tesla. (2, 2, 2) represents
three parts of ARIMA: AutoRegressive (p), Integrated (d), and Moving Average (q). It indicates that the model uses the past two values or lags of the series to predict the current value, and needs to be differenced twice to make it stationary. The last “2” means the error terms of the past two predictions are used to predict the current value. Reflecting on the plot, as indicated by the blue line, the forecasted trend is expected to continue to be volatile, with a slight downward trend in the next 100 days. However, there’s a widening grey-shaded confidence interval, which reflects that the forecasted values are less certain. The actual high price might vary more significantly from the predicted average.

4.2. BYD

The increasing and fluctuating mean and standard deviation of BYD also indicate that it is not stationary. The p-value from the Dickey-Fuller test is 0.811, which is greater than 0.05. The null hypothesis was rejected, proving that this group of data is non-stationary. For the red line, which is the historical closing price, BYD’s trend can reflect more than Tesla’s. The values after the sharp increase seem to plateau with smaller fluctuations around a new level, suggesting that whatever caused the sharp rise has stabilized to some extent. So less differencing is needed for BYD. The forecasted values are provided by the ARIMA(0, 1, 0) model. (0, 1, 0) suggests that the forecast is essentially a naive or random walk model, taking only 1 time differencing. It does not take into account any past values (as p=0) or any moving average components (as q=0). The model simply uses the last observed value and assumes that future values will vary around it with a certain amount of random noise, which is reflected in the wide confidence interval. The blue line is flat, suggesting that the model predicts that the future values of the series will continue at around the same level as the last observed value. Even though the predicting line is horizontal, it does not show the fluctuation during the 100 days. So it might still oscillating in future 100 days but in general, it keeps the same or similar value at the end of the 100 days. This prediction is also uncertain because the shaded area is still wide.

5. Conclusion

Based on the ARIMAL model’s prediction, Tesla might display a slight increasing trend and BYD might continue at around the same level after Oct. 31, 2023. In general, Tesla and BYD show that in future 100 days, the EV market may exhibit an upward trend, yet such a tendency is not pronounced and it is highly probable that it will maintain a level similar to that of October 31, 2023. For investors seeking stability, the EV market may present a favorable option. This research provides an outlook on the EV stock market, offering valuable insights for investors. Given the prominence of electric vehicles as a trending topic in recent years, this study facilitates further exploration of the EV market. Future research could expand upon this by conducting more comprehensive background analyses of various EV stocks. However, as the ARIMA model primarily reveals general trends and the results indicate considerable uncertainty, future studies could attempt to employ a wider range of models for more precise forecasting.

References


