Proceedings of ICEMGD 2025 Symposium: Innovating in Management and Economic Development
DOI: 10.54254/2754-1169/2025.LH23955

Analysis of the Walt Disney Company's Stock Using the ARIMA Model

Zejun Ni

East China University of Science and Technology, Shanghai, China
22013162@mail.ecust.edu.cn

Abstract: Nowadays, the entertainment industry, especially major companies like The Walt Disney Company, faces rapid market changes, and their stock prices are hugely influenced by global trends. Therefore, understanding and forecasting stock price movements has become crucial for investors and experts. This study uses the Auto Regressive Integrated Moving Average (ARIMA) model to analyze and forecast the stock price of the Walt Disney Company. Using daily stock data from September 2024 to March 2025, the research identifies that the original series is non-stationary and achieves stationarity after second-order differencing. Subsequently, based on an analysis of the Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF), coupled with model selection employing the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), ARIMA(0,2,1) was identified as the best model. The model passes residual and coefficient significance tests, and its short-term prediction yields a Mean Absolute Percentage Error (MAPE) of only 1.85%, indicating high accuracy. This paper demonstrates that ARIMA provides a effective method for short-term forecasting in industry stocks, while future improvements could also integrate more variables and hybrid models.

Keywords: ARIMA, Disney, stock price, forecasting

1. Introduction

As one of the biggest Hollywood production companies, The Walt Disney Company has been influencing the global entertainment industry. The price of its stock is affected by tons of factors, such as the watching habits of audience, trends in the film production industry, developments in technology applied to production, competitive dynamics, broader market movements and so on.

Recent years have seen the rapid changes of the whole entertainment industry. The rapid rise of streaming platforms and fluctuation in box office have caused increased volatility in Disney's stock price. Therefore, it is important to understand and forecast these fluctuations so that financial professionals and investors can make better and reliable decisions.

While there are plenty of methods that can be applied to analyze the stock data, the ARIMA model is chosen for this paper because of its strong performance and effectiveness in financial time series forecasting. Owing to its function for making short-term predictions, ARIMA model is considered a widely adopted tool in both economic and financial analyses[1, 2]. Traditional forecasting models are valued for their simplicity and reliability when working with historical data [3]. Although machine learning methods like LSTM are powerful, they often require larger datasets and deal better with

DOI: 10.54254/2754-1169/2025.LH23955

long-term and nonlinear patterns [4, 5]. On the contrary, ARIMA models are considered to be reliable for concise and precise short-term predictions.

Indeed, some studies have already applied time series models to perform stock price prediction, but many failed to adequately address the challenges of model identification and stationarity, or they neglect the importance of residual diagnostics and forecast accuracy evaluation. The aim of this study is to fill those gaps by applying the ARIMA model to Disney's stock data.

This paper is organized in the following manner: basic explaination of ARIMA model, and then step-by-step modeling process, finally the conclusion and discussion for future analysis.

2. ARIMA model

Time series models analyze historical data in order to forecast future values [6]. Various financial indicators like stock prices and exchange rates can be represented according to time series data. These models aim to find the patterns of data and make prediction to the future values based on historical data.

Due to its effectiveness, the ARIMA model sees widespread application in addressing non-stationary data. The model is characterized by three fundamental parameters: p, d, and q. Specifically, the parameter p specifies the number of autoregressive terms, d represents the degree of differencing needed to make the time series stationary, and q corresponds to the number of lagged forecast errors incorporated in the prediction formula. Therefore, the ARIMA model is considered to consist of three principal components:

2.1. Autoregressive (AR)

This part explains the variable using its own lagged values. The AR(p) model is given by:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t, \quad (1)$$

where X_t is the value at time t, ϕ are parameters, ε_t is white noise.

2.2. Moving Average (MA)

Associated with the variable, the error is modeled in this section as a linear combination of its historical error terms. The MA(q) model is formulated by the following equation:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \quad (2)$$

where μ is the mean of the series, $\theta_1 ... \theta_n$ are the parameters, ε_t is white noise.

2.3. Integrated (I)

This part involves differencing the series to make it stationary. For example, first-order differencing is:

$$\Delta X_t = X_t - X_{t-1} \quad (3)$$

If this is d-order differencing, it can be given as $\Delta^d X_t$.

By integrating these three components, the complete ARIMA (p, d, q) model can be formulated as:

$$\Delta^d X_t = c + \phi_1 \Delta^d X_{t-1} + \cdots + \phi_p \Delta^d X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}. \quad (4)$$

This formulation makes ARIMA both powerful and flexible, allowing it to effectively model many real-world time series data.

DOI: 10.54254/2754-1169/2025.LH23955

3. Model building

3.1. Data collection

Reliable financial data sources are used to acquire the data on Disney's stock price. The dataset is in .xlsx format and contains daily data from September 30, 2024 to March 28, 2025. It can be shown as plot in R. Figure 1 below shows the dataset. To assess whether the data is stable or not, Augmented Dickey-Fuller (ADF) test is performed, then the p-value returns as 0.44, which indicates that the data is unstable because p-value is more than 0.05. For its non-stationarity, data preprocessing would be required.

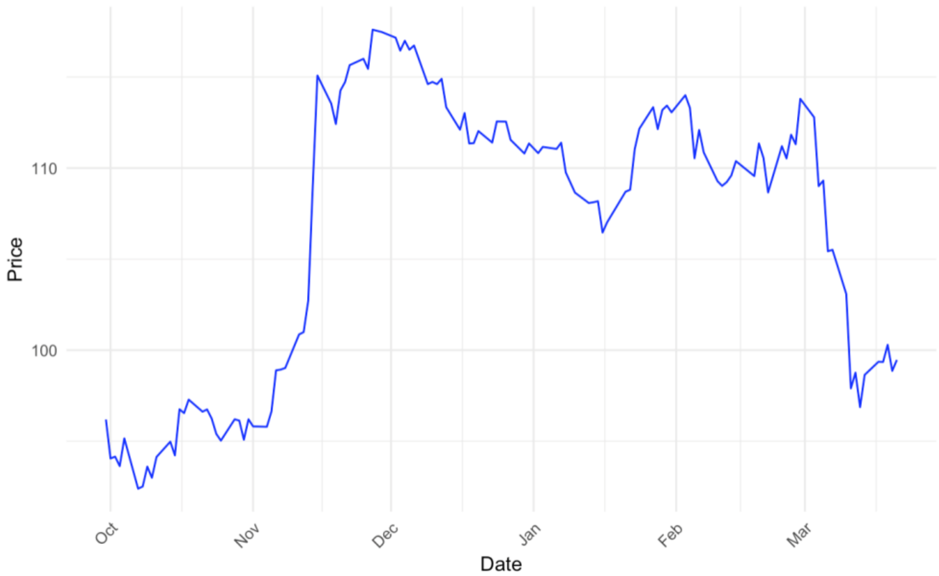

Figure 1: Disney stock daily data from 2024-09-30 to 2025-03-28

3.2. Data preprocessing

Before building the model, it is important to ensure that data is appropriate for time series modeling before applying the ARIMA model, or it would not work in further steps. Therefore, data preprocessing would be the first key step.

3.2.1. Stationarity test and differencing

The concept of stationarity indicates that the statistical attributes of a series, including its mean and variance, are stable across different time periods. For the data of Disney stock, which is originally non-stationary, it must be differenced to make it stationary. After differenced, ADF test is performed again to test if differenced data is stationary. After the second-order differencing, p is shown as 0.01, which is less than 0.05. So d is set as 2.

3.2.2. White noise test

The purpose of this preprocessing step is to determine if the time series can be characterized as a white noise series. If it turns out to be purely random, like white noise, it lacks patterns that can be predict, which makes modeling futile. The Ljung-Box test is applied to check autocorrelations at various lags. Table 1 shows the Ljung-Box test result of the second-time-differenced data, which suggests that it passes white noise test, thus this dataset can start to build the model.

Table 1: Ljung-Box test for the second-time differenced data

Lagging order	X-squared	p-value
6	47.04	0.00
12	61.59	0.00
18	64.67	0.00

3.3. Model identification

With the data successfully transformed into a stationary series and validated as ready for modeling, the subsequent objective becomes the precise identification of the parameters p and q that define the ARIMA structure. To guide this selection process towards the most suitable model configuration, the initial step involves generating and examining plots of ACF and PACF. These graphical tools offer valuable clues regarding the potential orders of the autoregressive and moving average components. To refine this preliminary identification and finalize the model selection, quantitative measures like AIC and BIC are subsequently applied, favoring the model that minimizes these information criteria.

3.3.1. ACF and PACF plots

Successfully identifying an appropriate ARIMA model depends on accurately recognizing the patterns present in the autocorrelation and partial autocorrelation functions [7]. The ACF and PACF plots are introduced to provide some visual and preliminary suggestions for the model order. o identify the order of the Moving Average (MA) component, the ACF plot is employed, while the PACF plot is introduced to determin the order of the Auto-Regressive (AR) part. Figure 2 and 3 shows the plot of each one. For the ACF plot, there is a significant spike at lag 1, then it drops off quickly. Thus their values are both around 1.

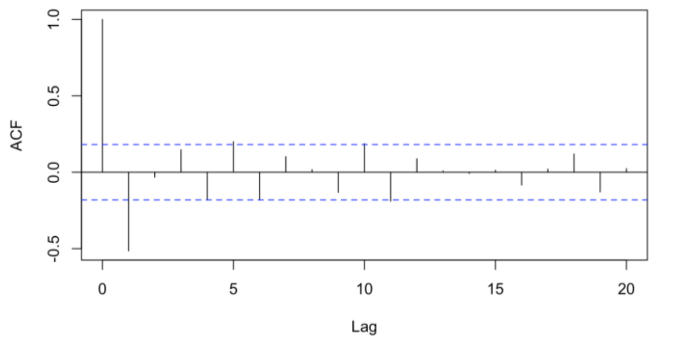

Figure 2: ACF after differential

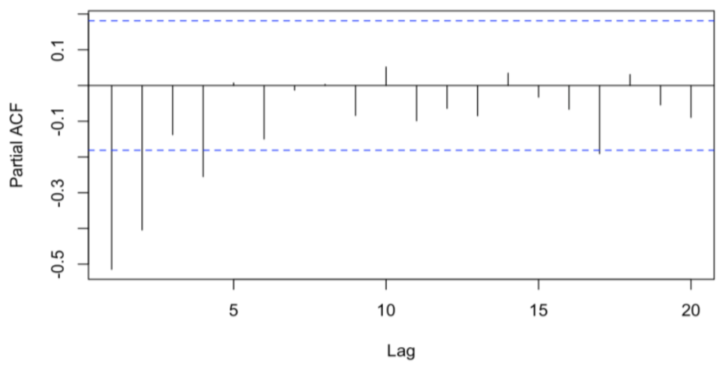

Figure 3: PACF after differential

DOI: 10.54254/2754-1169/2025.LH23955

3.3.2. AIC and BIC criteria

In order to further determine the best model, AIC and BIC are introduced. Both AIC and BIC assess how well a model fits the data, simultaneously incorporating a penalty for its complexity. The criterion for selecting the optimal model is based on identifying the one that yields the minimum values for both the AIC and BIC. Table 2 and 3 show the details of both AIC and BIC value with each model, and the best fit was achieved with ARIMA(0,2,1).

Table 2: AIC function value of ARIMA model

p	q	AIC	p	q	AIC
0	0	510.38	0	1	442.41
0	2	444.39	0	3	445.92
0	4	444.14	0	5	446.11
1	0	475.98	1	1	444.39
1	2	Inf	1	3	447.16
1	4	445.78	2	0	456.64
2	1	445.81	2	2	446.94
2	3	Inf	3	0	455.87
3	1	444.04	3	2	444.69
4	0	450.12	4	1	451.80

Table 3: BIC function value of ARIMA model

p	q	BIC	p	q	BIC
0	0	513.14	0	1	447.93
0	2	452.68	0	3	456.97
0	4	457.95	0	5	462.68
1	0	481.51	1	1	452.68
1	2	Inf	1	3	460.97
1	4	462.36	2	0	464.93
2	1	456.86	2	2	460.75
2	3	Inf	3	0	466.92
3	1	457.85	3	2	461.27
4	0	463.93	4	1	468.38

3.4. Residual analysis

The Ljung-Box test is then conducted on the residuals. Table 4 shows the result of test. Since the p-value are all more than 0.05, the residuals should resemble white noise, suggesting that the model successfully captured the essential structure within the data.

Table 4: Ljung-Box test for the differenced data

Lagging order	X-squared	p-value
6	2.13	0.91
12	6.82	0.87
18	11.01	0.89

3.5. Coefficient significance test

To further test the selected model and make sure it is not overfitted, a significance test is conducted on its estimated coefficients. Analysis of the results reveals an estimated coefficient of -0.94 for

DOI: 10.54254/2754-1169/2025.LH23955

MA(1), accompanied by a standard error of 0.07. This calculation produced a t-value of -13.40 and a p-value close to 0.00. This extremely low p-value suggests that the coefficient is highly statistically significant. Therefore, it can be confirmed that the model contains only meaningful parameters and is not overfitted with unnecessary terms.

3.6. Forecasting

The selected model is used to forecast the next five days of Disney's stock prices. Figure 4 shows the predicted data in plot. These forecasted values are compared with the actual observed prices. Then Mean Absolute Percentage Error (MAPE) is calculated to test prediction accuracy. Table 5 shows the result of the predicted values and the comparation. The model predicts next five days' prices with a MAPE of 1.85%, which is well below the 20% that is commonly used to define a good forecast. Therefore, this indicates excellent short-term predictive accuracy [8]. Meanwhile, the true value scenario shows that Disney's stock is trending down over the next few days, and the forecast results reflect this trend. Therefore, based on the above results, the model has good predictive ability.

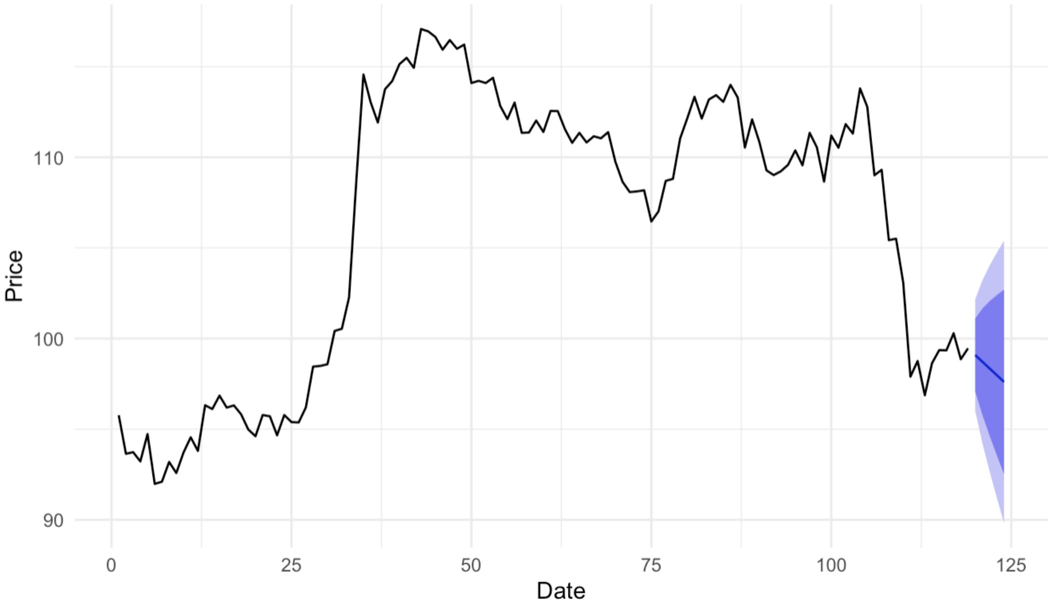

Figure 4: ARIMA forecast for Disney next 5 days

Table 5: Analysis of the forecast result

Date	Actual Value	Predicted Value	MAPE
2025-03-24	100.18	99.09	1.85%
2025-03-25	101.61	98.72	
2025-03-26	100.78	98.35	
2025-03-27	100.45	97.98	
2025-03-28	98.07	97.61	

4. Conclusion

This study explores the application of time series analysis in forecasting stock prices, with The Walt Disney Company's stock data as a case study. By collecting stock price data and conducting a series of processing steps, from which an ARIMA(p, d, q) model is able to be developed. With ACF and PACF analysis and AIC-based model selection, the ARIMA(0, 2, 1) model is considered as the most

DOI: 10.54254/2754-1169/2025.LH23955

suitable one. This model is then used to forecast the price in next five days, and the accuracy was evaluated by comparing predicted prices with actual values. The resulting MAPE was 1.85%, which is well below the commonly accepted threshold of 20%, indicating the good predictive performance.

This research still has several limitations. First, the analysis only used the stock's closing prices, those potentially useful variables such as trading volume, high and low prices, and macroeconomic indicators are not considered. These additional features might also contain some valuable signals, and it may enhance prediction accuracy. Second, external factors such as market news, geopolitical events, or industry-specific developments are not considered, despite their known influence on stock movements. Also, comparing different algorithms based solely on error may not be sufficient, as factors like calculation time, complexity, and required parameters also need consideration. This implies that while a model might show good performance in one metric (like error), it might have limitations in terms of computational cost or ease of use [9]. Finally, the study is limited to a relatively short time period, which may affect the robustness of the model in capturing long-term trends.

In future work, it would be valuable to extend the dataset to a longer time and consider other relevant financial and external features. At the same time, hybrid models like ARIMA+LSTM can be explored to capture both linear and nonlinear patterns [5]. Newer generative AI models may offer even better accuracy when combined with traditional techniques [10]. These hybrid models will have the potential to further improve forecasting accuracy and provide deeper insights into the dynamics of stock price behavior.

References

[1] *Khan, S. and Alghulaiakh, H. (2020) ARIMA Model for Accurate Time Series Stocks Forecasting. International Journal of Advanced Computer Science & Applications, 11*

[2] *Suripto, S. (2023) Decision-Making Model to Predict Auto-Rejection: An Implementation of ARIMA for Accurate Forecasting of Stock Price Volatility during the Covid-19. Decision Science Letters, 12, 107-116.*

[3] *Dey, B., Roy, B., Datta, S. and Ustun, T. S. (2023) Forecasting Ethanol Demand in India to Meet Future Blending Targets: A Comparison of ARIMA and various Regression Models. Energy Reports, 9, 411-418.*

[4] *Reddypogu, V. B. and Prasad, D. U. (2025) Demand Forecasting in E-Commerce Fashion Retail: A Comparative Study of Generative AI, LSTM and ARIMA Models. Journal of Information Systems Engineering and Management, 10, 23-28.*

[5] *Lahboub, K. and Benali, M. (2024) Assessing the Predictive Power of Transformers, ARIMA, and LSTM in Forecasting Stock Prices of Moroccan Credit Companies. Journal of Risk and Financial Management, 17, 293.*

[6] *Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. and Lachhab, A. (2018) Forecasting of Demand using ARIMA Model. International Journal of Engineering Business Management, 10, 1-9.*

[7] *Chodakowska, E., Nazarko, J., Nazarko, Ł., Rabayah, H. S., Abendeh, R. M. and Alawneh, R. (2023) ARIMA Models in Solar Radiation Forecasting in Different Geographic Locations. Energies, 16, 5029.*

[8] *Chicco, D., Warrens, M. J. and Jurman, G. (2021) The Coefficient of Determination R-Squared is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. PeerJ Computer Science, 7, e623.*

[9] *Shahvaroughi Farahani, M., Razavi Hajiagha, S. H. (2021) Forecasting Stock Price using Integrated Artificial Neural Network and Metaheuristic Algorithms Compared to Time Series Models. Soft Comput 25, 8483–8513.*

[10] *Preetha, K. G., Babu, K. R. R., Sangeetha, U., Thomas, R. S., Saigopika, Walter, S. and Thomas, S. (2022) Price Forecasting on a Large Scale Data Set using Time Series and Neural Network Models. KSII Transactions on Internet and Information Systems, 16, 3923-3942.*