Forecasting Weekly Stock Price of Microsoft Corporation (MSFT) Using ARIMA Box-Jenkins Method

Nur Farida[1, *] and Edwin Setiawan Nugraha[2]

[1,2] *Study Program of Actuarial Science, School of Business, President University, Bekasi, 17550, Indonesia*

[*] *Corresponding author: nur.farida@student.president.ac.id*

Abstract— The dynamic movement of stock prices in the financial market demands the ability to forecast future trends, particularly for investors and institutions aiming to make strategic decisions. This study focuses on forecasting the weekly stock price of Microsoft Corporation (MSFT) using the Autoregressive Integrated Moving Average (ARIMA) method, following the Box-Jenkins approach. The data used in this research consist of 54 weekly closing prices from April 29, 2024 to April 29, 2025. Several ARIMA models were tested, and the ARIMA(0,1,1) model was identified as the most suitable based on its accuracy metrics. This model achieved the lowest error values, including a Mean Absolute Percentage Error (MAPE) of 2.37%, indicating a high level of prediction accuracy. The results of this study provide useful insights for investors and financial analysts to better anticipate stock price fluctuations and enhance data-driven decision-making in portfolio management and investment strategies.

Keywords— Forecasting, Stock Price, ARIMA, Microsoft Corporation (MSFT)

I. INTRODUCTION

The stock market is a vital part of contemporary economic structures, as it is a mechanism by which companies mobilize capital and investors get a return on their investment. However, the stock market is also uncertain and volatile. Stock prices are not only influenced by a company's home-country financial performance but also by numerous external influences such as macroeconomic realities, political occurrences, global economic trends, and moods of investors. This dynamic nature makes it challenging for market participants to precisely predict price action. Among the various listed stocks, Microsoft Corporation (MSFT) is a leading global player as a diversified technology leader with reliable revenue streams, established brand recognition, and a solid financial performance. Consequently, its share price is a significant indicator of both technology sector health and overall market trends.

Because of the nature of speculation in stock prices, one of the largest challenges for investors and analysts is projecting future stock value based on historical data. Effective forecasting can enable reduction of the amount of risk for investments and strategizing more informed investments. Investors, financial institutions, and business clients utilize such forecasts to determine the best time to buy, sell, or hold shares. The randomness of stock prices increases the necessity for analytical techniques that can project and explain past trends to predict future directions. Although advanced machine learning techniques have been widely introduced in financial modeling, conventional statistical techniques such as the Autoregressive Integrated Moving Average (ARIMA) model remain highly relevant. ARIMA works best for short-term univariate time series forecasting and has been widely praised for its transparency, simplicity, and robust theoretical foundation.

Its application in economic and financial time series modeling in higher education forecasting has been tested and proven in past studies, warranting the Box-Jenkins methodology as a credible and formal statistical approach [1]. Another research used ARIMA to forecast the stock price of Apple Inc. and obtained that the model generated constant and readable short-term forecasts [2]. Research also used the ARIMA model to predict import values in Central Java, corroborating its effectiveness in discovering linear temporal patterns [3]. In banking, ARIMA was used to forecast stock prices of PT Bank Central Asia Tbk, and the result indicated a high correlation of model accuracy and error minimization [4]. PT Aneka Tambang Tbk was researched with the finding that ARIMA was suitable for modeling short-term weekly stock price variations and detected trend patterns with effectiveness [5]. Other research on YG Entertainment stock indicated that ARIMA produced accurate forecasts irrespective of brief data lengths, noting its usability in media and entertainment stock prediction [6]. Similarly, TSMC research depicted that ARIMA was a great model for predicting monthly stock price within the technology sector [7].

The goal of this study is to use the ARIMA Box-Jenkins method to forecast the weekly share price of Microsoft Corporation (MSFT) using past values. The goal is to develop a stable forecasting model that captures the underlying patterns in the time series and provides short-term forecasts. This research aims to serve investors and financial decision-makers by giving them insights on possible future price directions, thereby allowing more data-driven and informed strategies. Through the comparison of different ARIMA models and the testing of their forecasting capabilities in terms of error metrics such as MSE, RMSE, MAE, and MAPE, this research also further confirms the validation of ARIMA as a good forecasting tool for stock prices in the current financial world.The research is trying to forecast the closing Microsoft Corporation (MSFT) stock price for a four-week period using past data from April 29, 2024, to April 29, 2025, with 54 observations. The aim is to decide the best ARIMA model on the basis of statistical accuracy and residual analysis in order to provide short-run forecasts that can be valuable tools for investment decision making.

II. METHODOLOGY

A. Box-Jenkins Method

The Box-Jenkins ARIMA (AutoRegressive Integrated Moving Average) method is a type of forecasting and analyzing methodology for time series data [8]. This approach was developed by George Box and Gwilym Jenkins. The Box-Jenkins Model forecasts data using three principles: autoregression, differencing, and moving average. These three principles are known as p, d, and q, respectively. Each principle is used in the Box-Jenkins analysis; together, they are collectively shown as ARIMA (p, d, q) [9].

1. *Autoregressive (AR)*

The autoregressive (AR) component involves modeling the relationship between an observation and several lagged observations (previously observed points). This component gives us the idea that the current value of the time series is related to the previous values of the series [8]. Autoregressive (AR) component is denoted by **p** which expressed as:

$$Yt = \phi 1Yt-1 + \phi 2Yt-2 + \cdots + \phi pYt-p + et \quad (1)$$

2. *Moving Average (MA)*

This component represents the effect of past error terms on the current value of the time series. The moving average component can be represented as q, which is also known as the order of moving average [8]. And this Moving Average (MA) represented as :

$$Yt = et - \theta 1et-1 - \theta 2et-2 - \cdots - \theta qet-q \quad (2)$$

3. *Autoregressive Integrated Moving Average (ARIMA)*

Box-Jenkins is a type of autoregressive integrated moving average (ARIMA) model that gauges the strength of one dependent variable relative to other changing variables. The model's goal is to predict future securities or financial market moves by examining the differences between values in the series instead of through actual values [8]. Autoregressive (AR)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (3)$$

Forecasting nonstationary data can be done with the ARIMA (p,d,q) method by combining the AR and MA methods and using differencing $Wt = \nabla^d Yt$ [10]. ARIMA equation can be written as:

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (4)$$

Y_t : Data at time *t*

Y_{t-p} : Data at *p* period before time *t*

e_t : Error at time *t*

e_{t-q} : Error at *p* period before time *t*

ϕ_p : AR coefficient at order *p*

θ_q : MA coefficient at order *q*

W_t : Differencing process

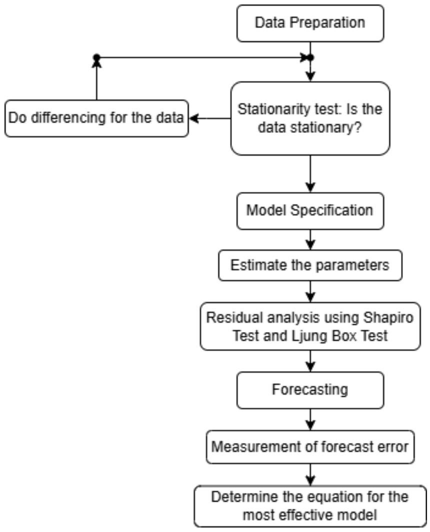

Figure. 1 ARIMA forecasting steps

B. Regression Model Evaluation Metrics

Regression metrics of measurement are applied to check the performance of regression models, which predict numerical continuous values. Metrics present a measure of precision and efficiency of models by comparing the predicted value with the actual value. Some common metrics of regression measurement are discussed below:

1. *Mean Absolute Error (MAE)*
   MAE estimates the average absolute differences between actual and predicted values. It is defined as:
$$MAE = \sum \frac{|Y' - Y|}{n}$$
2. *Mean Squared Error (MSE)*
   MSE estimates the mean squared differences between actual and predicted values. It is calculated as:
$$MSE = \sum \frac{(Y' - Y)^2}{n}$$
3. *Root Mean Squared Error (RMSE)*
   RMSE is the square root of the average of squared differences between actual and predicted values. It is calculated as:
$$RMSE = \sqrt{\sum \frac{(Y' - Y)^2}{n}}$$
4. *Mean Absolute Percentage Error (MAPE)*
   The Mean Absolute Percentage Error (MAPE) quantifies the precision of a forecast by taking the mean of the absolute percentage error between actuals and model forecasts. It is calculated as:
$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{|Y' - Y|}{Y} x\ 100\%$$

III. **RESULT AND DISCUSSION**

A. *Data Preparation*

The data used in this analysis is Microsoft Corporation (MSFT) weekly stock price data from April 29, 2024 to April 29, 2025 with a total of 54 data. 50 data as data training and 4 data as data test. Data shown in TABLE 1. The data is taken from Yahoo Finance website and is processed using R Studio. and I use library(readxl), library(ggplot2), library(tseries), library(forecast), library(scales) in solving forecasting weekly stock price Microsoft Corporation (MSFT).

TABLE 1
Data of The Stock Price of Microsoft Corporation (MSFT) Tbk

Date	Close	Date	Close
29/04/24	406.66	4/11/24	422.54
6/05/24	414.74	11/11/24	415
13/05/24	420.21	18/11/24	417
20/05/24	430.16	25/11/24	423.46
27/05/24	415.13	2/12/24	443.57
3/06/24	423.85	9/12/24	447.27
10/06/24	442.57	16/12/24	436.6
17/06/24	449.78	23/12/24	430.53
24/06/24	446.95	30/12/24	423.35
1/07/24	467.56	6/01/24	418.95
08/07/24	453.55	13/01/24	429.03
15/07/24	437.11	20/01/25	444.06
22/07/24	425.27	27/01/25	415.06
29/07/24	408.49	3/02/25	409.75
05/08/24	406.02	10/02/25	408.43
12/08/24	418.47	17/02/25	408.21
19/08/24	416.79	24/02/25	396.99
26/08/24	417.14	3/03/25	393.31
02/09/24	401.7	10/03/25	388.56
09/09/24	430.59	17/03/25	391.26
16/09/24	435.27	24/03/25	378.8
23/09/24	428.02	31/03/25	359.84
30/09/24	416.06	7/04/25	388.45
7/10/24	416.32	14/04/25	367.78
14/10/24	418.16	21/04/25	391.85
21/10/24	428.15	28/04/25	394.04
28/10/24	410.37	29/04/25	394.04

B. *Stationary Check*

Before processing the data, it is essential to confirm whether the data is stationary using the Augmented Dickey-Fuller (ADF) Test method in R. Stationarity is a crucial assumption in time series modeling, especially for ARIMA, and is determined by checking the p-value from the ADF test. If the p-value is below the significance level of 0.05, the data can be considered stationary. In Figure 1, the plot of Microsoft Corporation (MSFT) weekly stock price is shown. The result of the ADF test on the original time series produces a p-value of 0.2571. Since this value is above the threshold of 0.05, the data is classified as non-stationary, indicating that differencing is needed to transform the data into a stationary time series before further ARIMA modeling.

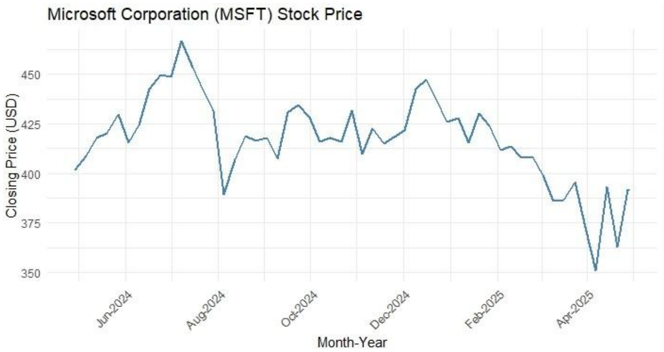

Figure. 1 Plot of PT Astra International Tbk Stock Price

To make the data stationary, differencing is required using the diff () function in R Studio. In Figure 2, the plot of the first difference of Microsoft Corporation (MSFT) weekly stock price is shown after applying first-order differencing. The Augmented Dickey-Fuller (ADF) Test conducted on the differenced data yields a p-value of 0.01263. Since this value is less than 0.05, it confirms that the differenced data is stationary. Therefore, the order of differencing d = 1 is appropriate for modeling the series using the ARIMA method.

Figure. 2 Plot of First Difference Microsoft Corporation (MSFT) Stock Price

C. *Model Specification*

The ARIMA model can be specified from the Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) Plot of the data. Figure 3 show the ACF and PACF plot for Microsoft Corporation (MSFT).

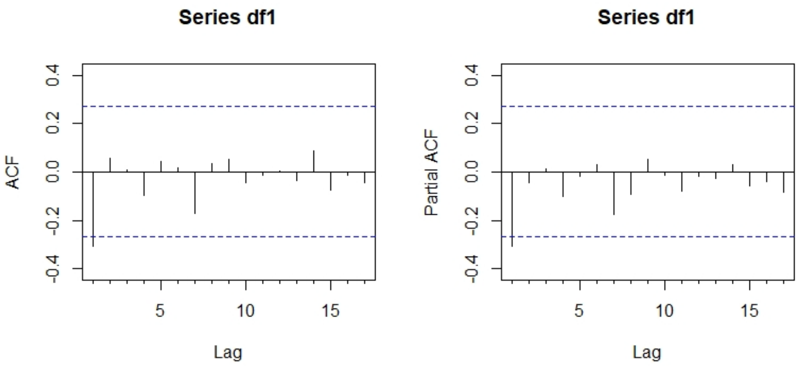

Figure. 3 Plot of ACF and PACF first Difference of Microsoft Corporation Stock Price

Based on the information provided by figure 3, this pattern indicates the presence of an autoregressive (AR) component in the time series. Meanwhile, the PACF plot shows a sharp cutoff after lag 1, indicating that only the first lag has significant partial autocorrelation and ACF cut off at a lag time 1. Thus, the order of P is determined as 1 and order of q is determined as 1. TABLE 2 shows the ARIMA model arranged based on these parameters.

TABLE 2

ARIMA Model Specification

No	ARIMA Model	p	d	q
1	ARIMA (0,1,0)	0	1	0
2	ARIMA (0,1,1)	0	1	1
3	ARIMA (1,1,0)	1	1	0
4	ARIMA (1, 1, 1)	1	1	1

D. *Parameter Estimation*

The parameter of the Autoregressive (AR) formula are denoted as phi with order p. while the parameters of the Moving Average (MA) formula are denoted as teta with orde q. TABLE 3 shows that by using summary (model) code, the value of each parameter as well as log likelihood and Akaike Information Criterion (AIC) can be obtained.

TABLE 3

Parameter Estimation for ARIMA Model

Model ARIMA	AR1	MA 1	Log Likelihood	AIC
(0, 1, 0)			-218.73	439.45
(0, 1, 1)		0.1308	-216.11	436.22
(1, 1, 0)	0.1295		-216.12	436.24
(1, 1, 1)	0.4856	0.4949	-216.06	438.13

E. *Residual Analysis*

The selection of the best ARIMA model is based on the results of the Shapiro Test and the Ljung-Box Test. These tests are used to evaluate the adequacy of the model residuals. An ARIMA model is considered acceptable if the p-value of both tests is greater than 0.05, which indicates that the residuals are normally distributed and free from autocorrelation. Table 4 presents the results of these diagnostic tests. Only models that pass both tests will proceed to the forecasting stage, where the forecasting accuracy will be evaluated using error measurement metrics.

TABLE 4

Result of Residual Analysis

No	Model ARIMA	Shapiro Test	Ljung-Box Test	AIC	Result
1	ARIMA (0, 1, 0)	0.71908	0.57165	439.45	Passed
2	ARIMA (0, 1, 1)	0.54525	0.97542	436.22	Passed
3	ARIMA (1, 1, 0)	0.54064	0.97082	436.24	Passed
4	ARIMA (1, 1, 1)	0.54205	0.97625	438.13	Passed

F. *Forecast*

TABLE 5

Forecasting Value of The ARIMA (0,1,0)

Date	Actual Data	Predicted Data	Lower Bound	Upper Bound
2025 – 04 – 14	393.22	391.3	361.9029	420.6971
2025 – 04 – 21	362.82	391.3	349.7262	432.8738
2025 – 04 – 28	391.96	391.3	340.3827	442.2173
2025 – 04 – 29	391.30	391.3	332.5058	450.0942

TABLE 6

Forecasting Value of The ARIMA (0, 1, 1)

Date	Actual Data	Predicted Data	Lower Bound	Upper Bound
2025 – 04 – 14	393.22	389.2784	361.0554	417.5015

2025 – 04 – 21	362.82	389.2784	355.0157	423.5412
2025 – 04 – 28	391.96	389.2784	349.8915	428.6654
2025 – 04 – 29	391.30	389.2784	345.3611	433.1957

TABLE 7
Forecasting Value of the ARIMA (1, 1, 0)

Date	Actual Data	Predicted Data	Lower Bound	Upper Bound
2025 – 04 – 14	393.22	391.5005	363.2719	419.7291
2025 – 04 – 21	362.82	391.4396	357.0427	425.8364
2025 – 04 – 28	391.96	391.4581	350.4875	432.4287
2025 – 04 – 29	391.30	391.4525	345.1983	437.7066

TABLE 8
Forecasting Value of The ARIMA (1,1,1)

Date	Actual Data	Predicted Data	Lower Bound	Upper Bound
2025 – 04 – 14	393.22	390.2334	361.7603	418.7066
2025 – 04 – 21	362.82	390.4048	355.9559	424.8537
2025 – 04 – 28	391.96	390.3773	350.1101	430.6445
2025 – 04 – 29	391.30	390.3817	345.1440	435.6194

Based on the forecasting value on TABLE 5, TABLE 6, TABLE 7 and TABLE 8, the best model for forecasting using ARIMA method can be determined by comparing the error among models as shown below.

TABLE 9
Error Measured from ARIMA Models

Model ARIMA	MSE	RMSE	MAE	MAPE
ARIMA (0, 1, 0)	203.8081	14.24614	7.765	2.12%
ARIMA (0, 1, 1)	181.7156	13.48019	8.776	2.37%
ARIMA (1, 1, 0)	205.5781	14.33799	7.748	2.12%
ARIMA (1, 1, 1)	193.2973	13.90314	8.268	2.25%

TABLE 7 shows that the ARIMA (0, 1, 1) model achieved the lowest error among the ARIMA models considered, indicating that ARIMA (0, 1, 1) is the most effective model for forecasting the stock price of Microsoft Corporation (MSFT). Refers to TABLE 3, the best model that is ARIMA (0, 1, 1) can be expressed in the form of equation as shown below.

$$Wt = 0.1308\ e_{t-1} + e_t$$

Let

$$Wt = Yt - Yt{-}1$$

Then the equation become

$$Yt - Yt{-}1 = 0.1308 \cdot et{-}1 + et$$

$$\mathbf{Yt = Yt - 1 + 0.1308 \cdot et{-}1 + et}$$

The accuracy of this model, as demonstrated by a **MAPE of 2.37%**, confirms its reliability for short-term forecasting of MSFT stock prices. This model can be used strategically by investors to support informed decision-making regarding stock buying or selling. Additionally, businesses and financial planners may utilize the forecasts to align their strategies with anticipated market movements.

Figure 4 below shows the visualization plot of forecasting from ARIMA (0, 1, 1) of Microsoft Corporation (MSFT) stock price forecasting from April 29, 2024 to April 29, 2025

Figure. 4 Forecasting Plot from ARIMA (0, 1, 1)

IV. CONCLUSION

Stock price forecasting has been carried out for Microsoft Corporation (MSFT) for the next 4 weeks using past weekly stock price data from April 29, 2024, to April 29, 2025, using 54 data points obtained from Yahoo Finance. Forecasting was carried out using the Box-Jenkins Method and implemented through R Studio. Based on the analysis, it was determined that the ARIMA (0,1,1) model had passed residual diagnostic tests and yielded the lowest forecast error among the considered models. This model was therefore selected as the best for predicting MSFT stock prices and can be expressed in the equation:

$$Yt = Yt - 1 + 0.1308 \cdot et-1 + et$$

However, note that even if the ARIMA (0,1,1) model produces relatively good and interpretable outcomes, it does have some disadvantages. One is its sensitivity to lagged values, which makes it less responsive to sudden market shocks or outside factors such as geopolitical events, earnings announcements, or macroeconomic events. Since the ARIMA model operates purely on historical price information, it lacks explanatory variables from outside that could have a significant effect on stock price movement.

For future study, it is recommended to consider more sophisticated models such as ARIMAX, which includes exogenous variables, or other machine learning models such as LSTM and XGBoost. Combining such models could potentially enhance accuracy in forecasting as well as deliver stronger insights when there are unstable or rapidly varying market conditions.

ACKNOWLEDGMENT

We would like to thank Causal Productions for permits to use and revise the template provided by Causal Productions. Original version of this template was provided by courtesy of Causal Productions (www.causalproductions.com).

REFERENCES

[1] M. A. Din, "ATINER ' s Conference Paper Series ARIMA by Box Jenkins Methodology for Estimation and Forecasting Models in Higher Education Marilena Aura Din," no. March, 2016, doi: 10.13140/RG.2.1.1259.6888.

[2] C. Jin, "Forecasting Models for Apple Inc . Stock Price Using Regres- sion Smoothing and Box Jenkins Time Series Analysis," vol. 2, no. Eass, pp. 1–11, 2024.

[3] R. A. Pitaloka, S. Sugito, and R. Rahmawati, "Perbandingan Metode Arima Box-Jenkins Dengan Arima Ensemble Pada Peramalan Nilai Impor Provinsi Jawa Tengah," *J. Gaussian*, vol. 8, no. 2, pp. 194–207, 2019, doi: 10.14710/j.gauss.v8i2.26648.

[4] A. Olivia and E. S. Nugraha, "Forecasting PT Bank Central Asia Tbk Stock Price Using ARIMA Model," vol. 2, no. 1, 2023.

[5] A. Intan, S. Wardhani, and M. R. Yudhanegara, "Forecasting Weekly Stock Price of PT . Aneka Tambang Tbk (ANTM) Using ARIMA Box-Jenkins Method," vol. 3, no. November 2021, pp. 20–31, 2025.

[6] N. Galuh Ramadhanty and E. Setiawan Nugraha, “Forecasting of YG Entertainment Stock Prices February 2022-August 2022 Using Arima Model,” *J. Actuar. Financ. Risk Manag.*, vol. 1, no. 2, pp. 1–9, 2022.

[7] G. M. Singgih and E. Setiawan Nugraha, “Forecasting the Monthly Stock Price per Share of Taiwan Semiconductor Manufacturing Company Limited (TSMC) using ARIMA Box-Jenkins Method,” *J. Actuar. Financ. Risk Manag.*, vol. 2, no. 1, pp. 1–9, 2023.

[8] “Box-Jenkins Methodology for ARIMA Models,” GeeksforGeeks. [Online]. Available: https://www.geeksforgeeks.org/box-jenkins-methodology-for-arima-models/

[9] Gordon Scott, “Box-Jenkins Model: Definition, Uses, Timeframes, and Forecasting,” Investopedia. [Online]. Available: https://www.investopedia.com/terms/b/box-jenkins-model.asp