Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 18, 2023, page 121 – 126

Predicting Stock Price Movement Statistically

Radedo Jeremy Siagiain, Adina Najwa Kamarudin*
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
*Corresponding author: adina.najwa@utm.my

Abstract

This study aims to assess and compare different statistical models for effectively predicting stock price movements. Specifically, we examine the performance of the ARIMA (Autoregressive Integrated Moving Average) model and the Exponential Smoothing method. To illustrate the analysis, we use the stock of TLKM as a case study. Our analysis involves applying the ARIMA and Exponential Smoothing methods to the time series data of TLKM stock prices. The findings reveal that the ARIMA model yields superior accuracy in predicting stock price movements, especially for TLKM stock, when compared to the Exponential Smoothing method. However, it is crucial to acknowledge that factors beyond these models should be considered when predicting stock prices. This research aims to provide practical guidance for individuals interested in making informed decisions regarding stock buying and selling, including selecting the most suitable method for predicting future stock prices. It is important to note that this study does not serve as an invitation or recommendation to engage in any securities transactions. The outcomes of this research are expected to benefit investors in their stock selection process. Furthermore, it is anticipated that potential investors will gain confidence in initiating investments within the Indonesian Capital Market. In the context of the companies analyzed, PT. Telekomunikasi Indonesia can leverage the insights gained from this study to prepare appropriate strategies in response to anticipated stock price fluctuations. In conclusion, the findings support the superior accuracy of the ARIMA model compared to Exponential Smoothing in predicting stock price movements, specifically for TLKM stock. Nonetheless, it is essential to consider additional factors when conducting stock price predictions.

Keywords: Stock Price Prediction; ARIMA Model; Exponential Smoothing; TLKM Stock

1. Introduction

Stocks or shares represent a means of participating in the capital of a limited liability company, providing advantages like dividends and capital gains. Shareholders also acquire influence, pride, and voting rights in shaping company operations. Stock exchanges serve as platforms for trading securities, including stocks. In Indonesia, the securities market is referred to as the Indonesia Stock Exchange (IDX) or Bursa Efek Indonesia (BEI), playing a vital role as an investment avenue for the general public. The IDX's composite stock price index, known as the Jakarta Stock Exchange Composite Index (JCI), has demonstrated positive growth and holds a favorable position within ASEAN and Asia-Pacific regions. Nevertheless, the IDX's market capitalization is comparatively smaller than the world's largest stock exchanges due to reduced investor interest. This study concentrates on capital gains, encompassing profits derived from price discrepancies between buying and selling shares influenced by market fluctuations. Accurate predictions of future stock prices can aid investors in making well-informed decisions concerning stock purchases.

2. Box Jenkins Method

The Box-Jenkins methodology consists of a collection of models, not just one model. It combines autoregressive (AR) and moving average (MA) models. The ARMA (p,q) model, a combination of AR(p) and MA(q) models, is well-suited for analyzing single time series data. The AR(p) model can be mathematically represented as follows:

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \emptyset_3 y_{t-3} + \emptyset_4 y_{t-4} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t$$

In contrast to the AR(p) model, the MA(q) model utilizes past errors as independent variables. The mathematical representation of the MA(q) model is as follows:

$$y_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \theta_4 \varepsilon_{t-4} + \cdots + \theta_p \varepsilon_{t-q} + \varepsilon_t$$

The General Equation for the Box-Jenkins method is represented as

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \emptyset_3 y_{t-3} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \cdots + \theta_p \varepsilon_{t-q}$$

The equation includes *y_t* for the actual value and εt for the residuals or random error terms at a specific time, *t*. The θi represents the coefficients for the nonseasonal MA component with order *i*, while $\emptyset i$ represents the coefficient for the nonseasonal AR component with order *i*. It is crucial to emphasize that AR and MA models are suitable for analyzing univariate stationary time series data. To determine the stationarity of a time series, it is necessary to check for the presence of a unit root. If the series' level is not stationary, differencing is applied to achieve stationarity. In such cases, the time series is referred to as an ARIMA (*p, d, q*) model, where p represents the order/degree of Autoregressive (AR), d represents the order/degree of differencing, and *q* represents the order/degree of Moving Average. In this particular scenario, the value of *d* will be equal to 1.

3. Exponential Smoothing Method

There are three models of exponential smoothing methods available. The single exponential smoothing model is used to forecast data when there is no seasonality or trend, or when the data is stationary. In the absence of seasonality but with a trend in the data, the double exponential smoothing model is applied for forecasting. Lastly, when both trend and seasonality are present in the time series data, Holt's-Winter exponential smoothing method is used to forecast the data.

The double exponential smoothing model involves forecasting the next period, *F_(t+1),* by considering the previous forecast for period *t*, *F_t*, and the observed value of the series in period *t*, *y_t*. The smoothing constant α, which determines the level of smoothing, must be within the interval of 0 < α < 1.

The double exponential smoothing model is an extension of the single exponential smoothing approach. It is suitable for time series data that displays a steady rate of growth or decline. Holt's Smoothing method incorporates two smoothing constants, *α* and *β*. By considering both the level and trend components, the double smoothing model enhances the accuracy of forecasts.

$$F_{t+m} = L_t + m b_t$$

$$L_t = \alpha y_{t+} (1-\alpha)(L_{t-1} - b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta) b_{t-1}$$

In the double exponential smoothing model, the forecast for the next period, F_(t+1), is based on the previous forecast for period t, F_t. The model considers the level of the series, represented as L_t, and the trend estimates denoted as b_t. It allows for forecasting future periods, where the number of periods to be forecasted is typically set to one (m = 1). The observed value of the series in period t is y_t. The smoothing constants α and β are used for smoothing the level and the trend, respectively, and they should fall within the range of 0 < α, β < 1.

4. Error Measurement

The purpose of error measurement is to gauge the efficiency of the forecasting model. This study employs four widely utilized criteria to evaluate effectiveness: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error, and Mean Absolute Percent Error (MAPE).

$$MSE = \sum_{t=1}^{n} \frac{(y_t - F_t)^2}{n}$$

$$RMSE = \sqrt{\sum_{t=1}^{n} \frac{(y_t - F_t)^2}{n}}$$

$$MAE = \sum_{t=1}^{n} \frac{|y_t - F_t|}{n}$$

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left|\frac{y_t - F_t}{y_t}\right|$$

The t-th sample's forecasting value, Ft, is compared to the corresponding actual value, yt. The total number of test sample data is represented by n.

5. Analysis

The line graph presented in Figure 1 illustrates the stock price of PT. Telkom Indonesia (TLKM) and its corresponding trend over the period spanning from December 31, 2021, to October 31, 2022, as visualized using Minitab software.

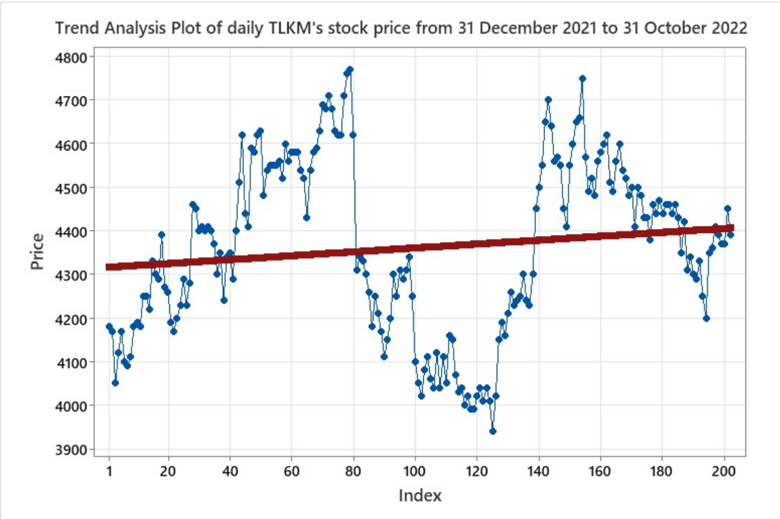

Figure 1 PT. Telkom Indonesia stock's price from 31 December 2021 until 31 October 2022

5.1 Box Jenkins Method

In the Box-Jenkins Methods, three models are employed: ARMA (*p,q*), ARIMA (*p,d,q*), and SARIMA (*p,d,q*)(*P,D,Q*)s. The choice of model depends on the presence of trend and seasonality in the data. When the data is stationary at the outset, the ARMA model is utilized for forecasting. In cases where the data exhibits a trend without seasonality, the ARIMA model is employed. Finally, if the data demonstrates both trend and seasonality, the SARIMA model is applied for forecasting.

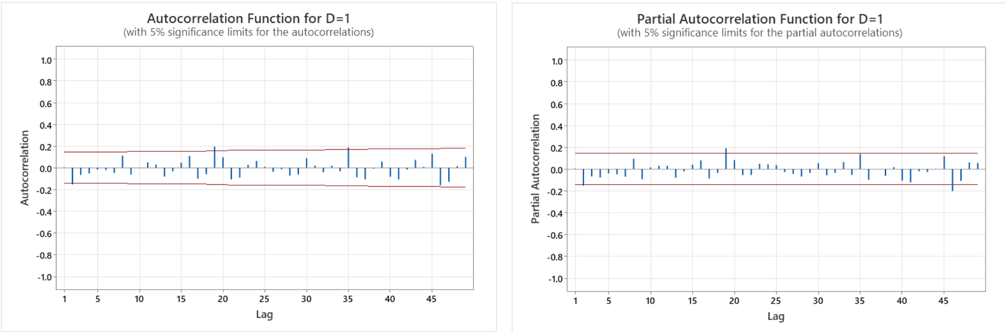

Figure 2 ACF and PACF after Differencing

Table 1 Possible ARIMA Model

	T test	Q test	White Noise	AIC	BIC
ARIMA (1,1,1)	Yes	Yes	Yes	-1.835	-1.7848
ARIMA (1,1,0)	No	No	No	-1.8229	-1.7893
ARIMA (0,1,1)	No	No	No	-1.8228	-1.7893

For our analysis, we will employ the ARIMA(1,1,1) method. This selection encompasses autoregressive (AR), differencing (I), and moving average (MA) components. The AR component captures temporal dependencies, the differencing component ensures stationarity, and the MA component accounts for past forecast errors. By incorporating these elements, the ARIMA(1,1,1) model offers a concise and effective approach to predict data dynamics while effectively addressing both trend and seasonality.

5.1 Exponential Method

The value of the optimal parameters which aims to minimize the sum of square error for the smoothing constant, α (smoothing constant for level series) and β (smoothing constant for trend) are automatically from the Minitab software output. Based on the Minitab software output, the smoothing parameter for α and β are 1.26785 and 0.01370 respectively.

6. Comparison Of the Accuracy Of Models

To evaluate the forecasting accuracy and performance of each model, four commonly used tools will be employed, namely, Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). To perform this evaluation, a total of 49 data points from 25th October until 30th December 2022 will be used as an out-of-sample dataset. This evaluation will be conducted to determine the forecasting performance and accuracy of all four models in the dataset.

Table 2 Error Measurement for Forecasting Model

Model	MSE	RMSE	MAE	MAPE
ARIMA (1,1,1)	285421	534.2481	469.5972	12.2302%
Double Exponential Smoothing	386296	621.5274	557.739	14.4846%

Based on the Error Measurement for Forecasting Model as presented in Table 2, it is evident that the ARIMA (1,1,1) model outperforms the exponential smoothing models, as it demonstrates the lowest values of MSE, RMSE, MAE, and MAPE. Therefore, it can be concluded that the ARIMA model is the most suitable model for forecasting the TLKM stock price.

Conclusion

The objective of this study is to utilize statistical techniques to forecast the price of PT Telkom Indonesia Tbk (TLKM) shares. The research aims to uncover concealed patterns and trends within TLKM's historical stock price data. By employing statistical time series methods, the study endeavors to construct a precise predictive model that can aid investors and market participants in enhancing their decision-making process. The findings from this research are anticipated to empower investors to optimize their investment strategies and minimize risks associated with TLKM stock trading. Based on the aforementioned analysis, the ARIMA model demonstrates superior predictive performance compared to exponential smoothing in forecasting TLKM stock prices.

Reference

[1] Guha, B., Bandyopadhyay, G., (2016). "Gold Price Forecasting Using ARIMA Model". ISSN: 2315-4462 (researchgate.net)

[2] Fattah, J., Ezzine, L., Aman, Z., Moussami, H.E., Lachhab, A., (2018). "Forecasting of demand using ARIMA model". https://journals.sagepub.com/doi/full/10.1177/1847979018808673

[3] Meher, B.K., Hawaldar, I.T., Spulbar, C., Birau, R., (2021). "Forecasting stock market prices using mixed ARIMA model: a case study of Indian pharmaceutical companies". https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3771334

[4] Duan, G., Niu, R., (2018). "Lake Area Analysis Using Exponential Smoothing Model and Long Time-Series Landsat Images in Wuhan, China". https://www.mdpi.com/2071-1050/10/1/149/htm

[5] Guleryuz, D., (2021). "Forecasting outbreak of COVID-19 in Turkey; Comparison of Box–Jenkins, Brown's exponential smoothing and long short-term memory models". https://reader.elsevier.com/reader/sd/pii/S0957582021001464?token=A1192579C33AD8667A9A6DA6378F28CCE48ED9B43F60068231102BA2AA516D77A049AC69CE545C9F38580C51B3BA1F8B&originRegion=us-east-1&originCreation=20221116034114

[6] Sinaga, H. D. E. , Irawati, N., (2018). "A Medical Disposable Supply Demand Forecasting By Moving Average And Exponential Smoothing Method". https://eudl.eu/pdf/10.4108/eai.24-1-2018.2292378

[7] Zhou, C., Wang, B., Qiu, J., (2020). "Forecasting stock prices with long-short term memory neural network based on attention mechanism". https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0227222&type=printable

[8] Alshara, M.A., (2022). "Stock Forecasting Using Prophet vs. LSTM Model Applying Time-Series Prediction". https://koreascience.kr/article/JAKO202213042317199.pdf

[9] Naik, P., Reddy, Y.V., (2021). "Stock Market Liquidity: A Literature Review". Stock Market Liquidity: A Literature Review - Priyanka Naik, Y. V. Reddy, 2021 (sagepub.com)

[10] Brennan, M. J., Chordia, T., Subrahmanyam, A., Tong, Q., (2012). "Sell-order liquidity and the cross-section of expected stock returns". Sell-order liquidity and the cross-section of expected stock returns - ScienceDirect

[10] Bustos, O., Quimbaya, A., (2020). "Stock market movement forecast: A Systematic review". Stock market movement forecast: A Systematic review - ScienceDirect

[11] Zulkernine, F., Shah, D., Isah, H., (2019). "Stock Market Analysis: A Review and Taxonomy of Prediction Techniques". Stock market movement forecast: A Systematic review - ScienceDirect