JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 – 8369
VOLUME 8 (ISSUE 1) JUNE 2024: 128 – 137
https://doi.org/10.21009/JSA.08111

A MULTIVARIATE APPROACH: FORECASTING JAKARTA COMPOSITE USING PROPHET FACEBOOK

Shafa Amalia Iskandar[1], Arum Handini Primandari[2*]

[1,2]Statistics Department, Faculty of Mathematic and Natural Science, Universitas Islam Indonesia
Kaliurang St., Sleman, Yogyakarta, 55584, Indonesia

*Corresponding author's e-mail: * **primandari.arum@uii.ac.id***

ABSTRACT

The Jakarta Composite Index (JCI, Composite Stock Price Index / IHSG) presents the average share price movement of companies listed on the Indonesia Stock Exchange (BEI/IDX), which can reflect the stock market performance. JCI forecasting can provide benefits for investors regarding risk management. On the other hand, gold is a low-risk asset with no credit risk and maintains its value over time. During the pandemic, gold prices increased significantly while stock prices decreased sharply, so gold prices can be used as a regressor in forecasting the JCI. A multivariate approach for time series analysis looks at how several variables relate to one another over time in order to reveal complex patterns and dependencies in the data. It also involves regressor. Researchers obtained historical data on the JCI and gold prices (dollars/ounce) from January 1, 2018, to December 31, 2022. The approach used in this research is Prophet model with regressor. The Prophet model uses a procedure to estimate time series data based on an additive model with trends that can be adjusted for annual, weekly, and daily seasonality. Based on the analysis results, the Prophet's multivariate approach is the best method for predicting the JCI compared to the univariate approach. The parameters used in the model are as follows: yearly seasonality, multiplicative seasonality mode, seasonality prior scale, namely 0.5, and changepoint prior scale, namely 0.001. The Mean Absolute Percentage Error (MAPE) obtained from the model is 2.78%.

Article History:
Received: 16 May 2024
Revised: 4 June 2024
Accepted: 9 June 2024
Published: 30 June 2024
Available online.

Keywords:
Jakarta Composite Index, Prophet, Forecasting

How to cite this article:
S.A. Iskandar, A.H. Primandari, "FORECASTING JAKARTA COMPOSITE USING PROPHET FACEBOOK", Jurnal Statistika dan Aplikasinya, vol. 8, iss. 1, pp. 128 – 137, June 2024

Journal homepage: *https://journal.unj.ac.id/unj/index.php/statistika*
Journal e-mail: *jsa@unj.ac.id*
Research Article · Open Access

1. INTRODUCTION

As our needs increase, our financial needs also grow. To save our assets or funds, we can invest our money, as suggested by Komaria, Hardianti, & Lestari in 2021 [1]. Investing means saving our existing funds in the hope of obtaining future profits, as mentioned by Darmawan, Kurnina, & Rejeki [2]. When we invest, we hope that the value of our investments will increase over time. Investment can also be seen as future savings, and anyone who invests is called an investor [3].

Investing is a great way to grow your wealth. There are several types of investments you can make, such as owning a private house, education, business, gold, property, land, and more. However, one of the most popular investments is shares, also known as stocks or equities. It refers to a unit of equity ownership in the capital stock of a corporation or units of real estate investment trusts, limited partnerships, and mutual funds on the financial markets. Share capital comprises the entirety of the enterprise's shares. The individual who owns shares is referred to as a shareholder or stockholder. By owning shares, a shareholder has the right to the company's assets and profits.

If you want to start investing in shares, you can observe the Jakarta Composite Index (JCI or IHSG). JCI is an indicator that presents the average share price value of companies listed on the Indonesian Stock Exchange. It reflects the stock market's performance, and investors should conduct a capital market analysis using the JCI indicator before investing, such as deciding whether to buy, sell, or hold shares [4].

Apart from shares, gold is also a valuable precious metal that investors can invest in. When world gold prices increase, investors will be more interested in investing in gold than shares. One of the studies conducted found that the world gold price had a significant effect on the Composite Stock Price Index. Therefore, investors often use gold as an asset because the gold price can prevent them from high investment risks to protect the value of their investment when JCI fell, when COVID-19 pandemic, and was not affected by inflation.

Investors must be careful in making investment decisions by paying attention to or anticipating changes in share prices in the JCI. Making forecasts to know changes in the JCI movement in advance will benefit in reducing risks and help with decision-making.

A technique for analyzing time-dependent data with numerous variables (simultaneously multiple time series) over time is multivariate time series analysis [5]. This methodology facilitates an exhaustive investigation of correlations, trends, and interdependencies across multidimensional data, culminating in an enhanced comprehension of the fundamental mechanisms impacting the data. Multivariate time series analysis improves forecasting accuracy and facilitates well-informed decision-making by studying numerous factors at once. We call $Z_t = [Z_{1,t}; Z_{2,t}; \dots Z_{m,t}]'$ a multivariate time series or a vector time series, where the first subscript refers to a component and the second subscript refers to the time [6].

A research of forecasting electricity consumption uses Prophet for the first step obtaining seasonality and multivariate data involving Prophet residual and meteorological data (temperature, precipitation, wind speed, humidity, sunshine duration, cloud cover) are inputted in GRU [7]. The proposed approach outperforms both the Prophet and GRU models, reducing prediction errors and offering valuable insights into electricity consumption patterns. Another multivariate time series analysis employs Gated Res2Net (GRes2Net) with four time series datasets that better performances over the state-of-the-art methods thus indicating the superiority [8]. Regression and tree model including lasso regression, random forest, and LSTM networks are implemented to forecast based on demand history and multiple predictors [9].

Prophet is an open-source package published by Facebook's Data Science core team for estimating time series data based on an additive model where non-linear trends fit annual, weekly, and daily seasonality, plus holiday effects. Forecasting Prophet can be done using Python and R [5]. Prophet creates models based on trend pattern approaches, seasonal patterns, and holiday effects. A univariate and multivariate approach using Prophet to forecast the depth to groundwater data show promising result on multivariate approach having higher accuracy [11].

It can handle time series data with seasonal solid patterns, as shown in JCI and gold prices during a pandemic. Prophet automatically models seasonal components and can detect seasonality in the data. The Facebook Prophet forecasting tool has been utilized in the realm of business and economics to detect seasonal patterns in revenue growth, thereby demonstrating its suitability for revenue forecasting in the e-commerce sector [6]. Moreover, the Prophet algorithm has been implemented in various

domains, including the prediction of carbon trading prices, water availability, and non-weather-related electric power outages, demonstrating its versatility in meeting various forecasting requirements [7] [8] [9].

This research will forecast JCI utilizing the univariate and multivariate approaches in Prophet methods. In the univariate approach, the JCI is forecasted employing its historical data. On the other hand, in the multivariate approach, we will use gold price as a regressor. The comparison uses Mean Absolute Percentage Error (MAPE) as a measurement to select the best Prophet approach.

2. METHODS

The data used in this research is Jakarta Composite Index (JCI) data and world gold prices from January 1 2018 – December 31 2022 which are daily data. The data used is secondary data, JCI is obtained from the Yahoo Finance website with the link https://finance.yahoo.com/, and world gold prices are obtained on the investing.com website with the link https://id.investing.com/. The measurement for JCI is Indonesian Rupiah and the measurement for gold price is dollar per ounch.

Research Method

The following are the steps involved in this research:

1. Firstly, data pre-processing is carried out which includes checking for missing data. If any data is missing, the moving average method is used to replace it with a value. After this, the cleaned data is stored in object form for time analysis.
2. Next, the data is analyzed descriptively to identify trend patterns and possible seasonality.
3. The data is then split into two parts: the training set and the testing set. The training set is used to create a model while the testing set is used to evaluate the performance of the model.
4. The parameters in Prophet are adjusted to obtain the best model. The Mean Square Error (MSE) is used to select the best model from multiple parameter combinations. The best model is determined for each approach in this stage.
5. To evaluate the performance of the model, Mean Absolute Error Percentage (MAPE), Mean Square Error (MSE), and Root Mean Square Error (RMSE) are used. The value obtained is used to select the best model generated from the univariate and multivariate analysis using the following formula.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}\left(y(t)-\hat{y}(t)\right)^2}{n}}$$

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{y(t)-\hat{y}(t)}{y(t)}\right|}{n} \times 100\%$$

Finally, the best model is analyzed using charts with annual trend patterns, weekly seasonality, and seasonality within a year to interpret the model. Using the best model, forecasting is done for the next 31 days, specifically from the 1st to the 31st of January 2023.

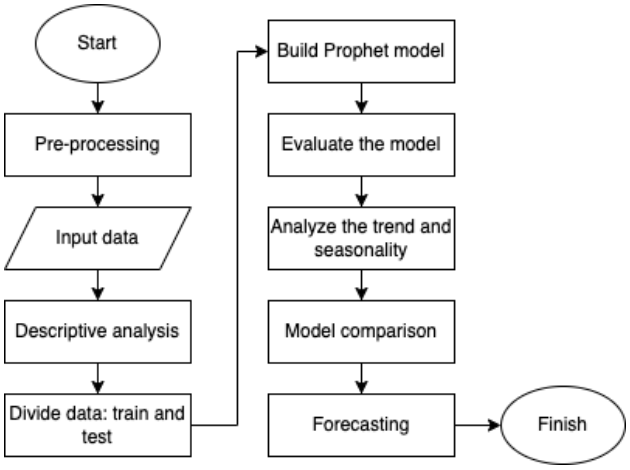


Figure 1. Research Flowchart

Prophet

Prophet is a package developed by Sean J. Taylor and Ben Letham from the Data Science Team at Facebook in 2017 [10]. It is an open-source tool that can be used for forecasting time series data in both Python and R [11]. Initially, the prophet model was created to predict daily business data using three features: piecewise trends, multiple seasonality, and floating holidays. The algorithm identifies existing pattern trends and seasonal influences in the time series data to make predictions. The formula for the model is shown below [12] [19].

$$y(t) = g(t) + s(t) + h(t) + \in_t \tag{1}$$

Where $g(t)$ is a trend model that describes the increase or decrease in long-term data, $s(t)$ which is modeling seasonal that can be interpret how the data can be influenced by factors seasonal weekly or annual , $h(t)$ is modeling impact day holidays (big events) that influence business calendar, and $\in_t$ is term error, i.e., the discrepancy between actual data and predicted values which are cannot be explained by the model.

The following is linear mode equation:

$$g(t) = (k + a(t)^T\delta)t + (m + a(t)^T\gamma) \tag{2}$$

The following is logistic mode equation:

$$g(t) = \frac{C(t)}{1+\exp\left(-(k+a(t)^T\delta)(t-(m+a(t)^T\gamma))\right)} \tag{3}$$

Where $C(t)$ is capacity varying time (time carrying capacity), k is growth rate, m is a balancing parameter (offset), γ is a vector of adjustment of changepoints, δ is a vector of adjustment of growth rates, T is total time, and t is running time. On the other hand, the seasonality is handled by using the Fourier series.

$$s(t) = \sum_{n=1}^{N}\left[a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)\right] \tag{4}$$

Where N is the number of parameters (the Fourier order), T is the expected regular periods of time series (e.g., 365.25 for yearly seasonality, 7 for weekly seasonality), and $[a_n, b_n]$ are coefficients estimated from the data.

In the Facebook Prophet model, several hyperparameters can be adjusted to optimize model performance [20]:

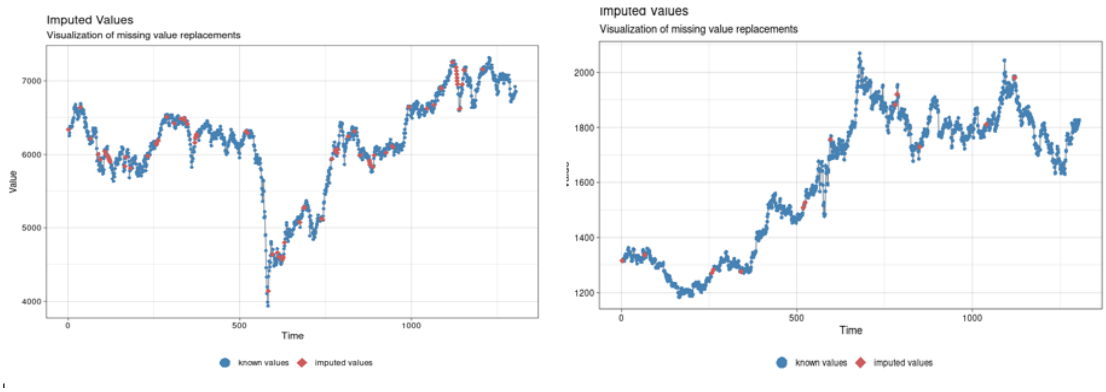
1. Changepoint Prior Scale (`changepoint_prior_scale`): Controls the flexibility of the model in fitting trend changes. Lower values make the trend smoother.

2. Seasonality Prior Scale (`seasonality_prior_scale`): Controls the strength of seasonality components. Higher values allow more seasonality variations.
3. Seasonality Mode (`seasonality_mode`): Can be 'additive' or 'multiplicative'. 'Multiplicative' is useful when the seasonal effect scales with the level of the time series.
4. Holidays Prior Scale (`holidays_prior_scale`): Adjusts the influence of holiday effects.

These hyperparameters can be tuned using cross-validation to find the best set of parameters for the specific dataset.

3. RESULTS

Initially, the imputation method (moving average) imputed missing data on the JCI data and world gold prices. The moving average worked for order equall to two (k=2). In Figure 2, the red dots are imputed data and the blue dots are actual data. Based on the figure, the imputed values still follow the data fluctuation.



(a) (b)

Figure 2. Imputation results for (a) JCI; (b) Gold Price

Based on the graph, we can observe a significant drop in the IHSG value and a substantial increase in the price of gold in early 2020 (the 500th data), coinciding with the onset of the Covid-19 pandemic. Throughout 2020, the IHSG value was at its lowest compared to previous years. The graph also illustrates that in 2022 (the 1,000th data), the JCI value has notably increased, reaching a level of 7,000, while the price of gold during this time also saw an increase to approximately US$2,000 per troy ounce. This indicates that the IHSG value and the price of gold have been highly volatile over the 5-year period from 2018 to 2022.

Descriptive Analysis

The following Table 1 is a descriptive analysis that describes the characteristics of the JCI data and world gold prices from January, 1 2018 to December, 30 2022.

Table 1. Descriptive Analysis

Measurement	JCI(IDR)	Gold ($ per ounce)
N	1,305	1,305
Min	3,937.6	1,184
1st Qu.	5,931.3	1,331.2
Median	6,207.2	1,710.7
Mean	6,160.2	1,610.4
3rd Qu	6,562.9	1,813.5
Max	7,318	2,069.4

The number of actual data for the JCI and its regressor are the same. The mean and median for both data are slightly difference indicating the symmetrical distribution.

Prophet Analysis

When running for Prophet model, the data needs to be converted into columns with the names “ds” and “y” to meet the required input. Column “ds” contains the date/time component and column “y” contains JCI data.

Table 2. Prophet input data in the univariate approach

No.	etc.	y
1.	2018-01-01	6339.24
2.	2018-01-02	6339.24
3.	2018-01-03	6251.48
…		
1305.	2022-12-30	6850.62

When carrying out Prophet using a multivariate approach, it is necessary to add a regressor variable, namely the price of gold, with no specific provisions for the regressor column.

Table 3. Prophet input data in the multivariate approach

No	etc.	y	Gold
1.	2018-01-01	6339.24	1316.1
2.	2018-01-02	6339.24	1316.1
3.	2018-01-03	6251.48	1318.5
…			
1305.	2022-12-30	6850.62	1826.2

The data was split into two parts - a training set and a testing set. Between 2018 and 2022, there were 1,305 observations, with 90% of these being allocated to the training set (about 1,174 observations), and 10% to the testing set (131 observations). Researchers used the training set to build a model and then evaluated the model's performance using the testing set.

In this research, the best Prophet model was created by adjusting hyperparameters for trend and seasonal changes. The parameters used in the model can influence the prediction results. The hyperparameters used in this study included yearly seasonality, seasonality mode, seasonality prior scale, and changepoint prior scale. As the stock market is closed during holidays, the holiday effect was not considered in this study. Table 4 shows the hyperparameter adjustments, along with the MAPE and RMSE values for univariate and multivariate approaches.

Table 4. Hyperparameter Adjustment and Error Rate Results

No	Yearly seasonality	Fashion seasonality	Seasonality prior scale	Change-point prior scale	Univariate		Multivariate	
					MAPE	RMSE	MAPE	RMSE
1.	True	Additives	0.01	0.1	7.41	582.47	5.29	464.03
2.	True	Additives	0.1	0.01	5.20	477.43	4.73	443.48
3.	True	Additives	0.1	0.5	6.80	552.63	6.12	521.01
4.	True	Additives	0.1	1	6.01	509.49	5.38	477.96
5.	True	Additives	0.5	0.01	4.97	459.87	4.72	442.64
6.	True	additives	1	0.01	5.08	468.75	4.75	445.18
7.	True	additives	0.5	0.001	11.2 2	830.15	2.78	193.71
8.	True	multiplicative	0.01	0.1	7.35	580.22	5.01	448.92
9.	True	multiplicative	0.1	0.01	4.48	413.56	4.23	390.66
10	True	multiplicative	0.1	0.5	7.22	585.41	6.56	554.73
11	True	multiplicative	0.1	1	6.57	552.74	6.13	531.14
12	True	multiplicative	0.5	0.01	4.68	425.91	4.06	375.03
13	True	multiplicative	1	0.01	4.86	438.77	4.30	398.75
14	True	multiplicative	0.5	0.001	11.83	871.91	3.33	281.40

Based on the results in Table 4, the Prophet method with the Multivariate approach has the smallest error rate calculation compared to the Prophet method with the univariate approach, namely MAPE of 2.78% and RMSE of 193.71. These results are obtained when the model has the yearly seasonality parameter with the '*True'* option, the seasonality mode parameter is '*additive'*, the seasonality prior scale parameter is 0.5 and the changepoint prior scale parameter is 0.001.

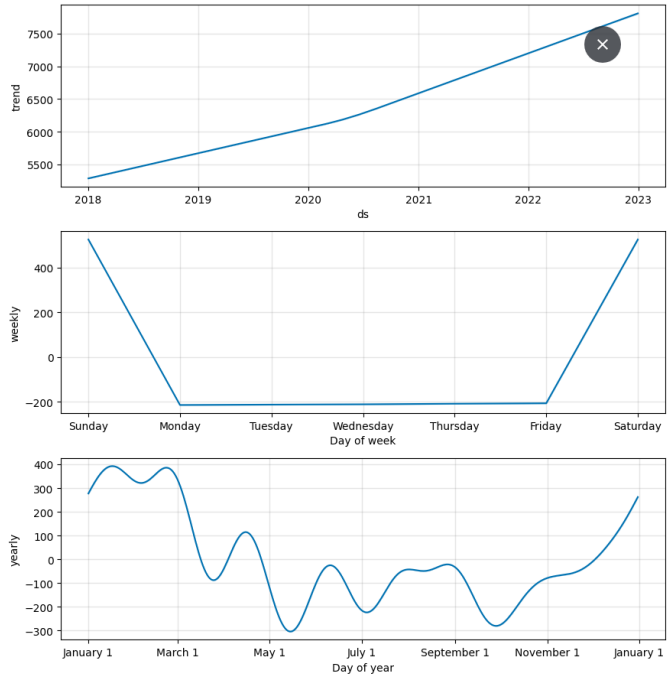


Figure 3. Trend plot, weekly seasonality, and annual seasonality using the multivariate prophet method.

Based on Figure 3., JCI has upward trend of 2018 to 2023. On weekly diagram, the flat line during weekdays indicates the absence of seasonality, in contrast with weekend value. Moreover, the yearly index of seasonality show decline in the middle years and strengthening at the end until beginning year.

The graph in Figure 4 is comparison of actual data and predicted data for JCI data from January 1, 2018 to December 30, 2022.

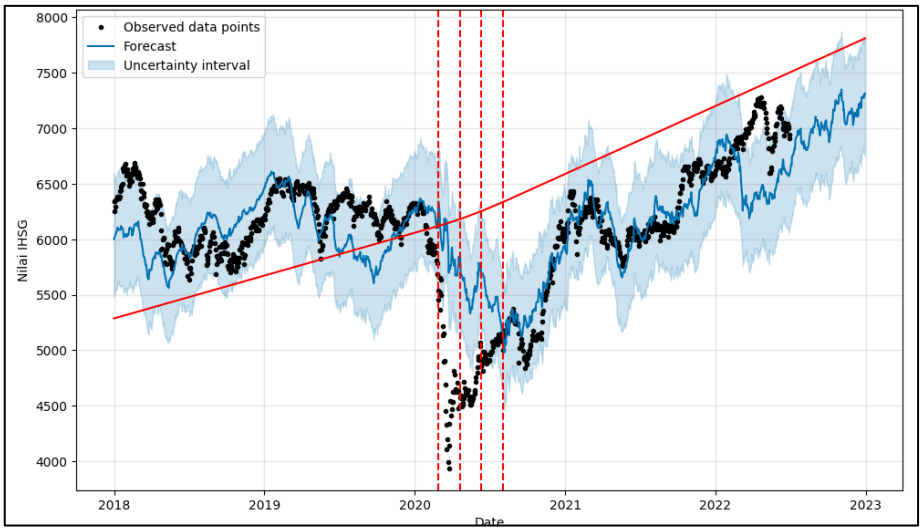


Figure 4. Graph plot comparison of actual data and predicted data with multivariate prophet method

Based on Figure 4, the fitted value and the forecasting from Prophet are able to follow the actual data pattern quite well. When significant changes occur, for example in early to mid-2020 (pandemic), the model is less able to follow sharp declines. Trend change points are marked with red vertical dotted lines which can be set with the changepoint prior scale parameter. Based on trials with the smallest MSE values, changes in trends were detected quite significantly during the pandemic.

Below in Table 5 it is presented results testing Prophet's best performance with using testing data.

Table 5. Comparison of Actual Data and Predicted Data along with Error Rates using the Prophet Multivariate Method.

Date	Actual Data	Prediction Data	MAPE	MSE	RMSE
01/07/2022	6794.33	6604.70	2.79	35959.65	189.63
04/07/2022	6639.17	6679.99	0.61	1666.16	40.82
05/07/2022	6703.27	6671.31	0.48	1021.54	31.96
06/07/2022	6646.41	6751.55	1.58	11054.55	105.14
07/07/2022	6652.59	6751.10	1.48	9705.14	98.51
…	…	…	…	…	…
12/26/2022	6835.81	7289.98	6.64	206268.23	454.17
12/27/2022	6923.03	7273.44	5.06	122784.48	350.41
12/28/2022	6850.52	7305.42	6.64	206930.31	454.90
12/29/2022	6860.08	7297.69	6.38	191499.56	437.61
12/30/2022	6850.62	7312.93	6.75	213734.49	462.31

Following is manual calculations using Prophet method with approach Multivariate with seasonality mode being additive, for the period 2 January 2023:

$$y(t) = g(t) + s(t)_{weekly} + s(t)_{yearly} + regressor)$$

$$y(t) = 7590.21 + (-180.29) + 193.74 + (-433.01) = 7170.65$$

4. DISCUSSIONS

Multivariate time series analysis involves examining datasets with multiple variables observed over a period of time. The error measurement result using Prophet's multivariate approach is almost half as much as the univariate, which indicates the success of using regressors. The relationship between these two data is shown during COVID-19 that they have contrast movements. Investors tend to invest in other alternative securities whenever the JCI does not have a promising outcome; in this case, they may choose gold.

The result of this research is about the same as that of Riyantoko et al., in which the multivariate Prophet outperformed the univariate approach [7]. In forecasting water availability, Riyantoko did not necessarily find a notable difference in error measurement between the two approaches. Comparing the computed elapsed time will help in choosing the best model. Meanwhile, our research may have significant differences. On the other hand, Prophet also successfully decomposes data into some components: the trend, the weekly seasonality, yearly seasonality and holiday effect, which help us understand the data well. The research by Navratil said the decomposition of the data and prediction using Prophet has an advantageous result in business economy cases [6]. The results show that JCI has yearly seasonality that peaks at the year's beginning and end. In addition, even with the drastic decline in the pandemic, JCI has an upward trend.

5. CONCLUSION

JCI forecasting for the last five years, including during the pandemic, can be done using Prophet by setting *yearly seasonality* parameters , *seasonality mode parameters , seasonality prior scale parameters* and *changepoint prior scale parameters.* Each parameter plays a role important in determining accuracy of results forecasting. On the other hand, usage *regressors* in approach Multivariate can lower error, in case This MAPE value decreases almost 50% compared to the approach univariate. Selection of appropriate regressors help increase accuracy forecasting on the Prophet model.

6. ACKNOWLEDGMENTS

We gratefully acknowledge the financial support for this publication research provided by Statistics Department.

7. REFERENCES

[1] R. Komaria, R. S. Hardianti and W. Lestari, "Analisis Faktor Pengetahuan, Motivasi, Dan Pelatihan Pasar Modal Terhadap Minat Berinvestasi Mahasiswa," *Juremi: Jurnal Riset Ekonomi*, vol. 1, no. 3, pp. 185-192, 2021.

[2] A. Darmawan, K. Kurnina and S. Rejeki, "Pengetahuan Investasi, Motivasi Investasi, Literasi Keuangan dan Lingkungan Keluarga Pengaruhnya Terhadap Minat Investasi di Pasar Modal," *Jurnal Ilmiah Akuntansi Dan Keuangan*, vol. 8, no. 2, pp. 44-56, 2019.

[3] B. T. Cahya, "Pengaruh Motivasi dan Kemajuan Teknologi Terhadap Minat Investasi Saham," *Am Masharif: Jurnal Ilmu Ekonomi dan Keislaman*, vol. 7, no. 2, pp. 192-207, 2019.

[4] G. A. Danardono and M. , "Analisis Determinan Perubahan Indeks Harga Saham Gabungan (IHSG) di Bursa Efek Indonesia (BEI)," *Jurnal Manajemen Bisnis Indonesia (JMBI)*, vol. 5, no. 6, pp. 626-640, 2016.

[5] R. S. Tsay, Multivariate Time Series Analysis: with R and Financial Applications, Chicago: John Wiley & Sons, 2014.

[6] W. W. Wei, Multivariate Time Series Analysis and Applications, Philadelphia: John Wiley & Sons, 2019.

[7] N. Son and Y. Shin, "Short- and Medium-Term Electricity Consumption Forecasting Using Prophet and GRU," *Sustainability*, vol. 15, no. 15860, pp. 2-16, 2023.

[8] C. Yang, M. Jiang, Z. Guo and Y. Liu, "Gated Res2Net for Multivariate Time Series Analysis," 19 September 2020. [Online]. Available: arXiv:2009.11705. [Accessed 30 July 2024].

[9] M. Motamedi, J. Dawson, N. Li, D. G. Down and N. M. Heddle, "Demand forecasting for platelet usage: From univariate time series to multivariable models," *PLoS ONE,* vol. 19, no. 4, p. https://doi. org/10.1371/journal.pone.0297391, 2024.

[10] B. Jange, "Prediksi Harga Saham Bank BCA Menggunakan Prophet," *Journal of Trends Economics and Accounting Research,* vol. 2, no. 1, pp. 1-5, 2021.

[11] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani, A. Muhaimin and T. Trimono, "Water Availability Forecasting Using Univariate and Multivariate Prophet Time Series Model for ACEA (European Automobile Manufacturers Association)," *International Journal of Data Science, Engineering, and Analytics* , vol. 1, no. 2, 2021.

[12] M. Navrátil and A. Kolková, "Decomposition and Forecasting Time Series in Business Economy Using Prophet Forecasting Model," *Central European Business Review,* vol. 8, no. 4, December 2019.

[13] P. A. Riyantoko, T. M. Fahrudin, K. M. Hindrayani, A. Muhaimin and T. Trimono, "Water Availability Forecasting Using Univariate and Mul- tivariate Prophet Time Series Model for ACEA (European Automobile Manufacturers Association)," *Internasional Journal of Data Science, Engineering, and Anaylitics,* vol. 1, no. 2, 2021.

[14] F. Meng and R. Dou, "Prophet-LSTM-BP Ensemble Carbon Trading Price Prediction Model," *Computational Economics*, 2023.

[15] O. O. Owolabi and D. A. Sunter, "Bayesian Optimization and Hierarchical Forecasting of Non-Weather-Related Electric Power Outages," *Energies*, vol. 15, no. 6, 2022.

[16] G. Rafferty, Forecasting Time Series Data with Facebook Prophet, Birminghan: Packt Publishing Ltd, 2021.

[17] N. Zhao, Y. Liu, J. K. Vanos and G. Cao, "Day-of-week and seasonal patterns of PM2.5 concentrations over the United States: Time-series analyses using the Prophet procedure," *Atmospheric Environment,* vol. 192, pp. 116-127, 2018.

[18] A. H. Primandari, A. Kesumawati and A. K. Thalib, "Analysis of Changes in Atmospheric CO2 Emissions Using Prophet Facebook," *Enthusiastic : International Journal of Applied Statistics and Data Science,* vol. 2, no. 1, pp. 1-9, 2022.

[19] A. H. Primandari and A. Kesumawati, "The hybrid prophet-ELM approach for athmospheric methane analysis," in *The 4th International Conference on Chemical Processing and Engineering (4th IC3PE),* Yogyakarta, 2024.

[20] P. Facebook, "Forecasting at Scale," 2023. [Online]. Available: https://facebook.github.io/prophet/docs/quick_start.html. [Accessed 30 July 2024].