IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.x, No.x, July xxxx, pp. 1~5
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.xxxx

Machine Learning Approaches for Predicting Seasonal Stock Trends

Jason Miracle Gunawan*[1], Christopher Andreas[2], Theresia Ratih Putri[3]
[1,2,3]Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Ciputra Surabaya, Surabaya, Indonesia
e-mail: *[1]jmiracle@student.ciputra.ac.id, [2]theresia.ratih@ciputra.ac.id

Abstrak

Pasar keuangan sangat penting untuk pertumbuhan ekonomim tetapi sering mengalami volatilitas, terutama pada sektor transportasi Indonesia Dimana harga saham sangat dipengaruhi oleh flutuasi musiman. Metode peramalan konvensional sering mengabaikan pola berulang ini sehingga mengakibatkan rendahnya akurasi prediktif. Penelitian ini mengevaluasi kemampuan algoritma Machine Learning untuk mengenali pola musiman pada prediksi harga saham, dengan menggunakan data bulanan PT Garuda Indonesia (Persero) Tbk (GIAA.JK) dari Agustus 2019 hingga Mei 2025 yang didapatkan dari Yahoo Finance. Empat model, yaitu Linear Regression, Extreme Gradient Boosting (XGBoost), Gated Recurrent Unit (GRU), dan Long Short-Term Memory, dilatih dan diuji performanya dengan menggunakan Root Mean Square Error (RMSE), Mean Absolute Error (MAE), dan Mean Absolute Percentage Error (MAPE). Hyperparameter tuning diterapkan pada XGBoost, LSTM, dan GRU, sedangkan validasi statistik menggunakan tes Kruskal-Wallis. Hasil peneltian menunjukan model tuned GRU mengungguli model lain dengan MAE 5.90, RMSE 7.33, dan MAPE 9.67%, menunjukkan akurasi 'excellence' dalam memodelkan dinamika jangka pendek dan musiman. Penemuan ini menekankan keunggulan GRU dalam memodelkan fluktuasi jangka pendek dan ketergantungan musim jangka Panjang dalam harga saham. Hasilnya memberikan ilmu praktis bagi investor dan menekankan pentingnya mengintegrasikan musiman dalam model prediktif untuk sektor yang fluktuatif.

Kata kunci—*Machine Learning, Gated Recurrent Unit, Prediksi Harga Saham, Analisis Pola Musiman, Sektor Transportasi Indonesia*

Abstract

The financial market is vital for economic growth yet it often experiences volatility, particularly in Indonesia's transportation sector where stock prices are strongly affected by seasonal fluctuations. Conventional forecasting methods often neglect these recurring patterns, lowering predictive accuracy. This study assesses the capability of Machine Learning algorithms to capture seasonality in stock price prediction, using PT Garuda Indonesia (Persero) Tbk (GIAA.JK)'s monthly data from August 2019 to May 2025, retrieved from Yahoo Finance. Four models–Linear Regression, Extreme Gradient Boosting (XGBoost), Gated Recurrent Unit (GRU), and Long Short-Term Memory (LSTM)–were trained and tested, with performance evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Hyperparameter tuning was applied to XGBoost, LSTM, and GRU, while statistical validation employed the Kruskal-Wallis test. Results showed that the tuned GRU outperformed other models, achieving MAE of 5.90, RMSE of 7.33, and MAPE of 9.67%, demonstrating 'excellent' accuracy in modelling both short-term and seasonal dynamics. These findings highlight the superiority of GRU in modelling both short-term fluctuations and long-term seasonal dependencies in stock prices. The results contribute practical insights for investors and emphasize the importance of integrating seasonality in predictive models for volatile sectors.

Received June 1st,2012; Revised June 25th, 2012; Accepted July 10th, 2012

Keywords**—*Machine Learning, Gated Recurrent Unit, Stock Price Forecasting, Seasonal Pattern Analysis, Indonesia Transportation Sector*

1. INTRODUCTION

Financial market has a significant contribution in global economic by facilitating capital allocation and strengthening economic growth [1]. In fact, financial market can be one of the indicators of a country's economy that all the financial experts are able to estimate the economic health of a country currently [2]. The COVID-19 outbreak in 2020 exemplifies how unexpected global events can severely affect economic sectors. The government policy in terms of facing the phenomena by conducting isolation inflicted bad impacts in economy sector, such as supply chain disruption, decreased consumption and investment, and a sharp decline in financial market activity [3]. Moreover, the pandemic caused a continuous decline on stock price due to decline in company performance and the high uncertainty that led some investors to panic. One of the sectors that was affected significantly is transportation, where social distancing and the economic uncertainty caused decline in operational and demands [4].

Indonesia's transportation sector experiences the high fluctuation due to its strong seasonal nature, particularly during public holidays such as Christmas, Eid al-Fitr, and school breaks. These periods trigger an escalation in public mobility across land, air, and sea transportation, often accompanied by shifts in mode preference, such as from buses and trains to airplanes or private vehicles [5]. This occurrence significantly affects the performance of transportation-related stocks, including both public transport providers and supporting infrastructure companies such as ports and airports [6]. The constant ascendancy of particular transport modes during holiday periods emphasizes how seasonal patterns extremely influence this sector [7]. However, this mobility dependency and periodic cycles also leads to volatility in the transportation stock prices. As a result, the forecasting stock prices accuracy in this specific domain remains challenging, specifically when the existing approaches often ignore or insufficiently model these seasonal dynamics. In order to improve the accuracy of predictive model, it is important to build models that explicitly account for such recurring fluctuations.

Recent studies have investigated multiple artificial intelligence approaches to analyse seasonal fluctuations in stock prices. Linear Regression is applied to the energy sector in North Lampung to model the relationship between stock prices, fuel prices, seasonal trends, and consumer behaviour [8]. The research highlights how well the model captured general stock trends without requiring de-trending or de-seasonalizing. However, the model's accuracy varied across months and depended greatly on the input data and external adjustments. In contrast, Bidirectional Long Short-Term Memory (BiLSTM), one of deep learning models, have highlighted better capability in modelling complex time-series patterns. The PT Garuda Indonesia (GIAA.JK) case demonstrated that BiLSTM achieved low prediction errors by learning temporal dependencies from both forward and backward sequences [9]. From a wider perspective, LSTM-based models are known to perform well, yet they often need careful hyperparameter tuning to achieve optimal results [2], [10]. Also, research on Tesla Inc. (TSLA) showed that combining GRU with VADER sentiment analysis boosted forecasting performance, reaching a Mean Absolute Error (MAE) of 6.93 [11]. GRU captured both long-term and short-term dependencies more effectively and computationally efficiently than LSTM because of its simpler architecture. However, the research highlighted that GRU also required proper tuning to perform optimally. Aside from stock prediction, the combination of XGBoost and Bidirectional Gated Recurrent Unit (BiGRU) has been applied to seasonal photovoltaic energy production [12]. While XGBoost itself achieved strong predictive accuracy, the research showed that its error patterns differed significantly from BiGRU's, suggesting that XGBoost lacks the ability to fully model temporal dependencies. By combining both models, the performance improved which emphasizing the XGBoost's limtitation and the benefit of ensemble methods for capturing time-based patterns more effectively.

Regardless of the advancements, the research that explicitly incorporates seasonality into stock price prediction models for Indonesia's transportation sector is still limited. The existing researches concentrate on general time-series trends or particular occasions, such as the COVID-19 pandemic, yet ignore the unique seasonal patterns that affect stock volatility in this sector–such as holiday-driven demand escalation or fuel consumption cycles. This gap highlights the need for forecasting approaches that not only model market dynamics but also account for recurring seasonal factors to improve prediction accuracy.

Therefore, this study investigates how Machine Learning algorithms can be utilized to capture seasonality patterns in forecasting stock prices within Indonesia's transportation sector. Specifically, this research compares four predictive models—Linear Regression [8], [13], XGBoost [2], [12], LSTM [10], and GRU [11]–that have each demonstrated strong individual performance in prior studies. While some previous research employed BiLSTM to leverage bidirectional temporal dependencies [10], this study adopts standard LSTM to align with the unidirectional nature of real-world forecasting, where future data is not accessible at prediction time. LSTM still offers strong capabilities in modelling sequential data, particularly relevant for capturing seasonal patterns. This study provides a comparative approach, consistent pre-processing, and evaluation conditions, focusing on its ability to catch stock price behavior that is caused by seasonal trends of short-term and long-term fluctuations as novel. The aim of this research to identify which method delivers the most robust and accurate predictions for practical stock management in a seasonal context.

2. METHODS

This research is an applied quantitative study focusing on stock price prediction in Indonesia's transportation sector by incorporating seasonal patterns. The research design consists of three main phases: data collection and understanding, data preprocessing, and model generation. The dataset, comprising secondary stock price entries, was retrieved from Yahoo Finance using Python libraries. The first phase, data collection and understanding, involves performing Exploratory Data Analysis (EDA) to observe the dataset, identify seasonal patterns, and detect outliers. The second phase, data preprocessing, involves value interpolation, feature engineering, train–test splitting, data scaling, and reshaping the data for sequential models to establish compatibility with the machine learning algorithms. The third phase, model generation, involves constructing, training, evaluating, and fine-tuning models to determine the best-performing approach. The fourth phase, model evaluation, performs statistical evaluation using Kruskal-Wallis test to determine the best model based on their evaluation metrics. A clearer flow of the research design is illustrated in Figure 1.

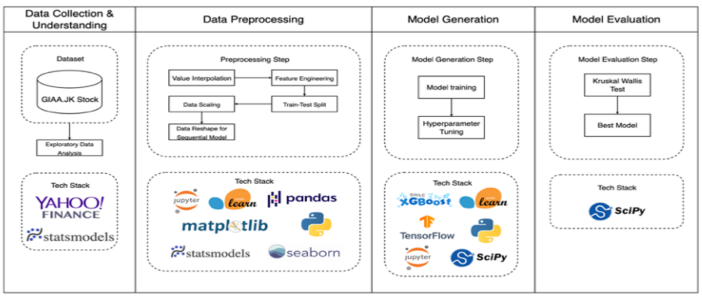

Figure 1. Proposed Methodology

2.1 Data Collection and Understanding

This study utilized monthly historical stock data from PT Garuda Indonesia (Persero) Tbk (GIAA.JK) spanning August 2019 to May 2025. The dataset contained trading information such as date, month, year, open, low, high, volume, and close. As publicly available financial data, it was retrieved using the Yahoo Finance library in Python. For analysis, the data was divided into three subsets: 2019 data, training data, and testing data. The 2019 data was specifically employed to construct lag features, ensuring that the earliest records in the training set had sufficient historical context for generating time-dependent features. Figure 2 presents an excerpt of the dataset, which highlights periods of stagnation between mid-2021 and late-2022 that later required preprocessing to ensure data quality. To obtain further comprehension into the dataset's features, an Exploratory Data Analysis (EDA) will be conducted.

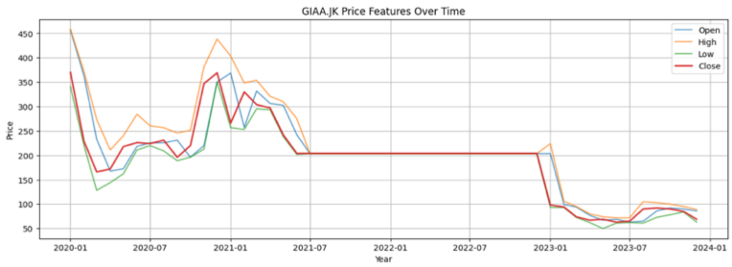

Figure 2. Price Features Over Time

2.2 Data Preprocessing

This phase aims to transform raw data into a clean, structured, and model-ready format. This step is important to improve prediction accuracy, optimize model performance, and eliminate irrelevant or redundant information. The first procedure involves detecting stagnant values by interpolation. Then, feature engineering will include the Autocorrelation Function (ACF) and generate new features.

After the feature construction, the dataset is split into training and validation subsets based on temporal sequence. Data prior to 2023 is assigned for training, while the rest of the data is reserved for validation. Since the study utilizes both traditional machine learning and deep learning models, MinMaxScaler is used to standardize feature values, ensuring compatibility with deep learning architectures. Lastly, sequence construction is carried out using a sliding window approach based on the identified lag.

2.3 Model Generation

This involves the construction and training of forecasting models. This study compares the performance of traditional machine learning models, Linear Regression and XGBoost, with deep learning approaches, LSTM and GRU. The Linear Regression [14] and XGBoost [15], [16] models are implemented using their default parameters, serving as a basic approach without manual tuning. In contrast, the LSTM and GRU models are designed with custom architectures adapted from previous studies that have shown effectiveness in the case of time seires forecasting [2]. The training process for the learning models incorporates regularization techniques such as dropout and early stopping to reduce overfitting.

After the baseline training, hyperparameter tuning is applied to XGBoost, LSTM, and GRU to optimize model performance. The configuration of parameters for each model is summarized in Table 1. For XGBoost, the selection of parameters is guided by previous research examining the effects of individual parameter values [17] and tuning is performed using Grid Search. For LSTM and GRU, the tuning process follows established practices from the existing literature [10], [18] and is performed using Random Search in Keras. All models are trained and evaluated on their respective training and validation sets , with performance measured using Root

Mean Square Error (RMSE) [10], [12], Mean Absolute Error (MAE) [11], [12], and Mean Absolute Percentage Error (MAPE) [12], [19].

Table 1. Model Parameter Configuration

Model	Parameter	Description
XGBoost	Learning_rate	[0.0001, 0.001, 0.01]
	N_estimators	[100, 200, 300]
	Max_depth	[3, 5, 7]
LSTM	Units_1	[64, 128, 192, 256]
	Units_2	[32, 64, 96, 128]
	Dropout	[0.2, 0.3, 0.4, 0.5]
	Dense_units	[9, 25, 40, 56]
	Activation	[relu, tanh, leaky_relu]
	Learning_rate	[0.001, 0.0005, 0.0001]
GRU	Units_1	[64, 128, 192, 256]
	Units_2	[32, 64, 96, 128]
	Dropout	[0.2, 0.3, 0.4, 0.5]
	Dense_units	[9, 25, 40, 56]
	Activation	[relu, tanh, leaky_relu]
	Learning_rate	[0.001, 0.0005, 0.0001]

2.4 Model Evaluation

The last phase focuses on selecting the best forecasting model through statistical evaluation. To achieve this, the Kruskal-Wallis test [20] was utilized as a non-parametric method to examine whether the observed differences in model performance across multiple evaluation metrics–RMSE, MAE, and MAPE–were statistically significant. This approach was chosen because it does not assume a normal distribution of errors, making it suitable for comparing models trained on financial time-series data.

If a statistically significant difference is detected for any of the evaluation metrics, that metric is prioritized in identifying the best-performing model. The model demonstrating superior performance under this criterion is then selected for further testing. Once the best model is determined, a forecasting test is conducted using the designated unseen dataset to assess its generalization capability.

3. RESULTS AND DISCUSSION

This part presents the results obtained from the proposed methodology and provides an profound discussion of their implications. The results are reported through both quantitative metrics and visualizations to illustrate the models' performance more clearly. Each model, including its tuned variant, is compared based on standard evaluation measures such as RMSE, MAE, and MAPE. Beyond the numerical outcomes, this section also discusses the significance of the outcomes, interprets the differences in model performance, and highlights their relevance for forecasting seasonal stock price behavior in Indonesia's transportation sector.

3.1 Exploratory Data Analysis

After conducting the EDA on the dataset, several insightful patterns were identified. Figure 3 illustrates the seasonality index, showing recurring peaks at the end and beginning of each year. Meanwhile, Figure 4 presents the seasonal decomposition, revealing a downward trend in stock prices, the presence of seasonal components, and residual fluctuations that capture irregular variations.

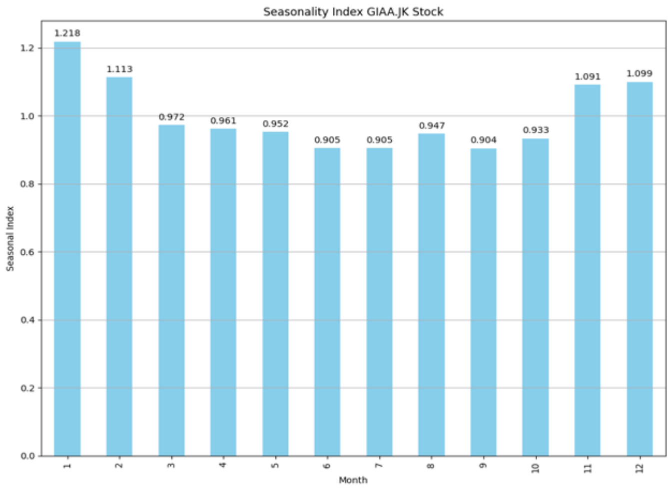

Figure 3. Seasonality Index

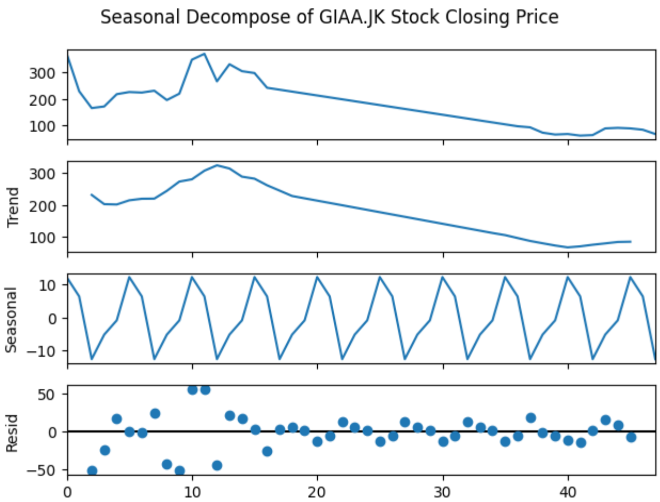

Figure 4. Seasonal Decomposition

3.2 Data Preprocessing

Based on the characteristics of the dataset, multiple preprocessing methods were applied to improve the model's performance. Handling stagnant values through interpolation proved to be crucial. An experiment was conducted without interpolation, which resulted in poor performance across all models. Specifically, the best-performing model under this setting produced an MAE of 35.51, RMSE of 36.77, and MAPE of 6942.88%. Therefore, interpolation was deemed necessary. A first-order polynomial interpolation was applied, replacing prolonged constant values with linear trends that better preserved the underlying data behavior. As shown in Figure 5, the interpolated closing price exhibits a more natural slope compared to the unprocessed values.

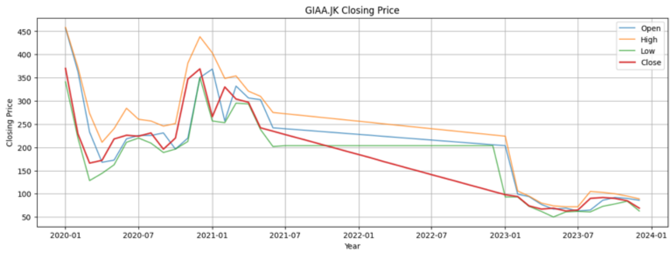

Figure 5. Value Interpolation Result

Feature engineering in this research involved two main steps: applying the ACF and generating lag features. Based on the ACF analysis, the optimal lag was determined to be five. A multivariate analysis was also conducted to assess the relationships among features. As illustrated in Figure 6, the variables *high*, *low*, and *open* showed strong associations with the *close* price. However, only the *close* and *high* features were used to construct lag features due to their highest correlation values. Both past and future lags (five periods each) were incorporated, and the features were generated using the 2019 dataset and the test data to ensure alignment with the model's input requirements. For the deep learning models, the input sequences were constructed using five consecutive time steps, consistent with the identified lag structure.

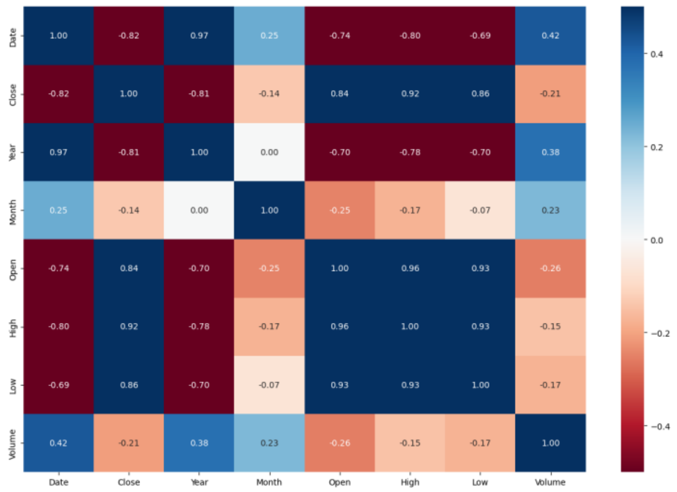

Figure 6. Multivariate Heatmap

3.3 Model Performance Evaluation

In this research, both traditional machine learning models and deep learning were trained using the same preprocessed dataset. All models, including their tuned variants, were evaluated on the validation dataset using RMSE, MAE, and MAPE as evaluation metrics. The training outcomes are presented in Table 2. Among traditional machine learning approaches, XGBoost achieved an RMSE of 6.70, MAE of 44.87, and MAPE of 63.71%, outperforming Linear Regression (RMSE of 7.46, MAE of 55.60, MAPE of 70.11%) and its tuned variant (RMSE of 7.75, MAE of 60.03, MAPE of 84.29%), demonstrating its capability to capture nonlinear relationships in the data. Interestingly, the tuning process did not improve XGBoost's performance, possibly due to a limited hyperparameter search space or suboptimal parameter values selected during tuning. In contrast, the deep learning models showed stronger results: the baseline LSTM recorded an RMSE of 5.98, MAE of 35.70, and MAPE of 55.79%, while the tuned LSTM slightly worsened to RMSE of 6.69, MAE of 44.79, and MAPE of 70.03%. The

GRU achieved a remarkable improvement over all other models, with an RMSE of 0.89, MAE of 0.79, and MAPE of 1.25%. Its tuned variant further enhanced performance, achieving the best overall results with RMSE of 0.51, MAE of 0.26, and MAPE of only 0.42%, clearly demonstrating its strong capability to capture seasonal trends of both short-term and long-term fluctuations.

For XGBoost, the tuned variant used a max_depth of 3, n_estimators of 300, and learning_rate of 0.01. The number of estimators and depth were intended to capture more complex nonlinearities, while the lower learning rate was meant to stabilize convergence. However, these adjustments may have caused overfitting or insufficient exploration of parameter space, which explains why the tuned model performed worse than the baseline.

For the LSTM, the tuned model employed two hidden layers with 64 units each, the *tanh* activation function, and a dropout rate of 0.2. These parameters were chosen to balance representational capacity with regularization. Nevertheless, the tuning increased model complexity, which may have amplified overfitting and degraded predictive performance compared to the baseline LSTM.

For the GRU, the tuned configuration consisted of 128 and 96 units in the first and second hidden layers, a dropout rate of 0.3, a learning rate of 0.0001 and the *tanh* activation function. The larger number of units allowed the model to learn richer temporal dependencies, while dropout provided regularization against overfitting. The choice of *tanh* stabilized training dynamics, and the small learning rate ensured gradual convergence. This combination resulted the most accurate and generalizable outcomes among all tested models.

Table 2. Model trained Performance

Model	RMSE	MAE	MAPE
Linear Regression	7.46	55.60	70.11%
XGBoost	6.70	44.87	63.71%
Tuned XGBoost	7.75	60.03	84.29%
LSTM	5.98	35.70	55.79%
Tuned LSTM	6.69	44.79	70.03%
GRU	0.89	0.79	1.25%
Tuned GRU	0.51	0.26	0.42%

3.4 Kruskal-Wallis Test

The research did not directly select the best model for testing. Instead, all models were evaluated across three metrics using the Kruskal–Wallis test to determine whether the performance differences were statistically significant. For this test, RMSE, the MAE, and MAPE values obtained from the forecasting results on the training dataset (Table 2) were used as input. The results, presented in Table 3, show that among the three metrics, only MAPE resulted a p-value of 0.01, which is below the significance threshold of 0.05, showing a statistically significant difference in model performance for this metric. In contrast, both RMSE ($H = 9.96$, $p = 0.13$) and MAE ($H = 9.96$, $p = 0.13$) showed no significant differences across models, likely due to their low variance. Given that MAPE demonstrated statistical significance, it was chosen as the deciding metric for model selection. Consequently, the tuned GRU, which achieved the lowest MAPE, was selected for testing.

Table 3. Kruskal Wallis Test Result

Model	RMSE	MAE	MAPE
Linear Regression	7.46	55.60	70.11%
XGBoost	6.70	44.87	63.71%
Tuned XGBoost	7.75	60.03	84.29%

3.5 Testing

Based on the Kruskal-Wallis test resulr, the tuned GRU model emerged as the most suitable candidate and was applied to unseen testing dataset. The testing data underwent the same preprocessing steps as the training and validation sets to ensure consistency. After preprocessing, six valid entries were available for prediction. The tuned GRU demonstrated promising performance, achieving an MAE of 5.90, RMSE of 7.33, and MAPE of 9.67%. The MAPE score places the model in the 'excellent' category of predictive accuracy. Table 4 presents a detailed comparison between stock price predictions and actuals, while Figure 7 provides a visualization of the results. The visual confirms that predicted and actual values align closely, showing that the model generalizes well to unseen data. In practical terms, these results suggest that implementing the tuned GRU model for real-world stock price forecasting may provide reliable guidance without imposing significant risk to investors.

Table 4. Predicted and Actual Stock Prices

Predicted Stock Price	RMSE
56	71
58	55
59	62
58	56
57	50
55	60

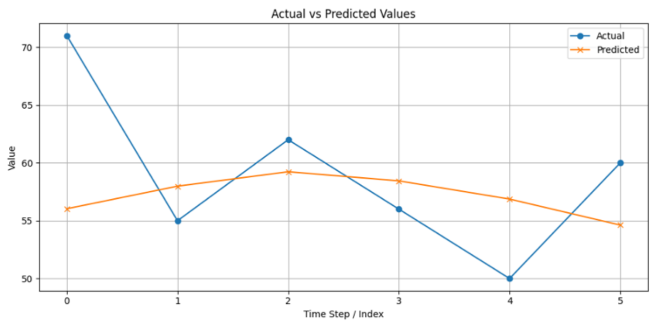

Figure 7. Actual and Predicted Stock Price Comparison

4. CONCLUSIONS

Based on the research results, Indonesia's transportation sector, particularly PT Garuda Indonesia (Persero) Tbk (GIAA.JK), is shown to exhibit seasonal patterns. By comparing Linear Regression, XGBoost, LSTM, and GRU, and applying hyperparameter tuning alongside statistical validation through the Kruskal-Wallis test, the tuned GRU model exhibited the best performance. The tuned GRU achieved superior performance on both validation and testing datasets, with a MAE of 5.90, RMSE of 7.33, and MAPE of 9.67% on unseen data, which falls within the "excellent" category of forecasting performance. Thus, the research demonstrates the effectiveness of deep learning models–particularly GRU–in capturing seasonal trends of short-term and long-term fluctuations in transportation sector stock prices. The integration of seasonality into model construction, combined with thorough statistical evaluation, strengthens the reliability of the forecasting outcomes and underscores their practical applicability for investment decision-making.

Nevertheless, the research is limited by its focus on a single stock, the use of monthly frequency, a relatively small testing dataset, and the presence of stagnant values that affect performance. Future research should explore using different data frequencies, expanding the testing dataset, and including multiple companies in the transportation sector that do not exhibit data anomalies. Hybrid or ensemble models that combine traditional and deep learning approaches are also recommended. Additionally, incorporating external factors such as public sentiment, global events, or political and legislative actions is suggested to enrich the seasonal pattern analysis. Such recommendations would further validate the robustness and generalizability of the proposed methodology. Ultimately, the findings suggest that the tuned GRU model holds strong potential as a reliable tool for stock forecasting in seasonal and volatile sectors such as transportation.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Lawrence, G. S. Thattil, T. Antony, and J. Ittoop, “Stock Market Movements - A Comparative Analysis on Growth Patterns in India’s Leading Stock Exchange,” *International Journal of Economics and Financial Issues* , vol. 15, no. 2, pp. 235–243, Feb. 2025, doi: 10.32479/ijefi.17865.

[2] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, “Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications,” Sep. 01, 2023, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/ijfs11030094.

[3] L. Chin, Y. S. Foo, K. S. Chen, F. Taghizadeh-Hesary, and W. L. Lin, “Sustainability of Stock Market Against COVID-19 Pandemic,” *International Journal of Economics and Management*, vol. 16, no. SpecialIssue1, pp. 33–43, 2022, doi: 10.47836/IJEAMSI.16.1.003.

[4] Tiara, Munawarah, and M. Yunus Kasim, “NAVIGATING FINANCIAL TURBULENCE: HOW DISTRESS AFFECTS STOCK PRICES IN INFRASTRUCTURE, UTILITIES, AND TRANSPORTATION,” *Jurnal Keuangan dan Bisnis (JKB) ISSN*, vol. 22, no. 2, pp. 160–176, 2024, doi: https://doi.org/10.32524/jkb.v22i2.1205.

[5] S. Perkasa, “Beda Mudik Zaman Dulu dan Masa Kini: Pergeseran Moda hingga Esensi,” https://www.metrotvnews.com/read/KYVC4nVB-beda-mudik-zaman-dulu-dan-masa-kini-pergeseran-moda-hingga-esensi.

[6] K. Maharani, “Mengenal Saham Transportasi dan Potensinya di Pasar Saham,” https://reku.id/campus/mengenal-saham-transportasi-dan-potensinya-di-pasar-saham.

[7] A. A. Putri, “Transportasi Pilihan Masyarakat untuk Mudik Lebaran 2023,” https://data.goodstats.id/statistic/transportasi-pilihan-masyarakat-untuk-mudik-lebaran-2023-hQaqh.

[8] A. A. Afandi, “Forecasting Pertalite Stock Expenditures Using Exponential Smoothing and Linear Regression,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 18, no. 4, Oct. 2024, doi: 10.22146/ijccs.98849.

[9] D. A. Prasetyo and R. Rokhim, “Indonesian Stock Price Prediction using Deep Learning during COVID-19 Financial Crisis,” vol. 3, no. 2, pp. 64–70, 2022, doi: https://doi.org/10.46336/ijbesd.v3i2.273.

[10] H. Dhake, Y. Kashyap, and P. Kosmopoulos, “Algorithms for Hyperparameter Tuning of LSTMs for Time Series Forecasting,” *Remote Sens (Basel)*, vol. 15, no. 8, Apr. 2023, doi: 10.3390/rs15082076.

[11] A. Mamillapalli, B. Ogunleye, S. Timoteo Inacio, and O. Shobayo, “GRUvader: Sentiment-Informed Stock Market Prediction,” *Mathematics*, vol. 12, no. 23, Dec. 2024, doi: 10.3390/math12233801.

[12] G. Xiong, J. Zhang, X. Fu, J. Chen, and A. W. Mohamed, “Seasonal short-term photovoltaic power prediction based on GSK–BiGRU–XGboost considering correlation of meteorological factors,” *J Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/s40537-024-01037-x.

[13] J. P. Halawa, A. Hermawan, and . J., “Implementation of Linear Regression Algorithm to Predict Stock Prices Based on Historical Data,” *bit-Tech*, vol. 5, no. 2, pp. 103–112, Dec. 2022, doi: 10.32877/bt.v5i2.616.

[14] K. Qu, “Research on linear regression algorithm,” *MATEC Web of Conferences*, vol. 395, p. 01046, 2024, doi: 10.1051/matecconf/202439501046.

[15] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A Pedestrian Detection Method Based on Genetic Algorithm for Optimize XGBoost Training Parameters," *IEEE Access*, vol. 7, pp. 118310–118321, 2019, doi: 10.1109/ACCESS.2019.2936454.

[16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[17] H. Dong, D. He, and F. Wang, "SMOTE-XGBoost using Tree Parzen Estimator optimization for copper flotation method classification," *Powder Technol*, vol. 375, pp. 174–181, Sep. 2020, doi: 10.1016/J.POWTEC.2020.07.065.

[18] S. khan *et al.*, "Optimizing deep neural network architectures for renewable energy forecasting," *Discover Sustainability*, vol. 5, no. 1, Dec. 2024, doi: 10.1007/s43621-024-00615-6.

[19] I. Muthahharah, S. M. Meliyana, and Z. Mar'ah, "Forecasting Indonesia's Wholesale Price Index (WPI) Using the Holt's Exponential Smoothing Method," *Quantitative Economics and Management Studies*, vol. 6, no. 2, 2025, doi: 10.35877/454RI.qems.

[20] R. A. K. Sherwani, H. Shakeel, W. B. Awan, M. Faheem, and M. Aslam, "Analysis of COVID-19 data using neutrosophic Kruskal Wallis H test," *BMC Med Res Methodol*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12874-021-01410-x.