Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

Web Application for IHSG Prediction Using Machine Learning Algorithms

Andryan Kalmer Wijaya[1], Henry Lucky[2], Samsul Arifin[3,*]

[1]*Department of Statistics, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia*
[2]*Department of Computer Science, School of Computer Science, Bina Nusantara University, Jakarta, 11480, Indonesia*
[3]*Department of Data Science, Faculty of Engineering and Design, Institut Teknologi Sains Bandung, Bekasi, 17530, Indonesia*
[*]*Corresponding author: samsul.arifin@itsb.ac.id*

©2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License-(CC-BY-SA) (https://creativecommons.org/licenses/by-sa/4.0/)

Article Info

Article history:
Received: March 27, 2025
Revised: May 3, 2025
Accepted: May 4, 2025
Available online: June 1, 2025

DOI:
https://doi.org/10.71385/idjams.v2i1.21

Abstract

This study investigates the effectiveness of the Long Short-Term Memory (LSTM) method in predicting the stock price of the Composite Stock Price Index (CSPI). LSTM, a variant of Recurrent Neural Networks, is designed to overcome challenges such as the vanishing gradient problem and long-term dependencies in time-series data. Given the dynamic and volatile nature of financial markets, accurate stock price prediction is crucial for investors and analysts. The data set used in this study consists of daily CSPI prices from January 2000 to December 2023, which serve as both training and testing data for model development. The LSTM model is trained to forecast the next day's stock price, and its performance is compared with traditional statistical models, particularly the Autoregressive Integrated Moving Average (ARIMA) model and linear regression. Performance evaluation is based on the Mean Absolute Percentage Error (MAPE), a widely used metric for assessing predictive accuracy. The results indicate that while the ARIMA model achieves a lower MAPE of 0.7%, demonstrating slightly superior accuracy, the LSTM model also performs well, with a MAPE of approximately 1%. These findings suggest that while statistical models like ARIMA remain highly effective for stock price forecasting, deep learning approaches such as LSTM still offer promising predictive capabilities, especially when handling large and complex datasets. The ability of LSTM to capture non-linear patterns and temporal dependencies makes it a viable alternative for financial forecasting, potentially benefiting traders and market analysts seeking data-driven decision-making tools.

Keywords: LSTM; Recurrent Neural Network; Price Forecasting; Composite Stock Price Index; Linear Regression

I. Introduction

Stocks or shares are one of the popular financial instruments in society. Stocks are one way for companies to raise funds by selling a portion or all the company's ownership. Through this, the public or potential investors can easily own a large company that has gone public. Quoting Kompas, throughout the year 2021, the Indonesia Stock Exchange recorded a 10.08 percent year-to-date increase in the Composite Stock Price Index (IHSG). The sectors that supported IHSG's performance were the basic industry and financial sectors, with increases of 53.2 percent and 50.52 percent, respectively. There were also sectors that experienced a decrease, namely the consumer and property sectors, with values of -43.59 percent and -42.01 percent [1] [2]. The IHSG is a key measure reflecting the performance of the stock market in the Indonesia Stock Exchange (BEI). By mirroring the stock price movements of various

companies listed on the Indonesian stock market, the IHSG plays a crucial role. The IHSG serves as an indicator of the health of the capital market, providing insight into investor optimism or uncertainty regarding the economy and corporate growth in Indonesia. Furthermore, the IHSG serves as a reference for investment decision-making by individuals, financial institutions, and fund managers. The movement of the IHSG can provide insights into investor sentiment toward the national and global economic conditions. Therefore, the IHSG is not just a number but also a reflection of the dynamics of the Indonesian economy and stock market that is relevant to analysts, market participants, and the government [3] [4].

The Central Bureau of Statistics (BPS) reported that Indonesia's economy in the second quarter of 2021 experienced a year-on-year growth of 7.07 percent. Furthermore, Indonesia's economy in the second quarter of 2021 saw a 3.31 percent quarter-to-quarter growth compared to the previous quarter. The improvement in Indonesia's economy was mainly driven by enhanced performance in exports, household consumption, investments, and government consumption. This economic improvement indicates that Indonesia has successfully rebounded after facing pressure [5] [6]. In the modern era, computerization rapidly evolved, including the field of artificial intelligence. One branch of this field is machine learning, which is an algorithm that can be learned similarly to humans. Artificial intelligence has several branches or applications, including machine learning, natural language processing, expert systems, vision, speech, planning, and robotics. Machine learning technology is developed to learn independently without guidance from its creator, based on statistical principles. It learns from existing data to predict future activities or outcomes [7] [8].

Machine learning was initially discovered by mathematicians, including Adrien Marie Legendre, Thomas Bayes, and Andrey Markov in the 1920s. It was further developed by the American company IBM with the Deep Blue project in 1996. The Deep Blue project is an example of machine learning application, where IBM tested it to play chess against professional chess champions and successfully won. Long Short-Term Memory (LSTM) networks are a modification of Recurrent Neural Networks (RNNs). One of LSTM's advantages is its ability to store information for a long time, addressing the vanishing gradient issue in RNNs and improving prediction accuracy [9] [10]. Time series data can be categorized into two types: cross-sectional data collected at a specific time and time series data collected over a period. Time series data is used in time series analysis to model data and predict future values by utilizing past data. One model that can be used for this purpose is the Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model assumes data stationarity and ignores variable independence. The forecasts generated by this model can assist in decision-making and policymaking since the effectiveness of a decision depends on many factors that are not apparent at the time the decision is made [11] [12].

In recent years, Machine Learning technology has become extremely popular in various fields, including financial market analysis. One popular Machine Learning technique for time series forecasting is LSTM, a type of Recurrent Neural Network (RNN) that excels at forecasting time series due to its ability to recognize complex patterns and relationships in data [13]. This research aims to apply Machine Learning technology, particularly the LSTM model, to predict the future movement of the IHSG. Additionally, this research will compare the performance of the LSTM model with the ARIMA model, which is a traditional statistical method widely used in time series forecasting [14]. The results of this research will be implemented in a web-based application that is user-friendly. The implementation will use Rshiny, which is based on R. This application aims to provide insights to users regarding the accuracy of LSTM predictions for stocks, and users will have the power to adjust machine learning learning variables [15] [16].

RShiny was chosen as the web application development platform in this study due to its tight integration with the R programming language, which is often used in statistical analysis and data modeling. RShiny allows interactive visualization and model implementation with a user-friendly interface without requiring in-depth expertise in web development. However, other alternatives such as Flask and Django (based on Python) offer greater flexibility in the back end and can be easily combined with various machine learning libraries. Meanwhile, Streamlit is a lighter and faster choice in building

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

machine learning-based applications without the need for complex configurations. Therefore, the selection of RShiny should be supported by stronger justification, such as consideration of compatibility with the research ecosystem or ease of integration with the analysis tools used in this study [17] [18].

LSTM is chosen in this study because of its ability to handle time series data with long-term dependencies, which are often a challenge in IHSG prediction. Compared with other methods such as Transformer or XGBoost, LSTM has the advantage of remembering long-term historical patterns through gating mechanisms, which helps in overcoming the vanishing gradient problem that often occurs in conventional RNN models. Although Transformer is increasingly popular in time-series forecasting, its implementation and training require larger computational resources than LSTM, especially for datasets with daily frequencies such as IHSG. On the other hand, XGBoost as a decision tree-based model tends to be more effective for tabular data with diverse features, while stock price prediction relies more on data sequence patterns that are better handled by RNN-based architectures such as LSTM. Therefore, LSTM is chosen as the main model in this study because of its balance between the ability to capture long-term patterns and better computational efficiency than Transformer [19] [20]. Figure 1 contains an illustration of 2 models used in this study.

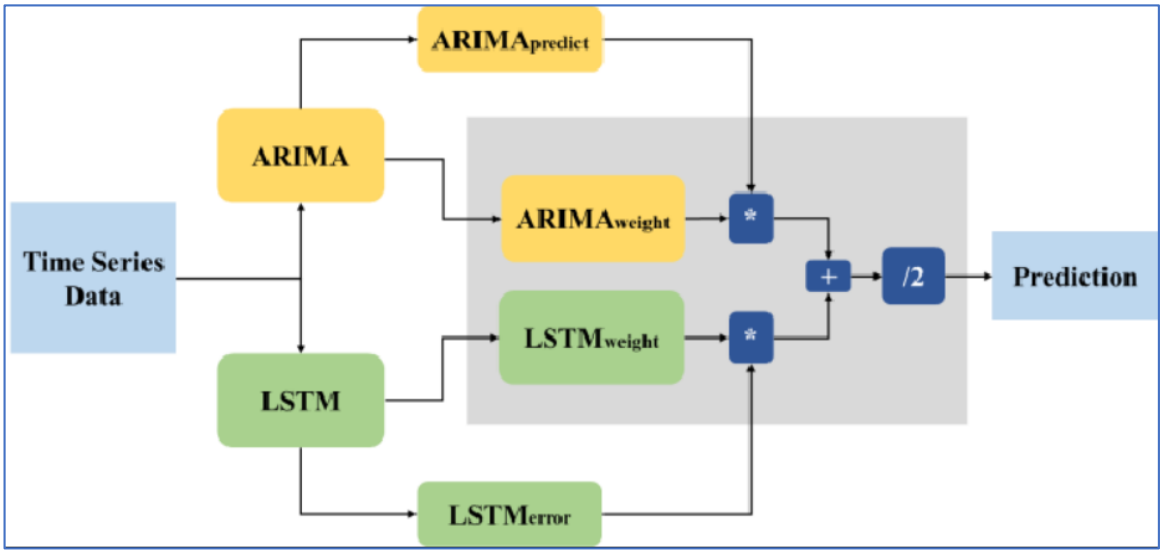


Figure 1. Illustration of 2 Models Used in This Study [21]

II. Research Methods

The methodology used in this research follows a structured problem-solving framework consisting of five stages: the Communication Stage (Light Blue), where the problem is identified, objectives are defined, and initial discussions take place to understand the scope; the Planning Stage (Light Green), which involves gathering data, selecting appropriate machine learning models, and designing the system architecture; the Modeling Stage (Yellow), where machine learning models are developed, trained, and evaluated based on performance metrics; the Construction Stage (Orange), focusing on implementing the models into a functional web-based application with necessary coding and system integration; and finally, the Deployment Stage (Red), where the application is tested, deployed for real-world use, and monitored for performance improvements. Each stage plays a crucial role in ensuring an effective and efficient implementation of machine learning for IHSG prediction on a web-based platform [22] [23]. The framework of thinking used in this study can be seen in Figure 2.

The Stage of Communication (Light Blue) focuses on determining the speed and accuracy of artificial intelligence predictions in forecasting time series data. Given the rapid advancements in artificial intelligence, this research aims to evaluate how well AI can compete with traditional methods in terms of accuracy and speed. The study employs the LSTM method as an artificial intelligence approach and

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

ARIMA as a comparative method to analyze their respective advantages and limitations. To achieve this, the Analysis Stage consists of several key steps: Data Collection, where IHSG data is gathered from reliable sources such as the Indonesia Stock Exchange (BEI) or leading financial services, including IHSG closing prices over the past few years; Data Preprocessing, which involves cleaning the data, handling missing values, and performing normalization before using it in the forecasting models; LSTM Modeling, where an LSTM model is developed with three hidden layers—two LSTM layers with 50 neurons each and one dense layer; ARIMA Modeling, where an ARIMA model is created using the auto.arima function from the forecast library; RMSE Testing, where the Root Mean Square Error (RMSE) is calculated for both models, with LSTM using the last RMSE value from training and ARIMA obtaining RMSE from the summary function; MAE Testing, where the Mean Absolute Error (MAE) is calculated for LSTM using the MAE function and for ARIMA using the summary function; and MAPE Testing, where the Mean Absolute Percentage Error (MAPE) is computed by multiplying the MAE value by 100 for LSTM and obtaining it directly from the summary function for ARIMA. These steps ensure a thorough evaluation of both methods, providing insights into their practical applications in IHSG forecasting [24] [25].

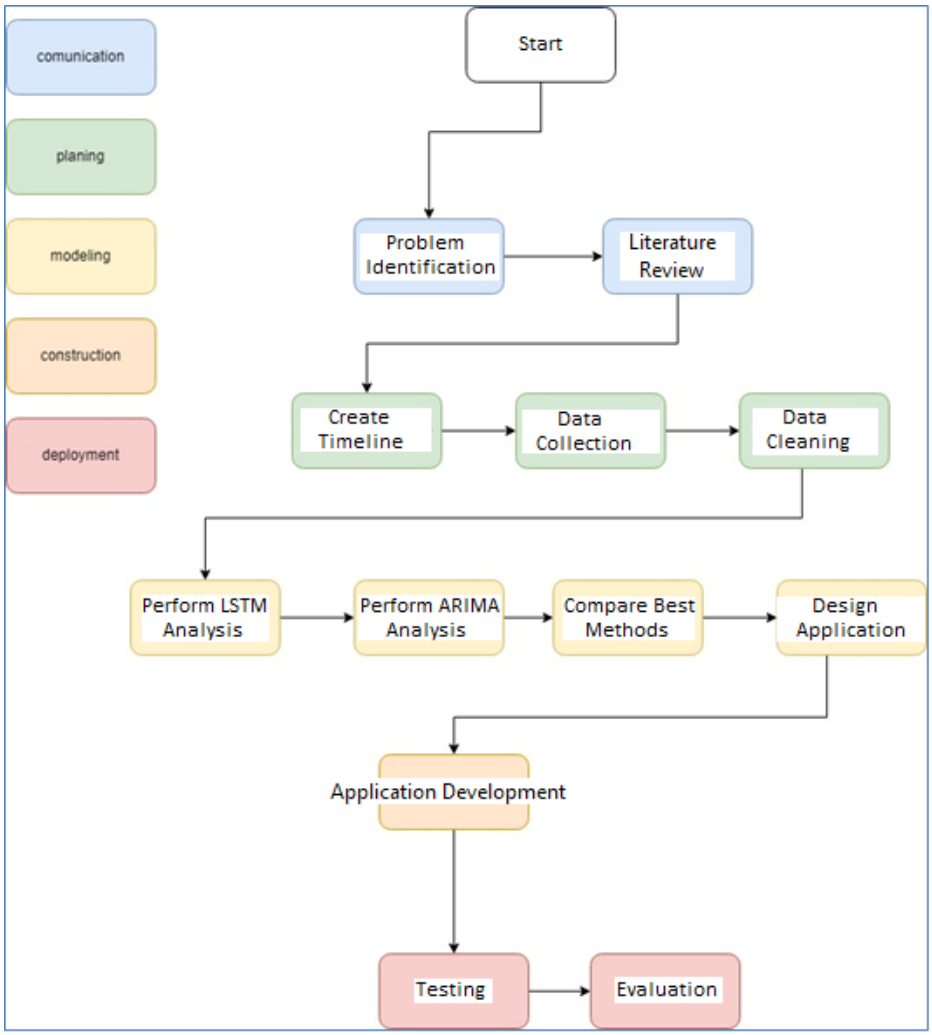


Figure 2. The Framework of Thinking Used in This Study [26]

Although we utilized the auto.arima() function for parameter estimation, we ensured that the model assumptions were met. Stationarity was tested using the Augmented Dickey-Fuller (ADF) test prior to

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

modeling. Additionally, residual diagnostics were conducted including the Ljung-Box test to examine autocorrelation and visual inspections of residual plots to verify white noise behavior. These steps were crucial to ensure the robustness and interpretability of the ARIMA model's forecasts, particularly in complex time series scenarios [27] [28]. The flowchart used in this study can be seen in Figure 3.

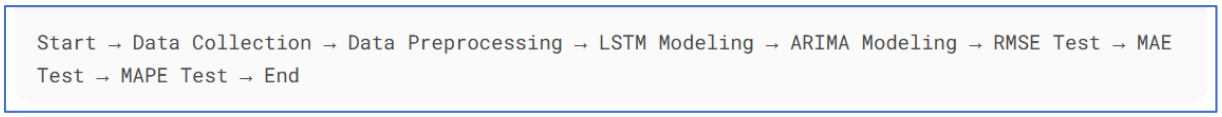


Figure 3. The Flowchart [29]

In the Planning Stage (Light Green), the author develops a timeline outlining the application's development, which is expected to take approximately 3.5 months. The application will be built on the R platform and deployed on the web, ensuring accessibility across various devices. Additionally, this stage involves studying the characteristics of LSTM, allowing users to understand its application in appropriate cases and gain insights into the nature of the data being processed. Proper planning in this phase ensures a structured and efficient development process, laying the foundation for the subsequent implementation stages. In the Modeling Stage (Yellow), the author conducts an in-depth analysis of the LSTM method by utilizing processed data and applying it to generate predictive results. These results are then systematically compared with those obtained from the ARIMA method, using the same analytical approach to evaluate performance differences. Through this comparative analysis, the best-performing method is identified based on key evaluation metrics. Once the optimal method is determined, the author proceeds with designing the application to ensure seamless integration of the selected machine learning model into a functional web-based platform [30] [31].

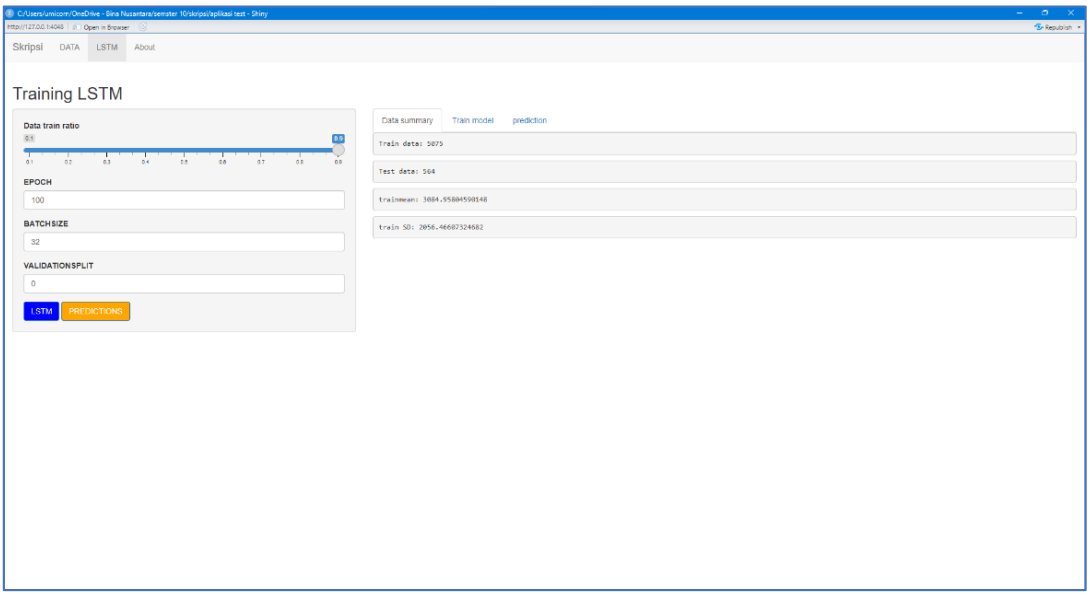


Figure 4. LSTM Data Summary [32]

The Construction Stage (Orange) focuses on the application development phase, starting with a low-fidelity design where the layout and buttons are structured to be interactive and user-friendly. At this stage, the author ensures that the interface is intuitive and accessible for users. Additionally, module integration begins to guarantee seamless functionality across different components of the application, ensuring that machine learning models, data processing, and visualization tools work together

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

effectively. This phase is crucial in refining the application's usability and performance before deployment [33] [34]. Figure 4 provides an overview of the resulting LSTM Data Summary.

On this page, users will have the freedom to create their own LSTM models with the aim of understanding the nature of LSTM and experimenting within the application. Users can utilize their time series data for prediction using LSTM. The accuracy of LSTM will be assessed through standard error metrics such as RMSE, MAE, and MAPE [35]. Furthermore, Figure 5 provides an overview of the resulting LSTM Prediction [36].

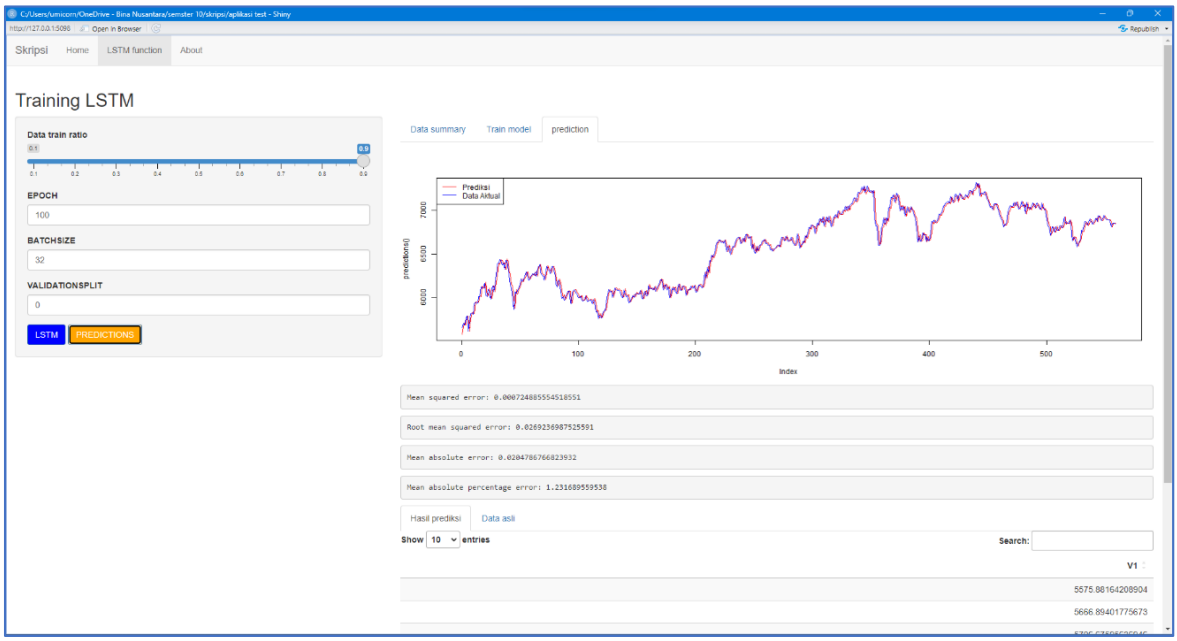


Figure 5. LSTM Prediction [37]

The page will display the prediction results and the standard errors of the LSTM predictions, enabling users to gauge the accuracy of LSTM in predicting the data. The application is built on existing libraries, including library(shiny), library(keras), library(tensorflow), library(caret), library(forecast), library(DT), library(ggplot2), library(shinycssloaders), library(dplyr), and library(shinythemes) [38] [39]. Some of the few lines of R code when creating an LSTM model used in this study can be seen in part in Figure 6.

```
# Membuat model LSTM
model <- keras_model_sequential() %>%
  layer_lstm(units = 50, input_shape = c(window_size, 1), return_sequences = TRUE) %>%
  layer_lstm(units = 50) %>%
  layer_dense(units = 1)
```

Figure 6. Few Lines of R Code When Creating an LSTM Model [40]

In this research, the LSTM model is created with 3 layers - 2 layers with 50 units each and 1 dense layer with 1 unit. The "return sequence" attribute is active, meaning the model can perform backpropagation during training, which leads to higher accuracy. During training, artificial intelligence's performance is measured by the mean squared error (MSE) value it produces. The smaller the value, the more accurate the predictions. The Deployment Stage (Red) is the final phase, where the web application and its integrated LSTM module are officially launched for real-world use. At this stage, the application

Indonesian Journal of Applied Mathematics and Statistics (IdJAMS)
ISSN 3064-4062
Volume 2 Issue 1 June, 2025, pp. 29-40
https://idjams.com/index.php/jms/index

undergoes thorough testing and evaluation based on the eight golden rules and blackbox testing principles to ensure its functionality, usability, and reliability. This process helps identify any potential issues and ensures that the application meets the intended requirements. Ultimately, the goal of this stage is to deliver a well-performing and user-friendly platform that effectively supports users in IHSG prediction and analysis [26] [39].

Moreover, the proposed method does not require the data to follow a specific distribution such as normality. However, for models like ARIMA, the underlying assumption is that the data should be stationary. For this reason, preliminary checks such as the ADF test were conducted to assess stationarity. If needed, differencing was applied to transform non-stationary series into stationary ones. LSTM, being a deep learning model, is more flexible and does not require strict distributional assumptions. This study incorporates the ARIMA model, a traditional time series forecasting technique based on the Box-Jenkins approach, as one of the baseline methods. The ARIMA model was selected and configured using the auto.arima() function from the forecast package in R, enabling automated identification of optimal parameters (p,d,q). The ARIMA model is mathematically expressed as ARIMA(p,d,q), where p denotes the number of autoregressive terms, d is the number of differences needed to make the series stationary, and q represents the number of lagged forecast errors in the prediction equation. The LSTM network is formulated using gate-controlled memory cells, allowing it to capture long-term dependencies in sequences [26] [36]. For model evaluation, we used the following metrics: The MAPE, MAE, and RMSE, defined as follows:

- "MAE = $\frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$"
- "RMSE = $\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$"
- "MAPE = $\frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|$""

III. Results and Discussion

This section presents the results and discussion of the research, focusing on the analysis of the LSTM model and its performance in predicting IHSG values. The evaluation includes an in-depth examination of Prediction According to Data Split Ratios, where different training and testing proportions are assessed to determine their impact on prediction accuracy. Additionally, the study explores The Influence of Batch Size on MSE, along with other key performance metrics such as RMSE, MAE, and MAPE. These analyses aim to provide insights into the effectiveness of the LSTM model and the factors that contribute to optimal forecasting performance [27].

Training the LSTM model occurs using the backpropagation algorithm, a machine learning technique designed to minimize errors, which represent the difference between the model's predictions and actual data values. The training process follows these steps: the LSTM model first receives input data from the training set, then predicts output values based on this input. It calculates the error between its predictions and the actual values, using this error to update the weights and biases of its neurons. This cycle—input data, prediction, error calculation, and weight/bias updates—continues iteratively until the error reaches a desired minimum. During each iteration, the LSTM model continuously adjusts its weight and biases to improve accuracy, as these parameters determine how it processes input data. The smaller the error, the better the model's performance in making predictions [41].

The research began with the collection of IHSG data from 2000 to 2023, sourced from Yahoo Finance. The dataset was then filtered to focus solely on closing price data, followed by a cleansing process to remove NULL and ERROR values. After data cleansing, a scaling process was applied as previously described in earlier chapters. A crucial phase in this study was data splitting, aimed at determining the optimal data split ratio for training and testing. The ratios tested included 50:50, 60:40, 70:30, 80:20, and 90:10, all applied to the IHSG dataset. During this stage, the model was trained with 100 epochs, a batch size of 32, and a validation split of 0. Since machine learning performance is influenced by batch size and epochs, this study also investigated the most suitable batch size while

maintaining a fixed epoch count of 100. After training the data using the LSTM model, the prediction accuracy was evaluated using RMSE, MAE, and MAPE. This chapter aims to identify the most effective data split ratio and batch size, ultimately determining the best combination for accurately predicting IHSG prices.

The influence of batch size on MSE was analyzed using a 90:10 data split ratio, where batch sizes of 32, 64, and 128 were examined to determine their impact on model performance. The results indicate that a batch size of 32 produces the best performance, as reflected in the lowest RMSE value of 0.02735312. This finding suggests that a smaller batch size allows for more effective model updates, leading to improved prediction accuracy in the LSTM model for IHSG forecasting. Table 1 provides an illustration of some of the effects of batch size on MSE.

Table 1. Effect of Batch Size on MSE

Batch size	MSE	RMSE
32	0.000748193	0.02735312
64	0.0008152588	0.02855274
128	0.0008634777	0.02938499

Once the plots are obtained, the next step is to calculate the error values, which will then be compared with ARIMA to assess the performance of the proposed model. This comparison is crucial in determining whether the model is over-trained or not. The evaluation is based on three key metrics: RMSE, MAE, and MAPE. Among these, MAPE serves as a benchmark for model accuracy, where a value close to 0 indicates that the predictions closely align with the original data. By analyzing these metrics, the study aims to determine the effectiveness of the LSTM model, and its potential advantages or limitations compared to ARIMA. Some scenarios on LSTM used in this study are described in Table 2.

Table 2. Some Scenarios on LSTM

Ratio	RMSE	MAE	MAPE
50:50	53.23046	38.24452	0.73023
60:40	53.05742	38.32594	0.6863142
70:30	55.37353	39.90018	0.680201
80:20	60.10875	43.72405	0.7345304
90:10	55.07792	41.87749	0.6423141

After experimenting with different data split ratios and generating the plot results, the next step involves validating the models using RMSE, MAE, and MAPE to assess their performance. The RMSE and MAE values for the ARIMA model are considerably larger compared to the values obtained for the LSTM model. This difference is primarily due to the nature of the data used by ARIMA, which remains in its original form without scaling, leading to larger numerical values in the error metrics. Some scenarios on ARIMA used in this study are described in Table 3.

Table 3. Some Scenarios on ARIMA

Ratio	RMSE	MAE	MAPE
50:50	0.3024161	0.2309936	5.189663
60:40	0.05622855	0.04227333	1.545754
70:30	0.04431479	0.03416687	1.544992
80:20	0.03552541	0.02753447	1.653176
90:10	0.02748326	0.0210417	1.259786

LSTM shows a higher MAPE than ARIMA, indicating that this model has a higher error rate in predicting the JCI. One possible cause is overfitting, where the model overfits the training data and is

less able to generalize the test data. This can happen because the number of epochs is too large or because of the lack of regularization mechanisms such as dropout layers. In addition, the choice of hyperparameters also plays an important role in model performance. The use of different batch sizes can affect model convergence, and tuning the number of neurons or layers in the LSTM can improve accuracy. Data quality can also be a major factor—if there is noise or anomalies in the historical JCI data, deep learning-based models such as LSTM can have difficulty in capturing the correct pattern. Therefore, more stringent data cleaning and exploration of additional features can improve model performance.

Based on the analysis, we conclude that the MAPE serves as the primary metric for evaluating model performance, and upon comparison, ARIMA consistently outperforms LSTM with significantly lower MAPE values. This demonstrates that ARIMA remains superior in terms of accuracy. Furthermore, these results are reinforced by the plotted predictions, which closely follow the trends of the original data across all data split ratios. The consistency of these findings further solidifies ARIMA's reliability in forecasting IHSG values. With this analysis, we conclude the discussion presented in this paper.

IV. Conclusions

Based on the research conducted in the previous chapters, the best LSTM model, as indicated by MAPE and RMSE, is the 90:10 ratio; however, this ratio results in predictions that are higher than the actual values. In the context of IHSG predictions, the goal is to have values that closely approximate the actual values without exceeding them, making the 70:30 ratio more suitable. ARIMA still outperforms LSTM in terms of accuracy, as demonstrated by the significantly smaller MAPE values. The best batch size is determined to be 32, as it results in the smallest RMSE and requires the same amount of time compared to other batch sizes. In terms of application, all functions work well, verified through blackbox testing, and the application functions adequately, though there is room for further development. Users can experiment without concerns about its capabilities since the application is web-based, allowing access from various devices. Based on these conclusions, several recommendations can be considered: the data used in this research serves as an example of how the LSTM method can be applied, and it is advisable to explore its application in different output types, such as classification. Additionally, for future application implementations, it is recommended to explore the possibility of using Python to expand capabilities beyond those offered by Rshiny.

One limitation of this study is the lack of discussion regarding the limitations of the study, especially related to data aspects, model complexity, and potential improvements in the future. The data used in this study is limited to IHSG with a certain time span, so it may not fully reflect broader market changes or other external factors that affect stock movements. In addition, although LSTM offers advantages in handling time-series data, this model has challenges such as higher computational requirements and sensitivity to hyperparameter tuning, which are not discussed in depth in this study. In the future, research can be expanded by exploring feature engineering techniques, using hybrid models, or comparing results with other approaches such as Transformer or XGBoost to improve prediction accuracy.

Acknowledgment

The authors would like to express their sincere gratitude to all parties who have contributed to the completion of this research on the Implementation of Machine Learning for IHSG Based on Web. Special appreciation goes to our academic advisors for their invaluable guidance, insightful feedback, and continuous support throughout the research process. We also extend our thanks to the institutions and platforms that provided the necessary data and computational resources, enabling the successful implementation of the machine learning models. Additionally, we acknowledge the contributions of our colleagues and peers for their constructive discussions and encouragement. Finally, we are grateful to our families and friends for their unwavering support and motivation, which played a crucial role in the completion of this study.

References

[1] Ł. Górski and S. Ramakrishna, "Explainable Artificial Intelligence, Lawyer's Perspective," in *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, New York, NY, USA: ACM, Jun. 2021, pp. 60–68, DOI: 10.1145/3462757.3466145.

[2] H. A. Khoiri and E. Arghawaty, "Analyzing the IHSG Value and Influencing Factors in the Covid-19 Pandemic Era," *JAD J. Ris. Akunt. Keuang. Dewantara*, vol. 3, no. 2, pp. 110–121, 2020.

[3] A. Salah, M. Bekhit, E. Eldesouky, A. Ali, and A. Fathalla, "Price Prediction of Seasonal Items Using Time Series Analysis," *Comput. Syst. Sci. Eng.*, vol. 46, no. 1, pp. 445–460, 2023, DOI: 10.32604/csse.2023.035254.

[4] D. Maulani and D. Riani, "Autoregressive Integrated Moving Average (ARIMA) in Munufacturing Industry at Bursa Efek Indonesia (Case Study PT. Gudang Garam Tbk.)," *Pros. LPPM UIKA BOGOR; 2020*, 2020, [Online]. Available: http://pkm.uika-bogor.ac.id/index.php/prosiding/article/view/624

[5] T. Wang, T. Liu, and Y. Lu, "A Hybrid Multi-step Storm Surge Forecasting Model Using Multiple Feature Selection, Deep Learning Neural Network and Transfer Learning," *Soft Comput.*, vol. 27, no. 2, pp. 935–952, 2023, DOI: 10.1007/s00500-022-07508-8.

[6] Asriani, D. Herdhiansyah, W. Embe, and M. F. Aksara, "Forecasting Model of Corn Commodity Production in Indonesia: Production and Operation Management-Quantitative Method (POM-QM) Software," *J. Comput. Sci.*, vol. 20, no. 4, pp. 454–464, 2024, DOI: 10.3844/jcssp.2024.454.464.

[7] G. Yang, S. Du, Q. Duan, and J. Su, "Deep Reinforcement Learning-Based Trading Strategy for Load Aggregators on Price-Responsive Demand," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–14, 2022, DOI: 10.1155/2022/6884956.

[8] R. M. Nefdt, "Are machines radically contextualist?," *Mind Lang.*, vol. 38, no. 3, pp. 750–771, 2023, DOI: 10.1111/mila.12429.

[9] R. D. Agustin, "Want to Invest in Stocks? IHSG is Important Information to Know!," Jurnal.Id. Accessed: Oct. 10, 2023. [Online]. Available: https://www.jurnal.id/id/blog/mau-investasi-saham-ihsg-adalah-informasi-penting-untuk-diketahui/

[10] S. Arifin, A. K. Wijaya, R. Nariswari, I G. A. A. Yudistira, Faisal, and D. Wihardini, "Long Short-Term Memory (LSTM): Trends and Future Research Potential," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 13, no. 5, pp. 24–35, 2023, DOI: 10.46338/ijetae0523_04

[11] M. S. Nurjamil, "ARIMA Analysis Method for Forecasting the Number of Train Passengers on Java Island," Rpubs.com. Accessed: May 25, 2023. [Online]. Available: https://rpubs.com/MaritaNurjamil/prakkomstatG

[12] S. Arifin, M. M. Manurung, S. Jonathan, M. Effendi, and P. W. Prasetyo, "Trend Analysis of the ARIMA Method: A Survey of Scholarly Works," *Recent Eng. Sci. Technol.*, vol. 2, no. 03, pp. 1–14, 2024.

[13] Ruangmenyala, "LQ45 Stocks Are: Criteria and List of Stocks for the 2022 Period." Accessed: May 22, 2023. [Online]. Available: https://www.ruangmenyala.com/article/read/saham-lq45-adalah-kriteria--daftar-saham-periode-2022

[14] N. Adib and A. Ghofar, *Stock Investment: Basic Concepts and Practical Tips in Stock Business*. Universitas Brawijaya Press, 2021.

[15] A. N. Fadila and C. Nuswandari, "What are the factors that influence stock prices?," *E-Bisnis J. Ilm. Ekon. Bisnis*, vol. 15, no. 2, pp. 283–293, 2022.

[16] S. Arifin, I. B. Muktyas, P. W. Prasetyo, and A. A. Abdillah, "Unimodular Matrix and Bernoulli Map on Text Encryption Algorithm Using Python," *Al-Jabar J. Pendidik. Mat.*, vol. 12, no. 2, pp. 447–455, 2021.

[17] N. R. Haddaway, M. J. Page, C. C. Pritchard, and L. A. McGuinness, "PRISMA2020: An R Package and Shiny App for Producing PRISMA 2020- Compliant Flow Diagrams, with Interactivity for Optimised Digital Transparency and Open Synthesis," *Campbell Syst. Rev.*, vol.

18, no. 2, p. e1230, 2022.

[18] L. Jia *et al.*, "Development of Interactive Biological Web Applications with R/Shiny," *Brief. Bioinform.*, vol. 23, no. 1, p. bbab415, 2022.

[19] Z. Shi, Y. Hu, G. Mo, and J. Wu, "Attention-based CNN-LSTM and XGBoost Hybrid Model for Stock Prediction," *arXiv Prepr. arXiv2204.02623*, 2022.

[20] R. Zhu, Y. Yang, and J. Chen, "XGBoost and CNN-LSTM Hybrid Model with Attention-based Stock Prediction," in *2023 IEEE 3rd Int. Conf. on Electron. Technol. Comm. and Inform. (ICETCI)*, IEEE, 2023, pp. 359–365.

[21] C. Kasemset, K. Phuruan, and T. Opassuwan, "Shallot Price Forecasting Models: Comparison among Various Techniques," *Prod. Eng. Arch.*, vol. 29, no. 4, pp. 348–355, Dec. 2023, doi: 10.30657/pea.2023.29.40.

[22] D. E. Waluyo, H. W. Kinasih, C. Paramita, D. Pergiwati, R. Nohan, and F. A. Rafrastara, "Comparison and Implementation of Machine Learning Regression Algorithms for Composite Stock Price Index Prediction," *J. Inform. J. Pengemb. IT*, vol. 9, no. 1, pp. 12–17, 2024.

[23] A. A. Abdillah, Azwardi, S. Permana, I. Susanto, F. Zainuri, and S. Arifin, "Performance Evaluation of Linear Discriminant Analysis and Support Vector Machines to Classify Cesarean Section," *Eastern-European J. Enterp. Technol.*, vol. 5, no. 2–113, pp. 37–43, 2021, DOI: 10.15587/1729-4061.2021.242798.

[24] A. F. Rasyid, D. Agushinta, and D. T. Ediraras, "Deep Learning Methods in Predicting Indonesia Composite Stock Price Index (IHSG)," *Int. J. Comput. Inf. Technol.*, vol. 10, no. 5, pp. 209–217, 2021.

[25] S. Arifin, I. Bayu Muktyas, and K. Iswara Sukmawati, "Product of Two Groups Integers Modulo m,n and Their Factor Groups Using Python," *J. Phys.: Conf. Ser.*, vol. 1778, no. 1, 2021, DOI: 10.1088/1742-6596/1778/1/012026.

[26] G. Li and N. Yang, "A Hybrid SARIMA-LSTM Model for Air Temperature Forecasting," *Adv. Theory Simulations*, vol. 6, no. 2, 2023, DOI: 10.1002/adts.202200502.

[27] M. A. Ibrahim, S. Arifin, I. A. A. Yudistira, R. Nariswari, A. A. Abdillah, N. P. Murnaka, and P. W. Prasetyo, "An Explainable AI Model to Hate Specch Detection on Indonesian Twitter," *CommIT-Communication Inf. Technol. J.*, vol. 16, no. 2, 2022, DOI: 10.21512/commit.v16i2.8343

[28] S. Tarigan, N. P. Murnaka, and S. Arifin, "Development of Teaching Material in Mathematics 'Sapta Maino Education' on Topics of Plane Geometry," in *AIP Conference Proceedings*, American Institute of Physics Inc., 2021, DOI: 10.1063/5.0041650

[29] M. Gui and X. Xu, "Technology Forecasting Using Deep Learning Neural Network: Taking the Case of Robotics," *IEEE Access*, vol. 9, pp. 53306–53316, 2021, DOI: 10.1109/ACCESS.2021.3070105

[30] D. E. Waluyo, C. Paramita, H. W. Kinasih, F. A. Rafrastara, and D. Pergiwati, "Optimizing Investment in the Indonesian Stock Market: Improving Investment Decisions with IHSG Prediction using Decision Tree," *ABDIMASKU J. Pengabdi. Masy.*, vol. 7, no. 1, pp. 403–409, 2024.

[31] B. Rahman, S. Arifin, and I. Muktyas, "All Cyclic Subgroups in Group (Zm × zn, +) Using Python," *Int. J. Sci. Technol. Res.*, vol. 8, no. 9, pp. 2282–2285, 2019.

[32] D. Fan, H. Sun, J. Yao, K. Zhang, X. Yan, and Z. Sun, "Well Production Forecasting Based on ARIMA-LSTM Model Considering Manual Operations," *Energy*, vol. 220, p. 119708, 2021.

[33] P. K. Aritonang, S. K. Wiryono, and T. Faturohman, "A Comparative Analysis of Deep Learning and Traditional Statistics for Stock Price and Return Forecasting," *Glob. Bus. Financ. Rev.*, vol. 29, no. 8, p. 1, 2024.

[34] I. B. I. B. Muktyas, Sulistiawati, and S. Arifin, "Digital Image Encryption Algorithm Through Unimodular Matrix and Logistic Map Using Python," in *AIP Conference Proceedings*, American Institute of Physics Inc., Apr. 2021, DOI: 10.1063/5.0041653.

[35] I. R. Munthe, B. H. Rambe, F. Hanum, A. T. Amanda, and A. S. R. Hutagaol, "Implementation

of Stacking Technique Combining Machine Learning and Deep Learning Algorithms Using SMOTE to Improve Stock Market Prediction Accuracy," *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 2079–2091, 2024.

[36] D. N. Melati *et al.*, "A Comparative Evaluation of Landslide Susceptibility Mapping Using Machine Learning-based Methods in Bogor Area of Indonesia", *Environ. Earth Sci.*, vo. 83, no. 3, 2024, DOI: 10.1007/s12665-023-11402-3

[37] G. Abdoli, *Comparing the Prediction Accuracy of LSTM and ARIMA Models for Time-Series with Permanent Fluctuation*. in Periódico do Núcleo de Estudos e Pesquisas sobre Gênero e DireitovCentro de Ciências Jurídicas - Universidade Federal da Paraíba, vol. 9, no. 2, 2020. [Online]. Available: https://books.google.co.id/books?id=yhD2zgEACAAJ

[38] J. A. A. Moral-Muñoz *et al.*, "SoftwareTools for Conducting Bibliometric Analysis in Science: An Up-to-date Review," *Prof. la Inf.*, vol. 29, no. 1, p. 4, 2020.

[39] I. Assagaf, A. Sukandi, A. A. Abdillah, and S. Arifin, "Machine Predictive Maintenance by Using Support Vector Machines," *J. Recent Eng. Sci. Technol.*, vol. 1, no. 1, pp. 1–5, 2023.

[40] D. Stiawan *et al.*, "An Improved LSTM-PCA Ensemble Classifier for SQL Injection and XSS Attack Detection," *Comput. Syst. Sci. Eng.*, vol. 46, no. 2, pp. 1759–1774, 2023, DOI: 10.32604/csse.2023.034047.

[41] S. Arifin, I. B. Muktyas, W. F. A. Maki, and M. K. B. M. Aziz, "Graph Coloring Program of Exam Scheduling Modeling Based on Bitwise Coloring Algorithm Using Python," *J. Comput. Sci.*, vol. 18, no. 1, pp. 16-32, 2022, DOI: 10.3844/jcssp.2022.26.32.