Matrix: Jurnal Manajemen Teknologi dan Informatika
Volume 16 Issue 2 Year 2026 Pages 86-100
DOI: http://dx.doi.org/10.31940/matrix.v16i2.86-100
URL: http://ojs2.pnb.ac.id/index.php/matrix

Daily USD-IDR exchange rate prediction using Long Short-Term Memory (LSTM) with macroeconomic indicators and recursive multi-step prediction

Ni Luh Gede Ina Ari Richardi [1]*, Putu Indah Ciptayani [2], I Putu Bagus Arya Pradnyana [3]

[1,2,3] *Software Engineering Technology Study Program, Politeknik Negeri Bali, Indonesia*

*Corresponding Author: *luhgedeiar@gmail.com*

Abstract: The foreign exchange market's high volatility and non-linear dynamics pose significant challenges for accurate currency price forecasting. This study develops a predictive model for the daily USD-IDR exchange rate using Long Short-Term Memory (LSTM) method integrated with macroeconomic indicators. Historical price data, global oil prices, interest rates, and the US Dollar Index (DXY) were collected by fetching from Yahoo Finance and FRED. A comprehensive experimental design comprising 432 configurations was evaluated across varying data periods, train-validation-test splits, feature combinations, and hyperparameters. Results indicate that a 10-year historical dataset yields the most stable and accurate performance, achieving a test MAPE of 0.52% and R^2 of 0.904. The integration of macroeconomic variables improved predictive capability, though the impact of DXY was found to be conditional on data split ratios and hyperparameter settings. The optimal model was deployed as an interactive web application using Streamlit, enabling users to forecast USD-IDR rates up to seven days ahead via recursive multi-step forecasting. The proposed system provides a practical, accessible tool for retail traders and financial analysts to support informed decision-making in volatile forex markets

Keywords: Long Short-Term Memory, USD-IDR, deep learning, foreign exchange.

History Article: Submitted 25 May 2026 | Revised 18 July 2026 | Accepted 30 July 2026

How to Cite: N. L. G. I. A. Richardi, P. I. Ciptayani, and I. P. B. A. Pradnyana, "*Daily USD-IDR exchange rate prediction using Long Short-Term Memory (LSTM) with macroeconomic indicators and recursive multi-step prediction,*" *Matrix: Jurnal Manajemen Teknologi dan Informatika*, vol. 16, no. 2, pp. 86–100, 2026, doi: 10.31940/matrix.v16i2.86-100

Introduction

Foreign Exchange (forex) is the largest currency trading market in the world [1]. According to the Bank of International Settlements, the average daily trading volume in forex market reached approximately $9.6 trillion (Rp161.81 quadrillion) in April 2025 [2]. In general, forex refers to the mechanism for trading one country's currency against another country's currency, such as the Euro (EUR) against the United States Dollar (USD), the United States Dollar (USD) against the Indonesian Rupiah (IDR), and others [3], [4].The forex market is uniquely characterized by its global operation across four different time zones for 24 hours a day, 5 days a week [5][6]. Initially, this market mostly dominated by large institutions, however it has recently attracted increasing interest from individual investors due to its high liquidity, flexibility and potential for substantial profits within a short period of time [4], [5], [7].

Despite the significant profit opportunities, forex also involves substantial risks. Forex is characterized by high volatility, irregularity, non-linearity, and fluctuations, making it a challenging for novice traders [6]. Furthermore, exchanges rate fluctuation in forex are influenced by various macroeconomic factors such as interest rates, global oil prices, and others economic indicator, which require more complex analysis than basic technical analysis [8], [9].These complexities often make it difficult for novice traders to accurately interpret and predict market movements, thereby increasing the likelihood of erroneous trading decisions and financial losses [6].

To mitigate these risks, developing a currency price prediction system using the LSTM (Long Short-Term Memory) method, which integrates historical data with macroeconomic indicators, can be an effective solution. The LSTM is selected due its capability of process long time series data, retain long-term patterns, capturing non-linear relationships in the data, and integrate macroeconomic variables into the model learning process [8], [9].

Several previous studies have demonstrated the effectiveness of LSTM in predicting currency exchange rates. For instance, Wijaya et al., successfully predicted the EUR-USD and GBP-USD exchange rates using the LTSM method with high accuracy [4]. Similarly, Poernomo achieved accurate prediction of the Indonesian Rupiah against various Southeast Asia currencies using LSTM with a high level of accuracy (MAPE < 10%) [8]. In a relates study, Dita Ardriani et al., successfully predicted the USD-IDR exchange rate using daily historical data with LSTM, but acknowledged the model's limitations in handling volatility due to unexpected events [9]. However, these studies primarily focused on model experimentation and predictive evaluations without developing an integrated system that combines historical data with macroeconomic data in a single prediction model. To address this gap, this study proposes the development of a website-based exchange rate prediction system that integrate historical data with macroeconomic data, aiming to enhance prediction accuracy while ensuring user-friendly accessibility.

The development of this web-based exchange rate prediction system using LSTM, incorporating both historical and macroeconomic data, is expected to provide valuable insights into currency market movements, particularly for novice traders. Moreover, the web-based implementation ensures broad accessibility, facilitating widespread adoption among users. This research also aims to pave the way for future advancements in integrating predictive technologies with broader digital financial solutions.

Methodology

This study employs a quantitative experimental approach to determine the optimal LSTM configuration for daily USD-IDR exchange rate prediction. The research pipeline consists of six main stages: data collection, data preprocessing, experimental setup, model training, model evaluation, and web implementation, as illustrated in Figure 1.

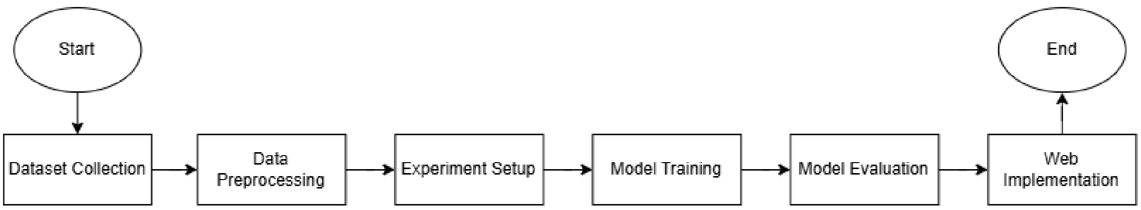

Figure 1. Research workflow

Data Collection

Daily frequency data were retrieved from Yahoo Finance and the Federal Reserve Economic Data (FRED). Four datasets were utilized: USD-IDR closing price (primary target), global oil prices, interest rates, and the US Dollar Index (DXY). Data were collected using Python libraries across three time periods: short period (2020–2025), medium period (2015–2025), and long period (2010–2025).

Data Preprocessing

This stage is conducted to ensure that the data is ready for use in the model development process. The collected raw data undergoes several procedures as follows:

Resampling

The resampling stage is conducted to ensure that all utilized data shares the same frequency [10]. In this study, the interest rate data undergoes resampling, where it is converted from a monthly frequency to a daily frequency to align with the other datasets. This process generates new data entries for each trading day within a month, and the newly created rows are filled using the forward fill (ffill) method. Consequently, the same monthly interest rate value is propagated to every trading day until the next monthly observation becomes available. For example, in the short-period dataset, the interest rate data originally consisted of 72 monthly observations. After the resampling process, it expanded to 2,192 daily observations while preserving the same monthly interest rate values across all trading days within each month. As shown in Figure 2 (a), the excerpts interest rate data before resampling is presented, while Figure 2 (b) shows the excerpts interest rate data after resampling.

	DATE	RATE
67	1 Aug 2025	6.07
68	1 Sep 2025	5.76
69	1 Oct 2025	5.53
70	1 Nov 2025	5.47
71	1 Dec 2025	5.46

(a)

	Date	RATE
2187	27 Dec 2025	5.46
2188	28 Dec 2025	5.46
2189	29 Dec 2025	5.46
2190	30 Dec 2025	5.46
2191	31 Dec 2025	5.46

(b)

Figure 2. (a) Interest rate data before resampling (b) Interest rate data after resampling

Check and Fill Missing Value

This stage is conducted to verify that all utilized data is complete and maintains chronological order. In this study, the forward fill and backward fill methods are employed for data imputation, as these techniques help preserve the continuity of the time-series data [11], [12].

Check and Drop Duplicate Entries

This stage is conducted to ensure that there no duplicate records exist to maintaining data quality. Duplicate records may arise when merging data from multiple sources [13]. In this study, duplication checks are conducted on the date column to guarantee data validity.

Data Splitting

This stage is conducted to divide the dataset into several parts, each of which will be used for a specific purpose. In this study, the data is divided into three parts: training set, validation set, and testing set. The training set is used to train the model, the validation set is employed for hyperparameter tuning and performance model evaluation, and the testing set is reserved for assessing the model's final performance. Dataset partitioning is crucial to ensure the model operates optimally on unseen data and avoids overfitting to the training data [14].

Feature Scaling

At this stage, the data is transformed from its original range to a standardized scale between 0 and 1. The Min-Max Normalization method is applied for feature scaling in this study. This technique ensures that all attribute values are constrained within a consistent range, enabling fair comparison and unbiased analysis [15]. The scaling process was performed only on the training data to prevent data leakage. After the minimum and maximum parameters were obtained from the training data, the transformation process was applied to the training, validation, and testing datasets. As shown in Figure 3 (a), the data before the scaling process, while Figure 3 (b) shows the data after the scaling process.

Date	USD_IDR	OIL_PRICE	RATE
2015-01-01	12390.000000	56.419998	7.122027
2015-01-02	12390.000000	56.419998	7.122027
2015-01-05	12480.000000	53.110001	7.122027
2015-01-06	12625.000000	51.099998	7.122027
2015-01-07	12635.000000	51.150002	7.122027

(a)

✅ Scaling selesai (MinMax 0-1)

Date	USD_IDR	OIL_PRICE	RATE
2015-01-01	0.000000	0.341371	0.680309
2015-01-02	0.000000	0.341371	0.680309
2015-01-05	0.021872	0.310907	0.680309
2015-01-06	0.057111	0.292407	0.680309
2015-01-07	0.059541	0.292867	0.680309

(b)

Figure 3. (a) Data before scaling (b) Data after scaling

Create Sequence

Sequence creation in this study is performed using the sliding window method, where time-series data is extracted according to the time step parameter value to predict the price for the subsequent day [16]. This approach enables the model to effectively learn temporal patterns and trends within the data.

Experiment Setup

To determine the best model configuration, experiments involving several parameters and hyperparameters were required. The search for the optimal model configuration was conducted using the Grid Search approach. This method systematically evaluates all possible combinations of predefined hyperparameter values. In this research, a total of 432 experimental combinations were conducted by combining various configurations of data split ratios, historical periods, macroeconomic feature sets, and hyperparameters such ad learning rate, units, time step dan batch size, as detailed in Table 1.

Table 1. Experiment configuration

Parameter	Value	Total	Description
Data Splitting			
Train:Val:Test	Scenario 1: 70%:15%:15% Scenario 2: 80%:10%:10%	2	Data splitting scenarios
Dataset Configuration			
Time Period	5 years, 10 years, 15 years	3	Historical data length
Feature Selection			
Macroeconomic Variable	Scenario 1: Oil Price & Interest Rate Scenario 2: Oil Price, Rate & DXY	2	External feature combinations
Hyperparameters			
Learning Rate	0.001, 0.005, 0.01	3	Adam optimizer learning rate
Units	50, 100	2	Hidden state in LSTM layer
Time Step	30, 60, 90	3	Sliding window size (days)
Batch Size	32, 64	2	Training Batch Size

Model Training

The model was developed using the LSTM which is designed to overcome the problem of vanishing gradient and exploding gradient in RNN by using memory cells regulated by three gates: input gate, forget gate and output gate which control the entry and exit of information from the memory cell [17]. The architecture of the model built with consisting of an LSTM layer as the input layer, followed by a dropout layer to prevent overfitting by randomly deactivating a portion of the neurons, and two consecutive dense layers: a fully connected layer with 32 neurons and a layer without an activation function containing 1 neuron that acts as the output layer.

Each experiment was validated using 3-fold time series cross-validation with Time Series Split method. This method preserves the chronological order of the data by applying an expanding window strategy, where the training data gradually increases in each fold, while validation is conducted on the subsequent unseen time period. This approach was chosen to prevent information leakage from future data and to produce a more realistic evaluation of model performance on time-series data. After identifying the optimal configuration, the model was retrained on separate training and validation sets and evaluated using some evaluation matrix on the unseen test set.

Model Evaluation

After training the model with the optimal configuration, evaluation was performed using data from the test set, which the model had not previously encountered. In this study, two evaluation metrics were employed to assess model performance: Mean Absolute Percentage Error (MAPE) and R-Squared (R^2). MAPE was used to quantify the average prediction error of the model, as formulated in Equation (1) [18]. Meanwhile, R-Squared was utilized to measure the proportion of data variance explained by the model and to determine whether the model successfully captures the underlying patterns and trends in exchange rate movements, as expressed in Equation (2) [6], [9]. The best-performing model was selected based on the lowest MAPE and the highest R^2 values.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| x\ 100\% \tag{1}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_t)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

Web Implementation

After obtaining the best model configuration, the final stage was to implement the model into an interactive website that can be widely used by users. In this study, the user interface was developed using the Streamlit framework, which accepts date inputs from users and generates exchange rate predictions along with graphical visualizations of price movements.

The Recursive Multi-step Forecasting method was applied in the system and operates iteratively, where the prediction result from the first day is reused as an input feature to predict the following day [19]. This approach enables users to generate forecasts for several days ahead. Considering the error accumulation characteristics of recursive forecasting, the prediction scope in this system is limited to a maximum horizon of 7 days ahead in order to maintain the reliability and stability of prediction accuracy.

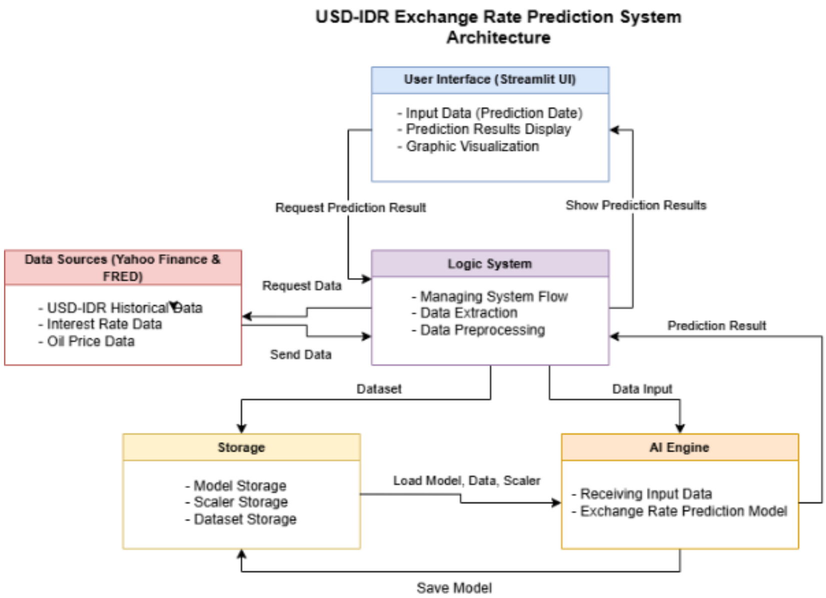

Figure 4. System architecture

Figure 4 illustrates the architecture of the system. The system consists of five main components. The first component is the user interface, which functions as the medium through which users interact with the system. The second component is the system logic, which manages the entire system workflow, starting from data retrieval from external sources, data extraction, data preprocessing, data integration, data storage, sending user inputs to be processed by the AI model, and forwarding the prediction results to be displayed to users.

The third component is the data source module, which provides the data required by the system. The fourth component is data storage, which serves as a repository for important system components such as prediction models, datasets, and scalers. Finally, the AI engine acts as the core module of the prediction system, where the prediction process is carried out. The model used contains parameters learned during the training process and is utilized to predict exchange rate values based on the prepared dataset.

Results and Discussions

Results

Model Experiment

After conducting 432 experiments validated using 3-fold time series cross-validation using Time Series Split method by combining variations in data periods, data split ratios, macroeconomic feature configurations, and hyperparameter combinations, the results showed that using 10 years of data provided the best and most stable performance across both data-splitting scenarios. As summarized in Table 2, the experiments with IDs EXP-3 and EXP-4 under the 70:15:15 data split scenario produced the best testing performance, with test MAPE values of 0.52% and 0.48%, and test R^2 values reaching 0.904 and 0.906.

Table 2. Experiment result

Best Exp ID	Data Period	Macro	Best Hyperparameter				Validation		Test	
			Units	Learning Rate	Time Step	Batch Size	MAPE	R2	MAPE	R2
Data Split Scenario 1: 70:15:15										
EXP-1	5 Year	2 (Oil & Rate)	100	0.005	30	64	0.43 ± 0.026	0.872 ± 0.030	0.40	0.735
EXP-2		3 (Oil, Rate & DXY)	50	0.005	30	64	0.54 ± 0.08	0.840 ± 0.01	1.80	-1.12
EXP-3	10 Year	2	100	0.001	30	64	0.69 ± 0.23	0.907 ± 0.07	0.52	0.904
EXP-4		3	50	0.005	30	32	0.64 ± 0.11	0.923 ± 0.03	0.48	0.906
EXP-5	15 Year	2	100	0.001	90	64	6.28 ± 5.38	-1.13 ± 1.55	5.68	-4.546
EXP-6		3	100	0.001	60	64	7.32 ± 5.36	-1.55 ± 1.30	2.14	0.175
Data Split Scenario 2: 80:10:10										
EXP-7	5 Year	2	100	0.01	30	64	0.51 ± 0.09	0.888 ± 0.022	0.40	0.715
EXP-8		3	50	0.01	30	64	1.03 ± 0.57	0.571 ± 0.381	0.84	0.258
EXP-9	10 Year	2	50	0.005	60	32	0.56 ± 0.13	0.926 ± 0.034	0.54	0.588

EXP-10		3	50	0.001	60	64	0.58 ± 0.07	0.908 ± 0.049	0.71	0.538
EXP-11	15 Year	2	50	0.001	30	64	7.20 ± 5.81	-1.389 ± 2.010	4.67	-2.63
EXP-12		3	50	0.001	90	64	5.25 ± 3.38	-0.551 ± 1.191	3.94	-1.737

Based on Table 2, the comparison of data period lengths shows a significant impact, particularly on the R^2 values. The dataset with a 10-year period (2015–2025) produced the best and most stable accuracy, whereas the dataset with a 5-year period (2020–2025) achieved good test MAPE and R^2 values under several configurations but was still unable to effectively capture the underlying patterns and trends in the data. Meanwhile, the dataset with a 15-year period (2010–2025) produced the worst performance, with the test MAPE increasing to as high as 5.68% and negative R^2 values.

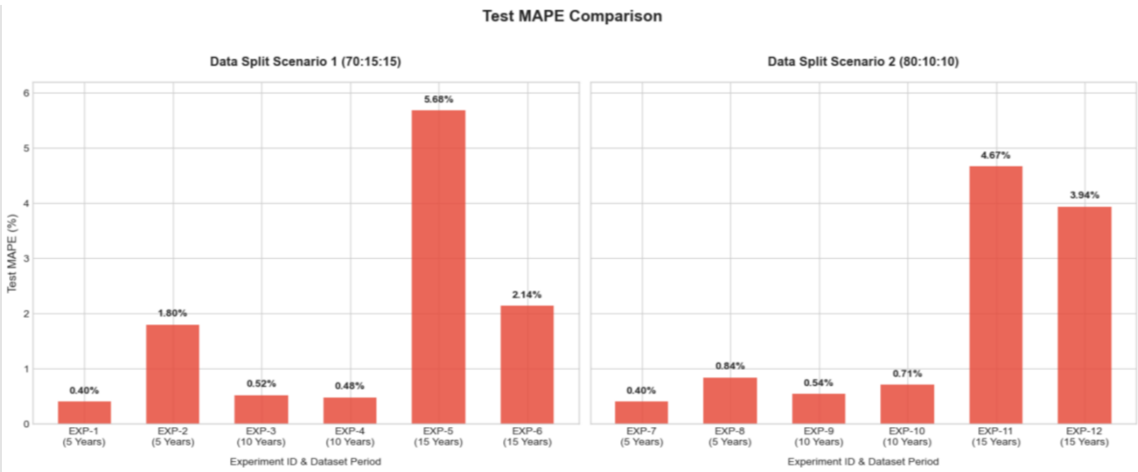

Figure 5. MAPE comparison chart

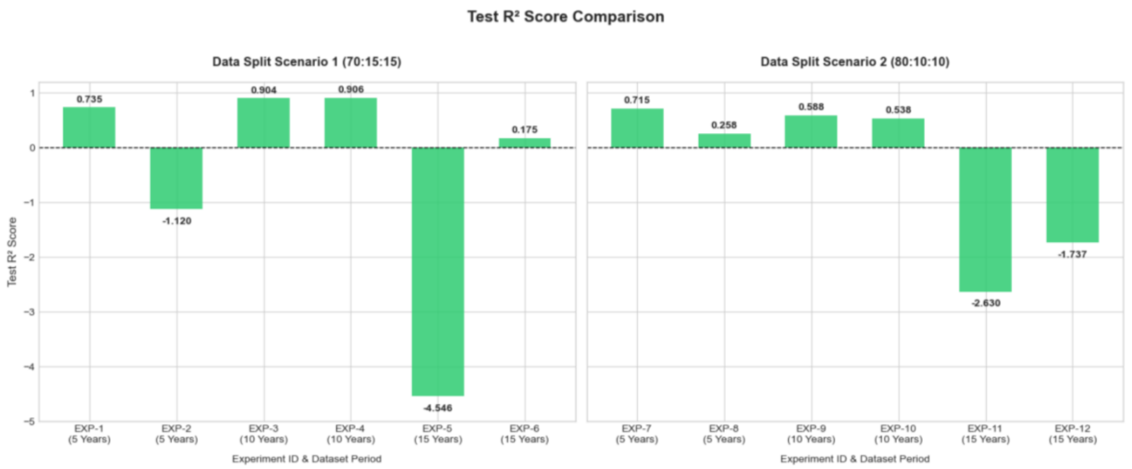

Figure 6. R^2 comparison chart

The addition of the Dollar Index (DXY) as a third macroeconomic feature did not consistently improve test accuracy. In the first data-splitting scenario, the integration of the DXY feature in experiment ID EXP-4 successfully reduced the test MAPE from 0.52% to 0.48% and increased the R^2 value. However, in the second scenario, the configuration with only two macroeconomic features resulted in a lower test MAPE of 0.54% compared to 0.71% for the configuration with three macroeconomic features. This indicates that the contribution of the DXY feature is conditional and depends on the combination of hyperparameters and the data split ratio.

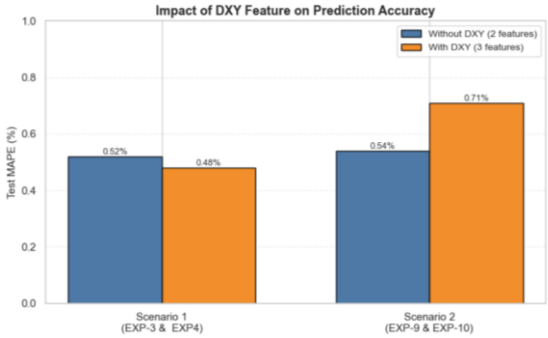

Figure 7. DXY impact chart

There were also significant differences between the validation and test accuracy results in several configurations. Although some experiments achieved very low validation MAPE values (<1%) with small standard deviations, the performance on the test data declined considerably, particularly in the 15-year dataset and several configurations of the 5-year dataset. The negative test R^2 values in EXP-2, EXP-5, EXP-11, and EXP-12 indicate possible overfitting to the validation data or the model's inability to generalize patterns to entirely new time periods.

Based on the tuning results, learning rates of 0.001 and 0.005 showed the most stable performance with the lowest errors. A time step of 30 days was sufficiently effective in capturing daily patterns, while longer time steps of 60–90 days tended to increase validation error variance. LSTM unit sizes of 50 and 100 produced comparable performance, whereas batch sizes of 32–64 did not show statistically significant differences. The parameter combination that most consistently produced high accuracy consisted of a 10-year data period, learning rates of 0.001–0.005, and a 30-day time step.

To evaluate the specific impact of macroeconomic indicators on prediction accuracy, a baseline experiment was conducted excluding macroeconomic data. This comparison was applied to the best-performing models, EXP-3 and EXP-4, utilizing identical datasets, data split configurations, and architectural workflows. As summarized in Table 3, the integration of macroeconomic variables substantially enhances the forecasting accuracy.

Table 3 Comparison of model configuration with and without macroeconomic features.

Feature Configuration	Test MAPE (%)	Test R^2
Historical USD/IDR Only (Baseline)	1.50	0.531
Historical USD/IDR + Interest Rate + Oil Price (EXP-3)	0.52	0.904
Historical USD/IDR + Interest Rate + Oil Price + DXY (EXP-4)	0.48	0.906

Furthermore, to evaluate the model practical viability for extended forecasting horizons, the trained models were subjected to a Recursive Multi-step Forecasting evaluation. Since the LSTM models were inherently trained on a single-step-ahead prediction framework using the sliding window approach, generating a multi-step forecast required the system to iteratively feed its own prior predictions back into the feature matrix as input for subsequent days [19].

Testing was conducted across various forecasting horizons up to a maximum of seven days ahead. The limitation of selecting prediction dates to a maximum of seven days ahead is due to the recursive multi-step forecasting mechanism, where the model predicts the exchange rate for the first day using actual historical data. However, to predict the second day, the model must use its own prediction result from the first day as an input feature, treating it as actual data. As a result, any prediction error generated in the first step will propagate through subsequent iterations and gradually increase over time [20]. Therefore, a seven-day limit was established as the maximum forecasting horizon in order to maintain the accuracy and stability of the LSTM model.

Recursive evaluation was performed on the two best models (EXP-3 and EXP-4), where the model with the best performance was selected and implemented in the interactive web application, ensuring reliable short-term forecasting within the limited 7-day forecasting window. Based on the recursive evaluation of the models presented in Table 5 and Table 5, the model with experiment ID EXP-3 was selected to be implemented in the website.

Table 4. EXP-3 model recursive evaluation result

Day	MAPE	R2
+1	0.5020	0.9023
+2	0.5910	0.8768
+3	0.6837	0.8462
+4	0.7859	0.8070
+5	0.8763	0.7628
+6	0.9669	0.7157
+7	1.0538	0.6661
AVG	**0.7799**	**0.7967**

Table 5. EXP-4 model recursive evaluation result

Day	MAPE	R2
+1	0.6933	0.7733
+2	0.9587	0.5601
+3	1.2144	0.2824
+4	1.4705	-0.0454
+5	1.6905	-0.4059
+6	1.9029	-0.7829
+7	2.0926	-11.728
AVG	**1.4318**	**0.1130**

Web Implementation

The best model identified was implemented into a website developed using the Streamlit framework. The developed system implements the Recursive Multi-Step Forecasting method, which is used to predict the USD-IDR exchange rate for several days ahead in a recursive manner. In this study, the forecasting horizon was limited to a maximum of 7 days ahead. The process is carried out by using the prediction result from the first day as a new input feature to predict the following day. The system consists of two main pages: the homepage and the prediction results page.

The homepage is the first page viewed by users when accessing the website. On this page, the system provides a date input feature that can be filled in by users, along with an action button labeled *"Mulai Prediksi"* to run the prediction model. In addition, the page also displays a historical data table containing USD-IDR exchange rate information for the past 7 days, as well as macroeconomic indicators including global oil prices and interest rates, as shown in Figure 8.

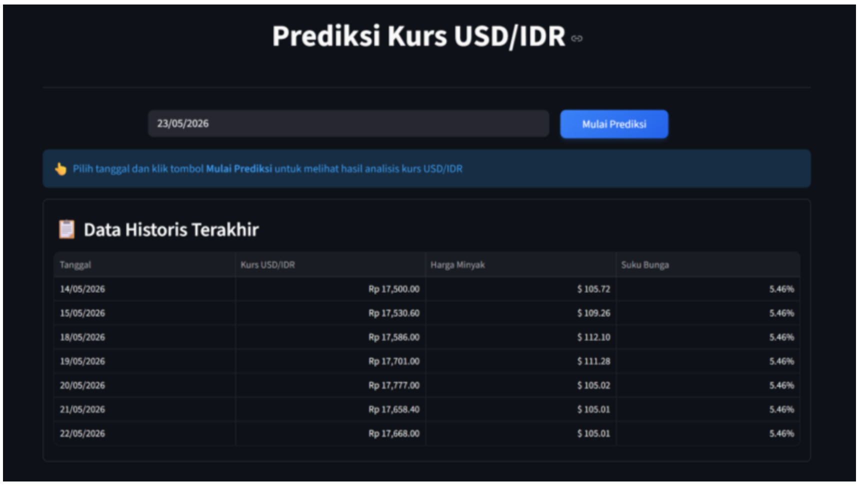

Figure 8. Homepage of the USD-IDR exchange rate prediction system

The second page is the prediction results page. After users select a date and click the "Start Prediction" button, the system processes the data and displays the results in the form of a summary block consisting of four sections, which provide information regarding the predicted exchange rate, model accuracy, the latest global oil price, and the latest interest rate as shown in Figure 9. If users select a prediction horizon of more than one day, the system will automatically display the prediction results in tabular form as shown in Figure 10. In addition, the page also includes an exchange rate chart that displays the exchange rate values from the past 7 days along with the prediction results to visualize the direction of currency movements.

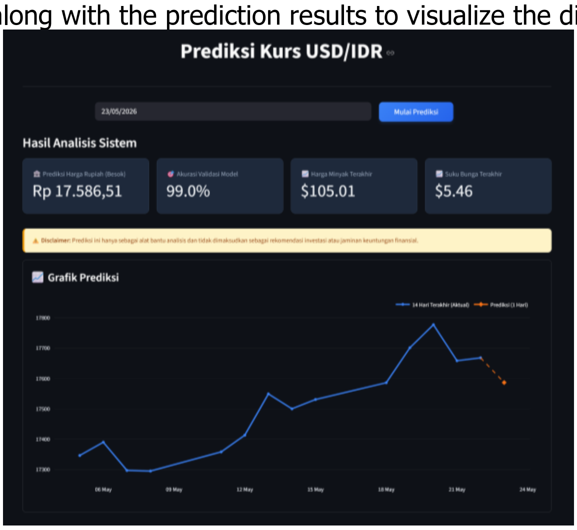

Figure 9. Single prediction result page

Figure 10. Multiple prediction result page

Discussions

The comparison of data period lengths shows a significant impact, particularly on the R^2 value. The dataset with a 5-year period (2020–2025) produced good test accuracy in terms of MAPE and R^2 under several configurations, but it was not sufficient to capture the underlying patterns and trends in the data. This indicates that the model experienced underfitting due to the relatively short data period. This finding is consistent with previous studies conducted by Poernomo [8] and Wijaya et al. [4], which showed that deep learning models require sufficiently long historical data to effectively capture non-linear market dynamics.

The dataset with a 15-year period (2010–2025) produced the worst accuracy, with the test MAPE increasing to 5.68% and a negative R^2 value. This demonstrates that a longer data period does not always improve prediction accuracy [21]. Data from 15 years ago may no longer be relevant to current exchange rate volatility characteristics, where including overly outdated data contributes minimally or may even negatively affect model performance [22].

The dataset with a 10-year period (2015–2025) produced the best and most stable accuracy because this period provides an optimal balance between volatile data variations caused by unexpected events such as the COVID-19 pandemic and economic policy changes, as well as more stable data and sufficient data quantity. This is important because the model must be able to capture patterns and trends across various market conditions, enabling it to achieve a balanced learning process and avoid overfitting to a specific condition [23], [24].

Compared to the baseline experiment that relied solely on historical data, the integration of macroeconomic factors substantially improved the model's performance, reducing the test MAPE from 1.50% to 0.52% and increasing the R2 score from 0.531 to 0.904. Furthermore, this approach noticeably outperforms existing models in the literature, such as the study conducted by Pahlevi et al. [1], which reported a MAPE value of 2.5%, and the forecasting model by Wahyudi et al. [25], which achieved a MAPE value of 0.81%. This internal head-to-head comparison empirically demonstrates that embedding macroeconomic indicators provides critical external shock signals that historical data alone cannot fully resolve.

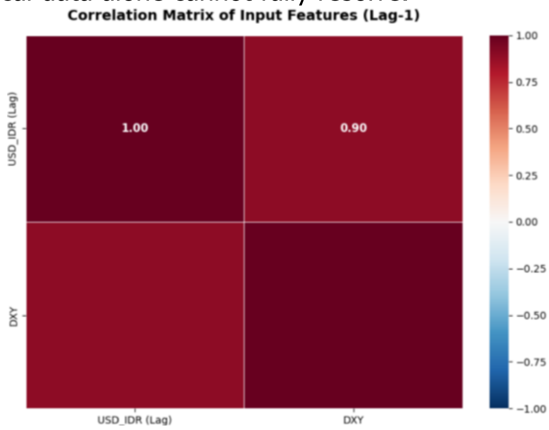

Figure 11 Heatmap correlation historical data and DXY macro indicator

Based on Figure 11, the addition of the US Dollar Index (DXY) did not consistently improve accuracy, this is because the DXY feature has a very high correlation value of 0.9 with the USD-IDR as lag feature. This indicating severe multicollinearity, which leads to information redundancy, meaning that DXY fails to provide significant novel information to the dataset [26], [27]. Therefore, feature selection remains an important factor in exchange rate forecasting models. Consequently, this redundancy underscores that rigorous feature selection remains a critical factor in exchange rate forecasting models.

This phenomenon directly explains the significant decline in accuracy observed in the model with ID EXP-4. During the recursive forecasting phase, the model encountered a critical operational mismatch, the USD-IDR values were updated using the system’s predicted values, while the DXY values were filled using the forward fill method. This created a mismatch, where the USD-IDR exchange rate moved dynamically while the DXY remained static, thereby triggering substantial increases in prediction error.

Furthermore, the recursive multi-step forecasting evaluation further showed that prediction accuracy gradually decreased as the forecasting horizon increased. This occurred because prediction errors from previous steps were reused as inputs for subsequent predictions, causing error accumulation over time. Nevertheless, the EXP-3 model maintained relatively stable performance within the 7-day forecasting horizon as shown in Table 4, making it suitable for short-term exchange rate forecasting applications. The trend of increasing MAPE values and decreasing R^2 scores across the forecasting horizon can also be visually observed in Figure 12 and Figure 13.

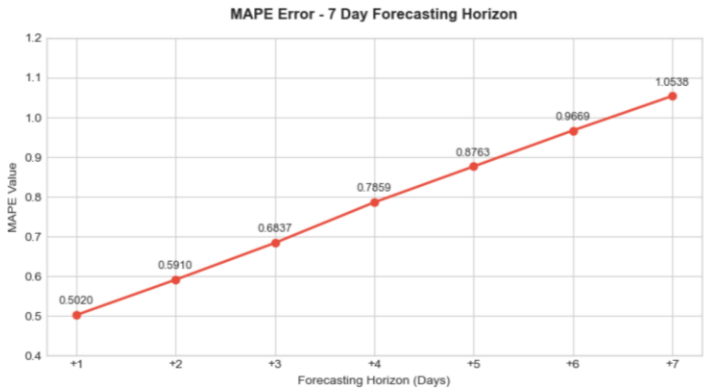

Figure 12. EXP-3 MAPE error chart

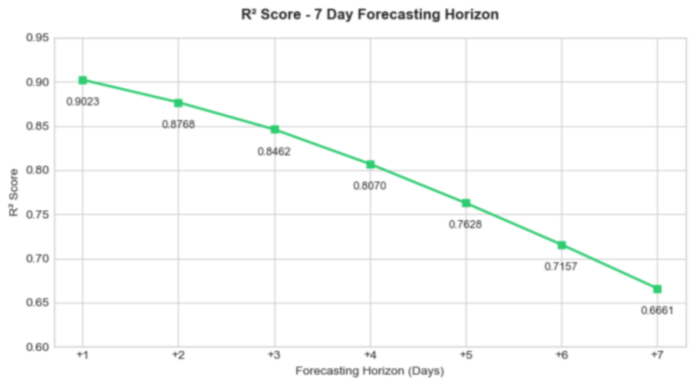

Figure 13. EXP-3 R^2 error chart

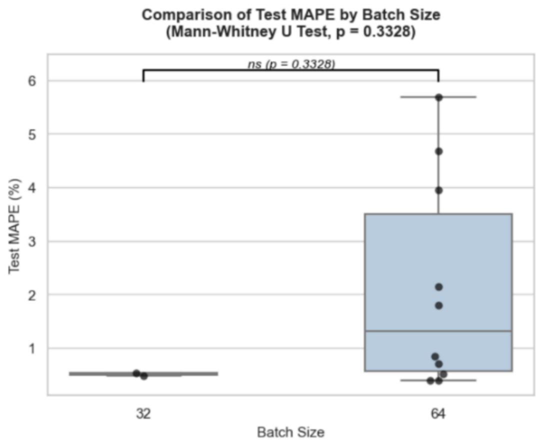

Figure 14 MAPE comparison by batch size

Based on the Mann-Whitney test shown in Figure 14 the comparison of test MAPE values between the batch size 32 and batch size 64 groups yielded a p-value of 0.3328. In standard statistical methodology, a p-value greater than the established significance level $\alpha = 0.05$. This outcome empirically demonstrates that the performance difference between the two batch sizes is not statistically significant. Research results also demonstrate that the best-performing models used different batch sizes, specifically the EXP-3 model with batch size 64: MAPE = 0.52%, R^2 = 0.904 and the EXP-4 model with batch size 32: MAPE = 0.48%, R^2 = 0.906), showing a difference of only 0.04% in MAPE and 0.002 in R^2.

The implementation of the model into an interactive web application using Streamlit demonstrates the practical applicability of the proposed system. The developed application enables users to perform short-term USD-IDR exchange rate forecasting through a user-friendly interface, making the system more accessible for traders and financial analysts. The developed application enables users to perform short-term USD-IDR exchange rate forecasting through a user-friendly interface, making the system more accessible for traders and financial analysts. In a real-world operational context, this deployment eliminates the need for users to possess advanced programming skills or perform manual data preprocessing. By utilizing a modular architecture, the system automatically pulls data, cleans it, and executes a recursive multi-step forecasting method to provide a restricted 7-day prediction horizon.

This specific 1-week window offers high specificity for short-term risk management, allowing financial officers to make data-driven decisions during high market volatility. However, for practical deployment, users must consider that the model relies heavily on historical patterns and macroeconomic indicators. Therefore, in real-world usage, it is highly recommended to utilize this application as a decision-support tool alongside human expert oversight to mitigate risks from sudden, unpredictable global economic shocks (black swan events).

Conclusion

This study successfully developed and evaluated an LSTM-based forecasting system for daily USD-IDR exchange rates by integrating historical price data with macroeconomic indicators. Through a comprehensive experimental design of 432 configurations, the optimal and stable model was obtained using a 10-year dataset, two macroeconomic features, and hyperparameters consisting of 100 units, a learning rate of 0.001, a 30-day time step, and a batch size of 64, which were determined through a hyperparameter tuning process. While the baseline model relying on historical data yielded a lower accuracy with a test MAPE of 1.50% and an R2 of 0.531, the proposed macro-integrated configuration significantly outperformed it, achieving an optimal test MAPE of 0.52% and an R2 of 0.904. The findings confirm that integrating external economic variables significantly improves short-term currency prediction accuracy compared to models using only historical data. The implementation of the model into an interactive web application using Streamlit provides an accessible data-driven forecasting tool for users.

Although the system demonstrated strong performance under normal and moderately volatile market conditions, several limitations remain, including the potential accumulation of errors in recursive multi-step forecasting, vulnerability to structural changes within extended historical windows, and the absence of real-time sentiment data or policy announcements. Future research should investigate dynamic retraining mechanisms and hybrid deep learning architectures to further improve the model's adaptability.

Acknowledgement

The author gratefully acknowledges the support and contributions from all parties involved in the completion of this article. Particular thanks are expressed to Putu Indah Ciptayani, S.Kom.,M.Cs. and I Putu Bagus Arya Pradnyana, S.Kom.,M.Kom, for their guidance and input throughout the research process.

References

[1] M. R. Pahlevi, "Prediksi Harga Forex Menggunakan Algoritma Long Short-Term Memory," *JNANALOKA*, pp. 69–76, Sep. 2023, doi: 10.36802/jnanaloka.2022.v3-no2-69-76.

[2] Bank of International Settlements, "OTC foreign exchange turnover in April 2025," Bank of International Settlements. Accessed: Jan. 04, 2026. [Online]. Available: https://www.bis.org/statistics/rpfx25_fx.htm

[3] M. Edi, E. Utami, and A. Yaqin, "Prediksi Harga pada Trading Forex Pair USDCHF Menggunakan Regresi Linear," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 13, no. 2, pp. 109–119, Sep. 2023, doi: 10.34010/jamika.v13i2.9826.

[4] A. J. Wijaya, W. Swastika, and O. H. Kelana, "PREDIKSI HARGA FOREIGN EXCHANGE MATA UANG EUR/USD DAN GPB/USD MENGGUNAKAN LONG SHORT-TERM MEMORY," *SAINSBERTEK Jurnal Ilmiah Sains & Teknologi*, vol. 2, Sep. 2021.

[5] Sudarmadji, "Trading Forex Online Individu: Masalah Syariah," *Labs: Jurnal Bisnis dan Manajemen*, vol. 28, 2023.

[6] M. R. Pahlevi, K. Kusrini, and T. Hidayat, "Comparison of LSTM and GRU Models for Forex Prediction," *Sinkron : Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 4, pp. 2254–2263, Oct. 2023, doi: 10.33395/sinkron.v8i4.12709.

[7] M. Tjitrayudha and Kosasih, "UJI EFEKTIVITAS METODE BREAKOUT SUPPORT RESISTANCE UNTUK PROBABILITAS PADA FOREIGN EXCHANGE MARKET," *MANAJEMEN : JURNAL EKONOMI USI*, vol. 6, no. 2, 2024.

[8] A. Poernomo, "RUPIAH EXCHANGE RATE PREDICTION WITH LONG SHORT-TERM MEMORY ALGORITHM," *Syntax Literate: Jurnal Ilmiah Indonesia*, vol. 10, no. 1, Jan. 2025.

[9] N. N. Dita Ardriani, P. Sugiartawan, and G. A. Santiago, "Deep Learning Approach for USD to IDR Forecasting with LSTM," *JSIKTI: Jurnal Sistem Informasi dan Komputer Terapan Indonesia*, vol. 8, no. 1, pp. 91–100, Jun. 2025.

[10] D. Safira, I. A. Sobari, and S. Rahayu, "PREDIKSI HARGA BBCA MENGGUNAKAN LSTM MULTIVARIAT DENGAN FITUR BI-RATE SEBAGAI INDIKATOR EKSTERNAL," *Information System Journal*, vol. 8, no. 02, pp. 119–128, Dec. 2025, doi: 10.24076/infosjournal.2025v8i02.2372.

[11] F. Kamalov and H. Sulieman, "Time series signal recovery methods: comparative study," Oct. 2021, doi: https://doi.org/10.48550/arXiv.2110.12631.

[12] J. Y. Lee *et al.*, "Comparison of Models for Missing Data Imputation in PM-2.5 Measurement Data," *Atmosphere (Basel).*, vol. 16, no. 4, Apr. 2025, doi: 10.3390/atmos16040438.

[13] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[14] H. Allam, L. Makubvure, B. Gyamfi, K. N. Graham, and K. Akinwolere, "Text Classification: How Machine Learning Is Revolutionizing Text Categorization," *Information (Switzerland)*, vol. 16, no. 2, Feb. 2025, doi: 10.3390/info16020130.

[15] A. Pranolo *et al.*, "Enhanced Multivariate Time Series Analysis Using LSTM: A Comparative Study of Min-Max and Z-Score Normalization Techniques," *ILKOM Jurnal Ilmiah*, vol. 16, no. 2, pp. 210–220, Aug. 2024, doi: 10.33096/ilkom.v16i2.2333.210-220.

[16] U. Brawijaya, I. R. Nashrullah, E. R. Widasari, and B. H. Prasetio, "Implementasi Metode Sliding Window Untuk Peningkatan Perekaman Real-Time Pada Sistem Deteksi Kesadaran Berbasis Sensor Electroencephalography Satu Kanal," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 3, pp. 2548–964, 2026, [Online]. Available: http://j-ptiik.ub.ac.id

[17] I. D. Mienye and T. G. Swart, "A Comprehensive Review of Deep Learning: Architectures, Recent Advances, and Applications," *Information (Switzerland)*, vol. 15, no. 12, Dec. 2024, doi: 10.3390/info15120755.

[18] M. K. Najib and S. Nurdiati, "Pemodelan Deret Waktu Menggunakan Non-linear Autoregressive Neural Network: Studi Kasus Prediksi Harga Saham Mandiri," *Jambura Journal of Mathematics*, vol. 7, no. 2, pp. 213–220, Aug. 2025, doi: 10.37905/jjom.v7i2.33397.

[19] H. T. A. Simanjuntak, A. Lumbanraja, G. Samosir, and Regita, "Prediksi Single-Step dan Multi-Step Data Cuaca Menggunakan Model Long Short-Term Memory dan Sarima," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 399–410, Apr. 2025, doi: 10.25126/jtiik.2025129444.

[20] N. Fitriyati, M. Liebenlito, and S. Hasna Tsabitah, "Pengaruh Hiperparameter dan Variabel Eksogen pada Prediksi Multi-Langkah Kecepatan Angin menggunakan LightGBM," *JURNAL FOURIER*, vol. 15, no. 1, pp. 1–16, Apr. 2026, doi: 10.14421/fourier.2026.151.1-16.

[21] A. Lazcano, P. Hidalgo, and J. E. Sandubete, "Walking Back the Data Quantity Assumption to Improve Time Series Prediction in Deep Learning," *Applied Sciences (Switzerland)*, vol. 14, no. 23, Dec. 2024, doi: 10.3390/app142311081.

[22] Q. Yu and B. Tolson, "How much historical data do we need? The role of data recency and training period length in LSTM-based rainfall-runoff modeling," *J. Hydrol. (Amst).*, vol. 668, Apr. 2026, doi: 10.1016/j.jhydrol.2026.135046.

[23] K. Jongseon, K. Hyungjoon, K. Hyun Gi, L. Dongjun, and Y. Sungroh, "A comprehensive survey of deep learning for time series forecasting: architectural diversity and open

challenges," *Artif. Intell. Rev.*, vol. 58, no. 7, Jul. 2025, doi: 10.1007/s10462-025-11223-9.

[24] G. Taneva-Angelova and D. Granchev, "Deep Learning and Transformer Architectures for Volatility Forecasting: Evidence from U.S. Equity Indices," *Journal of Risk and Financial Management*, vol. 18, no. 12, Dec. 2025, doi: 10.3390/jrfm18120685.

[25] A. Wahyudi, W. Setiawan, and Y. D. P. Negara, "PREDIKSI KURS MATA UANG RIYAL KE RUPIAH MENGGUNAKAN METODE SUPPORT VECTOR REGRESSION (SVR)," *Jurnal METHODIKA*, Sep. 2024.

[26] J. Y. Le Chan *et al.*, "Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review," Apr. 01, 2022, *MDPI*. doi: 10.3390/math10081283.

[27] S. F. Akintade, K. Roy, and S. T. Kim, "Hybrid Deep Machine Learning Feature Selection for High-Dimensional Cybersecurity Data," *IEEE Access*, vol. 13, pp. 172136–172156, 2025, doi: 10.1109/ACCESS.2025.3615582.

© 2026 by the author; licensee Matrix: Jurnal Manajemen Teknologi dan Informatika. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).