Available online at https://journal.rescollacomm.com/index.php/ijbesd/index

International Journal of Business, Economics and Social Development

e-ISSN 2722-1156
p-ISSN 27722-1164

Vol. 6, No. 1, pp. 132-138, 2025

Comparative Analysis of LSTM and GRU Models for Ethereum (ETH) Price Prediction

Moch Panji Agung Saputra[1*], Riza Andrian Ibrahim[2], Renda Sandi Saputra[3]

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang, Indonesia*
[2]*Doctoral Mathematics Study Programme, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang, Indonesia*
[3]*Department of Informatics, Faculty of Technology and Information, University of Informatics and Business, Bandung, Indonesia*

**Corresponding author email: moch16006@unpad.ac.id*

Abstract

The increasing use of cryptocurrencies has changed the dynamics of investment, presenting both opportunities and challenges for investors. Although various studies have compared the performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) in predicting financial asset prices, there are still differences in results regarding which model is superior. Therefore, this study aims to compare the performance of LSTM and GRU in predicting Ethereum prices using a hyperparameter tuning approach. The data used is historical data of Ethereum (ETH) shares from 2020 to 2025. The research methodology includes data preprocessing using Min-Max scaling, model development with various layer configurations, and comprehensive evaluation using several performance metrics. The results show that the GRU Model provides superior performance with a lower Root Mean Squared Error (RMSE) of 0.0234 and Mean Absolute Error (MAE) of 0.0168, compared to LSTM's RMSE of 0.0265 and MAE of 0.0193. While LSTM exhibits a slightly better Mean Absolute Percentage Error (MAPE) of 18.08% compared to GRU at 18.17%, the GRU model achieves a higher R^2 Score of 0.9442 compared to LSTM at 0.9282. Visual analysis of the prediction patterns and residual distributions further demonstrates GRU's more consistent and accurate performance in capturing Ethereum price movements. These findings suggest that while both models are effective for cryptocurrency price prediction, GRU offers slightly better overall performance and stability, especially in maintaining consistent prediction accuracy across different market conditions.

Keywords: Ethereum price prediction, deep learning, long short-term memory (LSTM), gated recurrent unit (GRU), cryptocurrency, hyperparameter tuning.

1. Introduction

The development of its message and the increasing use of cryptocurrencies have changed the dynamics of investment, bringing opportunities and challenges for investors (Qi et al., 2025). Currently in the era of disruption 4.0, one of the interesting things in Indonesia is related to the presence of Crypto Investments such as Bitcoin (digital currency), where this is one of the economic indicators that can influence the global economy (Jubaedah et al., 2022). Ethereum, as one of the popular digital currencies, enables global value transactions without dependence on a single authority due to its decentralized nature (Abdiwi, 2024).

As a decentralized digital asset, Ethereum enables transactions without intermediaries, making it an attractive option for investors and blockchain technology developers (Sumathy, 2023). However, many investors suffered huge losses with the collapse of the ethereum coin in 2022 (Kale, 2022). In this case, deep learning approaches such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) become potential solutions to capture historical Ethereum price patterns.

The results of Rahmadeyan's (2024) research show that the application of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) is able to build an accurate model. Based on the implementation results, GRU proved to be the best model with a Mean Squared Error (MSE) value of 4958.9168, Root Mean Squared Error (RMSE) of 70.4195, and Mean Absolute Percentage Error (MAPE) of 1.1699%. Penelitian Khairunisa dan Hendikawati (2024) menunjukkan bahwa GRU lebih akurat dibandingkan LSTM, dengan RMSE 34.4233 dan MAPE 1.27%, sementara LSTM memiliki RMSE 35.3775 dan MAPE 1.28%.

Although various studies have compared the performance of LSTM and GRU in predicting financial asset prices, there are still differences in results regarding which model is superior. In addition, several previous studies have not optimized the model with hyperparameter tuning thoroughly, so the best performance potential of both models has not been fully explored. Therefore, this study aims to compare the performance of LSTM and GRU in predicting Ethereum prices with a hyperparameter tuning approach.

2. Methodology

This study aims to compare the performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models in predicting Ethereum prices using a deep learning approach. The study was conducted through several main stages, namely data collection, preprocessing, model building, hyperparameter tuning, model evaluation, and result analysis.

2.1. Data Collection

The data used in this study were obtained through web scraping techniques using the *yfinance library* from Yahoo Finance. The data collected includes the closing price of Ethereum in the period 2020-01-01 to 2025-01-31. because the closing price is often used in market analysis and financial asset price prediction models. The closing price movement of Ethereum (ETH-USD) from 2020 to 2025 will be shown in figure 1.

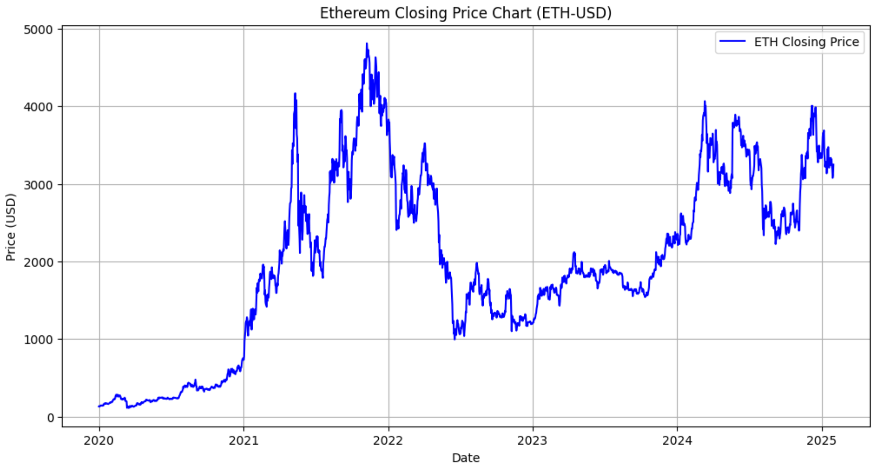

Figure 1: Ethereum (ETH-USD) closing price movement from 2020 to 2025

At the beginning of the period, the price of Ethereum experienced a gradual increase until mid-2021, where there was a sharp spike that peaked above 4000 USD. After that, the price experienced high volatility with several significant corrections, especially in 2022, where the price dropped drastically below 1000 USD. Entering 2023, the price began to recover and showed a moderate upward trend. In 2024 to 2025, the price of Ethereum again experienced fluctuations with several peaks and corrections, indicating a pattern of volatility that continued during that period.

2.2. Data Preprocessing

Before the data is processed for model training, a preprocessing stage is carried out to ensure that the data has optimal quality and consistency (Jo, 2019). The data is checked for the presence of missing values or anomalies, and if any are found, the linear interpolation method is applied to fill in the gaps. To prepare the data for both the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, it is transformed into a suitable format using the sliding window method, enabling the models to capture historical patterns in a time series format. The final step is normalization using the Min-Max Scaling method, which restructures data values within the range [0,1] (Raju et al., 2020). This normalization is applied using the following formula:

$$\acute{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where $\acute{X}$ is the data value after normalization, X is the original value, and X_{min} and X_{max} are the minimum and maximum values in the dataset.

2.3. Hyperparameter Tuning

In this study, hyperparameter tuning was performed to optimize the performance of both models. Keras Tuner was used with the Hyperband method, which is an optimization technique based on Successive Halving. Some of the main hyperparameters that were tuned include the number of units in each GRU layer, the dropout rate, and the learning rate of the Adam optimizer. The number of units in the first GRU layer was determined in the range of 32 to 256, while the second layer was in the range of 32 to 128, with a step of 32. In addition, dropout was applied to reduce the risk of overfitting, with the optimized value in the range of 0.1 to 0.5. The learning rate, which plays a role in determining the speed of model convergence, was tested with values of 0.01, 0.001, and 0.0001. To improve training efficiency, the Early Stopping and ReduceLROnPlateau strategies were used. Early Stopping stops training if the validation loss does not improve for five consecutive epochs, thus avoiding unnecessary training. Meanwhile, ReduceLROnPlateau reduces the learning rate by 50% if there is no performance improvement in three consecutive epochs, which helps the model achieve more stable convergence.

2.4. Model Development

Model development in this study was carried out based on the results of the hyperparameter tuning process using Keras Tuner with the Hyperband method, which aims to find the best combination of parameters for the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models. The model built consists of two GRU/LSTM layers with different units and dropout to avoid overfitting. The following table shows the results of hyperparameter tuning for both models:

Table 1: Hyperparameter tuning results for LSTM and GRU

Model	Layer 1 Units	Dropout 1	Layer 2 Units	Dropout 2	Learning Rate	Best Validation Loss
LSTM	96	0.2	64	0.4	0.01	0.0005149006
GRU	96	0.1	128	0.1	0.001	0.0005190244

Based on the tuning results, the LSTM model has 96 units in the first layer and 64 units in the second layer, with a dropout of 0.2 and 0.4, and a learning rate of 0.01. Meanwhile, the GRU model has 96 units in the first layer and 128 units in the second layer, with a dropout of 0.1 for both layers, and a learning rate of 0.001. Both models use the Adam Optimizer and the Mean Squared Error (MSE) loss function to minimize prediction errors. From the validation results, the LSTM model has a best validation loss of 0.0005149006, slightly better than the GRU model with a best validation loss of 0.0005190244.

2.5. Model Training

After obtaining the best hyperparameters, the LSTM and GRU models were trained using the processed dataset. Training was carried out for 50 epochs with a batch size of 16 and using Mean Squared Error (MSE) as the loss function. The optimizer used was Adam with a tuning learning rate. To prevent overfitting, early stopping was applied, which stops training if the validation loss does not improve after 5 epochs, and ReduceLROnPlateau to adjust the learning rate when the model stagnates. The training data consists of 80% of the data, while the remaining 20% is used for validation.

2.6. Model Evaluation

Model evaluation is performed using the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2 Score metrics to measure the model's performance in predicting Ethereum prices. These metrics are used to determine how far the prediction results are compared to the actual value (Chicco et al., 2021). RMSE measures the root of the mean squared error, where the smaller the RMSE value, the better the model's performance (Kambezidis, 2021). MAE calculates the average of the absolute errors between the predicted and actual values without considering the direction of the error (Karunasingha, 2022). Meanwhile, MAPE is used to determine the percentage of prediction error to the actual value, with smaller values indicating higher accuracy (Kim and Kim, 2016). The R^2 Score measures how well a model can explain variation in the data, with values closer to 1 indicating a better model in explaining stock price patterns (Spüler et al., 2015). Here is the formula used to calculate this evaluation:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \tilde{y}_i| \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \tilde{y}_i}{y_i}\right| \times 100\% \tag{4}$$

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2} \tag{5}$$

Where y_i is the actual value, $\tilde{y}_i$ is the predicted value, $\bar{y}_i$ is the average of the actual values and n is the number of samples.

3. Results and Discussion

3.1. Model Training Results

The comparison of loss during the training and validation process for the LSTM and GRU models will be shown in Figure 2.

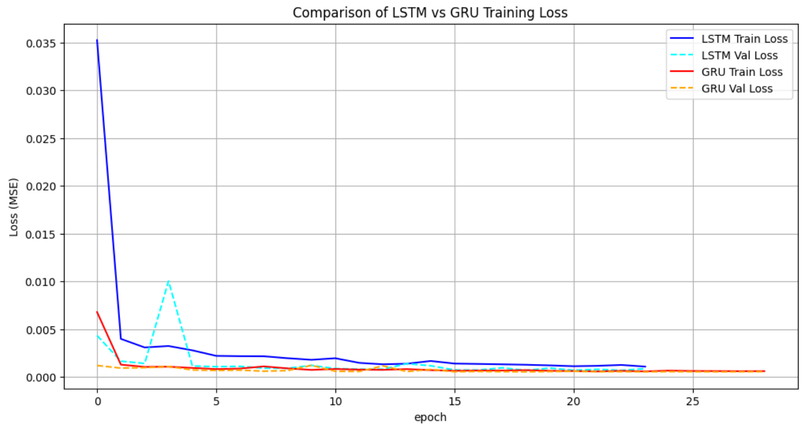

Figure 2: Comparison of training and validation loss on LSTM and GRU models

At the beginning of training (epochs 0–5), the training loss for both LSTM (blue line) and GRU (red line) models shows a drastic decrease, indicating that both models are rapidly learning patterns from the data. However, LSTM shows a sharper decrease than GRU. After epoch 5, the loss movement starts to stabilize and approaches a minimum value. However, the validation loss of LSTM (dashed cyan line) shows some spikes at the beginning of training, which could indicate temporary overfitting. In contrast, the validation loss of GRU (dashed orange line) is more stable throughout training. In the later epochs (10–25), both LSTM and GRU show almost flat losses, indicating that the models have reached a point of convergence. GRU appears to have a lower loss than LSTM, indicating better performance in minimizing errors.

3.2. Model Evaluation

The evaluation results of Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R^2 Score are presented in the following table:

Table 2: LSTM and GRU model evaluation results

Model	RMSE	MAE	MAPE (%)	R^2 Score
LSTM	0.0265	0.0193	18.08	0.9282
GRU	0.0234	0.0168	18.17	0.9442

From the evaluation results above, it can be seen that the GRU model has a smaller RMSE (0.0234) than LSTM (0.0265). The lower RMSE (Root Mean Squared Error) value indicates that the GRU model has a smaller prediction error rate than the LSTM model. In addition, the MAE (Mean Absolute Error) value of the GRU model of 0.0168 is also smaller than LSTM which has an MAE of 0.0193. This shows that on average, the absolute error in Ethereum price prediction is smaller in the GRU model compared to LSTM. In terms of MAPE (Mean Absolute Percentage Error), the LSTM model is slightly better than GRU, with a value of 18.08% compared to 18.17%.

Although the difference is small, this indicates that in some scenarios, the LSTM model is able to provide a lower relative error rate to the actual value than GRU. In terms of R^2 Score, which measures how well a model can explain variations in the data, the GRU model has a value of 0.9442, higher than LSTM which only reached 0.9282. A higher R^2 value indicates that the GRU model is better able to capture patterns from Ethereum's historical data, thus providing better predictions. Based on the results of this evaluation, the GRU model shows superior performance compared to LSTM in terms of RMSE, MAE, and R^2 Score, although it is slightly inferior in MAPE. Thus, the GRU model can be said to be a more accurate model in predicting Ethereum prices in this study.

3.3. Comparison of Actual Prices with Predictions

A comparison between the actual price and the predictions generated by the LSTM and GRU models in predicting stock price movements will be shown in Figure 3.

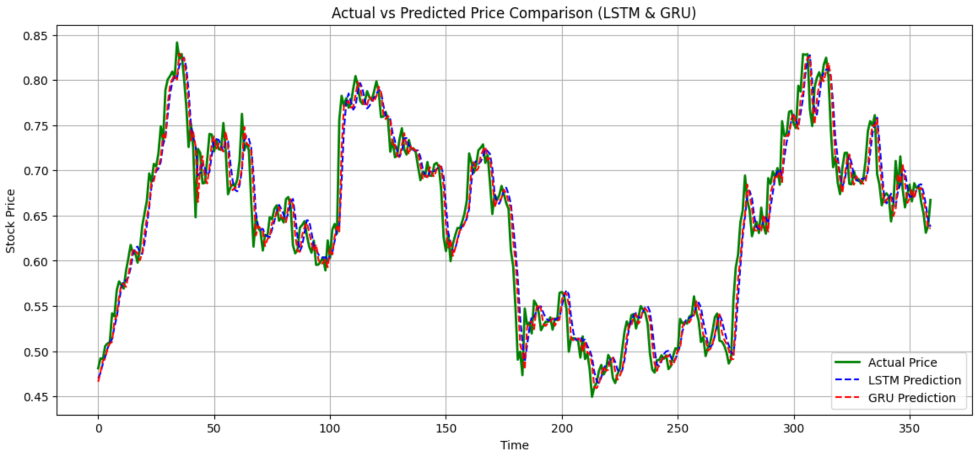

Figure 3: Comparison of actual and predicted price movements

Overall, the predicted price lines (LSTM in dashed blue and GRU in dashed red) have very similar movements to the actual price line (green). Both models are able to follow the up and down patterns of the actual price very well. Although there are some points where the predicted price is slightly different from the actual price, especially when there is a sharp change in trend, the difference is relatively small. This shows that both LSTM and GRU can capture the price movement pattern with high accuracy. In addition, to see the direct relationship between actual prices and predicted prices and the extent to which the model follows the actual price trend will be shown in Figure 4.

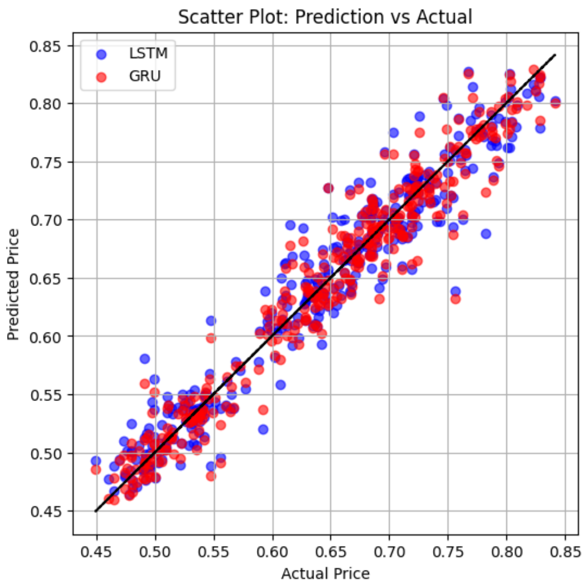

Figure 4: Relationship between model predictions and actual values

The results of Figure 4 show that the GRU Model has a denser distribution of points on the diagonal line compared to LSTM, indicating that GRU produces more accurate predictions with smaller errors. The prediction movement follows a linear pattern, indicating that the model is able to capture stock price trends well. The results of the prediction errors and to see if there are any patterns in the errors that indicate bias in the model will be displayed in Figure 5.

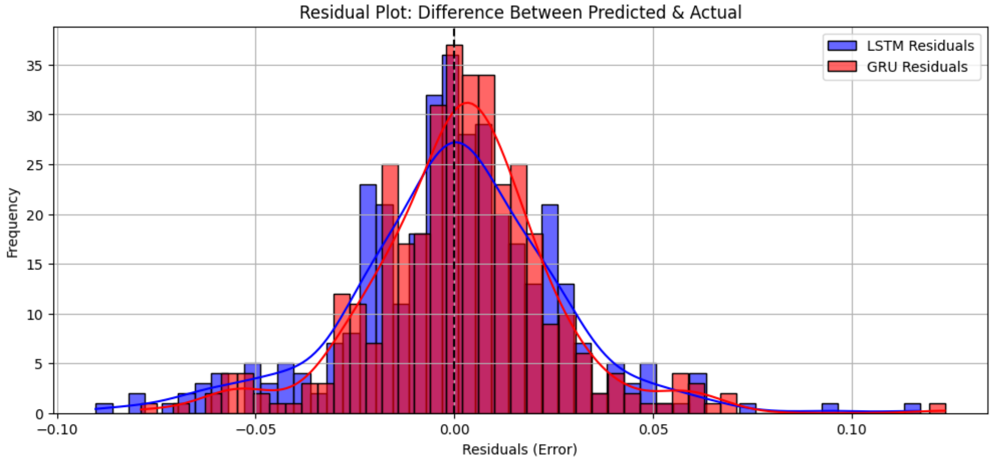

Figure 5: Model Prediction residual distribution

Based on Figure 5, most of the residuals are clustered around zero, which means the difference between the prediction and the actual price is small. The residual distribution of LSTM (blue color) looks slightly wider compared to GRU (red color), which indicates that the LSTM prediction has a larger error variation. The movement of GRU residuals is more centered around zero with a slimmer distribution, indicating that GRU is more consistent in producing accurate predictions. Both LSTM and GRU have residual distributions that resemble a normal distribution, indicating that the prediction errors are symmetrically distributed around the actual value.

4. Conclussion

Based on the research findings and analysis presented, this study demonstrates that both LSTM and GRU models are effective in predicting Ethereum prices, though GRU shows slightly superior performance overall. Through comprehensive hyperparameter tuning and evaluation using multiple metrics, the GRU model achieved better results in terms of RMSE (0.0234), MAE (0.0168), and R^2 Score (0.9442) compared to LSTM's performance (RMSE: 0.0265, MAE: 0.0193, R^2 Score: 0.9282). Although LSTM showed marginally better performance in terms of MAPE (18.08% vs 18.17%), the GRU model's more stable validation loss during training and denser distribution of predictions around actual values indicate its greater reliability and consistency in price prediction. The visual analysis of residuals further supports this conclusion, showing that GRU maintains a more concentrated distribution of errors around zero.

References

Abdiwi, F. G. (2024). Detection of Digital Currency Fraud through a Distributed Database Approach and Machine Learning Model. *TEM Journal*, *13*(4), 3025-3039.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, *7*, e623.

Jo, J. M. (2019). Effectiveness of normalization pre-processing of big data to the machine learning performance. *The Journal of the Korea institute of electronic communication sciences*, *14*(3), 547-552.

Jubaedah, D., Ahyani, H., Putra, H. M., Prakasa, A., & Mutmainah, N. (2022). Legal Analysis of Crypto Investment in Era 4.0 View from Credo Theory. *Diponegoro Law Review*, *7*(2), 262-278.

Kale, S. (2022). They couldn't even scream any more. they were just sobbing': The amateur investors ruined by the crypto crash [WWW document]. *The Guardian. URL https://www. theguardian. com/technology/2022/jul/12/they-couldnt-even-scream-any-m o re-they-were-just-sobbing-the-amateur-investors-ruined-by-the-crypto-cras h (accessed 7.14. 22)*.

Kambezidis, H. D. (2021). The Solar Radiation Climate of Greece. *Climate*, *9*(12), 183.

Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, *585*, 609-629.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669-679.

Qi, J., Zhang, Y., & Ouyang, C. (2025). Cryptocurrency Investments: The Role of Advisory Sources, Investor Confidence, and Risk Perception in Shaping Behaviors and Intentions. *Journal of Risk and Financial Management*, *18*(2), 57.

Rahmadeyan, A. (2024). Long Short-Term Memory and Gated Recurrent Unit for Stock Price Prediction. *Procedia Computer Science*, *234*, 204-212.

Raju, V. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., & Padma, V. (2020, August). Study the influence of normalization/transformation process on the accuracy of supervised classification. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 729-735). IEEE.

Spüler, M., Sarasola-Sanz, A., Birbaumer, N., Rosenstiel, W., & Ramos-Murguialday, A. (2015, August). Comparing metrics to evaluate performance of regression methods for decoding of neural signals. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1083-1086). IEEE.