

Ethereum Gas Fee Prediction Using LSTM's Bi-Directional Algorithm

Rizky Parluka¹, Irohito Nozomi², Muhamad Faizhal Musthafa³

^{1,3}Universitas Pembangunan Nasional "Veteran" Jawa Timur, Surabaya, Indonesia

²Universitas Putra Indoensia YPTK, Padang, Indonesia

¹rizkyparluka.if@upnjatim.ac.id, ²irohito_nozomi@upiyptk.ac.id, ³22081010032@student.upnjatim.ac.id

Article Info

Article history:

Received August 08, 2026

Revised August 25, 2026

Accepted August 26, 2026

Keywords:

Ethereum

Gas Fee

Bi-LSTM

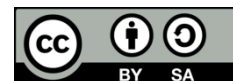
Deep Learning

Time Series Prediction

ABSTRACT

The Ethereum ecosystem is synonymous with high volatility dynamics, not only in the exchange rate of its assets but also in the transaction fees (gas fees). Because of these uncertainties, proper cost estimation is essential for transaction efficiency. In this study, the Bi-Directional Long Short-Term Memory (Bi-LSTM) algorithm is suggested to be used to model Ethereum's gas fee predictions. The Bi-LSTM architecture was chosen because of its advantages in processing sequential information from two directions (forward and backward). Thus, this architecture makes it possible to map the dependencies of complex long-term patterns more comprehensively compared to one-way models. Data is processed through the normalization and restructuring stages of the sliding window using hourly resolution on-chain datasets collected from Etherscan over a period of one year. The results of the performance evaluation showed that the model was quite responsive in following the trend of gas cost movements, as shown by the Root Mean Squared Error (RMSE) value of 9.3517 Gwei and Mean Absolute Error (MAE) of 3.7893 Gwei. However, a Mean Absolute Percentage Error (MAPE) value of 20% was obtained, which suggests that the model still struggles to anticipate extreme price spike anomalies, also known as spikes. Overall, the study concludes that Bi-LSTM has significant potential for blockchain time series data analysis; However, in order to handle the highly volatile characteristics of the market, this method needs to be further optimized.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Irohito Nozomi

1. INTRODUCTION

The emergence of cryptocurrencies, which have evolved from mere speculative assets to the basis of the global digital economy, has defined the development of financial technology in the last ten years. Although cryptocurrencies have a very high level of volatility and difficult market characteristics to predict, their ecosystem has grown rapidly to become one of the most in-demand financial instruments in the world. Over the years, constant price fluctuations have been the subject of much research, sparking the interest of academics and practitioners to dissect the dynamics of this decentralized market. To understand the patterns of market movements that often seem random (stochastic), a number of advanced computing techniques have been developed, one of which is by utilizing daily transaction data to form a patterned sample that can be analyzed [1]. In this context, the use of minimum, maximum, and average values in historical datasets is an important step to understand the basic nature of the fluctuations of these digital assets [2].

Ethereum plays an important role as a pioneer of smart contract technology that enables decentralized applications (Dapps) and Decentralized Finance (DeFi) for the development of blockchain platforms. However, this popularity has resulted in significant technical problems, especially related to the operating costs of the transaction and the exchange rate of the asset itself, which is technically known as gas fees. The supply and demand mechanism of block space within the Ethereum network is very dynamic; As network activity increases, competition between users to validate their transactions triggers a drastic spike in costs. This phenomenon often leads to network congestion and very high costs, which ultimately harm retail users and institutions. Therefore, transaction cost optimization in the Ethereum blockchain has become a very crucial and urgent issue, requiring an in-depth analytical approach to ensure the efficiency of funds as well as the sustainability of the scalability of the Ethereum blockchain network in the future [3].

Despite the fact that mitigating the risk of transaction fee volatility is crucial, a current literature review suggests that there is an imbalance in the research perspective. The majority of research is still dominated by research that focuses on predicting asset prices or commonly known as forecasting. Researchers are more likely to develop price prediction models of major cryptocurrencies such as Bitcoin, XRP, Tether, and Ethereum that use various variants of artificial neural networks and machine learning algorithms [4][5]. In contrast, there is very little specialized research on gas fee predictions and has not included asset price predictions. DeepAR's probabilistic methods for predicting gas prices and machine learning approaches to cost estimation are some of the initial initiatives [6][7]. In addition, there is also a comparative study on machine learning methods for transaction cost forecasting [8]. Conventional methods, however, are often unable to capture the full complexity of gas cost data, which has the characteristics of long-term or long-term dependencies, as well as high noise levels. Compared to classical statistical methods or standard machine learning, a more sophisticated and flexible model architecture is required.

Deep Learning algorithms have been proven to be empirically better than conventional methods such as ARIMA or linear regression in handling complex and non-linear time series data on the blockchain. In particular, the Long Short-Term Memory (LSTM) model has been recognized for its effectiveness in predicting the movement of virtual currencies due to its ability to recall historical information over a long period of time and overcome the vanishing gradient problem that often occurs in Recurrent Neural Networks (RNNs) [9]. Thus, LSTM architecture has become a standard where this architecture has structural limitations because it only processes data in one direction, namely running from the past to the future. The model may lose any contextual information that may be available from future data sequences during this training phase due to these limitations.

To overcome these limitations, the Bi-Directional LSTM (Bi-LSTM) architecture was developed as a more comprehensive evolution. Bi-LSTM has the ability to process data from two directions at once (forward and backward), which allows the model to capture the context of the pattern in a more complete, precise, and in-depth way. Previous studies comparing the performance of LSTMs, Gated Recurrent Units (GRUs), and Bi-LSTMs concluded that a two-way approach provides significantly more accurate forecasting results, especially in highly volatile crypto markets [10]. Bi-LSTM represents a novel model that is able to improve prediction performance by studying the data structure from both endpoints of time, making it more robust against data anomalies [11].

The application of Bi-LSTM in the cryptocurrency domain has shown very promising results in various experiments. The integration of the Bi-LSTM algorithm, both as a single model and in a hybrid architecture (such as CNN-BiLSTM), is capable of producing high accuracy with a determination coefficient value (R^2) reaching 0.9415 in the case of Ethereum price prediction. These advantages show that Bi-LSTM, when combined with data pre-processing and precise feature selection, is able to maintain trends and reduce prediction lag effectively. Departing from a real research gap, where the reliability of Bi-LSTM has been tested for validity in asset price prediction but has not been explored optimally for the case of transaction fees (gas fees), this study aims to fill the contribution gap. This study proposes the implementation and analysis of the Bi-Directional LSTM algorithm specifically to predict Ethereum gas fee fluctuations. In this study, the model is expected to be developed not only able to provide better mathematical accuracy, but also can be a practical instrument to anticipate spikes in network costs, so that it can be a strategic reference for DApps users and developers in conducting transaction efficiency in the ever-evolving blockchain ecosystem.

2. METHOD

The research methodology used in this study is systematically designed to achieve the main goal, which is to predict transaction fees (gas fees) on the Ethereum network with high accuracy using Deep Learning algorithms. This approach adapts a framework that has been tested in various recent literature studies on crypto asset forecasting and time series. The research flow consists of the stages of data collection, pre-processing, model architecture design, and performance evaluation.

2.1. Data Collection

The first step in this study is the collection of historical data on Ethereum transaction fees obtained from a trusted public source, namely Etherscan. The dataset covers the last one-year time period with hourly data resolution to capture intraday dynamics. The main variable used is the Average Gas Price measured in Gwei units (10⁻⁹ ETH).

Data retrieval is done automatically using the Etherscan API. This approach based on on-chain analysis is crucial because the data directly from the blockchain network provides the most accurate picture of network congestion and block demand. On-chain metrics have a stronger correlation to price volatility than external market data [12]. In addition, the selection of data periods also takes into account the post-implementation market mechanism of EIP-1559 (London Hard Fork), which changes the base fee mechanism to be more dynamic but predictable, as analyzed by Reijnders et al. in the transaction fee market efficiency study [13]. The data obtained is ensured to be real-time and covers all extreme fluctuations that occur during the observation period.

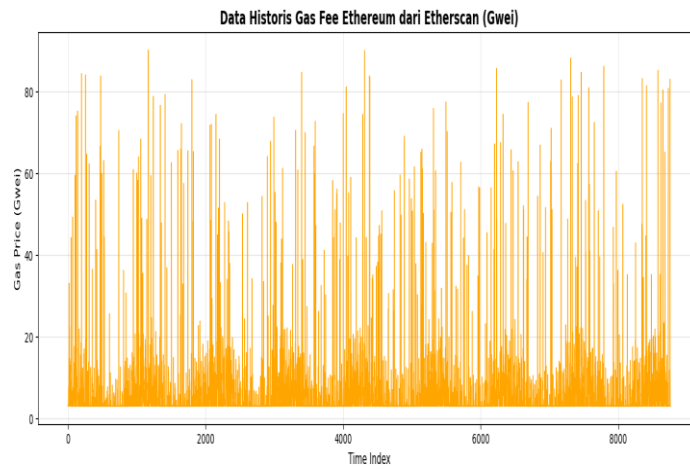


Figure 1. Gas Free Ethereum Historical Data

From Figure 1, it can be seen that Ethereum's historical Gas Fee data has extreme volatility characteristics (heteroscedasticity), where value can jump from the range of 10 Gwei to above 100 Gwei in a short period of time due to congested network activity (e.g. during NFT minting or DeFi volatility).

2.2. Pre-processing of Data

The pre-processing stage of data is a critical phase to ensure the quality of input for the Deep Learning model and reduce noise that can distort the learning process. In handling time series data, crypto assets require the right pre-processing techniques so that the Recurrent Neural Networks (RNN) model can work optimally without being trapped in local minima [14].

2.2.1. Data Normalization

Gas fee data has a highly variable range of values with high volatility. If raw data is directly fed into the neural network, the error gradient can explode (exploding gradient) or disappear (vanishing gradient). To accelerate the convergence of the model during the training process and balance the feature weight, normalization was carried out using the Min-Max Scaling method. This method transforms data x into the range $[0, 1]$ using the equation:

$$x^1 = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

2.2.2. Formation of Time Sequence Datasets

The normalized data is converted into a supervised learning format using the sliding window technique. This technique maps the sequence of past data (lagged values) as an input feature (X) to future values as targets (Y).

In this study, a time step (time window) of 48 hours was used. This means that the model will use a data sequence of the previous 48 hours ($t-48$ to $t-1$) to predict the price of gas fees in the next hour ($t+1$). The

choice of a 48-hour time window is based on empirical studies that suggest that intraday and daily cycle patterns in the crypto market often repeat over a 2-day period [15].

2.3. LSTM Bi-Directional Model Architecture

The proposed model is Bi-Directional Long Short-Term Memory (Bi-LSTM). Unlike standard LSTMs that only process information from the past to the future, Bi-LSTM processes time sequence data from two directions: forward pass (past to future) and backward pass (future to past). This mechanism allows the model to capture the richer context of the time series data structure. The final output of each time step is the result of the concatenation (usually concatenation) of the two directions:

$$y_t = \sigma(W_y[h_t, \bar{h}_t] + b_y) \quad (2)$$

The Bi-LSTM approach, especially when combined with proper feature weighting, is able to significantly improve the accuracy of crypto asset price forecasting compared to the one-way model. In addition, Bi-LSTM-based models are highly effective in capturing volatile market dynamics [16][17].

2.3.1. Layer Configurations

The model is built using the TensorFlow/Keras framework with the following architectural configuration:

- Input Layer: Accepts tensors with dimensions (Batch_Size, 48, 1) as per the time step.
- Bi-LSTM Layer 1: Consists of 128 neuron units with return_sequences=True. This arrangement is crucial for the entire output sequence to be passed to the next recurring layer, maintaining the temporal dimension of the data.
- Dropout Layer 1: Using a 0.1 ratio. This layer randomly deactivates 10% of neurons during training to prevent overfitting [18].
- Bi-LSTM Layer 2: Consists of 64 neuron units with return_sequences=False. This layer serves as a high-level feature extractor that condenses sequential information into a single context vector.
- Dropout Layer 2: Uses a 0.1 ratio for additional regularization.
- Dense Layer (Output): 1 unit of neurons with a linear activation function to predict the continuous value of gas fees.

2.3.2. Model Training

The model training process was carried out using the Adam optimization algorithm (Adaptive Moment Estimation). Adam was chosen as the lead optimizer because of his excellence in combining the heuristics of RMSProp and Momentum, which allows him to adaptively adjust the learning rate for each weight parameter. This adaptive mechanism is crucial for handling data with sparse gradients and high noise characteristics such as crypto asset transaction data [19]. To measure the performance of the model during iteration, the loss function is set as Mean Squared Error (MSE). The selection of the MSE is based on its mathematical properties which provide a quadratic penalty against large prediction errors. Given that Ethereum's gas fees often experience extreme spikes, the use of MSE forces the model to be more sensitive and try to minimize errors at those extreme values so as not to result in significant under-fitting.

The technical configuration of the training is set for up to 100 epochs, with a batch size of 32 epochs. This batch size was chosen to optimally compensate for the computational speed and stability of the stochastic gradient estimation. An early stop mechanism is used to prevent overfitting, when the model detects and memorizes noise from the training data and is unable to generalize. This mechanism will track the value of the validation loss and automatically stop the training process if there are no significant errors within a certain period of time. This method has proven successful in generating convergent, stable, and accurate prediction models for Ethereum without consuming excessive computing resources [20]. To ensure the validity of the evaluation, the dataset is partitioned in an 80:20 ratio. 80 percent of the data is distributed chronologically as a training set for pattern learning, and the remaining 20 percent is separated as a test set to measure and test the performance of the trained model on never-before-seen future data.

2.4. Model Evaluation

The performance of the model used is evaluated with a set of standard metrics in time series prediction to ensure the validation of the model's results. The use of a combination of absolute and relative metrics in

comparing the performance of Machine Learning and Deep Learning models on crypto data is more recommended [21]. Metrics to be used in the model evaluation include:

- Root Mean Squared Error (RMSE): Measures the average of prediction errors in original units (Gwei), giving more weight to the big errors that often occur during price spikes.
- Mean Absolute Error (MAE): Measures the average of the absolute difference between the predicted value and the actual value to find out the average error magnitude.
- Mean Absolute Percentage Error (MAPE): Measures the percentage of predicted average error against the actual value, providing an overview of relative accuracy.
- Coefficient of Determination (R2 Score): Added to measure how well the variability of the data can be explained by the model. A value close to 1 indicates an excellent model fit, as suggested in the evaluation of modern regression models [22]

The prediction results are then visualized in the form of a time series comparison graph between the actual data and the prediction data, as well as an error distribution histogram for further residual analysis.

3. RESULTS AND DISCUSSION

3.1. Data Descriptive Statistics

Before moving on to the Deep Learning model training stage, a comprehensive descriptive statistical analysis was carried out on the Ethereum gas fee dataset obtained from Etherscan. This stage is crucial to map the characteristics of the data distribution that the model will study. Based on the hourly data collected during the observation period (a total of 8,760 data points), there are strong indications that the Ethereum transaction fee market has extreme volatility (heteroscedasticity) with a very wide range of values, as summarized in Table 1.

Table 1. Descriptive Statistics of Ethereum Gas Fee Data

Metric Statistics	Value (Gwei)
Amount of Data	8760
Minimum Score	3.00
Maximum Value	90.09
Mean	6.50
Standard Deviation	9.28

An in-depth analysis of Table 1 reveals a significant gap between the mean value (mean) of 6.50 Gwei and the maximum value of 90.09 Gwei. This disparity, coupled with a high standard deviation value (9.28 Gwei)—which even exceeds its own average—indicates that the data distribution is highly skewed to the right (positive skewness). This phenomenon shows that although network conditions are generally stable at low numbers, the frequency of extreme price spikes (outliers) is quite significant. Data characteristics like this are a challenge in time series forecasting, because the model must be able to distinguish between normal trend patterns and momentary anomalies due to network congestion.

3.2. Model 2 Training Results

The Bi-Directional Long Short-Term Memory (Bi-LSTM) model training process was carried out for 100 epochs by utilizing Adam optimization. Monitoring of model performance during the learning phase was carried out by observing the Mean Squared Error (MSE) Loss Function decline curve, both in training data and validation data, as presented in Figure 2.



Figure 2. Graph Loss

Based on the visualization in Figure 2, a healthy convergence pattern can be seen where losses decrease sharply in the first 10 epochs, indicating that the learning model recognizes the dominant features of the data quickly. After that initial phase, the curve slopes and moves steadily until the end of training. The most crucial thing to note is the very thin gap between the training loss line and the validation loss. This phenomenon indicates that the model has good generalization capabilities and does not experience overfitting—a condition in which the model only memorizes training data but fails to predict new data.

3.3. Model Performance Evaluation

To quantitatively measure the reliability of the proposed Bi-LSTM architecture, an evaluation was conducted on the test data (test set) using three industry-standard metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The results of the calculation of these metrics are summarized in Table 2.

Table 2. Descriptive Statistics of Ethereum Gas Fee Data

Evaluation Metrics	Value (Gwei)	Remarks
RMSE	9.3517	Root Mean Squared Error (Gwei)
MAE	3.7893	Mean Absolute Error (Gwei)
MAPE	20%	Mean Absolute Percentage Error

The results of the evaluation show quite promising performance. The MAE value of 3.7893 Gwei means that on average, the model's prediction is only about 3 to 4 Gwei from the original price. Meanwhile, the higher RMSE value (9.3517 Gwei) compared to MAE shows that there are still some prediction errors of large magnitude. This is natural given that the RMSE formula gives a quadratic penalty weight to large errors, which usually occurs when the gas fee suddenly jumps dramatically (anomaly).

Relatively speaking, a MAPE score of 20% puts the model in the "Decent" to "Good" forecasting category according to the Lewis standard evaluation scale. This achievement is quite competitive compared to those that apply the GRU-LSTM hybrid model to Ethereum data and notes a comparable error rate in volatile market conditions. This MAPE value confirms that the model is able to capture daily price dynamics fairly accurately, although challenges remain at extreme points.

3.4. Prediction Visualization Analysis

Qualitative analysis was carried out by comparing the time series plot between the actual value (ground truth) and the estimated value of the Bi-LSTM model on the test data, as visualized in Figure 3. The graph provides in-depth insights into the dynamic behavior of the model in response to stochastic gas fee market fluctuations.

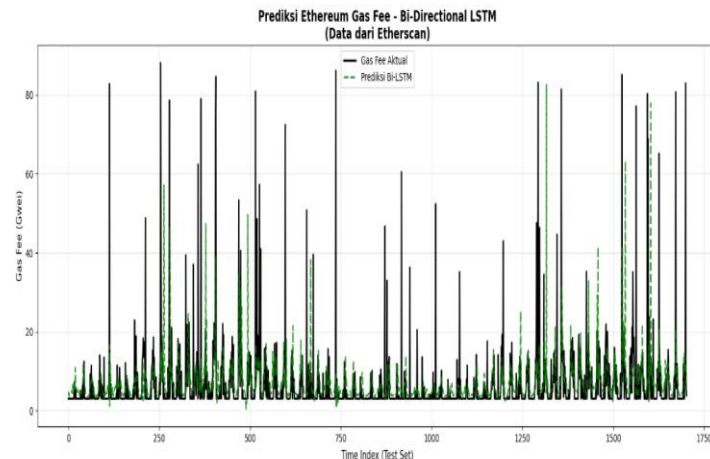


Figure 3. Ethereum Free Gas Prediction Chart

From Figure 4, it is clear that the Bi-LSTM model has managed to map global trend patterns; When the actual gas fee rises or falls, the model's prediction line is able to follow the direction of the movement quite responsively. However, a more detailed analysis of peak areas shows the existence of an under-prediction phenomenon, where the model's predictions tend to be lower than the original extreme price spikes.

In general, the Bi-LSTM model shows an impressive ability to map global trends (trend-following). When transaction costs increase or decrease gradually, the model's prediction lines are able to squeeze with enough precision to the actual data. prove that the bidirectional architecture of Bi-LSTM provides significant advantages over unidirectional LSTMs or GRUs, especially in its ability to capture the long-term temporal context of historical data sequences to minimize prediction lag [23].

However, closer observation at the peak point shows the phenomenon of under-prediction; Here, the model tends to provide a lower value estimate than the actual extreme price spike. For example, in the 150th and 300th hours of the test data, there was a huge increase in gas costs, reaching more than 70 Gwei. The model, which anticipates a range of 50-60 Gwei, can only respond to it.

These limitations can be explained through two theoretical perspectives:

Loss Function Properties: During training, the model was encouraged to minimize global mean errors by using mean squared error (MSE) values. As a result, to avoid large error penalties, models tend to "play it safe" by predicting values around the moving average. This causes the model to lose sensitivity to outliers. Deep Learning models on gas fee data often have difficulty capturing extreme volatility due to the highly skewed and noisy distribution of data [24].

Historical Feature Limitations: Gas price spikes are often triggered by real-time network congestion in the transaction queue, rather than simply being affected by past prices. Since this model is only trained using historical price data without including the mempool density feature in real-time, it is natural for the model to experience delays or conservative bias when there is a sudden spike due to "gas wars" during NFT minting or massive DeFi activity [25].

Despite the deviation in the extreme values, the MAPE value of 20% indicates that the model is still very reliable for general use, such as estimated daily transaction costs. In the context of highly volatile kipto predictions, maintaining an error rate below 25% is already categorized as optimal model performance and can be implemented for transaction time recommendation systems for users [26].

4. CONCLUSION

The study found that the implementation of the Bi-Directional Long Short-Term Memory (Bi-LSTM) algorithm is effective for projecting transaction fees or gas fees on the Ethereum network. Overall, as demonstrated by the pattern proximity between the prediction results and the actual data, the model has demonstrated the ability to handle complex time series data patterns and provide responses that correspond to the dynamics of network trends.

From a quantitative point of view, the results of the daily estimate are considered satisfactory, with a Mean Absolute Error (MAE) of 3.7893 Gwei and a Root Mean Squared Error (RMSE) of 9.3517 Gwei. However, the Mean Absolute Percentage Error (MAPE) value of 20% shows that the model still has limitations in handling the phenomenon of extreme volatility. This decrease in precision is especially observed when there

is a sudden price spike that deviates from the normal pattern. This is the normal nature of the crypto asset market that is difficult to avoid with just historical data.

Overall, this study validates that deep learning techniques that have a bidirectional architecture have very promising potential for blockchain time series analysis. Nonetheless, future model development is critical to achieving a higher level of resilience to market anomalies. To overcome uncertainty in volatile market conditions, it is recommended to use hybrid methods and incorporate external features, such as market sentiment or transaction volume.

REFERENCES

- [1] T. B. Sianturi, I. Cholissodin, and N. Yudistira, "Penerapan Algoritma Long Short-Term Memory (LSTM) berbasis Multi Fungsi Aktivasi Terbotol dalam Prediksi Harga Ethereum," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 3, pp. 1101–1107, 2023.
- [2] R. Parlika and B. Rahmat, "Minimum, Maximum, and Average Implementation of Patterned Datasets in Mapping Cryptocurrency Fluctuation Patterns," *INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION (JOIV)*, vol. 8, pp. 378–386, Mar. 2024.
- [3] A. Laurent, L. Brotcorne, and B. Fortz, "Transactions fees optimization in the Ethereum blockchain," *Blockchain: Research and Applications*, 2022.
- [4] R. Kaur et al., "Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum," *PeerJ Comput Sci*, 2025.
- [5] V. Tyshchenko, S. Achkasova, O. Naidenko, S. Kanyhin, and V. Karpova, "Development of recurrent neural networks for price forecasting at cryptocurrency exchanges," *Eastern-European Journal of Enterprise Technologies*, 2023.
- [6] A. Ferenczi and C. Bădică, "Prediction of Ethereum gas prices using DeepAR and probabilistic forecasting," *Journal of Information and Telecommunication*, 2023.
- [7] R. Mars, A. Abid, S. Cheikhrouhou, and S. Kallel, "A Machine Learning Approach for Gas Price Prediction in Ethereum Blockchain," in *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2021.
- [8] C. Butler and M. Crane, "Blockchain Transaction Fee Forecasting: A Comparison of Machine Learning Methods," *ArXiv*, 2023.
- [9] X. Li, Q. Liu, and Y. Wu, "Prediction on blockchain virtual currency transaction under long short-term memory model and deep belief network," *Appl. Soft Comput.*, 2021.
- [10] P. L. Seabe, C. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," *Fractal and Fractional*, 2023.
- [11] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms," *AI*, 2021.
- [12] N. Jagannath et al., "An On-Chain Analysis-Based Approach to Predict Ethereum Prices," *IEEE Access*, 2021.
- [13] D. Reijnders, S. Sridhar, B. Monnot, S. Leonardos, S. Skoulakis, and G. Piliouras, "Transaction Fees on a Honeymoon: Ethereum's EIP-1559 One Month Later," Apr. 2022, [Online]. Available: <http://arxiv.org/abs/2110.04753>
- [14] S. Hansun, A. Wicaksana, and A. Q. M. Khaliq, "Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches," *J Big Data*, 2022.
- [15] S. M. Raju and A. M. Tarif, "Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis,"
- [16] M. Rafi, Q. Mirza, M. I. Sohail, M. L. Aliasghar, A. Aziz, and S. Hameed, "Enhancing Cryptocurrency Price Forecasting Accuracy: A Feature Selection and Weighting Approach With Bi-Directional LSTM and Trend-Preserving Model Bias Correction," *IEEE Access*, 2022.
- [17] M. F. Bin Hossain, L. Z. Lamia, M. M. Rahman, and Md. M. Khan, "FinBERT-BiLSTM: A Deep Learning Model for Predicting Volatile Cryptocurrency Market Prices Using Market Sentiment Dynamics," *ArXiv*, 2024.
- [18] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.
- [19] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Jan. 2017, [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [20] H. Dzakwan Diash, V. Nathania, and M. Idhom, "Application of CNN-BiLSTM Algorithm for Ethereum Price Prediction," 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [21] K. Murray, A. Rossi, D. Carraro, and A. Visentin, "On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles," *Forecasting*, 2023.
- [22] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput Sci*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.
- [23] K. Murray, A. Rossi, D. Carraro, and A. Visentin, "On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles," *Forecasting*, 2023.
- [24] C. Jin and Y. Li, "Cryptocurrency Price Prediction Using Frequency Decomposition and Deep Learning," *Fractal and Fractional*, 2023.
- [25] D. Lan, H. Wang, C. Yin, L. Zhou, C. Ge, and X. Lu, Gas Price Prediction Based on Machine Learning Combined with Ethereum Mempool. 2022. doi: 10.1109/MASS56207.2022.00057.
- [26] I. P. B. Priadinata, I. G. I. Sudipa, N. P. S. Meinarni, I. M. L. Radhitya, and I. K. D. G. Supartha, "Comparative Analysis of LSTM, GRU, and Bi-LSTM Deep Learning Models for Time Series Cryptocurrency Price Forecasting," *sinkron*, vol. 9, no. 3, pp. 1024–1035, Jul. 2025, doi: 10.33395/sinkron.v9i3.14795.