International Journal of Advanced Multidisciplinary Research and Studies

Forecasting Silver Prices Using the Long Short-Term Memory (LSTM) Method

[1] Dian Kurniasari, [2] Stevanus Kenny, [3] Warsono, [4] Widiarti

[1, 2, 3, 4] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

Corresponding Author: **Dian Kurniasari**

Abstract

Silver is a versatile metal with numerous applications in human life, including industrial materials, ornaments, and investment assets. It serves as an alternative investment option due to its lower price compared to gold and its potential for value appreciation under specific market conditions. Historically stable, silver prices experienced a decline following the COVID-19 pandemic. Consequently, this study aims to forecast silver price trends to determine whether they will recover and stabilize. The forecasting is conducted using the Long Short-Term Memory (LSTM) method for a three-month period, from January 2023 to March 2023. The model's performance is evaluated using the Mean Absolute Percentage Error (MAPE), with the best-obtained MAPE value being 0.0547. The forecasting results indicate that silver prices will experience a brief decline before rebounding.

1. Introduction

The rapid advancement of technology has led to a future increasingly reliant on technological developments, particularly in forecasting. Forecasting is a compelling and relevant issue that integrates the need for analyzing time series data and predicting the most probable future outcomes [1].

Silver is one of the most renowned metals worldwide, serving multiple purposes, including its use as an alternative to gold in jewelry due to its relatively lower cost. Given its properties similar to gold, silver has also emerged as a competitive investment asset. Silver prices tend to remain stable and can appreciate under certain conditions. However, during the COVID-19 pandemic, silver prices declined significantly before recovering over time, eventually stabilizing. Thus, forecasting silver prices is crucial for assessing its viability as a sound investment. This study aims at forecasting silver prices for the period from January 2023 until March 2023 using historical data from January 2013 until January 2023, sourced from Investing.com.

Time series analysis offers various forecasting methods, including Auto-Regressive (AR), Moving Average (MA), and Auto-Regressive Moving Average (ARMA). Although linear statistical forecasting models are widely used, they often fail in providing accurate predictions for nonlinear time series data due to their underlying assumption of stationarity and linearity [2].

Recurrent Neural Networks (RNNs) have been introduced as a suitable forecasting approach for nonlinear time series data. However, RNNs struggle with long-term memory retention, leading to information loss over extended sequences. Long Short-Term Memory (LSTM) networks, a variant of RNN, improve upon this limitation by incorporating gating mechanisms that selectively store and discard information, ensuring long-term dependency retention [3]. The three gates in LSTM—input gate, forget gate, and output gate—enable effective filtering and updating of information, making it well-suited for long-term time series forecasting, including silver price predictions.

Previous studies have demonstrated the effectiveness of LSTM in various forecasting applications. Kumar *et al.* [4] applied LSTM for cloud datacenter forecasting, while Yadav *et al.* [5] used it in predicting the Indian stock market. Song *et al.* [6] employed LSTM in oil production forecasting, and Hong *et al.* [7] utilized it in air pollutant concentration prediction for managing port air quality. More recently, Pothuganti [8] applied LSTM in stock market exchange predictions.

This study aims at applying the LSTM model in forecasting silver prices in Indonesia over the next three months, evaluating its predictive performance and assessing its potential for investment decision-making.

2. Research Method

The data utilized in this study consists of secondary data obtained from https://id.investing.com/commodities/silver, specifically historical silver futures prices spanning a ten-year period from January 2013 to January 2023 on a weekly scale. The dataset comprises 522 entries, directly downloaded as a single CSV file. The price data is denominated in USD and measured in Troy ounces (31 grams).
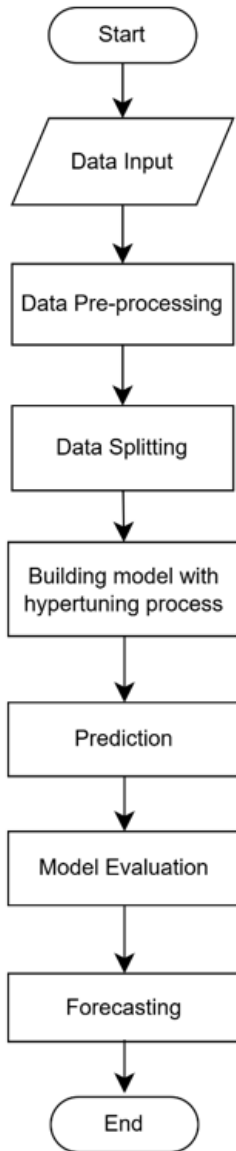

Fig 1: Research methods

2.1 Time Series

According to Kirchgässner et al. [9], a time series is a collection of quantitative observations arranged in chronological order. The data are systematically gathered over time, with intervals that may range from hours, days, and weeks to months, quarters, or years.

2.2 Data Mining

Data mining is a process that employs various analytical tools to identify patterns and relationships within data, facilitating the generation of valid predictions [10]. A crucial stage in data mining is data preprocessing, wherein raw data is refined to address inconsistencies, missing values, and redundancies that may hinder data processing. Several techniques exist for data preprocessing, including data cleaning, which eliminates noise and resolves inconsistencies, and data transformation, such as normalization, which enhances the accuracy and efficiency of algorithms [11]. One widely used normalization technique is Min-Max Normalization, which linearly transforms original data to maintain balance between pre- and post-processing values. The Min-Max Normalization formula is as Equation (1) [12].

$$x_{norm} = \frac{x - x_{min}}{r_{max} - r_{min}} \tag{1}$$

2.3 Machine Learning

As stated by Murphy [13], machine learning encompasses a range of methodologies that can autonomously detect patterns within data and subsequently utilize these patterns to predict future data or support decision-making under conditions of uncertainty, such as strategizing data collection efforts. Machine learning is generally classified into supervised, unsupervised, and reinforcement learning.

2.4 Deep Learning

Deep Learning was introduced in 2006 by Geoffrey E. Hinton, Simon Osindero, and Yee-Whye I, building upon the foundations of artificial neural networks. This breakthrough reignited interest in neural network research, leading to deep learning being regarded as a "new-generation neural network" [14].

According to Deng and Yu [15], three key factors contributed to the widespread adoption of deep learning: The rapid advancement of processing power, the substantial increase in training dataset sizes, and significant progress in machine learning and signal/information processing research.

2.5 Recurrent Neural Network

Recurrent Neural Networks (RNNs) represent a class of Artificial Neural Networks (ANNs) characterized by temporal loops in addition to input, hidden, and output units [16]. Their ability to retain previous inputs through internal memory makes them particularly effective for time-series data [17].

RNNs can also be considered a specialized form of feedforward networks, structured based on temporal layering, allowing them to process input sequences and generate corresponding output sequences [18]. However, a notable limitation of RNNs is their inability to retain long-term memory, leading to the loss of crucial information from earlier inputs.

2.6 Long Short-Term Memory

Long Short-Term Memory (LSTM) is a structured chain-based neural network model derived from RNNs, characterized by its intrinsic temporal memory and cyclical training feedback adaptation [19]. LSTM has the capability to learn which data should be retained and which should be discarded, facilitated by its specialized neuron architecture, where each neuron contains multiple gates that regulate its memory.

As noted by Saud and Shakya [20], LSTM is explicitly designed to mitigate the vanishing and exploding gradient problem. According to Karmiani et al. [21], LSTM memory

870

comprises an input gate, a forget gate, and an output gate. Each memory cell in LSTM consists of three sigmoid layers and one tanh layer, forming a structure that governs data retention and flow through these three essential gates.

Neural networks employ two primary activation functions: The sigmoid activation function and the tanh activation function. Activation functions play a crucial role in determining whether a neuron is activated or remains inactive [22].

1. The sigmoid activation function transforms input values ranging from -1 to 1 into an output between 0 and 1. It is mathematically represented as Equation (2) [23]:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

2. The tanh activation function operates within the range of -1 to 1 and serves as an alternative to the sigmoid function. Its mathematical formulation is given by Equation (3):

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

The following describes the gates within a single Long Short-Term Memory (LSTM) cell:

a. **Forget Gate:** This gate determines which information is deemed less relevant and should be discarded, utilizing the sigmoid activation function. The corresponding equation is:

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \tag{4}$$

b. **Input Gate:** This gate filters and selects specific information to be updated in the cell state. The sigmoid activation function is used for selection, while a new candidate vector is generated using the tanh function. The respective equations are:

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \tag{5}$$

$$C_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \tag{6}$$

Subsequently, the process of updating old information with new information is expressed as follows:

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{7}$$

c. **Output Gate:** This gate functions to determine the output value of the hidden state and regulate the cell state using the tanh activation function, based on input and memory block computations. After generating both sigmoid and tanh output values, their activations are multiplied, as represented by the Equations (8) and (9):

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = o_t * \tanh(C_t) \tag{9}$$
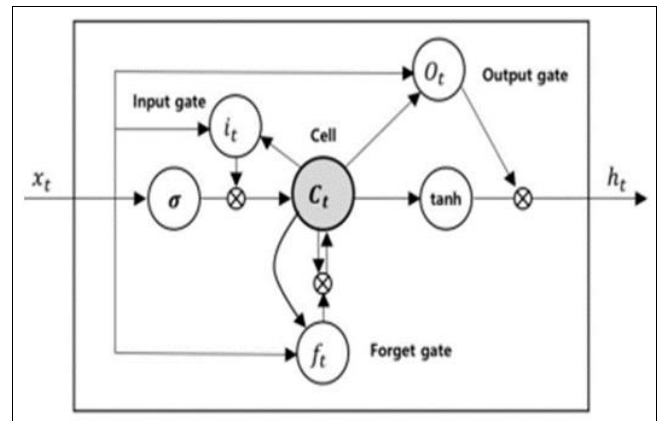
Source: Chung, H. & Shin, K [24]

Fig 2: LSTM Architecture

2.7 Metric Evaluation

In most forecasting applications, accuracy is regarded as the primary criterion for selecting a forecasting method. In many instances, the term "accuracy" is synonymous with "goodness-of-fit," which ultimately refers to how well a forecasting method can replicate known data. However, in forecasting, the accuracy of future predictions is of utmost importance. For a model, the "goodness-of-fit" pertains to the extent to which the model's known information—both quantitative and qualitative—has been considered [25]. One commonly used method to assess model accuracy is Mean Absolute Percentage Error (MAPE).

MAPE is employed to evaluate the precision of a model's estimates by expressing the average absolute percentage error. It is systematically formulated as Equation (10):

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{X_i - F_i}{X_i} \times 100\% \right| \tag{10}$$

3. Results and Discussion
3.1 Data Input
The variable utilized in this study is the Closing Price column, with data spanning from January 6, 2013, to January 1, 2023, at a weekly interval. The closing price is of particular significance as it serves as a reference for the opening price on the subsequent trading day. Table 1 illustrates the data input:

Table 1: Data Input

Dates	Closing Price
6/1/2013	
13/1/2013	
20/1/2013	
27/1/2013	
…	…
11/12/2022	23,328
18/12/2022	23,920
25/12/2022	24,040
1/1/2023	23,982

3.2 Data Visualization
This study utilizes acquired input data to analyze the evolution of silver prices over the past ten years through graphical representation. The objective is to determine whether the data exhibit fluctuations or remain stable and to

identify the periods when silver prices reached their highest and lowest points within the observed timeframe.
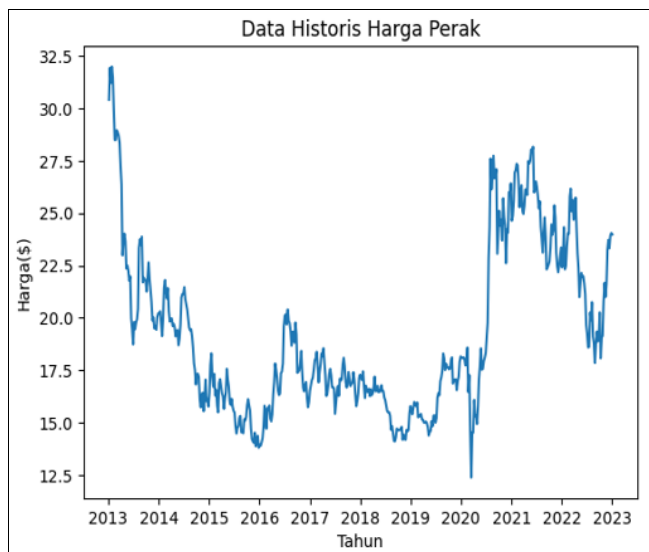

Fig 3: Data visualization

As illustrated in Fig 3, silver prices peaked in early 2013, reaching approximately $32. Between 2013 and 2015, the price fluctuated before entering a consistent downward trend until early 2016. Over the following three years, silver prices remained relatively stable, with the exception of a decline in 2019. The onset of the COVID-19 pandemic in late 2019 drove silver prices to their lowest level in a decade, reaching $12.5. However, a rapid recovery ensued, bringing prices back up to $27, followed by fluctuations until early 2022. Throughout 2022, silver prices declined for several months before rebounding toward the end of the year.

3.3 Handling Missing Value
A subsequent analysis is necessary to determine the presence of missing values within the dataset. If missing data account for only a small fraction (e.g., 1% of the total dataset), their impact is minimal. However, if a significant portion of data is missing, appropriate imputation techniques, such as replacing missing values with the mean or median of the respective column, should be applied to maintain data integrity.

```
#check missing value
print('Checking missing value')
print(df.isnull().sum())
print('Counting total missing value')
print(df.isnull().sum().sum())

Checking missing value
Harga Penutupan     0
dtype: int64
Counting total missing value
0
```

Fig 4: Check missing value

3.4 Data Splitting
The data splitting in this study follows three distinct schemes: (1) 70% training data, 20% validation data, and 10% testing data; (2) 75% training data, 20% validation data, and 5% testing data; and (3) 80% training data, 15% validation data, and 5% testing data. The training data is utilized for model training, the validation data is employed for model evaluation, and the testing data serves to assess the model's performance. Details of this data splitting are presented in Table 2.

Table 2: Data Splitting

Scheme	Subset	Data
1	Training 70%	364
	Validation 20%	104
	Testing 10%	49
2	Training 75%	392
	Validation 20%	104
	Testing 5%	21
3	Training 80%	418
	Validation 15%	78
	Testing 5%	21

3.5 Building the LSTM Model
LSTM networks involve several key parameters, including units, batch size, dropout, and epochs. In this study, a dropout rate of 0.1 and 200 epochs were employed. The number of units and batch size were determined through hyperparameter tuning, with initial parameter settings of 16, 32, 64, and 128 units, and batch sizes of 8, 16, and 32. The objective of this tuning process was to identify the optimal combination of parameters from the predefined variations. The process was conducted using early stopping to terminate training once the best-performing model was identified in Fig 5.

As shown in Fig 5, the optimal LSTM configuration consists of 128 units and a batch size of 8. This model will be utilized for predicting silver prices.

```
Best: -0.004085 using {'LSTM_unit': 128, 'batch_size': 8}
-0.045664 (0.017642) with: {'LSTM_unit': 16, 'batch_size': 8}
-0.051676 (0.021425) with: {'LSTM_unit': 16, 'batch_size': 16}
-0.060420 (0.026481) with: {'LSTM_unit': 16, 'batch_size': 32}
-0.044333 (0.027557) with: {'LSTM_unit': 32, 'batch_size': 8}
-0.047546 (0.015946) with: {'LSTM_unit': 32, 'batch_size': 16}
-0.058233 (0.023933) with: {'LSTM_unit': 32, 'batch_size': 32}
-0.020012 (0.024534) with: {'LSTM_unit': 64, 'batch_size': 8}
-0.049475 (0.027335) with: {'LSTM_unit': 64, 'batch_size': 16}
-0.049475 (0.014081) with: {'LSTM_unit': 64, 'batch_size': 32}
-0.004085 (0.003599) with: {'LSTM_unit': 128, 'batch_size': 8}
-0.026687 (0.031091) with: {'LSTM_unit': 128, 'batch_size': 16}
-0.055972 (0.027081) with: {'LSTM_unit': 128, 'batch_size': 32}
```

Fig 5: Hypertuning process

3.6 Model Evaluation
An evaluation of the model is conducted to analyze the loss plot and validation loss plot of the LSTM model. This process aims to ensure that the LSTM model intended for forecasting does not suffer from overfitting or underfitting. Overfitting occurs when the training loss continuously decreases to a minimum point while the validation loss starts increasing at a certain stage. Conversely, underfitting arises when the training loss fails to reach a minimum point and remains relatively high.

1. First Scheme
After obtaining the optimal parameters through hyperparameter tuning, the model's performance will be assessed using training loss and validation loss.

Based on Fig 6, the training loss decreases to its minimum point, while the validation loss, which initially rises during the first three epochs, subsequently declines to a level close to the training loss. This indicates that the LSTM model demonstrates satisfactory performance, as both the training loss and validation loss converge towards minimal values with no significant discrepancy between them.
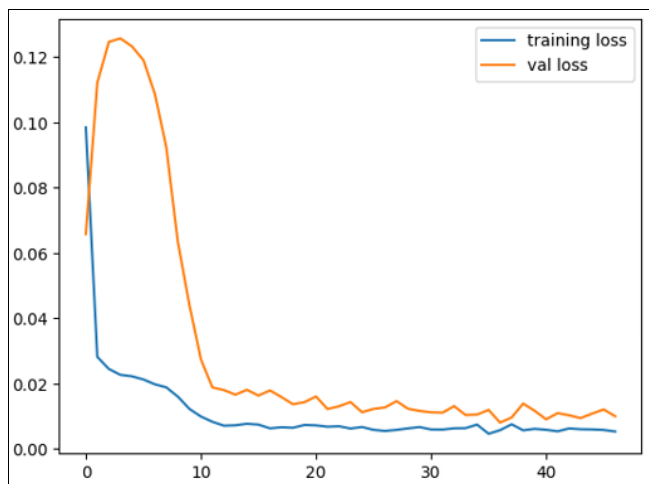

Fig 6: Loss graph of first scheme model

2. Second Scheme
In the second scheme, the same process as in the first scheme will be conducted. The model evaluation will be performed using training loss and validation loss plots to assess the model's feasibility.
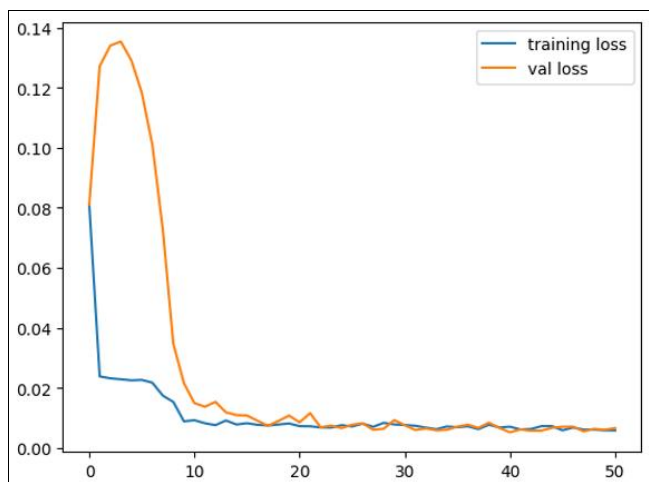

Fig 7: Loss graph of second scheme model

As depicted in Fig 7, the training loss and validation loss curves decrease to a stable point. The convergence of these two curves indicates that the model does not suffer from overfitting.

3. Third Scheme
The evaluation of the third scheme model was conducted using training loss and validation loss plots to assess the model's feasibility.
As depicted in Fig 8, both training loss and validation loss decrease to a stable point, indicating that the third-scheme LSTM model does not suffer from overfitting or underfitting.
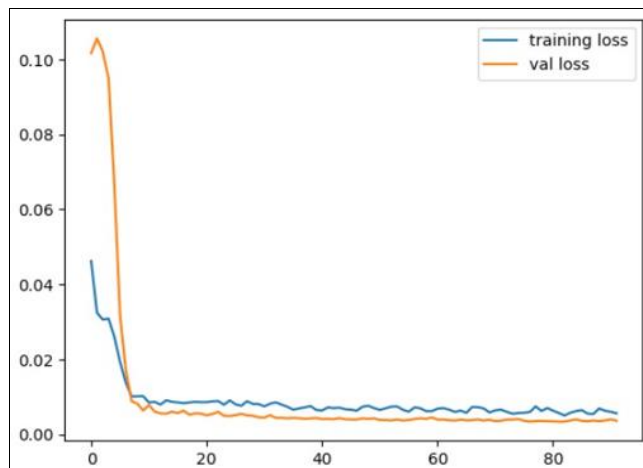

Fig 8: Loss graph of third scheme model

3.7 Silver Prices Prediction
This study aims to predict silver prices using normalized silver price data to assess the effectiveness of the developed Long Short-Term Memory (LSTM) model. The primary objective is to determine whether the model can accurately forecast silver prices. If the predicted values closely align with actual market data, the model can be considered suitable for silver price forecasting. The denormalized predicted values will be compared against actual data through graphical representations. Additionally, the accuracy of the predictions will be evaluated across different forecasting scenarios.

1. First Scheme
The predictive outcomes of training data, validation data, and testing data in the first scheme will be compared to actual data through visualized plots. Fig 9 above presents a plot visualization of the predicted and actual silver prices. Across all plots, it is evident that the predicted patterns closely align with the actual data trends, indicating the reliability and accuracy of the predictive results.
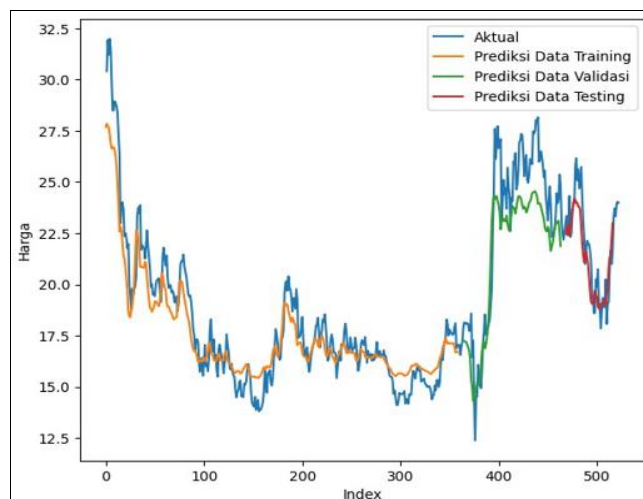

Fig 9: Prediction result of first scheme model

2. Second Scheme
A comparative analysis between the predictive outcomes of the second scheme and the actual data will be conducted through visualization plots. As illustrated in Fig 10, the predicted results successfully align with the actual data patterns to a considerable extent, indicating no significant differences between the observed and forecasted values.
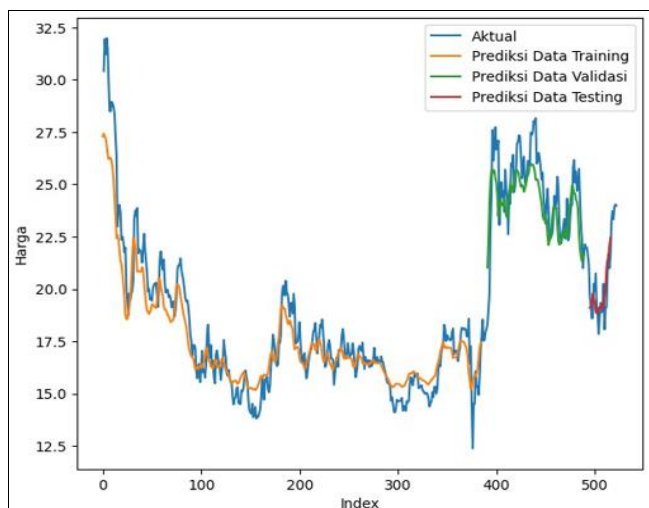
873

Fig 10: Prediction result of second scheme model

3. Third Scheme

The visualization of plots will be conducted to compare the obtained prediction results under the third scheme with the actual data. Fig 11 illustrates that the predicted data closely follows the pattern of the actual data. All prediction plots exhibit minimal deviation from the actual data trend, indicating the reliability of the predictions.
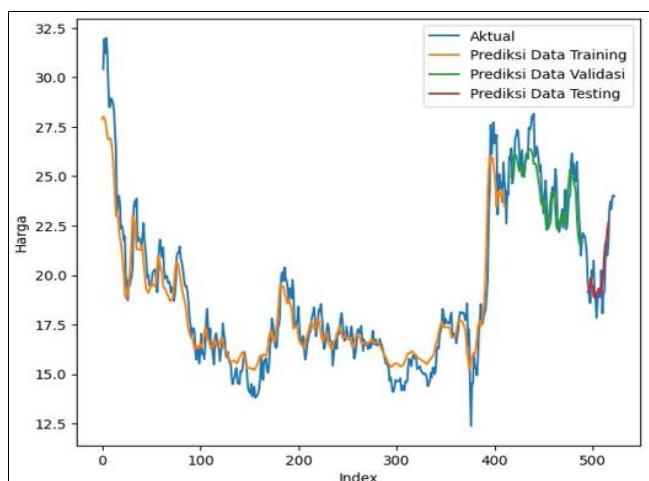

Fig 11: Prediction result of third scheme model

As presented in Table 3, the Mean Absolute Percentage Error (MAPE) improves as the size of the training dataset increases. The first scheme, which allocates 70% of the data for training, 20% for validation, and 10% for testing, yields a MAPE of 0.0626, corresponding to an accuracy of 99.9375%. The second scheme, with 75% training data, 20% validation data, and 5% testing data, results in a MAPE of 0.0618, achieving an accuracy of 99.9381%. The third scheme, which assigns 80% of the data for training, 15% for validation, and 5% for testing, attains the lowest MAPE of 0.0547, with an accuracy of 99.9452%.

Table 3: Model Evaluation

Scheme	MAPE	Accuracy
1	0,0626	99,9373%
2	0,0618	99,9381%
3	0,0547	99,9452%

Given that the MAPE values remain below 10%, the model's predictive performance can be deemed highly accurate, with the third scheme demonstrating the most optimal results.

3.8 Forecasting

Following the satisfactory evaluation of the predictive model, the most optimal model has been selected to forecast silver prices for the next three months, specifically for January, February, and March 2023. The forecasting results are as follows:
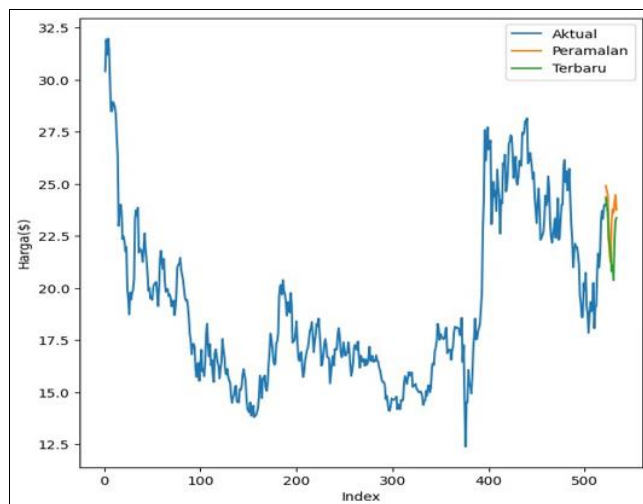

Fig 12: Forecasting of first scheme model

As illustrated in Fig 12, the orange line indicates that silver prices initially hover around $24.9 before declining to $21.2 and subsequently rebounding to $23.75. Similarly, the green line depicts an initial price of approximately $24.3, followed by a decrease to $20.3, before recovering to $23.3. These projections demonstrate a strong alignment with the latest data.

Fig 13 presents a similar trend, where the orange line begins at $23, declines to $20.7, and then rises to $22.5. Meanwhile, the green line starts at $24.3, falls to $20.3, and rebounds to $23.3. These results further affirm the accuracy of the forecasting model.
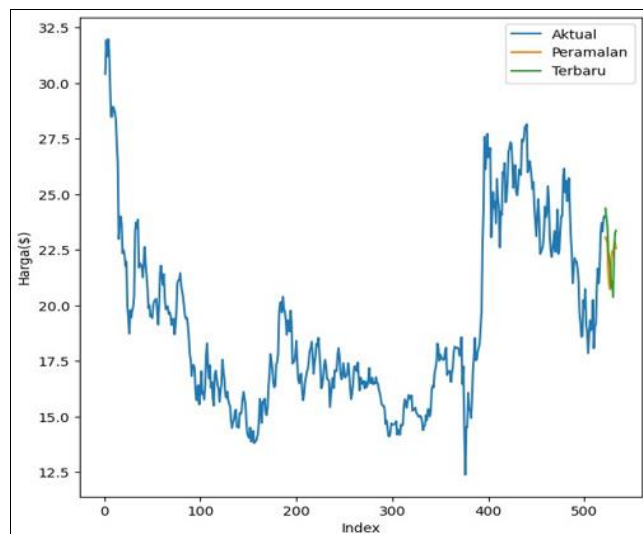

Fig 13: Forecasting of second scheme model

In Fig 14, the orange line represents silver prices fluctuating from \$23.5, decreasing to \$20.7, and subsequently recovering to \$23.3. Likewise, the green line starts at \$24.3, declines to \$20.3, and then rebounds to \$23.3. These results indicate a high degree of accuracy in capturing actual price movements.
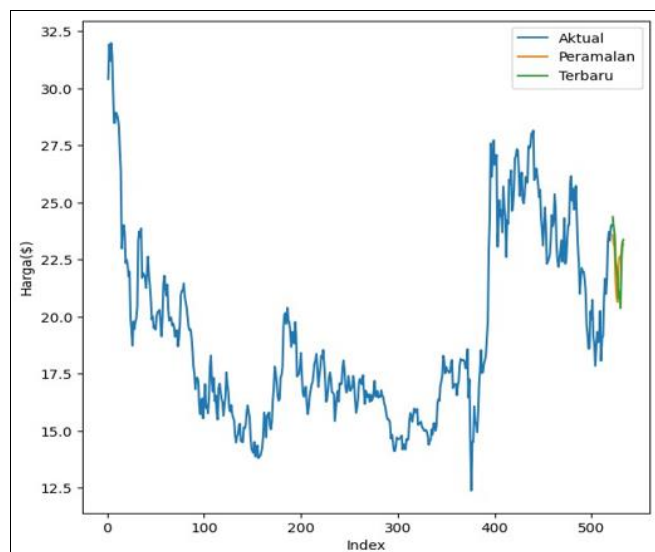

Fig 14: Forecasting of third scheme model

Across all three forecasting schemes, the model demonstrates strong predictive performance, closely following the most recent price trends with minimal discrepancies. The closer the forecasted values are to actual data, the more reliable the forecasting model becomes. Among the three schemes, the third forecasting approach exhibits the highest accuracy in mirroring actual market trends. Consequently, this model is deemed the most effective, predicting a temporary decline in silver prices followed by a subsequent recovery over the next three months.

4. Conclusion

Based on the findings and discussion, it can be concluded that silver price forecasting was conducted using an LSTM model with 128 units, a dropout rate of 0.1, a batch size of 8, and 200 epochs with early stopping. Model evaluation based on MAPE values indicates that the third forecasting scheme achieved the highest accuracy with a MAPE of 0.0547, outperforming the first and second schemes, which recorded MAPE values of 0.0626 and 0.0618, respectively. Given that all MAPE values are below 10%, it can be inferred that the predictive model performs well. The forecast for the period from January to March 2023 suggests that silver prices will experience a temporary decline before rising again.

5. References

1. Hahn Y, Langer T, Meyes R, Meisen T. Time Series Dataset Survey for Forecasting with Deep Learning. Forecasting. 2023; 5(1):315-335.
2. El-Houda Bezzar N, Laimeche L, Meraoumia A, Houam L. Data analysis-based time series forecast for managing household electricity consumption. Demonstr Math. 2022; 55(1):900-921.
3. Handari BD, Niman IMS, Hasan A, Purba JRP, Hertono GF. Comparation of elman neural network, long short-term memory, and gated recurrent unit in predicting dengue hemorrhagic fever at dki Jakarta. Commun Math Biol Neurosci. 2021; 2021.
4. Kumar J, Goomer R, Singh AK. Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model for Cloud Datacenters. In: Procedia Computer Science, 2018, 676-682.
5. Yadav A, Jha CK, Sharan A. Optimizing LSTM for time series prediction in Indian stock market. In: Procedia Computer Science, 2020, 2091-2100.
6. Song X, Liu Y, Xue L, Wang J, Zhang J, Wang J, *et al.* Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. J Pet Sci Eng. 2020; 186.
7. Hong H, Jeon H, Youn C, Kim HS. Incorporation of shipping activity data in recurrent neural networks and long short-term memory models to improve air quality predictions around busan port. Atmosphere (Basel). 2021; 12(9).
8. Pothuganti K. Long Short-Term Memory (LSTM) Algorithm Based Prediction of Stock Market Exchange. SSRN Electron J, 2021.
9. Kirchgässner G, Wolters J. Introduction to modern time series analysis. Introduction to Modern Time Series Analysis, 2007, 1-274.
10. Chakraborty S, Islam SH, Samanta D. Introduction to Data Mining and Knowledge Discovery. In: EAI/Springer Innovations in Communication and Computing, 2022, 1-22.
11. Ha J, Kambe M, Pe J. Data Mining: Concepts and Techniques. 4th ed. Data Mining: Concepts and Techniques, 2011, 1-703.
12. Shahi TB, Shrestha A, Neupane A, Guo W. Stock price forecasting with deep learning: A comparative study. Mathematics. 2020; 8(9).
13. Murphy KP. Machine Learning - A Probabilistic Perspective - Table-of-Contents. MIT Press, 2012.
14. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. Vol. 2, SN Computer Science, 2021.
15. Deng L, Yu D. Deep learning: Methods and applications. Vol. 7, Foundations and Trends in Signal Processing, 2013, 197-387.
16. Iskandar UP, Kurihara M. Long Short-term Memory (LSTM) Networks for Forecasting Reservoir Performances in Carbon Capture, Utilisation, and Storage (CCUS) Operations. Sci Contrib Oil Gas. 2022; 45(1):35-50.
17. Tariq H, Hanif MK, Sarwar MU, Bari S, Sarfraz MS, Oskouei RJ. Employing deep learning and time series analysis to tackle the accuracy and robustness of the forecasting problem. Secur Commun Networks, 2021.
18. Aggarwal CC. Neural Networks and Deep Learning: A Textbook. Neural Networks Deep Learn A Textb, 2023, 1-529.
19. Zhou H, Wang T, Zhao H, Wang Z. Updated Prediction of Air Quality Based on Kalman-Attention-LSTM Network. Sustain. 2023; 15(1).
20. Saud AS, Shakya S. Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE. In: Procedia Computer Science, 2020, 788-798.
21. Karmiani D, Kazi R, Nambisan A, Shah A, Kamble V.

Comparison of Predictive Algorithms: Backpropagation, SVM, LSTM and Kalman Filter for Stock Market. In: Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019, 2019, 228-234.

22. Nwankpa C, Ijomah W, Gachagan A, Marshall S. Activation Functions: Comparison of trends in Practice and Research for Deep Learning, November 8, 2018. Available from: http://arxiv.org/abs/1811.03378

23. Neapolitan RE, Jiang X. Neural Networks and Deep Learning. In: Artificial Intelligence, 2018, 389-411.

24. Chung H, Shin KS. Genetic algorithm-optimized long short-term memory network for stock market prediction. Sustain. 2018; 10(10).

25. Briggs WM, Makridakis S, Wheelwright SC, Hyndman RJ, Diebold FX. Forecasting: Methods and Applications. J Am Stat Assoc. 1999; 94(445):345.