e-ISSN: 2614-8404
p-ISSN:2776-3234

Short-Term and Long-Term Forecasting of Global Gold Prices Using LSTM and GRU Models

Fitria[1*)], Muhammad Syahid Pebriadi[2], Nitami Lestari Putri[3]
[1]Program Studi Sistem Informasi Akuntansi, Jurusan Akuntansi, Politeknik Negeri Banjarmasin
[2]Program Studi Sistem Informasi Akuntansi, Jurusan Akuntansi, Politeknik Negeri Banjarmasin
[3]Program Studi Sistem Informasi Kota Cerdas, Jurusan Teknik Elektro, Politeknik Negeri Banjarmasin
Email: [1]fitria@poliban.ac.id, [2]m.syahid@poliban.ac.id, [3]nitamiputri@poliban.ac.id

Abstract – Global gold prices exhibit high volatility and complex temporal patterns, making accurate forecasting a challenging task. This study aims to compare the performance of deep learning models for short-term and long-term gold price prediction using daily historical data. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) were selected because both models can capture temporal dependencies in financial time-series data, while having different architectural complexities and learning characteristics. Comparing these models is important to identify the most suitable approach for different forecasting horizons. The dataset consists of daily global gold prices denominated in USD obtained from an open financial data source covering the period from 2010 to 2024. The models were evaluated under two forecasting horizons, namely short-term prediction (1 day ahead) and long-term prediction (30 days ahead). Model performance was assessed using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Experimental results indicate that the GRU model outperforms LSTM in short-term forecasting by producing lower prediction errors, while LSTM demonstrates slightly better stability in long-term forecasting. These findings suggest that the effectiveness of deep learning models for gold price prediction is highly dependent on the forecasting horizon.

Keywords – Gold price prediction, LSTM, GRU, time series forecasting, deep learning

I. INTRODUCTION

Gold is often regarded as a 'safe-haven' asset, particularly in times of economic uncertainty, prompting a strong interest in its price forecasting [1]. Predicting gold prices not only aids investors in making informed decisions but also assists governments in economic planning [2]. Various studies have underscored the significance of timely and accurate gold price predictions given its role in market dynamics and investment strategies [3], [4].

Gold price forecasting is a complex time-series problem because price dynamics are often nonlinear and shaped by a wide range of global economic factors. Beyond short-term fluctuations driven by market sentiment, gold prices also exhibit long-run patterns associated with inflation, exchange rates, and broader macroeconomic conditions. These characteristics give rise to heterogeneous temporal dependencies across both short and long horizons, thereby increasing the difficulty of producing accurate price predictions.

Conventional approaches to time series forecasting, including classical statistical models, typically assume linear relationships and data stationarity [5], [6]. These assumptions are often violated in financial datasets that are inherently dynamic and volatile. Consequently, such models are limited in their ability to capture nonlinear patterns and complex temporal dependencies embedded in gold price movements. These limitations underscore the need for alternative methods that are more flexible and adaptive to the characteristics of modern financial data.

Alongside advances in computing technology and the increasing availability of large-scale datasets, machine learning and deep learning approaches have been increasingly adopted for forecasting financial asset prices. These methods can automatically learn complex patterns from historical data without relying on stringent linearity assumptions. In the context of gold price prediction, prior studies indicate that deep learning–based models often deliver higher predictive accuracy than traditional statistical techniques, particularly when the data exhibit nonlinear dynamics and high volatility.

One of the most widely adopted deep learning architectures for time series data is the Recurrent Neural Network (RNN), particularly its variants Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). LSTM was developed to address the vanishing gradient problem in conventional RNNs and is designed to capture long-range dependencies in sequential data. In contrast, GRU can be viewed as a streamlined alternative to LSTM, featuring a more compact gating structure and fewer parameters. As a result, GRU often provides better computational efficiency while maintaining competitive performance across a wide range of time series forecasting applications.

Although LSTM and GRU have been widely applied to gold price forecasting, most prior studies evaluate model performance at a single, fixed forecasting horizon. This approach does not explicitly distinguish the characteristics and challenges associated with short term versus long term prediction. In practice, short term forecasting emphasizes a model's ability to track daily price fluctuations, whereas long term forecasting prioritizes robustness and stability in capturing sustained trends and longer run patterns [7], [8], [9]. These differing objectives may substantially influence the suitability and effectiveness of the modeling approach used.

In the context of investment decision making and risk management, distinguishing between short term and long term forecasting is crucial. Short term investors require predictions that are sensitive to daily price movements, whereas long term investors need a more stable representation of price trends to support portfolio planning.

e-ISSN: 2614-8404
p-ISSN:2776-3234

Therefore, the evaluation of gold price forecasting models should explicitly account for the forecasting horizon so that the resulting performance assessment is better aligned with end user needs.

The dataset used in this study was collected from the Kaggle platform, which provides publicly accessible historical financial market data. Specifically, the dataset contains daily global gold prices denominated in USD covering the period from January 2010 to December 2024. The data were downloaded in CSV format and subsequently processed for forecasting experiments. The use of a long historical observation period enables the models to learn both short-term fluctuations and long-term temporal trends in gold price movements.

Building on the identified research gap, this study aims to compare the predictive performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models for forecasting global gold prices across two distinct time horizons: short-term (one day ahead) and long-term (thirty days ahead). Using historical daily gold price data, the study evaluates both models based on standard prediction error metrics, namely the Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). The findings are expected to provide a more comprehensive understanding of how forecasting horizon influences deep learning model performance, while also offering practical guidance for researchers and practitioners in selecting an appropriate gold price prediction model for analytical and decision-making purposes.

II. RESEARCH METHODOLOGY

This study was used a quantitative, experimental approach to forecast global gold prices using historical time-series data. The research procedure were comprised data collection, data preprocessing, predictive model development, and performance evaluation under both short-term and long-term forecasting horizons.

A. Dataset

The dataset used in this study was obtained from the Kaggle platform through the dataset entitled “Gold Price Historical Data” . The dataset contains historical daily gold prices denominated in United States Dollars (USD) and covers the observation period from January 2010 to December 2024.

The data were downloaded in CSV format and consist of several variables, including Date, Open, High, Low, Close, and Volume. Among these variables, the closing price was selected as the primary variable for forecasting because it represents the final market price at the end of each trading day and is commonly used in financial time-series prediction studies.

The dataset was chosen because it provides long-term sequential financial data with daily granularity, allowing the analysis of both short-term price fluctuations and long-term gold price trends using deep learning models.

B. Data Preprocessing

The historical gold price data used in this study were obtained as a daily time-series dataset. The first preprocessing step involved verifying that the time column (Date) was stored in a valid date format and that the observations were ordered chronologically.

Next, the closing price column was selected as the primary variable because it represents the final transaction price each day and constitutes the main target for price forecasting. To enhance the performance of the deep learning model, the data were subsequently normalized using Min-Max scaling, so that all values fell within the range of 0 to 1. This normalization facilitates faster convergence during model training.

C. Sequence Construction and Prediction Horizon

Historical gold price data were transformed into sequential samples using a 30-day window, where the preceding 30 days of prices were used as the model input. The study distinguishes two forecasting horizons, namely short-term and long-term prediction. Short-term prediction refers to one-day-ahead forecasting, whereas long-term prediction refers to 30-days-ahead forecasting. This separation is intended to examine how differences in the forecasting horizon affect model performance.

D. Forecasting Model

This study employs Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, both of which belong to the Recurrent Neural Network (RNN) architecture. The LSTM architecture incorporates memory cells and gates that regulate the flow of information. These gates include input, output, and forget gates, which manage the cell state and decide what information to keep or discard. This structured gating mechanism facilitates the learning of long-term dependencies in sequential data, making LSTMs especially suitable for tasks requiring retention of information over extended periods [10]. LSTMs have been found to generally outperform traditional RNNs when dealing with complex time series data. While RNNs may struggle with long-range dependencies, LSTMs maintain performance even over extended sequences, which is critical for capturing temporal patterns in financial datasets and economic indicators [11]. Recent comparative analyses have confirmed that LSTMs can yield better accuracy in forecast tasks as opposed to simpler models [12], [13].

GRUs, as a specific type of Recurrent Neural Network (RNN), are adept at processing variable-length sequences. Their design includes gating mechanisms that help mitigate the vanishing gradient problem, allowing the model to retain relevant information over longer periods, which is essential in financial data characterized by temporal dependencies [14], [15]. The advantages of GRUs over traditional RNNs and even Long Short-Term Memory (LSTM) networks stem primarily from their simpler architecture, which can lead to faster training times and fewer parameters to optimize [16], [17]. Recent studies emphasize that GRUs can maintain performance while being comparatively efficient, making them suitable for high-frequency financial data analysis [18].

e-ISSN: 2614-8404
p-ISSN:2776-3234

Empirical evaluations have indicated that GRUs frequently outperform other architectures when applied to stock price and cryptocurrency predictions. For instance, a study focusing on cryptocurrency price predictions indicated that GRUs outperformed both LSTM and bidirectional LSTM models in terms of prediction accuracy for Bitcoin, Litecoin, and Ethereum, leading to lower Mean Absolute Percentage Error (MAPE) values [19]. Furthermore, research exploring stock price predictions demonstrated that GRU models effectively captured intraday price movements better than competing architectures [20].

E. Data Splitting

The dataset was partitioned into two subsets: a training set and a test set. Specifically, 80% of the observations were used for model development, while the remaining 20% were reserved for evaluating the model's generalization performance. To mitigate the risk of data leakage during training, the split was performed chronologically based on the time order of the observations, ensuring that the test set contains only data points occurring after the training period.

F. Model Performance Evaluation

Model performance was evaluated using prediction error metrics, namely the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). RMSE quantifies the magnitude of prediction errors by assigning a higher penalty to larger deviations, whereas MAE measures the average absolute difference between the observed and predicted values. These two metrics were employed to compare the performance of LSTM and GRU models for both short-term and long-term forecasting.

G. Research Flow

In general, the research workflow begins with collecting a dataset of global gold prices, followed by data preprocessing and the construction of sequential input data. The processed data are then used to train LSTM and GRU models under two different forecasting horizons. The final stage involves evaluating and analyzing the prediction results to identify the most suitable model for each time horizon. The overall workflow is illustrated in Figure 1.

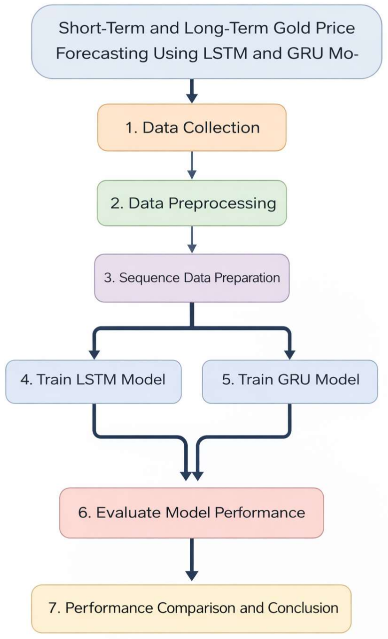

Figure 1. Research flowchart for short-term and long-term gold price forecasting using LSTM and GRU models

III. RESULTS AND DISCUSSION

This section presents the experimental results of global gold price forecasting using LSTM and GRU models under short-term and long-term horizons, and discusses the findings based on evaluation metrics and prediction visualizations.

A. Short-Term Forecasting Results

Short-term forecasting was conducted with a one-day-ahead horizon using historical daily gold price data. The LSTM and GRU models were trained on the training set and evaluated on the test set to assess each model's ability to track short-term fluctuations in gold prices. The evaluation results yielded the following prediction errors: the LSTM achieved an RMSE of 40.87 and an MAE of 31.12, while the GRU achieved an RMSE of 31.31 and an MAE of 23.26.

The lower RMSE and MAE values obtained by the GRU model indicate that GRU provides better performance for short-term gold price forecasting.

Figure 2. Short-term gold price prediction (1 day ahead) using the LSTM model.

e-ISSN: 2614-8404
p-ISSN:2776-3234

B. Long-Term Forecasting

Long-term forecasting was conducted using a 30-day ahead horizon to assess model stability in capturing gold price trends over an extended period. The evaluation results indicate that prediction errors increased substantially relative to short-term forecasts, a pattern commonly observed in time-series forecasting. The obtained performance metrics were as follows: the LSTM model achieved an RMSE of 113.76 and an MAE of 88.27, while the GRU model achieved an RMSE of 116.54 and an MAE of 88.42.

Although the differences are relatively small, the LSTM model exhibited a slightly lower RMSE than the GRU model, suggesting better stability for long-term prediction.

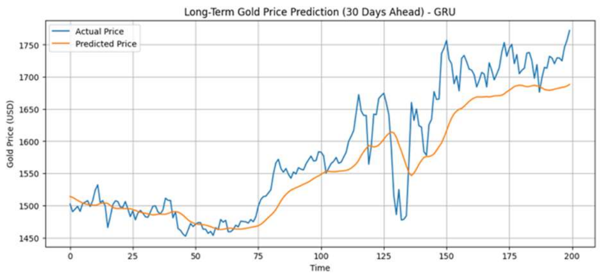

Figure 3. Long-term gold price prediction (30 days ahead) using the GRU model.

C. Comparison of Short-Term and Long-Term Forecasting Results

Table 1. Model Performance Comparison of LSTM and GRU

Horizon Prediksi	Model	RMSE	MAE
Short-Term (1 Day)	LSTM	40.87	31.12
Short-Term (1 Day)	GRU	**31.31**	**23.26**
Long-Term (30 Days)	LSTM	**113.76**	**88.27**
Long-Term (30 Days)	GRU	116.54	88.42

To provide a clearer quantitative assessment, the performance of the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models was compared using the root mean squared error (RMSE) and mean absolute error (MAE) across both forecasting horizons. The comparative results indicate that neither model consistently outperforms the other across all time horizons.

Based on the experimental findings, model performance appears to be strongly dependent on the forecasting horizon. For short-term forecasts, the GRU model yields lower prediction errors than the LSTM. This advantage may be attributable to the relatively simpler GRU architecture, which can be more responsive to short-term price fluctuations.

In contrast, for long-term forecasts, both models exhibit increased prediction errors, likely due to the accumulation of uncertainty over extended horizons. Nevertheless, LSTM demonstrates slightly more stable performance than GRU, suggesting a stronger capacity to capture long-range temporal dependencies.

Overall, these findings underscore that the selection of a gold price forecasting model should be aligned with the analytical objective. For short-term prediction tasks, GRU is recommended, whereas for longer-term forecasting, LSTM may represent a more stable option.

IV. CONCLUSION

This study aims to compare the predictive performance of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models in forecasting global gold prices over short-term and long-term horizons using daily historical data. The forecasting task is implemented under two horizon scenarios: one-day-ahead prediction for short-term forecasting and thirty-day-ahead prediction for long-term forecasting.

Experimental results indicate that model performance is strongly influenced by the forecasting horizon. For short-term forecasting, the GRU model outperforms LSTM, as reflected by lower RMSE and MAE values. This finding suggests that the comparatively simpler gating structure of GRU can respond more effectively to short-run fluctuations in gold prices.

In contrast, under long-term forecasting, both models exhibit a substantial increase in prediction error due to the accumulation of uncertainty inherent in multi-step time-series forecasting. Although the differences are relatively modest, LSTM achieves a slightly lower RMSE than GRU, indicating better stability in capturing long-range dependencies in gold price dynamics.

Overall, the results confirm that no single model is consistently superior across all forecasting horizons. Therefore, model selection for gold price forecasting should be aligned with the analytical objective. GRU is more suitable for short-term forecasting needs, whereas LSTM is more appropriate for long-term forecasting tasks that require stronger temporal stability.

REFERENCES

[1] Y. Wang and T. Lin, “A Novel Deterministic Probabilistic Forecasting Framework for Gold Price With a New Pandemic Index Based on Quantile Regression Deep Learning and Multi-Objective Optimization,” 2023. doi: 10.3390/math12010029.

[2] D. Makala and Z. Li, “Prediction of Gold Price With ARIMA and SVM,” 2021. doi: 10.1088/1742-6596/1767/1/012022.

[3] Z. Su, “Gold Price Forecast Based on ARIMA and ETS Models,” 2024. doi: 10.54254/2754-1169/2024.ga18972.

[4] S. Duman, S. Turnacıgil, E. Arık, and M. A. Aktaş, “The Role of International Variables in Predicting Gold Prices: Analysis With Machine Learning Algorithms,” 2025. doi: 10.17233/sosyoekonomi.2025.01.05.

[5] M. Madhavan, M. A. Sharafuddin, P. Piboonrungroj, and C. Yang, “Short-Term Forecasting for Airline Industry: The Case of

e-ISSN: 2614-8404
p-ISSN:2776-3234

Indian Air Passenger and Air Cargo," 2020. doi: 10.1177/0972150920923316.
[6] B. M. Pavlyshenko, "Forecasting of Non-Stationary Sales Time Series Using Deep Learning," 2022. doi: 10.48550/arxiv.2205.11636.
[7] P.-F. Dai, X. Xiong, and W.-X. Zhou, "The Role of Global Economic Policy Uncertainty in Predicting Crude Oil Futures Volatility: Evidence From a Two-Factor GARCH-MIDAS Model," 2020. doi: 10.48550/arxiv.2007.12838.
[8] G. Qiao, W. Cui, Y. Zhou, and C. Liang, "Volatility of Volatility and VIX Forecasting: New Evidence Based on Jumps, the Short-Term and Long-Term Volatility," 2024. doi: 10.1002/fut.22553.
[9] A. A. Salisu, R. Gupta, and R. Demirer, "Global Financial Cycle and the Predictability of Oil Market Volatility: Evidence From a GARCH-MIDAS Model," 2022. doi: 10.1016/j.eneco.2022.105934.
[10] T. M. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, "Meta-Learning in Neural Networks: A Survey," 2021. doi: 10.1109/tpami.2021.3079209.
[11] H. Xiao, "Enhanced Separation of Long-Term Memory From Short-Term Memory on Top of LSTM: Neural Network-Based Stock Index Forecasting," 2025. doi: 10.1371/journal.pone.0322737.
[12] K. Sako, B. N. Mpinda, and P. C. Rodrigues, "Neural Networks for Financial Time Series Forecasting," 2022. doi: 10.3390/e24050657.
[13] M. Elsaraiti and A. Merabet, "A Comparative Analysis of the ARIMA and LSTM Predictive Models and Their Effectiveness for Predicting Wind Speed," 2021. doi: 10.3390/en14206782.
[14] D. F. Godoy-Rojas *et al.*, "Attention-Based Deep Recurrent Neural Network to Forecast the Temperature Behavior of an Electric Arc Furnace Side-Wall," 2022. doi: 10.3390/s22041418.
[15] C. Li, L. Shen, and G. Qian, "Online Hybrid Neural Network for Stock Price Prediction: A Case Study of High-Frequency Stock Trading in the Chinese Market," 2023. doi: 10.3390/econometrics11020013.
[16] M. E. Karim, M. M. S. Maswood, S. Das, and A. G. Alharbi, "BHyPreC: A Novel Bi-LSTM Based Hybrid Recurrent Neural Network Model to Predict the CPU Workload of Cloud Virtual Machine," 2021. doi: 10.1109/access.2021.3113714.
[17] A. Lawi, H. Mesra, and S. Amir, "Implementation of Long Short-Term Memory and Gated Recurrent Units on Grouped Time-Series Data to Predict Stock Prices Accurately," 2022. doi: 10.1186/s40537-022-00597-0.
[18] M. Qin *et al.*, "Deep Learning for Multi-Timescales Pacific Decadal Oscillation Forecasting," 2022. doi: 10.1029/2021gl096479.
[19] M. J. Hamayel and A. Y. Owda, "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and Bi-LSTM Machine Learning Algorithms," 2021. doi: 10.3390/ai2040030.
[20] S. Tanwar, N. Patel, S. Patel, J. R. Patel, G. Sharma, and I. E. Davidson, "Deep Learning-Based Cryptocurrency Price Prediction Scheme With Inter-Dependent Relations," 2021. doi: 10.1109/access.2021.3117848.