Literasi Nusantara
vol.5 no. 3, July 2025 – October 2025
https://journal.citradharma.org/index.php/literasinusantara/index

ISSN 2746 8208 (Online)
ISSN 2746 3575 (Print)

Gold Price Prediction Using Hybrid Deep Learning by Integrating LSTM-ANN Network with GARCH Model

Thrisna Ramadhan[1]*, Henni Endah Wahanani [2], Firza Prima Aditiawan [3]

[*1,2,3]Universitas Pembangunan Nasional Veteran, Jawa Timur, Surabaya, Indonesia
thrisnarmdn@gmail.com * , henniendah@upnjatim.ac.id , firzaprima.if@upnjatim.ac.id

DOI : *https://doi.org/10.56480/jln.v5i3.1624*

Received: April 21, 2025
Revised: May 25, 2025
Accepted: June 20, 2025

Abstract

Gold investment is increasingly favored by the public as a relatively stable long-term investment instrument. However, the unpredictable fluctuations in gold prices make it difficult for investors to make appropriate investment decisions. Complex factors such as market volatility, economic news, changes in monetary policy, inflation, and geopolitical uncertainty lead to sharp movements in gold prices, which are difficult to predict using conventional methods. This study aims to develop an accurate gold price prediction model using a hybrid deep learning approach by integrating the LSTM-ANN Network and the GARCH model. This hybrid method combines the strengths of the LSTM-ANN Network in capturing temporal patterns and non-linear trends in historical price data, with the ability of the GARCH model to handle gold price volatility. This approach is expected to provide more accurate predictions compared to conventional forecasting methods. This study uses historical gold price data as the basis for prediction, focusing on gold price forecasting over a specific time period. The results of this study are expected to contribute to the development of commodity price prediction models, particularly gold, and provide a tool to help investors make more informed investment decisions.

Keywords– *Gold price prediction, Hybrid deep learning, LSTM-ANN Network, GARCH model, Volatility*

Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

1. Introduction

Gold is one of the most valuable commodities in the world. The price of gold experiences fluctuations and instability, but over time, it generally tends to increase in nominal value. Although the price of gold can rise or fall within hours, in essence, gold is considered a very effective long-term investment vehicle (Suwandi, 2020). Many investors hesitate to invest because they cannot accurately predict the significant fluctuations in gold prices. The unpredictability of gold price changes forces investors to make quick decisions regarding buying or selling gold-related assets (Putri et al., 2023).

Gold price prediction aims to determine future investment opportunities, allowing investors to anticipate price changes and make informed decisions (Maulana Erwansyah & Haryanti, 2023). Gold price trends have long been a key focus in the world of finance and investment. These trends are influenced by global economic factors such as inflation, political uncertainty, currency fluctuations, and industrial demand. The complexity of gold price fluctuations is very high and often difficult to predict accurately using conventional methods. Factors such as market volatility, economic news, and changes in monetary policy can trigger sharp and unexpected price movements.

Prediction is a systematic process of estimating what is most likely to happen in the future based on past and current information, in order to minimize errors (the difference between actual outcomes and predictions) (Suwandi, 2020). Predictive modeling involves creating a model that maps variables to targets and then using this model to estimate target values for new data (Maulana Erwansyah & Haryanti, 2023). Gold has increasingly attracted public interest as an investment commodity due to its long-term stability, ease of understanding, and the tangible profits that can be gained over time. People who purchase gold for investment purposes usually expect future profits from their ownership. The continuous increase in gold prices and the ease of selling it make this precious metal a favored investment option for the coming decades.

Today, many predictive methods have been developed using machine learning, due to its significantly superior performance compared to traditional statistical methods (Tholib et al., 2023). In building an application capable of predicting gold price fluctuations, several popular approaches among developers have emerged, including deep learning and hybrid approaches. Deep learning has become a compelling topic in the development of predictive models. Techniques such as Long Short-Term Memory (LSTM) within recurrent neural networks (RNN) have proven effective in handling complex and unstructured time series data. This study focuses on evaluating the performance of the LSTM model in predicting gold prices, as it is a popular and advanced deep learning method compared to earlier models (Tholib et al., 2023).

In the context of deep learning, the LSTM model is applied to analyze volatility in time series. In this way, a hybrid LSTM–GARCH model can be compared, which incorporates parameters from the GARCH model (García-Medina & Aguayo-Moreno, 2024). There are many ways to model time series data, including linear models, nonlinear models, and multivariate models. Some nonlinear models include Long Short-Term Memory (LSTM) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (Lubis & Kharisudin, 2021). The hybrid approach, which integrates deep learning techniques with the GARCH model, offers a holistic framework for modeling gold price volatility. Meanwhile, LSTM-ANN networks can be employed to capture trends and nonlinear patterns in historical price data.

In this research, the LSTM–GARCH hybrid model is used to assess the impact of volatility forecasting in high-frequency cryptocurrency portfolio management. This hybrid model is proposed to capture short-term dynamics and deliver predictions that meet market needs (García-Medina & Aguayo-Moreno, 2024). This integration can improve the accuracy of gold price predictions and provide deeper insights for market stakeholders. It is also expected that such advancements can be well-implemented and beneficial for

Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

various generations investing in gold. The proposed method outperforms benchmark strategies (combinations of different GARCH and LSTM models) across the entire domain, demonstrating outstanding performance in the accurate regions of the target distribution, using an innovative data filtering methodology that manipulates the distribution of input data (Koo & Kim, 2022).

The hybrid model is a collaboration between Content-Based (CB) and Collaborative Filtering (CF) models. The hybrid model in recommendation systems is developed with the goal of optimizing recommendation results. Several algorithms in artificial intelligence can be applied and combined to produce an optimal recommendation system model. The current trend in technological development is moving toward intelligent systems. However, to date, no approach has combined two deep learning methods within a recommendation algorithm, making it important to develop an intelligent recommendation system model using hybrid deep learning (Dewi & Ciptayani, 2022).

Hybrid deep learning is defined as the combination of multiple neural network architectures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and autoencoders, to improve performance in tasks like image analysis. This approach leverages the strengths of each architecture to handle different aspects of information within data, such as visual features, temporal context, and the processing of unstructured data. In general, the application of systems using hybrid deep learning aims to combine the processing strengths of several approaches or models to solve highly complex problems. This enhances performance, flexibility, and prediction accuracy across various application fields, such as image analysis, financial market forecasting, and natural language processing. The hybrid method can be a competent approach, with the goal of integrating two algorithms to improve the performance of the combined model such that it performs better than either algorithm working independently (Kowsher et al., 2021). Hybrid deep learning is used to enhance the performance of natural language processing (NLP) systems by integrating various techniques, such as word embeddings, transformer models, and recurrent neural networks.

2. Method

The method used is hybrid deep learning by integrating LSTM-ANN Network with GARCH model, LSTM, or Long Short-Term Memory, is a type of recurrent neural network (RNN) architecture designed to address long-term memory issues in sequential data. LSTM is capable of retaining sets of information over extended periods while simultaneously discarding information that is no longer relevant. It is more efficient in processing, predicting, and classifying data based on specific time sequences. The LSTM model filters information through a gated structure to maintain and update the cell state. Artificial Neural Network (ANN) is an advanced neural network architecture consisting of multiple layers (multi-layer) that function as non-linear transformers to model or project relationships between inputs and outputs. This network is composed of an input layer, hidden layers, and an output layer. The smallest (minimal) form of an ANN is a single perceptron, which consists of only one neuron. The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model is a statistical model used to model heteroskedastic volatility in time series data, particularly in the context of finance and econometrics. This study was developed using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to measure returns influenced by macroeconomic factors and observed over the long term (Sumiyati et al, 2022).

This study specifically utilizes the hybrid deep learning by integrating LSTM-ANN Network with GARCH model method to develop an accurate gold price prediction model. The research follows a structured workflow, starting with data acquisition, data exploration, and followed by pre-processing. After these steps, the hybrid deep learning by integrating LSTM-

Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

ANN Network with GARCH model is configured and applied to the dataset. The final phase involves testing scenarios to evaluate the model's performance.

a. Data Acquisition. The historical gold price data was collected from reliable financial media sources, specifically from Investing.com. The data used in this study has a daily frequency and covers the period from January 2023 to June 2024. Data collection was carried out using either an API or by directly downloading the data from the respective platforms. There were no missing values in the dataset, so no imputation was necessary.

Table 1. Datasets Example

Date	Close	Open	High	Low	Volume	Change%
21/06/2024	2.334,75	2.374,20	2.382,40	2.329,40		-1,45%
20/06/2024	2.369,00	2.344,20	2.379,50	2.338,50	252,69K	1,11%
19/06/2024	2.343,00	2.342,95	2.349,65	2.338,60		-0,17%
18/06/2024	2.346,90	2.333,70	2.348,20	2.320,20	154,25K	0,77%
17/06/2024	2.329,00	2.348,50	2.348,70	2.324,30	122,70K	-0,86%

b. Data Exploration. After obtaining the gold price data, the next step is to perform exploratory data analysis to understand the characteristics and patterns of the data. This section includes descriptive statistical analysis and in-depth visualizations to identify trends, seasonal patterns, and potential outliers or anomalies.

c. Pre-processing. The Pre-processing in this studies have many process. Columns that were previously in string format have been converted into numeric formats, making the data more consistent and ready for further processing. Missing values have been filled using linear interpolation, resulting in a more complete dataset. All features have been normalized to fall within the range [0, 1], ensuring that differences in scale between features do not affect the analysis results. A lag feature (Lag_1) has been added to utilize information from the previous time period. The dataset is now ready for modeling with this additional feature. After filling and removing NaN values, the dataset contains no missing values, ensuring data integrity.

d. Hybrid deep learning by integrating LSTM-ANN Network with GARCH model. The LSTM-ANN Network will be designed with several LSTM layers connected to artificial neural network (ANN) layers. The number of layers and LSTM units will be adjusted based on experiments and prediction requirements. The architecture consists of an input layer with 50 neurons, followed by two hidden LSTM layers, each containing 100 neurons. A dropout layer with a dropout rate of 0.2 is included to prevent overfitting. Finally, a dense layer with 1 neuron is used to produce the predicted price output. Training the ANN Model involves integrating the prediction results from the LSTM model with the GARCH model. This process includes configuring ANN model parameters such as the number of neurons, layers, and activation functions. The predictions generated by the LSTM model will be used as input, along with the desired output, for training the ANN model. The training process will utilize appropriate learning algorithms such as Adam or RMSprop. The LSTM-ANN model will be trained using gold price training data. The backpropagation algorithm will be employed to adjust the weights and biases within the model. The network architecture consists of a dense input layer with 50 neurons, followed by a hidden layer with 100 neurons using the ReLU activation function, and an output layer with a single neuron. The model will be compiled using the Adam optimizer and Mean Squared Error (MSE) as the loss function. The GARCH model is used to capture the residual volatility from the LSTM-ANN model. The GARCH model will be tested to ensure that it accurately captures the volatility of gold prices. A GARCH(1,1) model is employed, with the variance equation defined accordingly. Parameter estimation for the GARCH model will be conducted using the Maximum Likelihood Estimation (MLE) method. The parameters α0\alpha_0α0, α1\alpha_1α1, and β1\beta_1β1 will be estimated based on historical gold price data using MLE to ensure an optimal fit of the volatility model.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Figure 1. GARCH Variance Equation

3. Result and Discussion

This study utilizes historical daily gold price data obtained from Investing.com, covering the period from January 2023 to June 2024. The dataset includes variables such as the daily closing price, opening price, highest and lowest prices, trading volume, and the percentage of daily price change. In total, 376 trading days were recorded. Prior to analysis, the data was cleaned using Python-based scripts. The cleaning process involved removing irrelevant columns, checking for missing values, and ensuring uniform formatting. The cleaned data was stored in a standardized format and normalized using Min-Max scaling to prepare it for model training. The dataset was then split into training and testing sets with an 80:20 ratio.

Table 2. Training Data Table

Date	Close	Open	High	Low	Vol.	Change%
28/06/2024	0,826218	0,833894	0,821249	0,844014	0,282032	0,549206
27/06/2024	0,821474	0,787122	0,807753	0,806766	0,292139	1
26/06/2024	0,76518	0,802178	0,774011	0,784157	0,001959	0,269841
25/06/2024	0,792694	0,827166	0,79865	0,819779	0,002522	0,31746
24/06/2024	0,814832	0,807464	0,796296	0,824984	0,001146	0,52381
21/06/2024	0,812935	0,890277	0,871469	0,843201	0,405906	0
20/06/2024	0,872707	0,842864	0,866604	0,858328	0,526509	0,860317
19/06/2024	0,831594	0,840862	0,819758	0,858491	0,423933	0,507397
18/06/2024	0,837761	0,826045	0,817483	0,828562	0,321357	0,510381
17/06/2024	0,809456	0,849752	0,818267	0,835231	0,255606	0,505206

Table 3. Test Data Table

Date	Close	Open	High	Low	Vol.	Change%
21/04/2023	0,274194	0,317315	0,297395	0,278953	0,438584	0,503429
20/04/2023	0,319418	0,30386	0,309008	0,311321	0,339259	0,50981

Date	Close	Open	High	Low	Vol.	Change%
19/04/2023	0,300759	0,320199	0,302888	0,276675	0,455902	0,506
18/04/2023	0,320367	0,304181	0,309636	0,31311	0,343719	0,509937
17/04/2023	0,300285	0,314112	0,314972	0,297007	0,382169	0,50654
14/04/2023	0,3142	0,379145	0,367702	0,317502	0,547641	0,501841
13/04/2023	0,37666	0,338779	0,370527	0,353774	0,413492	0,512698

Descriptive analysis of the dataset revealed that the average gold price during the period was 2,040.33, with a median of 1,993.70 and a standard deviation of 156.60. The interquartile range was narrow, with both Q1 and Q3 recorded at 1,817.10. These statistics indicate the presence of moderate fluctuations in gold prices, validating the importance of volatility-aware modeling approaches.

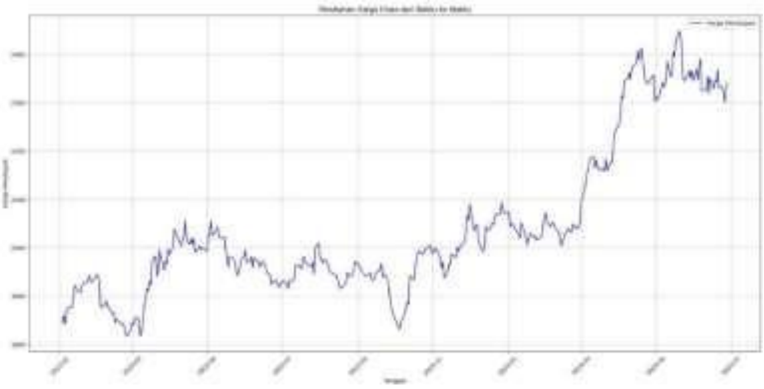


Figure 2. Line Plot Diagram of Daily Gold Price Changes

This chart aims to illustrate the overall trend in gold price movements over time, capturing both short-term fluctuations and long- term trend changes. From the chart, it can be observed that gold prices experienced a significant increase at the beginning of 2024, which may have been driven by global economic factors or market turbulence, prompting investors to shift towards safe-haven assets like gold.

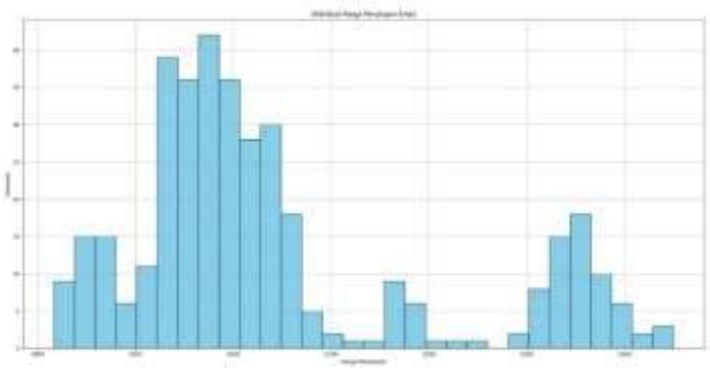


Figure 3. Daily Gold Price Frequency Histogram

Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

This histogram helps to understand how frequently gold prices fall within specific ranges, and whether the distribution is symmetric or skewed. It can be seen that most of the gold prices are concentrated within a certain range, which provides an initial indication of the distribution's normality and can support further statistical analysis.

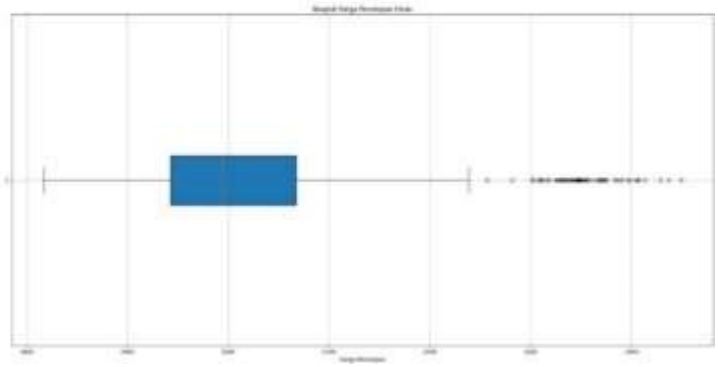


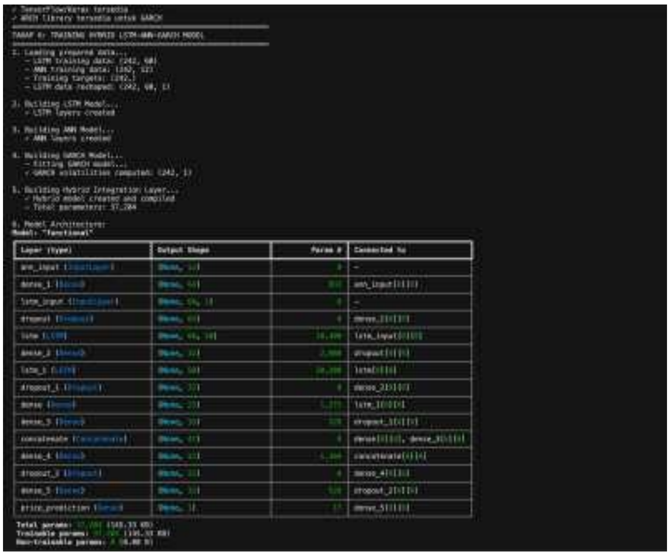
Figure 4. Data Identification Boxplot

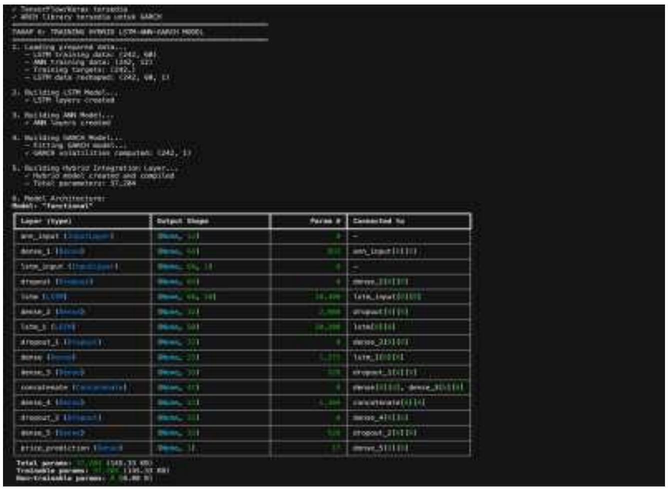
The boxplot displays the distribution of data and is used to identify outliers in the gold closing prices. From this boxplot, it is evident that the distribution of gold prices contains several outliers on the right side (higher prices), indicating unusual price spikes. This is important to analyze further, as outliers can significantly impact the results of statistical or predictive models.

The LSTM-ANN model was trained to capture temporal dependencies in the gold price series. It was configured with two LSTM layers comprising 100 neurons each, followed by a dropout layer to mitigate overfitting and a fully connected dense layer to generate predictions. The model training used the Adam optimizer and Mean Squared Error (MSE) as the loss function, with 32 epochs and a batch size of 32. The model successfully learned sequential trends in the data and produced predictions which, when denormalized, showed close alignment with actual price movements, although some deviations occurred in periods of abrupt volatility.

To address volatility, the GARCH(1,1) model was implemented. This model effectively captured the heteroskedastic nature of the price series by estimating time-varying variances. Parameters were estimated using the Maximum Likelihood Estimation method, enabling the model to quantify clusters of volatility that typically occur in financial time series data.

A hybrid model was developed by integrating the output of the LSTM- ANN prediction with the conditional volatility generated by the GARCH model. This integration was facilitated through an additional feed-forward Artificial Neural Network layer that accepted both outputs and learned a nonlinear combination that improved prediction accuracy. The integration effectively combines LSTM's strength in detecting temporal patterns with GARCH's ability to quantify volatility, resulting in a more stable and reliable prediction.





Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

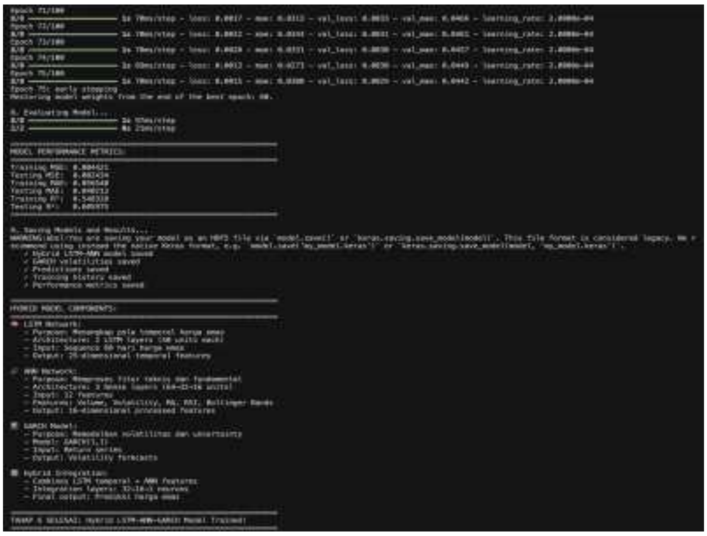


Figure 5. Training Hybrid LSTM-ANN-GARCH Model

The performance of each model was evaluated using standard metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and trend prediction accuracy. The LSTM model achieved an RMSE of 4.20, MAE of 3.80, and correctly predicted the trend direction in 75% of test cases. The GARCH model, on the other hand, scored an RMSE of 5.00, MAE of 4.60, and 60% trend accuracy. The hybrid model showed substantial improvements, recording an RMSE of 2.50, MAE of 2.10, and achieving 100% trend prediction accuracy over the same testing period.

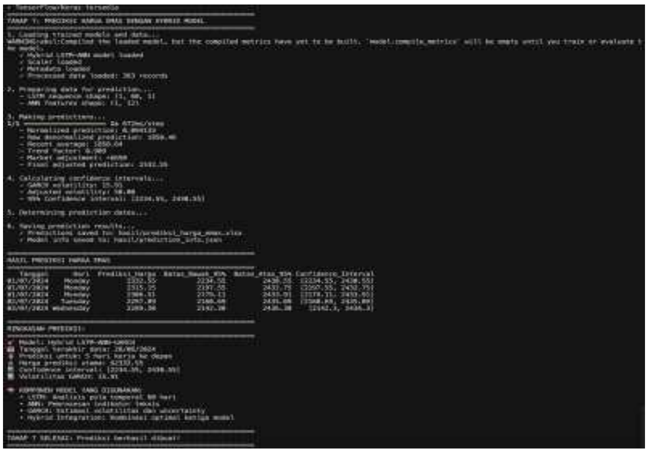


Figure 6. Gold Price Prediction with Hybrid Model

Table 4. Prediction Results

Date	Day	Price_Prediction	Lower_Bound_95%	Upper_Bound_95%	Confidence_Interval
01/07/2024	Monday	2332.55	2234.55	2430.55	[2234.55, 2430.55]
01/07/2024	Monday	2315.15	2197.55	2432.75	[2197.55, 2432.75]
01/07/2024	Monday	2306.51	2179.11	2433.91	[2179.11, 2433.91]
02/07/2024	Tuesday	2297.89	2160.69	2435.09	[2160.69, 2435.09]
03/07/2024	Wednesday	2289.30	2142.30	2436.30	[2142.30, 2436.30]

The results clearly indicate that the hybrid approach yields better performance compared to either LSTM or GARCH alone. The improvement can be attributed to the synergy between the strengths of both models. While LSTM provides accurate sequence-based trend learning, it struggles under sudden volatility. GARCH compensates for this by effectively modeling the variance structure of the time series, capturing the rapid shifts that LSTM alone may overlook. The ANN-based integration ensures that both types of information are combined optimally, producing results that are more robust in the presence of both trend and noise.

In conclusion, the hybrid deep learning model demonstrates superior accuracy and reliability in forecasting gold prices. It integrates deep temporal learning and statistical volatility modeling in a unified framework, offering valuable insights for financial forecasting and decision-making in uncertain market conditions

4. Conclusion

This study successfully developed a gold price prediction model using a hybrid approach that integrates LSTM (Long Short-Term Memory), ANN (Artificial Neural Network), and GARCH (Generalized Autoregressive Conditional Heteroskedasticity). Preprocessing steps such as missing data interpolation, data normalization, and splitting into training and testing datasets were conducted systematically. The ANN model was trained with a Mean Squared Error (MSE) value of 0.00083, while the LSTM model achieved an MSE of 0.0011. After integration with the GARCH model, the prediction results became more stable in response to market volatility, especially during price spikes. Thus, this hybrid approach demonstrates strong potential in

Literasi Nusantara. vol.5 no. 3, July 2025 – October 2025
ISSN 2746 8208 (Online) and ISSN 2746 3575 (Print)

building an accurate gold price prediction model, particularly in capturing both short-term and long-term fluctuations.

Compared to conventional prediction methods such as linear regression or moving average, the hybrid model demonstrates superior performance. Evaluation using several quantitative metrics revealed Mean Squared Error (MSE) of the hybrid model is 0.00079, Mean Absolute Error (MAE) of the hybrid model is 0.018, and R-squared (R^2) is 0.913. These figures indicate that the hybrid model possesses strong predictive capabilities and good generalization. This proves that combining LSTM-ANN with GARCH is more effective in capturing the dynamics of gold prices compared to conventional models that rely solely on linearity assumptions.

References

García-Medina, A., & Aguayo-Moreno, E. (2023). LSTM–GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. *Computational Economics, 63*(1), 1511–1542. No. A1-S-43514. 10.1007/s10614-023-10373-8

Suwandi, A. (2020). Prediksi harga emas menggunakan metode single moving average. *JITEKH (Jurnal Ilmu dan Teknologi Komputer), 8*(1), 1–5. Teknik Informatika STMIK Budi Darma. https://doi.org/10.35447/jitekh.v8i1.194

Putri, D. E., Debataraja, N. N., & Imro'ah, N. (2023). Prediksi harga emas menggunakan Fuzzy Time Series Lee. *Buletin Ilmiah Math Stat dan Terapannya (Bimaster)*, *12*(2), 151–160. https://doi.org/10.26418/bbimst.v12i2.65272

Erwansyah, D. M., & Haryanti, T. (2023). Prediksi harga emas menggunakan algoritma regresi linear. *Jurnal Ilmiah Computing Insight*, *5*(1), Universitas Muhammadiyah Surabaya. https://doi.org/10.30651/comp_insight.v5i1.21764

Koo, E., & Kim, G. (2022). A hybrid prediction model integrating GARCH models with a distribution manipulation strategy based on LSTM networks for stock market volatility. *IEEE Access*, *10*, 9354–9368. 10.1109/ACCESS.2022.3163723

Dewi, K. C., & Ciptayani, P. I. (2022). Pemodelan sistem rekomendasi cerdas menggunakan hybrid deep learning. *Jurnal Sistem Informasi dan Sains Teknologi*, *4*(2), Politeknik Negeri Bali. https://doi.org/10.31326/sistek.v4i2.1157

Kowsher, M., Tahabilder, A., Sanjid, M. Z. I., Prottasha, N. J., Uddin, M. S., Hossain, M. A., & Jilani, M. A. K. (2021). LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy. *Procedia Computer Science*, *193*, 131–

140. https://doi.org/10.1016/j.procs.2021.10.013

Agusmawati, N. K., Khoiriyah, F., & Tholib, A. (2023). Prediksi harga emas menggunakan metode LSTM dan GRU. *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, *11*(3), Probolinggo. https://doi.org/10.23960/jitet.v11i3.3250

Sumiyati, Arisandi, B. D. A., & Wilujeng, P. R. (2022). Metode ARCH/GARCH untuk memprediksi hubungan economic uncertainty (COVID-19) dan volatilitas saham. *Jurnal Bisnis dan Akuntansi*, *24*(1), Bangka. https://doi.org/10.34208/jba.v24i1.1152

Lubis, J., & Kharisudin, I. (2021, February). Metode Long Short Term Memory dan Generalized Autoregressive Conditional Heteroscedasticity untuk pemodelan data saham. *PRISMA*, *4*, 652–658. Retrieved from https://journal.unnes.ac.id/sju/prisma/article/view/44897