Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 9, Number 3, 2025
DOI: 10.30829/zero.v9i3.25962
Page: 1098 - 1108

Hybrid ARIMA-LSTM Model for Gold Price Forecasting at Pegadaian

[1] Erna Hudianti Pujiarini
Informatika, Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia

Article Info

Article history:

Accepted, 26 December 2025

Keywords:
ARIMA;
Forecasting;
Hybrid Model;
LSTM;
Pegadaian Gold Price.

ABSTRACT

Accurate forecasting of digital gold prices at PT Pegadaian is essential for managing volatility driven by macroeconomic factors, including exchange rates, inflation, and global gold prices. Conventional models present limitations: ARIMA effectively captures linear trends but fails to model non-linear patterns, whereas LSTM handles non-linearity but is prone to overfitting and poor generalization. This study proposes a hybrid ARIMA-LSTM model based on a quantitative time series approach. The analysis uses secondary data comprising daily digital gold prices from PT Pegadaian (2024–2025) and related macroeconomic indicators obtained from BPS and Bank Indonesia. Data are preprocessed to ensure stationarity and quality prior to modeling. The hybrid model combines linear forecasts from ARIMA with LSTM modeling of the resulting non-linear residuals. The hybrid model achieved MSE = 54,294.23, MAE = 113.56, and RMSE = 233.01 on the test set, representing reductions of approximately 75% in MSE, 54% in MAE, and 50% in RMSE relative to the standalone LSTM on testing data. The hybrid model outperforms both individual ARIMA and LSTM models in terms of generalization and accuracy. A primary limitation is the use of manual hyperparameter tuning; implementation of automated methods, such as grid search or Bayesian optimization, could further improve performance and robustness.

Corresponding Author:

Erna Hudianti Pujiarini,
Informatika,
Universitas Teknologi Digital Indonesia, Yogyakarta, Indonesia
Email: ernahudi@utdi.ac.id

1. INTRODUCTION

Gold has long been recognized as a safe haven asset against economic uncertainty, with its price influenced by factors such as global market trends, inflation, and geopolitical events [1],[2],[3],[4]. Pegadaian, a state-owned pawnshop company in Indonesia, plays a significant role in the gold market through its digital gold investment and pawning services. Accurate forecasting of digital gold prices at Pegadaian is essential for investors, financial institutions, and policymakers to make informed decisions amid market volatility.

Time series forecasting models, such as Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM), are commonly applied in financial predictions. ARIMA effectively models linear and stationary patterns but struggles with the nonlinear dynamics inherent in gold prices, often leading to issues like heteroskedasticity and non-normal residuals [13],[14],[15],[16],[26],[27],[28],[35],[36],[37]. In contrast, LSTM captures complex nonlinear relationships and temporal dependencies but is prone to overfitting, particularly with limited datasets, and requires substantial computational resources [17],[18],[29],[30],[31],[38],[39],[40].

Hybrid models combining ARIMA and LSTM have shown improved accuracy in various financial forecasting tasks by integrating linear and nonlinear components [19],[20],[21],[22],[41],[42],[43]. Related studies on gold pricing in Indonesia emphasize the role of local macroeconomic factors, such as USD/IDR exchange rates and inflation [44],[45]. Prior work on Pegadaian gold forecasting includes hybrid Double Exponential Smoothing with Neural Networks and Fuzzy Time Series methods [46],[47]. However, no study has applied a hybrid ARIMA-LSTM model specifically to Pegadaian's digital gold prices.

This research addresses this gap by developing a hybrid ARIMA-LSTM model to better capture both linear and nonlinear patterns in the data, thereby improving forecasting accuracy and generalization for Pegadaian's digital gold prices.

2. RESEARCH METHOD

This research was conducted using a quantitative approach with descriptive and predictive methods. The model employed is a hybrid of Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM). The ARIMA method is used to identify linear patterns, while LSTM is utilized to learn nonlinear patterns.

The research design for the time series model analysis for forecasting gold prices at Pegadaian using the hybrid ARIMA-LSTM model is presented in Figure 1.

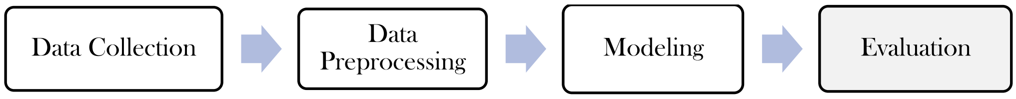

Figure 1. Research process stages

The dataset used in this study consists of secondary data obtained from official sources. Gold price data were collected from the official PT Pegadaian website for the period 2024 to 2025. Meanwhile, Indonesian macroeconomic data, including the IDR/USD exchange rate, Bank Indonesia (BI) reference interest rate, monthly inflation rate, and global gold prices, were sourced from the Central Statistics Agency (BPS) and Bank Indonesia. The inclusion of macroeconomic data aims to capture external influences.

The data preprocessing process in this study involves checking for duplicate data, inspecting missing values, and detecting outliers. For the ARIMA model, stationarity assessment and handling are required. Autocorrelation checks are conducted using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) to determine the ARIMA parameters (p, d, q). LSTM requires data to be scaled within the range [0,1] or [-1,1] for training stability.

Modeling is performed using a hybrid approach that integrates the ARIMA and LSTM models to capture both linear and nonlinear patterns in the gold price time series data. The first stage involves building the ARIMA model by identifying the order (p, d, q) through ACF and PACF analysis. The parameter d is determined based on the Augmented Dickey-Fuller (ADF) stationarity test, while parameters p and q are estimated using Maximum Likelihood Estimation (MLE). The second stage constructs the LSTM model with a three-layer architecture, an input layer receiving normalized gold price and macroeconomic data, a hidden layer, and an output layer for gold price predictions. The third stage combines both models, where ARIMA predictions capture linear components, and LSTM models the residuals to capture nonlinear patterns. The final result is the sum of ARIMA predictions and LSTM residual predictions. This hybrid model is evaluated using the same metrics and compared with individual models.

Model performance evaluation is conducted using three metrics, Root Mean Squared Error (RMSE) in the equation (1), Mean Absolute Error (MAE) in the equation (2), and Mean Absolute Percentage Error (MAPE) in the equation (3), with their formulas defined as follows [48]:

$$RMSE = \sqrt{\frac{1}{H}\sum_{t=1}^{H}(e_t)^2} \tag{1}$$

$$MAE = \frac{1}{H}\sum_{t=1}^{H}|e_t| \tag{2}$$

$$MAPE = \sum_{t=1}^{H}\frac{100|e_t|}{y_t} \tag{3}$$

where e_t is the error at time t and $y_t$data at time t.

2.1 ARIMA (p, d, q)

ARIMA (Autoregressive Integrated Moving Average) is a statistical model widely used for the analysis and forecasting of time series data [48]. ARIMA relies on autocorrelation (the correlation between y_t and y_{t-1}). This model combines three components:

1. The AR (Autoregressive) component assumes that the current value depends on previous values (lags) in the time series. The order p indicates the number of lags used. Mathematically, this can be expressed expressed in the equation (4).

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \tag{4}$$

2. The I (Integrated / Differencing) component is used to make the data stationary, ensuring a constant mean and variance over time by calculating the difference (differencing) between the current and previous observations. The order d indicates how many times differencing is performed. Mathematically, this can be expressed expressed in the equation (5).

$$y_t = y_t - y_{t-d} \tag{5}$$

3. The MA (Moving Average) component is used to model the error (residual) as a linear combination of past errors. The order q indicates the number of error lags used. Mathematically, this can be expressed expressed in the equation (6).

$$y_t = c + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_q y_{t-q} + \epsilon_t \tag{6}$$

2.2 Algoritma Long Short Term Memory

Long Short-Term Memory (LSTM) is an artificial neural network algorithm that belongs to the group of Recurrent Neural Networks (RNN). LSTM is designed to address challenges in modeling time series data or sequential data, which exhibit long-term dependencies [49]. LSTM utilizes a gating mechanism that enables the network to selectively remember or forget information over extended periods.

LSTM consists of recurrent units that have a cell state and three types of gates:

1. The Forget Gate is used to determine which information from the previous cell state will be forgotten. The Forget Gate employs a sigmoid function to generate values between 0 and 1 based on the current input and the previous hidden output.
2. The Input Gate is used to control the new information to be added to the cell state. This process involves using a sigmoid function to determine which parts of the input are relevant, followed by a tanh function to generate candidate updates for the cell state.
3. The Output Gate is used to determine which part of the cell state will be output at a specific time step. This output also serves as the hidden input for the next time step. The cell state functions as the network's memory, enabling LSTM to store information from previous time steps and use it to predict the next time step. This mechanism allows LSTM to capture nonlinear patterns and longterm dependencies in the data.

The following are the steps of the LSTM algorithm for a single time step (t_s):

1. For each time step (t) receiving input at the time (x_t), the previous hidden state, which is the output from the previous time step, represents the information that has been processed (h_{t-1}), the previous cell state is a longterm memory of the previous time step (C_{t-1}).
2. Calculate the forget gate value to determine the information to be forgotten using the equation (7).

$$f_t = \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \tag{7}$$

3. Determine the input gate to control the new information that will be added to the cell state, by determining the relevant information and calculating the update candidates using the equation (8) and (9).

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i) \tag{8}$$

$$\tilde{C}_t = tanh(W_C . [h_{t-1}, x_t] + b_C) \tag{9}$$

4. Updating Cell state by combining forgotten information and new information using the equation (10).

$$C_t = f_t . C_{t-1} + i_t \tilde{C}_t \tag{10}$$

5. Determining the output gate, which is part of the cell state that will be output. This is done by calculating the output gate and hidden state values using the equation (11) and (12).

$$o_t = \sigma(W_o . [h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t . \tanh{(C_t)} \tag{12}$$

6. Forecasting hidden state (h_{t-1}), then passed to the fully connected layer with a linear activation function to produce predictions using the equation (13).

$$\hat{y}_t = W_y h_t + b_t \tag{13}$$

3. RESULT AND ANALYSIS

The following section presents the results and discusses each step of this study.

3.1 Dataset

The dataset used in this study is secondary and was obtained from official sources. Gold price data was taken from the official website of PT Pegadaian for the period 2024 to 2025. Meanwhile, Indonesian macroeconomic data, including the IDR/USD exchange rate, Bank Indonesia's benchmark interest rate (BI Rate), monthly inflation rate, and global gold price, were obtained from the Central Statistics Agency (BPS) and Bank Indonesia. Dataset gold price, inflation, interest rate, dollar exchange rate as shown in the figure 2.

```
          Date  gold_price  Inflation  Interest_rate  Dollar_exchange_rate
0   2024-05-11       12380       2.84           6.25                 16453
1   2024-05-12       12590       2.84           6.25                 16453
2   2024-05-13       12590       2.84           6.25                 16453
3   2024-05-14       12590       2.84           6.25                 16012
4   2024-05-15       12540       2.84           6.25                 15914
..         ...         ...        ...            ...                   ...
357 2025-05-03       18510       1.17           5.75                 16453
358 2025-05-04       18410       1.17           5.75                 16460
359 2025-05-05       18410       1.17           5.75                 16968
360 2025-05-06       18630       1.17           5.75                 16432
361 2025-05-07       18890       1.17           5.75                 16492

[362 rows x 5 columns]
```

Figure 2. Dataset Gold Price, Inflation, Interest rate, Dollar Exchange Rate

3.2 Preprocessing

The data preprocessing process in the research was carried out by checking duplicate data, checking missing values, and detecting outliers as shown in Figure 3.

```
Number of rows after removing duplicates: 334

Missing Values:
gold_price              0
Inflation               0
Interest_rate           0
Dollar_exchange_rate    0
dtype: int64

Outliers Detected:
            gold_price  Inflation  Interest_rate  Dollar_exchange_rate
Date
2025-04-22       19710       1.17           5.75                 16858
```

Figure 3. Results of checking duplicate data, missing values, outliers

The results of data preprocessing showed that the number of rows after removing duplicates was 334 rows, down from the initial 363. There were no missing values in the Harga_emas_digital, Inflation, Interest Rate, and Exchange Rate features. One outlier was detected on April 22, 2025, with a digital gold price of 19,710, significantly higher than the average price in the dataset (ranging from 12,380 to 19,710).

ARIMA requires stationary data, so stationarity checks and handling are necessary. The Augmented Dickey-Fuller (ADF) test was performed to check for stationarity. Figure 4 shows an ADF p-value of 0.9989, indicating that the Harga_emas_digital data is non-stationary. After differencing, the value of ADF p-value = $2{,}26.10^{-06} < 0{,}05$, this means that the digital gold price data is stationary.

```
ADF p-value: 0.998929145922562
ADF p-value after differencing: 2.268464448453526e-06
```

Figure 4. ADF Gold Price

Autocorrelation checks are performed using ACF/PACF to determine the ARIMA parameters (p, d, q). Figure 5 shows the ACF plot showing a sharp cut-off at lag 1, so the order of q = 1. The PACF plot shows a sharp cut-off at lag 1, so the order of p = 1.

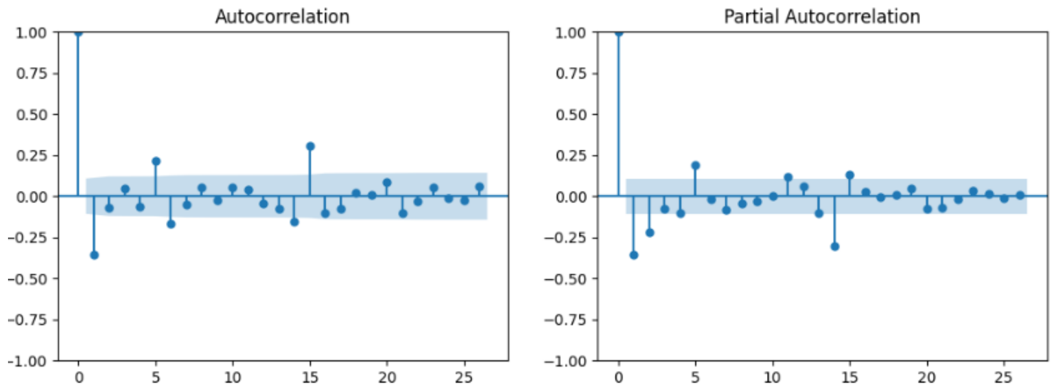

Figure 5. Autocorrelation and partial Autocorrelation Diagram

LSTM requires data in the range [0,1] or [-1,1] for training stability. Outliers such as the value of 19,710 can affect the scale, so normalization is performed with MinMaxScaler, as shown in Figure 6.

```
            gold_price  Inflation  Interest_rate  Dollar_exchange_rate
Date
2024-05-11    0.000000   1.000000            1.0              0.602090
2024-05-12    0.028649   1.000000            1.0              0.602090
2024-05-14    0.028649   1.000000            1.0              0.410100
2024-05-15    0.021828   1.000000            1.0              0.367436
2024-05-16    0.031378   1.000000            1.0              0.370048
...                ...        ...            ...                   ...
2025-05-02    0.836289   0.430034            0.0              0.602090
2025-05-04    0.822647   0.430034            0.0              0.605137
2025-05-05    0.822647   0.430034            0.0              0.826295
2025-05-06    0.852660   0.430034            0.0              0.592947
2025-05-07    0.888131   0.430034            0.0              0.619068

[334 rows x 4 columns]
```

Figure 6. Normalization

Feature engineering is done by adding 7-day lag features and moving average (MA-7) to capture temporal dependencies, then NaN data is removed so that the dataset appears in Figure 7.

```
           Price_lag_1  Price_lag_2  Price_lag_3  Price_lag_4  Price_lag_5  \
Date
2024-05-19    0.032742     0.043656     0.031378     0.021828     0.028649
2024-05-20    0.050477     0.032742     0.043656     0.031378     0.021828
2024-05-21    0.050477     0.050477     0.032742     0.043656     0.031378
2024-05-22    0.070941     0.050477     0.050477     0.032742     0.043656
2024-05-23    0.060027     0.070941     0.050477     0.050477     0.032742
...                ...          ...          ...          ...          ...
2025-05-02    0.860846     0.896317     0.903138     0.899045     0.903138
2025-05-04    0.836289     0.860846     0.896317     0.903138     0.899045
2025-05-05    0.822647     0.836289     0.860846     0.896317     0.903138
2025-05-06    0.822647     0.822647     0.836289     0.860846     0.896317
2025-05-07    0.852660     0.822647     0.822647     0.836289     0.860846

           Price_lag_6  Price_lag_7      MA_7
Date
2024-05-19    0.028649     0.000000  0.033912
2024-05-20    0.028649     0.028649  0.037030
2024-05-21    0.021828     0.028649  0.043072
2024-05-22    0.031378     0.021828  0.048529
2024-05-23    0.043656     0.031378  0.052816
...                ...          ...       ...
2025-05-02    0.903138     0.912688  0.885987
2025-05-04    0.903138     0.903138  0.874488
2025-05-05    0.899045     0.903138  0.862990
2025-05-06    0.903138     0.899045  0.856363
2025-05-07    0.896317     0.903138  0.854219
```

Figure 7. Result Feature engineering

Then the data was preprocessed as many as 327 divided into 80% for the train as many as 261 and 20% for the test as many as 66.

3.3 Modeling

This study uses a quantitative approach to time series analysis to forecast gold prices at Pegadaian. The method used is a hybrid model that integrates ARIMA for the linear component and LSTM for the non-linear component of digital gold price data.

The ARIMA model is used to model the linear component of the gold price time series.

.

```
Train size: 250, Test size: 63
                               SARIMAX Results
==============================================================================
Dep. Variable:             gold_price   No. Observations:                  250
Model:                 ARIMA(1, 1, 1)   Log Likelihood                 672.837
Date:                Tue, 23 Sep 2025   AIC                          -1339.674
Time:                        04:25:31   BIC                          -1329.121
Sample:                             0   HQIC                         -1335.426
                                - 250
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.1597      0.124      1.288      0.198      -0.083       0.403
ma.L1         -0.4722      0.129     -3.658      0.000      -0.725      -0.219
sigma2         0.0003   1.73e-05     15.193      0.000       0.000       0.000
===================================================================================
Ljung-Box (L1) (Q):                   0.14   Jarque-Bera (JB):               237.42
Prob(Q):                              0.70   Prob(JB):                         0.00
Heteroskedasticity (H):               2.18   Skew:                            -0.13
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.78
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Figure 8. Fitting ARIMA (1,1,1)

Interpretation of the output in Figure 8 is as follows:

1. Value Prob (Q) = 0.72 > 0,05 this means that the residuals do not have autocorrelation so that the model has captured the temporal pattern well.
2. Value Prob (JB) = 0.00 this means that the residuals are not normally distributed, possibly because there are outliers.
3. Value Heteroskedastisitas (H) = 0.00 this means the residual variance is not constant, there is a volatility cluster.

Overall, the model satisfies stationarity (differencing works) but has limitations in capturing volatility and external influences.

The LSTM model is used to capture non-linear patterns and long-term dependencies in the data. The LSTM model was trained for 50 epochs to predict digital gold prices based on the processed data. Table 1 shows the development of the loss values for the training and validation data (val_loss) for each epoch.

Table 1. Training results 50 epoch

Epoch	Loss	Val_Loss	Time of each Step	Information
1	0.0343	0.0136	104ms	Early, high losses
5	0.0013	0.0037	38ms	Rapid decline
10	0.00057674	0.00071406	40ms	Stabilization begins
25	0.00052891	0.0014	61ms	Fluktuasi val_loss
36	0.00051756	0.00048483	38ms	Val_loss lowest
50	0.00049954	0.00065552	38ms	Convergent, stable

The training process at epoch 1 shows a high initial loss (0.0343) and validation loss (0.0136), indicating that the model is still far from optimal at the start. Subsequently, the loss decreases sharply from 0.0343 at epoch 1 to 0.0013 at epoch 5, demonstrating that the model quickly learns the basic patterns in the data. After epoch 7, the loss stabilizes in the range of 4e-04 to 9e-04, with small fluctuations. The validation loss is sometimes higher than the training loss, indicating mild overfitting. By epoch 36, the model converges with stable performance, although no significant further reduction is observed.

The low loss and validation loss values suggest that the model is effective at minimizing errors on data normalized to a 0-1 scale. Fluctuations in validation loss, such as at epoch 25 (0.0014), may be caused by variations in the validation data or insufficient data for generalization. A dropout rate of 0.2 helps prevent severe overfitting, as the difference between loss and validation loss is not substantial.

The time per epoch varies from 0s 36ms/step to 1s 102ms/step, with earlier epochs being slower due to model initialization.

The hybrid ARIMA and LSTM model is developed by first fitting the ARIMA model to the training data and calculating its residuals. These residuals are then fed into the LSTM model and retrained. The residuals are normalized to model nonlinear patterns. The training process for the residual LSTM follows the same structure as the initial LSTM, with a dropout rate of 0.2 and 50 epochs. The training results for the residual LSTM show a similar pattern: a high initial loss of around 0.2072, a rapid decrease in early epochs, and stabilization in the range of 0.0070–0.0100 for training loss, with validation loss being higher, in the range of 0.0289–0.0300, indicating mild overfitting. The residual predictions from the LSTM are then added to the ARIMA predictions to form the hybrid prediction.

.

3.4 Evalution

In this research, the performance of the time series forecasting models was evaluated by comparing error metrics (MSE, MAE, and RMSE) between the standalone LSTM and the hybrid ARIMA-LSTM model, as presented in Figure 9 and summarized in Table 2.

```
Performance LSTM pada Data Training:
MSE: 13256.94, MAE: 80.61, RMSE: 115.14

Performance LSTM pada Data Testing:
MSE: 218443.40, MAE: 245.33, RMSE: 467.38
12/12 ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 0s 5ms/step

Performance Gabungan ARIMA-LSTM:
MSE: 54294.23, MAE: 113.56, RMSE: 233.01
```

Figure 9. Output the metrics MSE, MAE, MSE

Table 2. Model Performance Comparison Summary

Model	Dataset	MSE	MAE	RMSE
LSTM	Training	13256.94	80.61	115.14
LSTM	Testing	218443.40	245.33	467.38
Hybrid ARIMA-LSTM	Combined	54294.23	113.56	233.01

The standalone LSTM model exhibited strong performance on the training data, achieving MSE = 13,256.94, MAE = 80.61, and RMSE = 115.14, indicating effective capture of patterns during training. However, on the testing data, performance degraded substantially (MSE = 218,443.40, MAE = 245.33, RMSE = 467.38), reflecting classic overfitting.

The hybrid ARIMA-LSTM model decomposes the time series y_t into a linear component L_t modeled by ARIMA and a non-linear residual component N_t modeled by LSTM using the equation (14).

$$y_t = L_t + N_t \tag{14}$$

ARIMA efficiently captures the dominant linear trend with low bias and low variance for stationary components. By removing L_t, the LSTM is applied only to the residual N, which typically has lower amplitude and reduced systematic bias. This prevents the LSTM from overfitting to the strong linear signal, thereby reducing overall variance while maintaining low bias in modeling complex non-linear patterns. Consequently, the error structure is improved with residual autocorrelation and heteroskedasticity are reduced, leading to more stable and normally distributed forecast errors. The hybrid model achieved MSE = 54,294.23, MAE = 113.56, and RMSE = 233.01 on the test set, representing reductions of approximately 75% in MSE, 54% in MAE, and 50% in RMSE relative to the standalone LSTM on testing data.

The models incorporated macroeconomic inputs (USD/IDR exchange rate, inflation rate, and global gold price) as exogenous variables. However, this study did not conduct formal feature importance analysis (e.g., permutation importance or SHAP values) or ablation studies to quantify the individual contribution of each macroeconomic variable to forecast accuracy. This represents a limitation, as such analyses could reveal which factors most strongly influence digital gold price dynamics and potentially guide further model refinement.

Regarding model robustness, the hybrid approach showed reasonable stability across tested hyperparameter ranges. However, performance was sensitive to shorter training windows (<60 days) and to extreme macroeconomic shocks not adequately represented in the 2024–2025 training period. Sudden events, such as sharp rupiah depreciation or global crises, could degrade generalization if outside the historical distribution.

The findings are generalizable, within reasonable limits, beyond Pegadaian's digital gold prices to other financial time series in emerging markets that exhibit mixed linear and non-linear characteristics driven by macroeconomic factors. Examples include Antam gold prices, commodity prices (e.g., palm oil, nickel), or equity indices in Indonesia and similar economies (e.g., India, Brazil, South Africa). The core principle—decomposing linear trends with statistical models before applying deep learning to residuals—can improve generalization in volatile markets influenced by both global and local factors. Nevertheless, direct application requires retraining and validation on target-specific data due to differences in market microstructure and regulatory contexts.

4. CONCLUSION

The primary objective of this study was to develop and evaluate a hybrid ARIMA-LSTM model for accurate forecasting of digital gold prices at PT Pegadaian, overcoming the limitations of standalone ARIMA in handling non-linear patterns and standalone LSTM in generalization due to overfitting. This objective has been successfully achieved. The hybrid approach, which combines ARIMA's linear trend modeling with LSTM's processing of non-linear residuals, demonstrates significantly improved forecasting accuracy and generalization on unseen test data compared to the standalone LSTM baseline, with substantial reductions in error metrics (MSE, MAE, and RMSE).

In practical terms, the model provides substantial value for real-time decision-making at PT Pegadaian. It enables dynamic adjustment of digital gold pricing in response to market movements, supports proactive inventory management and hedging against volatility, and facilitates more informed risk assessment in gold investment and pawning services. By integrating macroeconomic influences such as USD/IDR exchange rates and inflation, the model allows Pegadaian's management and customers to anticipate price fluctuations more reliably, optimize investment timing, reduce exposure to adverse market shifts, and ultimately strengthen operational efficiency and customer trust in a highly volatile commodity market. Despite these advances, limitations remain, including mild overfitting in the LSTM component and reliance on manual hyperparameter tuning. Future research could address these through automated optimization techniques, expanded datasets, or additional regularization. From an applied mathematics perspective, future work could include theoretical analysis of error bounds in hybrid residual modeling or integration of GARCH-type structures into the linear component to explicitly account for time-varying volatility. Such developments would further enhance the model's robustness and applicability to financial time series forecasting in emerging markets.

5. REFERENCES

[1] V. M. Ngo, P. V. Nguyen, and Y. H. Hai, "The impacts of geopolitical risks on gold, oil and financial reserve management," *Resour. Policy*, vol. 90, p. 104688, Mar. 2024, doi: 10.1016/j.resourpol.2024.104688.

[2] A. A. Salisu, I. D. Raheem, and X. V. Vo, "Assessing the safe haven property of the gold market during COVID-19 pandemic," *Int. Rev. Financ. Anal.*, vol. 74, p. 101666, Mar. 2021, doi: 10.1016/j.irfa.2021.101666.

[3] J. Li and Y. Huang, "Dynamic spillovers of geopolitical risks and gold prices: New evidence from 18 emerging economies," *Resour. Policy*, vol. 70, p. 101938, Mar. 2021, doi: 10.1016/j.resourpol.2020.101938.

[4] Özgür Ergül and Tuba Karakaş, "Analysis of the relationship of gold prices with inflation and bitcoin in the post-tapering period," *Borsa Istambul review*, vol. 24, issue 4, p. 643-868, July, 2024, https://doi.org/10.1016/j.bir.2024.04.009.

[5] A. Sulityowati , B. S. Pitoyo, & M. Handayani, "Pengaruh Tingkat Inflasi dan Nilai Tukar Rupiah Terhadap Penyaluran Kredit Rahn PT. Pegadaian Syariah Periode Tahun 2011-2022," *Jurnal Ilmiah Ekonomi Islam,* vol. 9, no. 2, pp. 2510-2514, 2023, DOI: https://doi.org/10.29040/jiei.v9i2.9649.

[6] A. B. Kusuma, "The Effect of Inflation Rate and Gold Price on Gold Pawn Financing at Bank Syariah Indonesia," *West Sci. J. Econ. and Entrepreneurship*, vol. 3, no. 1, pp. 45–60, Feb. 2024, doi: https://doi.org/10.58812/wsjee.v2i01.542.

[7] M.F.M. Nasution, N. Aslami, Kusmilawaty, "Digital Innovation Strategy to Foster Customer Satisfaction Growth at Pegadaian Syariah," *Indonesian Journal of Innovation Studies,* vol. 26, no. 4, October, p. 1-19, 2025, doi: 10.12345/ijins.2025.4.2.1592.

[8] K. Kashif and R. Ślepaczuk, "LSTM-ARIMA as a hybrid approach in algorithmic investment strategies," *Knowledge-Based Systems*, vol. 320, art. no. 113563, 2025, doi: 10.1016/j.knosys.2025.113563.

[9] J. Zhang, H. Liu, W. Bai, and X. Li, "A hybrid approach of wavelet transform, ARIMA and LSTM model for the share price index futures forecasting," *The North American Journal of Economics and Finance*, vol. 69, art. no. 102022, 2024, doi: 10.1016/j.najef.2023.102022.

[10] C. Zhang, N. N. A. Sjarif, and R. Ibrahim, "Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022," WIREs Data Mining and Knowledge Discovery, vol. 14, no. 1, art. no. e1519, 2024, doi: 10.1002/widm.1519.

[11] M. Elamine and A. Ben Abdallah, "Predicting cryptocurrency prices with a hybrid ARIMA and LSTM model," *Journal of Telecommunications and the Digital Economy*, vol. 13, no. 1, pp. 98–117, 2025, doi: 10.18080/jtde.v13n1.1052.

[12] F. Pokou, J. S. Kamdem, and F. Benhmad, "Hybridization of ARIMA with learning models for forecasting of stock market time series," *Computational Economics*, vol. 63, no. 4, pp. 1349–1399, Apr. 2024, doi: 10.1007/s10614-023-10499-9.

[13] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of ARIMA and artificial neural networks models for stock price prediction,"*Journal of Applied Mathematics*., vol. 7, no. 1, pp1-7. 38–46, 2014, DOI:10.1155/2014/614342.

[14] X. Wang, Y. Ma, and W. Li, "The prediction of gold futures prices at the Shanghai Futures Exchange based on the MEEMD-CS-Elman model," *SAGE Open*, vol. 11, no. 1, pp. 1–12, Jan.–Mar. 2021, doi: 10.1177/21582440211001866.

[15] A. Prasetyo, N. Nurdin, H. A. K. Aidilof, "Comparison of Triple Exponential Smoothing and ARIMA in Predicting Cryptocurrency Prices," *International Journal of Engineering, Science and InformationTechnology*, vol. 4, no. 4, pp. 63-71, 2024, https://doi.org/10.52088/ijesty.v4i4.577.

[16] Y. Liu, Y. Zhang, and J. Wang, "International gold price forecast based on CEEMDAN and Stacking ensemble learning," *Discrete Dyn. Nature Soc.*, vol. 2022, art. no. 1511479, pp. 1–15, 2022, doi: 10.1155/2022/1511479.

[17] K. Alkhatib, H. Khazaleh, H. A. Alkhazaleh, A. R. Alsoud and L. Abualigah, "A New Stock Price Forecasting Method Using Active Deep Learning Approach," *Journal of Open Innovation: Technology, Market, and Complexity,* vol. 8, issue. 2, pp. 7123–7140, June. 2022, https://doi.org/10.3390/joitmc8020096.

[18] Bao W, Yue J, Rao Y, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory," *PLoS ONE*, vol. 12, no. 7, p. 1-14, Jul. 2017, doi :10.1371/journal.pone.0180944

[19] S. Saleti, L. Y. Panchumarthi, Y. R. Kallam, V. K. Reddy, and G. Supriya, "Enhancing forecasting accuracy with a moving average-integrated hybrid ARIMA-LSTM model," *SN Comput. Sci.*, vol. 5, no. 5, art. no. 704, Aug. 2024, doi: 10.1007/s42979-024-03060-4.

[20] X. Xiao, J. Xiao, and Y. Xiao, "Research on stock price time series prediction based on deep learning and autoregressive integrated moving average," Sci. Program., vol. 2022, art. no. 4758698, pp. 1–11, Apr. 2022, doi: 10.1155/2022/4758698.

[21] D. Li and Y. Li, "On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning," J. Forecast., vol. 39, no. 6, pp. 953–966, Sep. 2020, doi: 10.1002/for.2677.

[22] M. S. Kulkarni1, V. Bharathi, A Perdana, D. Kilari, "A quest for context-specific stock price prediction: A comparison between time series, machine learning and deep learning models," SN Comput. Sci., vol. 6, no. 335, p. 1-24, 2025, doi: 10.1007/s42979-025-03848-y.

[23] S. Setyowibowo, M. As'ad, Sujito, E. Farida, "Forecasting of Daily Gold Price using ARIMA-GARCH Hybrid Model," *Jurnal Ekonomi Pembangunan*, vol. 19, no. 2, December. 2021, doi: 10.1016/j.resourpol.2022.102847.
[24] C. Gil, A. Avishay, "Forecasting gold price using machine learning methodologies," *Chaos, Solitons and Fractals* , vol. 175, article id. 114079, Okt. 2023, doi : 10.1016/j.chaos.2023.114079.
[25] H. M. Zangana, S. R. Obeyd, "Deep Learning-based Gold Price Prediction: A Novel Approach using Time Series Analysis," *Sistemasi:Jurnal Sistem Informasi,* vol 13, no. 6, 2024, https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/view/4651.
[26] S. Galeshchuk and S. Mukherjee, "Deep networks for predicting direction of change in foreign exchange rates," *Intell. Syst. Account. Finance Manage.*, vol. 27, no. 2, pp. 100–110, Apr.–Jun. 2020, doi:10.1002/isaf.1404.
[27] T. C. Chiang and Z. Dai, "Time Series Analysis And Forecasting With ARIMA ModelGold Derivatives At MCX," *Educational Administration: Theory and Practice*, vol. 29(3), p. 357-365, 2023, doi: 10.54254/2754-1169/2025.LH23823.
[28] P. Garg1, R. Sharma, "Evaluation of ARIMA models for financial time series forecasting," *Int. J. Comput. Appl.*, vol. 43, no. 5, pp. 468–475, May 2021, doi: 10.53555/kuey.v29i3.4246.
[29] M. S. Amin, Y. K. Wong, and M. A. Razzaque, "Stock-Price-Prediction-Using LSTM ARIMA," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7111–7122, Oct. 2022, doi: 10.1016/j.jksuci.2021.08.004.
[30] H. Liu, Z. Liu, and X. Zhang, "Stock market forecasting using a multi-model architecture with XGBoost and LSTM," *Journal of Development Economics and Management Research Studies*, vol. 9(11), pp. 55-66, January 2022, doi:10.53422/JDMS.2022.91106.
[31] T. Chaudhary, S. Minirani , "Machine Learning and Time Series Approaches for Stock Price Prediction: A Case Study of Indian Banks," *Procedia Computer Science*, vol. 260, pp. 626-634, 2025, doi: 10.1016/j.procs.2025.03.241.
[32] M. Khani, S. Vahidnia, and A. Abbasi, "A deep learning based methods for forecasting gold price with respect to pandemics," *SN Comput. Sci.*, vol. 2, no. 4, art. no. 293, Jul. 2021, doi: 10.1007/s42979-021-00724-3.
[33] R. Ghule, A. Gadhave, "Gold Price Prediction using Machine Learning," *International Journal of Scientific Research in Engineering and Management,* vol. 6(6), June 2022, https://www.researchgate.net/profile/Jyoti-Kharade3/publication/362491642_Gold_Price_Prediction_using_Machine_Learning/links/62ecf7ee505511283e9102ce/Gold-Price-Prediction-using-Machine-Learning.pdf.
[34] B. A. Memon, R. Tahir, H. M. Naveed, and S. Memon, "Forecasting gold and platinum prices with an enhanced GRU model using multi-headed attention and skip connection," *Miner. Econ.*, vol. 38, no. 1, pp. 1–18, Mar. 2025, doi: 10.1007/s13563-025-00520-y.
[35] R. Elizabeth, S. Sitorus, "Forecasting The Indonesian Gold Price Using The ARIMA Method," *Jurnal Teknik Informatika dan Sistem Informasi,* vol. 11, no. 4, pp. 79-89, Dec. 2024, doi: 10.35957/jatisi.v11i4.8393.
[36] D. Li and Y. Li, "On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning," *J. Forecast.*, vol. 39, no. 6, pp. 953–966, Sep. 2020, doi: 10.1002/for.2677.
[37] K. A. Althelaya, E. M. El-Alfy, and S. Mohammed, "ARIMA Model for Accurate Time Series Stocks Forecasting," *International Journal of Advanced Computer Science and Applications* , vol. 11, no. 7, January 2020, doi:10.14569/IJACSA.2020.0110765.
[38] X. Zhang, "Which uncertainty measure better predicts gold prices? New evidence from a CNN-LSTM approach," *N. Am. J. Econ. Finance*, vol. 71, art. no. 102123, Jan. 2025, doi: 10.1016/j.najef.2024.102123.
[39] X. Zhang, L. Zhang, Q. Zhou, X. Jin, "A novel bitcoin and gold prices prediction method using an LSTM-P neural network model," *Comput. Intell. Neurosci.*, volume 2022, Article ID 1643413, May 2022, doi: 10.1155/2022/1643413.
[40] F. Bilgili, S. Kuşkaya, "Leveraging artificial intelligence for gold price forecasting: An LSTM-based analysis," *Research Article*, Sept. 2025, doi:10.21203/rs.3.rs-7488323/v1.
[41] K. Kashif and R. Ślepaczuk, "LSTM-ARIMA as a hybrid approach in algorithmic investment strategies," *Knowledge-Based Systems*, vol. 320, art. no. 113563, Jun. 2025, doi: 10.1016/j.knosys.2025.113563.
[42] F. Saâdaoui and H. Rabbouch, "Financial forecasting improvement with LSTM-ARFIMA hybrid models and non-Gaussian distributions," *Technological Forecasting and Social Change*, art. no. 123539, 2024, doi: 10.1016/j.techfore.2024.123539.
[43] M. H. Alharbi, "Prediction of the stock market using LSTM, ARIMA, and hybrid of LSTM-ARIMA models," *Journal of Knowledge Management Application and Practice*, vol. 7, no. 1, pp. 15–22, 2025, https://www.naturalspublishing.com/files/published/5qr8c1687avb59.pdf.
[44] U. Kalsum, R. Hidayat, and S. Oktaviani, "The Effect of Inflation, Interest Rates And World Oil Prices on Gold Prices in Indonesia With The US Dollar Exchange Rate as an Intermediary Variable," *Int. J. Finance Res.*, vol. 2, no. 1, pp. 1–12, Mar. 2021, doi: 10.47747/financeinvestmentderivative.v2i1.228.
[45] I. C. Rahmansyah and L. N. Rani, "Gold Price, Inflation, and Dollar Exchange Rate: The Case of Gold Investment in Indonesia," *Indonesian Capital Market Rev.*, vol. 13, no. 1, pp. 37–47, 2021, doi: 10.21002/icmr.v13i1.13238.

[46] D. M. Dewi, M. Z. Nafi', and N. Nasrudin, "Analisis Peramalan Harga Emas Bulanan dengan Holt Winter Exponential Smoothing," *Mediastima*, vol. 5, no. 2, pp. 38–50, Oct. 2023, doi: 10.36085/mediastima.v5i2.857.
[47] M. Risnawati, W. Syafmen, and B. Mardhotillah, "Prediksi penjualan emas di PT. Pegadaian Area Jambi menggunakan Fuzzy Time Series Cheng," Jurnal Ekonomi dan Statistik Indonesia, vol. 3, no. 1, pp. 70–84, Jun. 2023, doi: 10.11594/jesi.03.01.06.
[48] J. T. Manu Joseph, *Modern Time Series Forecasting with Python (Second Edition)*, Birmingham: Packt Publishing, 2024.
[49] J. Brownlee, *Deep Learning for Time Series Forecasting Predict the Future with MLPs, CNNs, and LSTMs in Python*, Machine Learning Mastery, 2018.