**Jurnal Teknik Informatika (JUTIF)** Vol. 7, No. 3, June 2026, Page. 2350-2375
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

# Comparative Evaluation of ARIMA, LSTM, Hybrid ARIMA-GARCH, and Hybrid GARCH-LSTM Models for Daily Bitcoin and Gold Price Forecasting

**Isna Nurul Fatatik*[1], Asyifa Nur Fadhilah[2], Irfan Adi Nugroho[3], Muhammad Muflih Affandi[4], Vriska Diah Novita Sari[5], Shaifudin Zuhdi[6]**

[1,2,3,4,5]Department of Data Science, Sebelas Maret University, Indonesia
[6]Department of Informatics, Sebelas Maret University, Indonesia

Email: [1]isnanurulfata@student.uns.ac.id

Received : Dec 10, 2025; Revised : Jan 5, 2026; Accepted : Jan 19, 2026; Published : Jun 15, 2026

**Abstract**

The volatile nature of digital financial markets poses major challenges for predictive modelling, particularly in developing accurate forecasting models that can address diverse asset characteristics such as Bitcoin, with its extreme fluctuations, and Gold, which is known for its stable movements. This study addresses this challenge by evaluating the robustness of linear, deep learning, and hybrid architectures in both high-volatility and stable asset environments. Utilizing Bitcoin and Gold closing price data from 2022 to 2025, the methodology adopts a comparative workflow that involves ARIMA, ARIMA-GARCH, LSTM, and LSTM-GARCH Hybrid models. Stationarity (ADF) and heteroskedasticity (ARCH-LM) diagnostics alongside AIC/BIC selection criteria were applied, followed by a walk-forward validation scheme to assess the model's performance. Results confirmed that the hybrid GARCH-LSTM model delivered the lowest Root Mean Squared Error (RMSE), significantly outperforming single models by integrating statistical variance and temporal neural learning. Therefore, this study contributes to the field of computational intelligence by validating an accurate Artificial Intelligence (AI) framework for volatility-based forecasting and proposing a scalable blueprint for engineers to develop models that are capable of capturing the dynamics of financial time series data.

***Keywords :*** *Bitcoin forecasting, GARCH-LSTM hybrid, Gold price prediction, time series volatility, RMSE evaluation.*



## 1. INTRODUCTION

Heightened economic uncertainty in the post-pandemic era has defined the trajectory of global financial markets [1], along with changes in monetary policy [2] and geopolitical dynamics [3] leading to direct impacts on the volatility of financial asset prices. These conditions have raised the need for price prediction models that are not only statistically accurate but also resilient to changes n data patterns and extreme market dynamics [4]. Asset price forecasting plays an essential role in investment decision-making, risk management, and the development of effective risk-hedging strategies [5],[6].

Bitcoin and Gold are assets that have distinct characteristics. However, they are both relevant in the realm of volatility differences [7], [8]. Bitcoin's reputation as a cryptocurrency with extreme price volatility stems from market sentiment [9], technology adoption [10], and global regulations [11]. In contrast, gold has served as a historically stable safe haven asset with relatively stable volatility, while remaining sensitive to inflation, interest rates, and global economic conditions [12], [13]. Their contrasting characteristics provide a useful basis for evaluating the robustness of price prediction models in dynamic and non-stationary market conditions [14].

Previous studies used conventional statistical models like Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) in

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

analyzing financial asset price predictions [15], [16]. ARIMA models effectively capture linear patterns and short-term autocorrelation; however, it has limitations regarding volatility clusters [17], [18]. This limitation leads to the development of GARCH models, which address ARIMA limitation by modeling conditional variability over time. GARCH models are widely used in finance volatility analysis [17], [19].

As computational capabilities evolve and large-scale historical data become more accessible. Deep Learning approaches such as Long Short-Term Memory (LSTM) are gaining its traction in financial time-series forecasting [20], [21]. LSTM excels at capturing non-linear patterns and long-term dependencies that traditional statistical models fail to capture [22], [23]. Nevertheless, several studies indicate LSTM performance may be unstable on data with extreme volatility if not accompanied by adequate volatility information [24].

Overcoming the limitations of single models, hybrid approaches such as ARIMA-GARCH and GARCH-LSTM have been developed, aiming for a combination of strengths from both statistical and deep learning models [25], [26]. Even so, most previous studies have only focused on one type of asset or simply compared one hybrid approach with another single model. Studies that simultaneously compare linear, non-linear, and hybrid models on two assets with contrasting volatility characteristics, particularly in periods of high volatility following a pandemic, remain limited.

Global financial market conditions during the 2022-2025 period market by various extreme shocks, ranging from COVID-19 pandemic, global economic recovery phases, geopolitical tensions, to significant monetary policy changes [27]. Period 2022-2025 poses a major challenge in terms of price prediction models due to increased instability in data patterns, changes in volatility structure, and regime shifts [28], [29]. Consequently, a comprehensive evaluation of prediction modelling is required to adequately address high volatility and non-linear data patterns in tandem.

In light of this research gap, the study conducted comparative analysis against four modelling approaches: ARIMA, ARIMA-GARCH, LSTM, and GARCH-LSTM, applied to two types of assets with contrasting characteristics, namely Bitcoin and Gold, in the post-pandemic period. This study aims to analyze and compare the performance of four single and hybrid models with parameter selection criteria based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), as well as evaluate model performance using Root Mean Squared Error (RMSE) within a walk-forward validation scheme. The findings of this study are expected to contribute significantly towards selecting an appropriate prediction model based not only on analysis objectives but also on asset characteristics. Furthermore, the results of this study can serve as a reference for developing risk management strategies and investment decisions.

## 2. METHOD

This research was conducted in several stages, namely: (1) collection of daily Bitcoin and Gold price data for the period 2022–2025, (2) data exploration by extracting the date and closing price columns and data visualization, (3) preprocessing stage which includes testing stationarity using Augmented Dickey-Fuller (ADF) test and differencing the data if it was not stationary, (4) dividing the data into training and testing data with a ratio of 80:20, (5) modeling using Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and Long Short-Term Memory (LSTM), (6) hybrid ARIMA-GARCH and GARCH-LSTM modeling, (7) model testing using the walk-forward validation method, and (8) evaluation and comparison of model performance based on the Root Mean Squared Error (RMSE) value on Bitcoin and Gold test data. The research process flow is shown in Figure 1.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

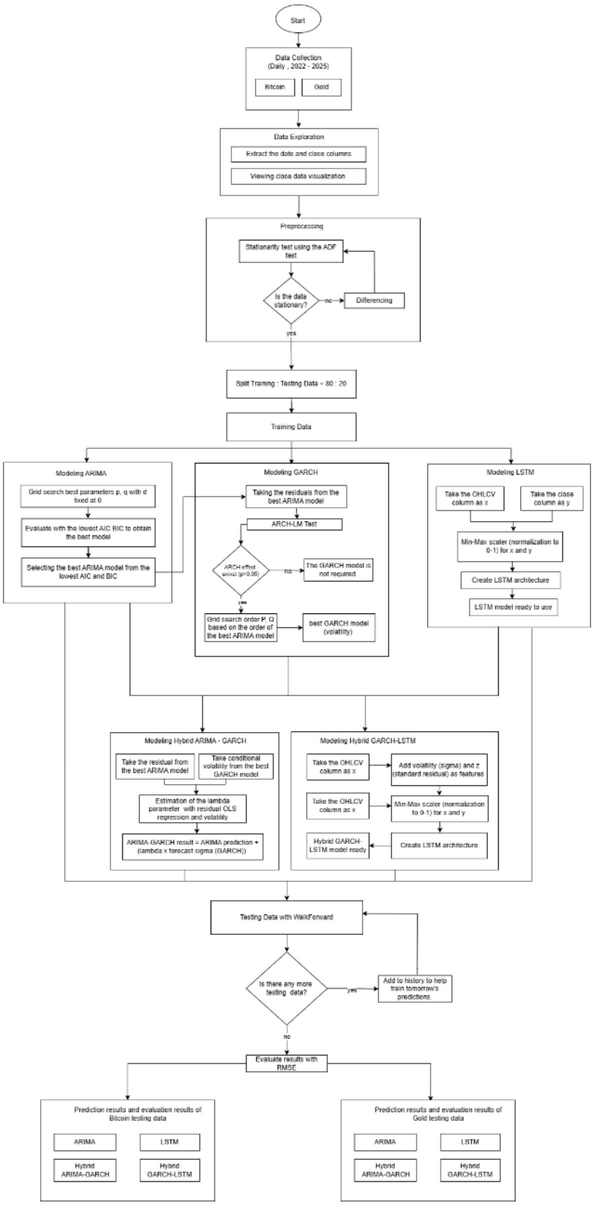



Figure 1. Research Workflow

## 2.1. Dataset

Datasets used in this study were obtained from Yahoo Finance through the yfinance library, which consists of two main financial assets: Bitcoin against the US Dolar (BTC-USD) and COMEX gold price (GC=F). Variables used include opening price (Open), highest price (High), lowest price (Low), closing price (Close), and trading volume (Volume). BTC-USD and GC=F were chosen as they are both alternative investment assets commonly used as hedges against inflation and market volatility as well as having complex and non-linear price movement characteristics [30]. This study utilized daily data spanning from January 1, 2022, to November 11, 2025, divided into training data (80% of total data) and testing data (20% of total data) as a means of evaluating prediction model performance.

Table 1 and Table 2 show sample datasets for BTC-USD and GC=F that will be used as training data in model development. BTC-USD and GC=F that will be used as training data in model development. BTC-USD dataset shows high volatility with closing prices ranging from USD 6985.47 to USD 7769.22 in early January 2022, while trading volume ranged from USD 18 billion to USD 28 billion. As for the GC=F dataset, gold prices show relatively more stable movements with closing prices ranging from USD 1524.50 to USD 1566.20 and lower trading volumes compared to Bitcoin.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

Table 1. Sample Dataset Bitcoin (BTC-USD)

| Date | Open | High | Low | Close | Volume |
|---|---|---|---|---|---|
| 2022-01-01 | 7202.551270 | 7212.155273 | 6935.270020 | 6985.470215 | 20802083465 |
| 2022-01-02 | 6984.428711 | 7413.715332 | 6914.996094 | 7344.884277 | 28111481032 |
| 2022-01-03 | 7345.375488 | 7427.385742 | 7309.514160 | 7410.656738 | 18444271275 |
| 2022-01-04 | 7410.451660 | 7544.497070 | 7400.535645 | 7411.317383 | 19876543210 |
| 2022-01-05 | 7410.452148 | 7781.867188 | 7409.292969 | 7769.219238 | 19725074095 |

Table 2. Sample Dataset Gold (GC=F)

| Date | Open | High | Low | Close | Volume |
|---|---|---|---|---|---|
| 2022-01-01 | 1518.099976 | 1528.699951 | 1518.000000 | 1524.500000 | 214.0 |
| 2022-01-02 | 1530.099976 | 1552.699951 | 1530.099976 | 1549.199951 | 107.0 |
| 2022-01-03 | 1530.099976 | 1552.699951 | 1530.099976 | 1549.199951 | 107.0 |
| 2022-01-04 | 1530.099976 | 1552.699951 | 1530.099976 | 1549.199951 | 107.0 |
| 2022-01-05 | 1580.000000 | 1552.699951 | 1560.400024 | 1566.199951 | 416.0 |

This closing price variable will be used for the main variable as it reflects an asset's final value at each trading period's end as well as being a key indicator of financial time series analysis [31]. Characteristics of the volatility and temporal patterns of these two assets lay the foundation for the use of ARIMA, ARIMA-GARCH, LSTM, and GARCH-LSTM methods to capture linear components, heteroskedasticity, and non-linear patterns in time series data

### 2.2. Augmented Dickey-Fuller Test (Uji ADF)

Prior to performing time series modelling using the ARIMA model, it is essential to ensure stationarity of the data, both in terms of mean and variance. Series are stationary if their mean and variance are constant over time. Dickey-Fuller tests are a generalization of ADF tests. ADF test adapted to address autocorrelation within errors, by adding lag values of the dependent variable [32]. Equations for ADF tests can be written as follows:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-1} + \varepsilon_t \quad (1)$$

Where $\Delta$ represents an operator of difference and $Y_t$ indicates Bitcoin or Gold closing price at time *t*. Parameter $\alpha$ is an intercept constant, $\beta$ is time trend coefficient. Parameter $\gamma$ is used to determine whether there is a unit root present, while $p$ indicates optimum lag length used in the test, and $\varepsilon_t$ is white noise error. Hypothesis zero ($H_0$) with $\gamma = 0$ suggest that data has a unit root (is non-stationary), thus if probability value (p-value) is smaller than its significance level ($\alpha = 0.05$), then $H_0$ is rejected, meaning the data is stationary [33].

### 2.3. Uji Efek ARCH (ARCH-LM Test)

An Largerange Multiplier (ARCH-LM) test is used to validate the assumption of homoscedasticity by detecting volatility patterns or conditional autoregressive heteroscedasticity effects in ARIMA model residuals [34]. It is essential to perform this test before deciding whether to use GARCH to capture non-constant error variance. Mathematically, it is formulated as follows:

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

$$\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \hat{\varepsilon}_{t-i}^2 + v_t \tag{2}$$

Where $\hat{\varepsilon}_t^2$ denotes squared residuals at time *t*, $\alpha_0$ is its intercept constant, and $\alpha_i$ coefficient of squared residuals at lags *i* through order *q*. $v_t$ represents residual error component. Hypothesis zero ($H_0$:$\alpha_1$ =…=$\alpha_q$ = **0** ) assumes that there is no ARCH effect. Hence, if results of statistical tests are significant, hypothesis zero (H0) is rejected, as it confirms the presence of heteroscedasticity within the data [34].

## 2.4. LSTM Model

The Long Short-Term Memory (LSTM) model is used as a learning method due to its ability to capture non-linear patterns and long-term dependencies within time series data. LSTM is frequently used in research related to Bitcoin and Gold price prediction given that its architecture can overcome vanishing gradients issues and exploding gradients that occur in Conventional Recurrent Neural Networks (RNN). Several studies indicate that LSTM can provide accurate predictions of highly volatile data with complex temporal patterns [31], [35]. Its computation process is performed using following equation:

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f) \tag{3}$$
$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \tag{4}$$
$$\hat{C}_t = tanh(W_c[h_{t-1}, X_t] + b_c) \tag{5}$$
$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \tag{6}$$
$$o_t = \sigma(W_o[h_{t-1}, X_t] + b_o) \tag{7}$$

Where $i_t$ is the input gate, $o_t$ is the output gate, and $f_t$ is the forget gate, and the cell state value ($\hat{C}_t$ ) functions to regulate the memory flow in each neuron [36].

## 2.5. ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) model is used as the basic method due to its ability to capture linear patterns and autocorrelation in financial time series data. ARIMA is often used in research on gold prices, financial indices, and cryptocurrencies because of the model's flexible structure in addressing trends and short-term autocorrelation. Several studies indicate that ARIMA can serve as a strong baseline for time series analysis, particularly for stationary data periods [37], [38]. The mathematical equation for the ARIMA model is:

$$Yt = \phi Y_{t-1} + \cdots .. + Y_{t-p} + \epsilon_t - \theta_{1\epsilon t-1} - \cdots - \theta_{q\epsilon t-q} \tag{8}$$

ARIMA equation illustrates how current values $Y_t$ is influenced by two main components: historical values and previous shocks. The autoregressive (AR) component ($\phi_i Y_{t-i}$ ) capturing influences from values of the past up to lag-$p$ , on the other hand moving average (MA) component ($-\theta_j \epsilon_{t-j}$) explains how errors from previous periods also affect its current value. $\epsilon_t$ component represents white noise, which accounts for shocks not explained by AR or MA [39]-[41].

## 2.6. GARCH Model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is used to model volatility that changes over time in time series data, especially in financial data that shows

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

heteroskedasticity, i.e., non-constant variance. This model is very useful for identifying volatility that often clusters [42]. GARCH can describe patterns of volatility that give rise to periods of high fluctuation followed by periods of low volatility [43]. Mathematicall, it is illustrated as follow:

$$\boldsymbol{\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2} \quad (9)$$

Where $\boldsymbol{\sigma_t^2}$ is conditional variance estimate (volatility) at time $\boldsymbol{t}$, $\boldsymbol{\varepsilon_{t-1}^2}$, residual square at time $\boldsymbol{t-1}$, and $\boldsymbol{\sigma_{t-1}^2}$ noted as volatility from previous period [42], [43]. It assumes current volatility rely on previous volatility and market errors or shocks. Parameter $\boldsymbol{\alpha_1}$ measures the influence of previous market shocks on current volatility, while $\boldsymbol{\beta_1}$ measures previous volatility influence on current volatility [42]. Parameter estimation in the GARCH model is generally performed using Maximum Likelihood Estimation (MLE), which seeks to maximize likelihood function based on existing data to obtain most optimal parameters for predicting volatility [42],[44].

## 2.7. Hybrid GARCH-LSTM Model

The GARCH-LSTM hybrid model is used to address two characteristics of modern financial markets: fluctuating volatility and long-term non-linear patterns. GARCH is effective in modelling volatility dynamics and volatility clustering, while LSTM is strong in capturing long-term relationships and non-linear patterns that are difficult to handle by traditional statistical models [40] ,[41], [44]. GARCH mathematically written as follow:

$$y_t = \mu + \varepsilon_t \quad (10)$$
$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (12)$$

This formula shows that the return $\boldsymbol{y_t}$ is composed of its mean $\boldsymbol{\mu}$ and error $\boldsymbol{\epsilon_t}$ , whose variance follows the GARCH process. Conditional variance $\boldsymbol{\sigma_t^2}$ is dynamic, influenced by previous shocks through the parameter $\boldsymbol{\alpha}$ , as well as by volatility persistence through the parameter $\boldsymbol{\beta}$ . This structure has been extensively used to model high-risk asset volatility since it can capture volatility clustering phenomena commonly found in crypto and commodity markets [41], [45]. It is calculated using these equations:

$$\boldsymbol{f_t = \sigma\left(W_f[h_{t-1}, X_t] + b_f\right)} \quad (13)$$
$$\boldsymbol{i_t = \sigma(W_i[h_{t-1}, X_t] + b_i)} \quad (14)$$
$$\boldsymbol{\widehat{C}_t = tanh(W_c[h_{t-1}, X_t] + b_c)} \quad (15)$$
$$\boldsymbol{C_t = f_t * C_{t-1} + i_t * \widehat{C}_t} \quad (16)$$
$$\boldsymbol{o_t = \sigma(W_o[h_{t-1}, X_t] + b_o)} \quad (17)$$

This set of formulas describes the internal mechanism of LSTM that works with gate-based memory control. The forget gate $f_t$ determines which information to discard, the input gate $i_t$ plays a role in regulating new information to be stored, while the $\hat{C}_t$ is the candidate memory. LSTM retains important information through the memory cell $C_t$, making it effective in learning long-term patterns in non-linear data. When features capture the relationship between volatility dynamics and price direction with stimuli, the model is able to learn well and improve prediction accuracy on highly fluctuating data assets [40], [46].

## 2.8. Hybrid ARIMA-GARCH Model

Hybrid ARIMA-GARCH models are a joint approach, designed to capture complex time series data, including linear patterns and high volatility. ARIMA models are used to capture linear patterns

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

and temporal dependencies in the data, however they have limitations in capturing non-linear patterns or non-constant variance. In contrast, GARCH models address heteroscedasticity or fluctuations in volatility among residuals which tend to change over time. Second-order information, conditional variance, was typically captured using an ARCH-based GARCH model. Using a combination of these two models provides more accurate modelling than a single model, especially in highly volatile financial markets such as stocks and commodities prices [47]-[49]. Mathematical equations for ARIMA model given by:

$$\omega_t = \Delta^d y_t = (1 - B)^d y_t \quad (18)$$

$$\widehat{\omega_t} = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p+\varepsilon_t} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad (19)$$

$$y_t^{\widehat{\text{ARIMA}}} = \Delta^{-d} \widehat{\omega_t} \quad (20)$$

Where $\omega_t$ represent ARIMA process after differencing order $d$, while $\widehat{\omega}_t$ as result of ARIMA model estimation. $\hat{y}_t^{ARIMA}$ value shows its prediction result at time $t$ , obtained through inverse differencing process ($\Delta^{-d}$ ) to return the data to its original scale. Furthermore, residuals produced from the ARIMA model are modeled using GARCH to capture volatility characteristics of data. Residuals are defined as follow:

$$e_t = y_t - y_t^{\widehat{ARIMA}} \quad (21)$$

The autoregressive (AR) component describes the influence of past values on current values, while the influence of past errors on present is explained by the Moving Average (MA) component [47], [48]. Mathematically GARCH is written ass follow:

$$\sigma_t^2 = \omega + \alpha_1 e_{t-1}^2 + \cdots + \alpha_q e_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 \quad (22)$$

The prediction of the GARCH component is expressed as:

$$y_t^{\widehat{GARCH}} = z_t \widehat{\sigma}_t \quad (23)$$

Where $z_t$ is a random variable that follows a standard normal distribution, $\omega$ is a constant, and$\widehat{\sigma}_t$ represents its onditional standard deviation at time $t$. In this equation, $\sigma_t^2$ (or $h_t$) represents the conditional variance or volatility in a period of time $t$ , while $\omega$ is constant. Finally, the forecast results from hybrid ARIMA–GARCH model are obtained by combining predictions from both models, namely [50]

$$y_t^{\widehat{hybrid}} = y_t^{\widehat{ARIMA}} + y_t^{\widehat{GARCH}} \quad (24)$$

### 2.9. Akaike Information Criterion (AIC) dan Bayesian Information Criterion (BIC)

Akaike Information Criterion (AIC) is a selection criterion intended to estimate the relative quality of statistical models, notably in the context of time series prediction. AIC serves as a metric for evaluating the balance between model fit and complexity [51]. It aims to prevent model overfitting by maintaining a balance between complexity (number of parameters) and model accuracy [52]. In the ARIMA or ARIMA-GARCH modelling process, AIC assists in determining the optimal model order.

Meanwhile, Bayesian Information Criterion (BIC), commonly known as Schwarz Criterion, refers to a selection criterion used to optimize time series prediction models. BIC also evaluates the trade-off between model fit and model complexity [51]. The difference is that BIC bases its evaluation on a stronger penalty for the number of parameters and sample size, making it more effective than AIC in finding models from actual data [52], [53], [54]. Model selection based on AIC-BIC evaluation yields

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

the model with the smallest AIC-BIC value [55], [56]. AIC and BIC are generally expressed by following equations.

$$AIC = 2 * k - 2 * \ln(L) \quad (25)$$
$$BIC = k * \ln(n) - 2 * \ln(L) \quad (26)$$

Where $k$ indicates the number of parameters used on model, $L$ noted as the value of model's likelihood function, and $n$ sample size from total number of observations used [57].

### 2.10. Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) refers to a model evaluation metric applied to assess a model's predictive performance [58]. RMSE measures the magnitude of difference between actual and predicted values, or prediction errors for time series models [59]. Errors in RMSE are computed by squaring the mean error between actual and predicted values. Since errors are squared, RMSE is highly sensitive to large errors or outliers [60].

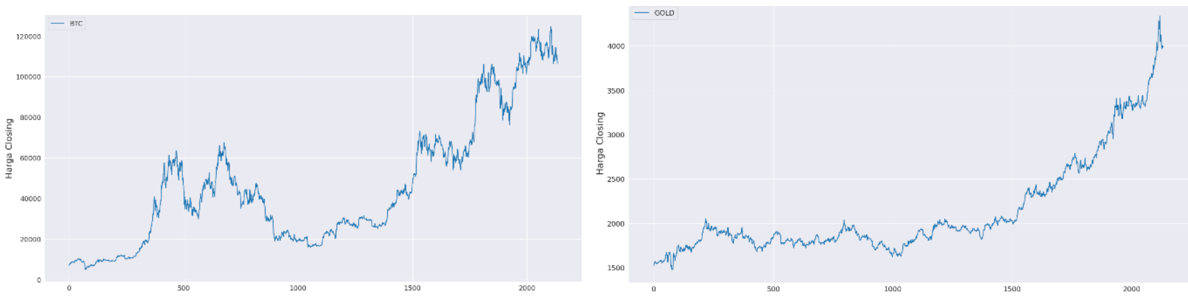
$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \overline{y_i})^2} \quad (27)$$

Where $y_i$ noted as actual value at observation $i$, $\overline{y_i}$ is the predicted value of the model at observation $i$ and $m$ is total number of observations used in the model evaluation process [61].

## 3. RESULT

### 3.1. Exploratory Data Analysis (EDA)

Two sets of data were utilized for this study, specifically Bitcoin and Gold. An preliminary analysis was conducted to examine the characteristics of the closing price movements of Bitcoin (BTC-USD) and Gold (GC=F) from January 1 to November 11, 2025.




(a) (b)

Figure 2. Closing Price Movements (a) Bitcoin dan (b) Gold

Figure 2 shows that the closing price movements of Bitcoin and Gold exhibit significantly different fluctuation patterns. Bitcoin shows very high volatility, with a sharp downward trend (bearish) throughout 2022 due to the crypto winter, followed by a recovery phase and a significant upward trend (bullish) from late 2023 to 2025. The extreme price surge in Bitcoin indicates a high degree of uncertainty and investment risk. In contrast, Gold price movements tend to be more stable but still experience a gradual upward trend, reflecting its function as a safe haven asset amid global economic uncertainty.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

The closing price data for these two assets shows a non-stationary pattern, where the mean and variance change over time. This indicates that the raw data cannot be directly used for modeling and requires a stationarity test first. In addition, there are indications of volatility clustering, particularly in Bitcoin prices. There are periods where high price fluctuations are followed by other high fluctuations, and vice versa. The existence of this volatility clustering is an early indication of the effect of heteroscedasticity (non-constant variance) in residuals' data. Therefore, an ARCH-LM test also needs to be performed.

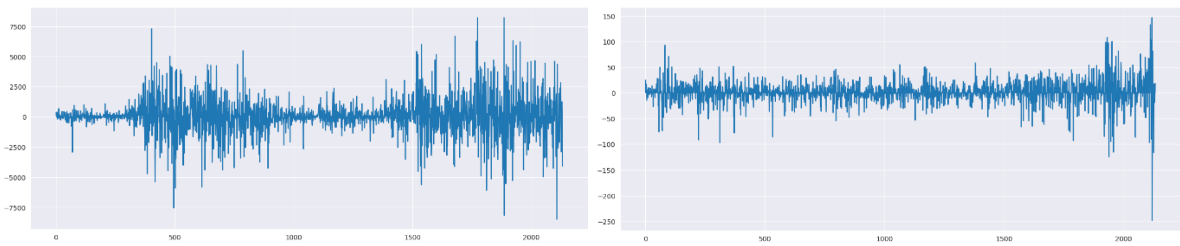
### 3.2. Augmented Dickey-Fuller (ADF) Test

Step one in time series modelling, especially ARIMA, require stationary data, meaning conditions where mean and variance remain constant over time. To test the assumption of stationarity, this study uses the Augmented Dickey-Fuller (ADF) test. Hypothesis zero ($H_0$) on ADF test states that the data has a unit root (is not stationary), while the alternative hypothesis ($H_1$) states that the data is stationary. If the probability value (*p-value*) is less than the significance ($\boldsymbol{\alpha = 0.05}$) or the ADF statistical value is smaller than *the critical value*. The results of the ADF test at the original data level and after first-order differencing.

Table 3. Stationarity Test Results with Augmented Dickey-Fuller (ADF)

| Aset | Tahapan Data | ADF Statistik | Critical Value | P-Value | Keterangan |
|---|---|---|---|---|---|
| Bitcoin | Original Data | -0.4978 | -2.863 | 0.8924 | Non-stationary |
| | *1st Difference* | -48.3767 | -2.863 | 0.0000 | Stationary |
| Gold | Original Data | 3.1675 | -2.863 | 1.0000 | Non-stationary |
| | *1st Difference* | -10.2679 | -2.863 | 0.0000 | Stationary |

Based on Table 3, testing on the original data shows that Bitcoin and Gold data are not stationary, as indicated by ADF statistical values greater than the critical value of 5% and p-values > 0.05 ($H_0$accepted). Stationarity is only achieved after first-order differencing (d=1), where the ADF statistic values for both assets drop sharply below the critical value with a p-value of 0.0000 ($H_0$rejected). This confirms that the *first-difference* data is stationary, as visualized in Figure 3.



(a) (b)

Figure 3. First-order Differencing Results (d=1) on (a) Bitcoin and (b) Gold Data

Based on Figure 3, after differencing, the trend in the data has disappeared and the data fluctuations move around zero (constant mean). Therefore, the data is ready to be used for the ARIMA model identification stage.

### 3.3. ARIMA Model

ARIMA modeling was performed to estimate linear patterns of the price movements of Bitcoin and Gold by utilizing autoregressive (AR), differencing (I), and moving average (MA) components.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

Based on the stationarity test and results of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), a single differencing process (d=1) was selected to stabilize the variance and eliminate strong trends. The model was then tested using various combinations of AR and MA parameters to obtain the most optimum configuration based on RMSE, AIC, and BIC metrics. Bitcoin and Gold closing prices, respectively, are used as target variables for model's prediction.

### 3.3.1. ARIMA Model on Bitcoin Data

ARIMA modeling was applied to the Bitcoin price dataset to capture daily price movement patterns from 2024 to the end of 2025. To determine the best parameters, several combinations (p, d, q) were tested in the range:

- p = 1 to 3
- d = 0 because it is already stationaire when differencing is performed once before modeling
- q = 1 to 3

The evaluation process is based on three metrics, namely RMSE to measure prediction error, AIC to determine model complexity penalty, and BIC to determine model complexity penalty that is stricter than AIC.

Table 4. ARIMA Model Evaluation on Bitcoin Data

| Ordo (p, d, q) | RMSE | AIC | BIC |
|---|---|---|---|
| (1, 0, 1) | 2035.8041 | 25265.3544 | 25286.5858 |
| (1, 0, 2) | 2036.4569 | 25265.7403 | 25292.2797 |
| (1, 0, 3) | 2038.1412 | 25267.1832 | 25299.0305 |
| (2, 0, 1) | 2036.5617 | 25265.8553 | 25292.3947 |
| (2, 0, 2) | 2044.0849 | 25259.6649 | 25291.5122 |
| (2, 0, 3) | 2040.9986 | 25261.5631 | 25298.7182 |
| (3, 0, 1) | 2037.0649 | 25267.2694 | 25299.1167 |
| (3, 0, 2) | 2044.0025 | 25269.7640 | 25306.9191 |
| (3, 0, 3) | 2045.2976 | 25263.7014 | 25306.1644 |

Based on Table 4, the best ARIMA model for Bitcoin data is provided by the ARIMA model with parameters p, d, q, respectively, which is the model with parameters 1, 0, 1, where this model shows the lowest RMSE , indicating the smallest prediction model error. Although some models have lower AIC, BIC shows a higher number, so it is considered less efficient in terms of complexity. The results of this modeling also explain that a larger combination of p and q does not provide a significant improvement and tends to increase the BIC penalty.

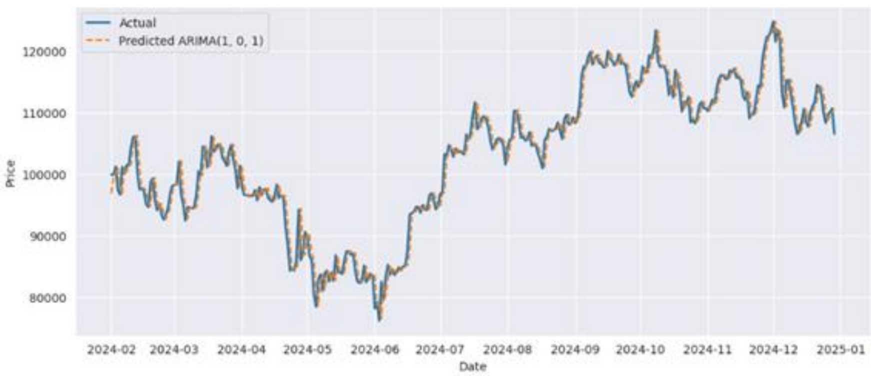



Figure 4. ARIMA (1, 0, 1) Model on Bitcoin Data

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

The results in Figure 4 show a comparison between the actual Bitcoin price and the ARIMA (1, 0, 1) prediction results, which have been returned to their original scale using the inverse differencing process. This figure shows that the ARIMA model with the optimal parameter combination is able to accurately capture the fluctuating patterns of actual Bitcoin data. This success is supported by the use of walk-forward validation, which allows the model to adapt to the latest data changes, as well as the ARIMA's ability to exploit the autocorrelation structure in historical data patterns.

### 3.3.2. Model ARIMA on Gold Data

ARIMA modeling was also applied to the Gold price dataset to evaluate the model's ability to predict assets with lower volatility compared to Bitcoin. The differencing process was also performed once to stabilize the trend pattern, followed by a search for the best parameters by testing several parameters (p, d, q) within the same range as Bitcoin, namely,

- p = 1 to 3
- d = 0 because it is already stationaire after being differentiated once before modeling
- q = 1 to 3

The evaluation process was also carried out using three metrics, namely RMSE, AIC, and BIC.

Figure 5. ARIMA Model Evaluation on Gold Data

| Model (p, d, q) | RMSE | AIC | BIC |
|---|---|---|---|
| (1, 0, 1) | 30.0767 | 12427.8068 | 12449.0383 |
| (1, 0, 2) | 30.1673 | 12422.7775 | 12449.3169 |
| (1, 0, 3) | 30.2030 | 12424.5651 | 12456.4124 |
| (2, 0, 1) | 30.1108 | 12422.7221 | 12449.2615 |
| (2, 0, 2) | 30.1795 | 12421.7716 | 12453.6189 |
| (2, 0, 3) | 30.1460 | 12423.7604 | 12460.9155 |
| (3, 0, 1) | 30.1626 | 12424.5657 | 12456.4130 |
| (3, 0, 2) | 30.2382 | 12423.7052 | 12460.8603 |
| (3, 0, 3) | 30.3111 | 12420.4071 | 12462.8701 |

Table 5 shows that the ARIMA (1, 0, 1) model has the lowest RMSE value of 30.0767, making it the best model based on prediction performance on Gold data. Although models with higher p and q values have lower AIC values, they were not selected because their BIC values are higher due to greater parameter complexity. Therefore, the ARIMA (1, 0, 1) model provides the best balance between model accuracy and efficiency and is selected as the final configuration for Gold price modeling.

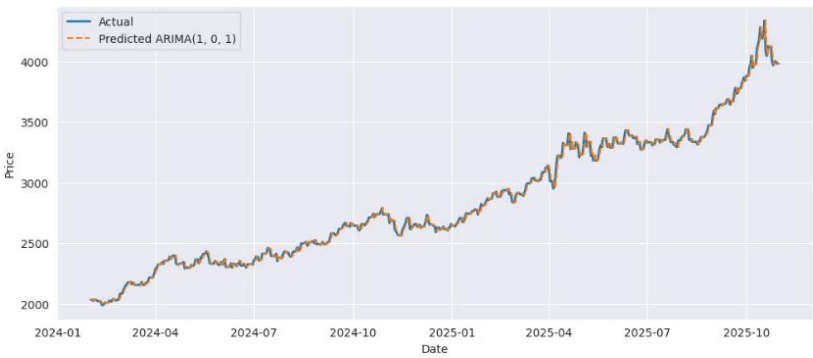



Figure 6. ARIMA (1, 0, 1) Model on Gold Data

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

Based on Figure 6, a comparison between the actual price of gold and the predicted results of the ARIMA (1, 0, 1) model on a native scale after the inverse differencing process. The blue line depicts the actual price of gold, which shows an upward trend with moderate fluctuations during the period from 2024 to the end of 2025. Compared to Bitcoin, the price pattern of gold is more stable and does not experience extreme spikes, but still shows clear fluctuations in line with changes in global economic conditions, inflation, pandemics, and market sentiment towards safe haven assets. Meanwhile, the dotted orange line represents the model's prediction results, which show fluctuating movements similar to the actual data. This occurs because the ARIMA model uses walk-forward validation, which enhances the model's ability to adapt to the volatility of the latest data based on historical patterns. The results of the tests conducted on gold and bitcoin data show that ARIMA with walk-forward validation is indeed capable of modeling data with non-linear patterns with precision

### 3.4. ARCH-LM Test

After modeling with ARIMA (1,1,1), residual values were obtained for the ARCH-LM test and GARCH modeling. ARIMA modeling requires that the residuals be homoscedastic (constant variance). If there is correlation in the residual squares, then heteroscedasticity (ARCH effect) occurs, which means that data volatility changes over time.

To assess any ARCH effect, this study uses the Lagrange Multiplier test (ARCH-LM). Hypothesis zero ($H_0$) in this test states that there is no ARCH effect (homoscedasticity), while the alternative hypothesis ($H_1$) states that there is an ARCH effect (heteroscedasticity). The following are the results of the ARCH-LM test:

Table 5. Heteroscedasticity Test Results (ARCH-LM Test)

| Aset | LM Statistik | P-Value ($X^2$) | F-Statistik | P-Value (F) | Keterangan |
|---|---|---|---|---|---|
| Bitcoin | 170.33 | 2.37 $10^{-31}$ | 18.42 | 9.98 $10^{-33}$ | Heteroscedasticity |
| Gold | 304.99 | 1.37 $10^{-59}$ | 35.43 | 1.70 $10^{-64}$ | Heteroscedasticity |

Based on Table 5, the ARCH-LM test results show significant evidence of heteroscedasticity in both assets, indicated by the LM Statistics values for Bitcoin (170.33) and Gold (304.99) which have *p-values* well below the significance level of 0.05. This condition leads to the rejection of the null hypothesis ($H_0$ ), meaning that there is volatility clustering in the ARIMA model residuals. Therefore, the ARIMA model alone is considered insufficient, and the application of the GARCH and GARCH-LSTM methods is needed to capture residual volatility.

### 3.5. GARCH Model

The GARCH model is used to model volatility that changes over time in *time series* data, especially in financial data that shows heteroscedasticity or non-constant variance. This model is useful for describing patterns of volatility fluctuations that often cluster, with periods of high volatility followed by periods of low volatility. GARCH assumes that current volatility is influenced by previous periods of volatility as well as market shocks in previous periods. The target variable in this modeling is the closing price that will be predicted by the model.

Table 6. GARCH Model Evaluation on Bitcoin Data

| Model (p, q) | AIC |
|---|---|
| (1, 1) | 10833.614 |
| (2, 1) | 10835.443 |
| (1, 2) | 10835.730 |

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

In Bitcoin price data, the models tested were GARCH (1, 1), GARCH (2, 1), and GARCH (1,2). The GARCH (1, 1) model was selected as the best model based on a lower AIC value of 10833.614.

Based on Table 6 with the AIC model selection criteria, the GARCH (1, 1) model was selected as the best model. The beta[1] parameter shows a value of (0.8674), which indicates a large influence on current volatility, with a very small *P*>|*t*| value of 2.061e-90, which shows that past volatility has a large impact on current price volatility. Conversely, alpha[1] shows a value of (0.1119) and a *P*>|*t*| value of 5.860e-02, which is slightly greater than the 5% level, indicating that the effect of past volatility on current volatility is not significant.

Table 7. Parameter Estimation of GARCH Model on Bitcoin Data

| | Coef. | Std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| Omega | 0.3661 | 0.144 | 2.544 | 1.094e-02 | [8.410e-02, 0.648] |
| Alpha[1] | 0.1119 | 5.914e-02 | 1.891 | 5.8602-02 | [-4.067e-03, 0.228] |
| Beta[1] | 0.8675 | 4.302e-02 | 20.163 | 2.061e-90 | [0.783, 0.952] |

Table 7 shows the results of GARCH model parameter estimation on Bitcoin data, where most volatility parameters show strong statistical significance. The constant parameter (ω) has an estimated value of 0.3661 with a p-value of 0.01094, indicating that the long-term variance component has a significant effect on Bitcoin volatility. The ARCH α(1) parameter has an estimated value of 0.1119 with a p-value of 0.0586, which is at a marginal level of significance and indicates the influence of short-term shocks on volatility. Meanwhile, the GARCH parameter β(1) shows a high estimated value of 0.8675 with a very small p-value ($2.061\times10^{-90}$), indicating a very strong level of volatility persistence. The relatively narrow confidence interval on the β(1) parameter also reinforces the stability of this estimate. Overall, these results show that Bitcoin volatility is highly persistent and significantly influenced by past variance, reflecting the highly volatile nature of the cryptocurrency market.

In the gold price data, the GARCH models tested include GARCH (1, 1), GARCH (2, 1), and GARCH (1, 2). The best model obtained is GARCH (1, 1), which has the lowest AIC of 5360.632, lower than the other models.

Table 8. GARCH Model Evaluation on Gold Data

| Model (p, q) | AIC |
|---|---|
| (1, 1) | 5360.632 |
| (2, 1) | 5362.363 |
| (1, 2) | 5362.382 |

Based on Table 8 with the AIC model selection criteria, the GARCH (1, 1) model was selected as the best model. The parameters *omega* 0.3898, *alpha[1]* 0.1217, and *beta[1]* 0.7544 have *P*>|*t*| values that are greater than 0.05, which means that the effect of past volatility on current volatility is not significant.

Table 9. Parameter Estimation of GARCH Model on Gold Data

| | Coef. | Std err | t | P>\|t\| | 95.0% Conf. Int. |
|---|---|---|---|---|---|
| Omega | 0.3898 | 1.076 | 0.362 | 0.717 | [-1.718, 2.498] |
| Alpha[1] | 0.1217 | 0.413 | 0.294 | 0.768 | [-0.688, 0.932] |
| Beta[1] | 0.7544 | 5.311 | 0.142 | 0.887 | [-9.655, 11.164] |
| Beta[2] | 0.1024 | 4.838 | 2.116e-02 | 0.983 | [-9.380, 9.584] |

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

Table 9 shows the results of the GARCH model parameter estimation on Gold data, where all estimated volatility parameters do not show statistical significance at the conventional significance level. The constant parameter (ω) has an estimated value of 0.3898 with a p-value of 0.717, while the ARCH parameter α(1) has an estimated value of 0.1217 with a p-value of 0.768, indicating that short-term shocks do not have a significant effect on volatility. In addition, the GARCH β(1) and β(2) parameters also show high p-values, 0.887 and 0.983, respectively, so that the persistence of volatility in Gold data is not strongly identified. The wide confidence intervals for all parameters reflect the high uncertainty of the estimates, which overall indicate that the dynamics of conditional variance in Gold prices are relatively stable and consistent with the characteristics of Gold volatility, which is lower than that of high-risk assets.

### 3.6. ARIMA-GARCH Model

ARIMA-GARCH model was utilized to the Bitcoin closing price dataset from 2024 to 2025. This data was used to capture closing price movement patterns and predict daily returns. In developing process, it is necessary to select most optimum parameters to obtain most optimal results. This is done by testing several combinations of ARIMA (*p, d, q*) and GARCH (*p, q*) orders. The combination of the ARIMA and GARCH models is based on the results of conducted LM-test to validate the presence of heteroscedasticity. Based on its results, Bitcoin and Gold are confirmed to have heteroscedasticity in the data.

The ARIMA-GARCH model was used on the Bitcoin closing price dataset to obtain the pattern of Bitcoin price movements throughout 2024 to the end of 2025. The model used was the model with the best parameters to produce the most optimal prediction and model evaluation results. This was done by testing the GARCH (*p, q*) combination and using the ARIMA (1,0,1) model, which had previously been tested to have the best model evaluation results. GARCH was tested on the combinations (1,1), (2,1), (3,1), (1,2), (2,2). The evaluation was based on the AIC metric to see the model complexity penalty.

Table 10. ARIMA-GARCH Model Evaluation onBitcoin

| Model ARIMA(p, d, q)-GARCH(p, q) | AIC |
|---|---|
| (1,0,1)-(1, 1) | 24003.654 |
| (1,0,1)-(2, 1) | 24005.655 |
| (1,0,1)-(3, 1) | 24046.323 |
| (1,0,1)-(1, 2) | 24042.653 |
| (1,0,1)-(2, 2) | 24027.873 |

Based on Table 10, the best GARCH model for Bitcoin data is the GARCH (1, 1) model. This model will be used for the hybrid model. The ARIMA-GARCH model for Bitcoin data is ARIMA (1, 0, 1) - GARCH (1, 1).

Based on Figure 7, it shows a comparison graph between actual price data and the predicted price of ARIMA (1, 0, 1) - GARCH (1,1). The figure shows that the ARIMA-GARCH prediction captures the pattern of Bitcoin closing price movements well with an RMSE of 2034.55.

The ARIMA-GARCH model was applied to analyze gold price movements from 2024 to 2025. The best model was selected by evaluating various combinations of ARIMA and GARCH parameters. The ARIMA (1, 1, 1) model was chosen based on previous test results that showed the best model performance in terms of prediction accuracy. On the other hand, the GARCH parameters were tested with various combinations, namely (1, 1), (1, 2), (2, 1), and (2, 2). The best model was selected based on the AIC value assessment, which describes the efficiency of the model by considering its complexity and predictive ability.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555


Figure 7. ARIMA-GARCH Model for Bitcoin

Table 11. ARIMA-GARCH Model Evaluation on Gold Data

| Model ARIMA(p, d, q)-GARCH(p, q) | AIC |
|---|---|
| (1,1,1)-(1,1) | 12432.644 |
| (1,1,1)- (1,2) | 12434.644 |
| (1,1,1)-(2, 1) | 12434.186 |
| (1,1,1)-(2, 2) | 12436.186 |

Based on the evaluation results using the model selection criteria metric with AIC in Table 11, the GARCH (1, 1) model shows the best results with the lowest AIC value compared to other GARCH model combinations. Therefore, the ARIMA-GARCH model for gold data is ARIMA (1, 1, 1) - GARCH (1, 1).

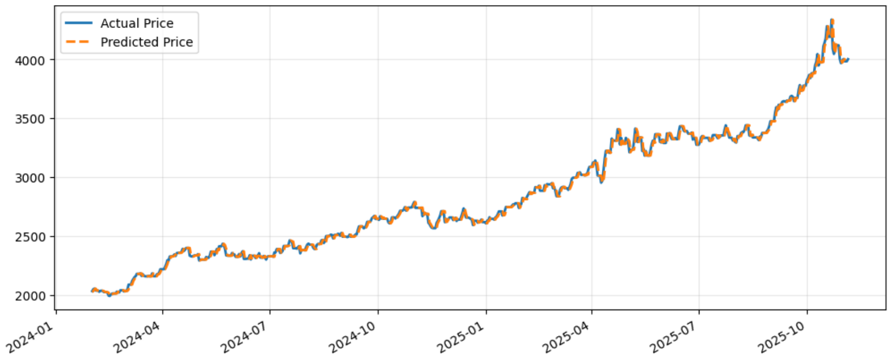



Figure 8. ARIMA-GARCH Model for Gold

In Figure 8, the graph shows a comparison between the actual price and the price predicted using ARIMA (1, 1, 1) - GARCH (1, 1). The blue line represents the actual price, while the orange line represents the predicted price. The two lines are very similar, indicating that this model successfully follows the price movement pattern well throughout the tested period with an RMSE of 30.0556.

### 3.7. LSTM Model

The LSTM (Long Short-Term Memory) model is a recurrent neural network architecture specifically designed to handle time series data with the ability to remember long-term patterns. Unlike traditional statistical methods, LSTM has the advantage of learning complex non-linear relationships between variables. In this study, LSTM was trained using price features such as Open, High, Low, Close, and Volume with Close as the prediction target. This approach allows the model to capture market dynamics directly from historical data without assuming a specific distribution.

This study evaluates the model using a walk-forward validation scheme, which provides a more realistic picture of the model's performance because at each time step, the model is only trained using historical data up to that point without looking at future data. This walk-forward validation is capable of producing predictions that are close to real market conditions, especially for highly dynamic assets such as Bitcoin and Gold.

Bitcoin exhibits sharp price swings, driven by rapid shifts in liquidity, market sentiment, and speculative behaviour. These characteristics make the dataset highly volatile and difficult to model using purely linear frameworks. To handle these characteristics, a model that is able to read actual market trends is needed, and LSTM, with its advantage in learning non-linear data patterns in time series, is the right choice. Through this approach, price features such as Open, High, Low, and Volume are entered as inputs so that the neural network can recognize price movement patterns both when the market is volatile and when price movements are relatively stable. The LSTM architecture makes predictions more adaptive and responsive to sudden market changes. As a result, the LSTM model is able to follow the actual Bitcoin price pattern quite well despite extreme volatility.


Figure 9. LSTM Model for Bitcoin

In Figure 9, the orange prediction line follows the actual data trend in blue with a fairly good distance. The model's predictions appear to be able to capture the general direction of movement despite some lag in periods of extreme change. This shows that LSTM has successfully learned both short-term and long-term Bitcoin price patterns directly from historical features. Overall, this model produces an RMSE value of 3795.9114.

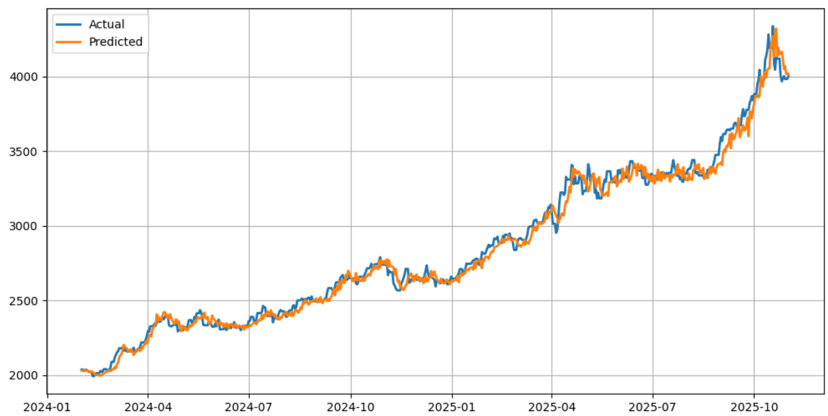



Figure 10. LSTM Model for Gold

Gold exhibits a noticeably smoother and more stable price trajectory compared with Bitcoin. Movements tend to unfold gradually, and major deviations are typically tied to macroeconomic shifts such as changes in interest rates or inflation expectations. Because of this stability, the LSTM primarily

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

focuses on learning gradual transitions and long-term structure rather than sudden shocks. Through this approach, price features such as Open, High, Low, Close, and Volume are included as inputs so that the neural network can recognize price movement patterns both when the market is volatile and when price movements are relatively stable. The LSTM architecture makes predictions more adaptive and responsive to sudden market changes. As a result, the LSTM model is able to follow the actual price patterns of Bitcoin quite well despite facing extreme volatility.

In Figure 10, it can be seen that the predicted line with orange follows the actual data with blue very well, especially in the long trending phase from early 2024 to mid-2025. The predictions look stable, smooth, and responsive, suggesting that LSTM is able to study the gradual but consistent growth patterns of gold prices.

When there was a surge in the price of gold in mid-2025 triggered by an increase in demand for hedge assets, the LSTM model managed to capture the acceleration of the trend quite accurately. This indicates that even though the movement of gold is calmer than Bitcoin, LSTM is still able to adjust its predictions to changes in market conditions through historical pattern learning. Overall, the LSTM model's performance on gold data is consistent, yielding an RMSE value of 58.0450.

### 3.8. GARCH-LSTM Hybrid Model

The GARCH-LSTM Hybrid Model combines the strengths of the GARCH statistical model and the non-linear learning capabilities of LSTM. In the first stage, the GARCH model will model daily volatility in extremely fluctuating data so that it can produce a series of volatility data that reflects the level of market uncertainty at each point in time. This volatility is then combined with other price features such as Open, High, Low, Close, and Volume, with Close used as the target model for the LSTM model. This approach is used for high-risk assets because it can capture both short-term and long-term volatility patterns.

This study evaluates the model using a walk-forward validation scheme, which provides a more realistic picture of the model's performance because at each time step, the model is only trained using historical data up to that point without looking at future data. This walk-forward validation is capable of producing predictions that are close to real market conditions, especially for highly dynamic assets such as Bitcoin and Gold.

In Bitcoin data, movements are highly volatile, so the ups and downs are influenced not only by long-term trends but also by short-term volatility dynamics that change very quickly. To handle these characteristics, a model is needed that can read actual market trends, and a hybrid GARCH-LSTM model combines the ability of GARCH to model volatility with the advantage of LSTM in learning non-linear data patterns in time series. Through this approach, the daily volatility of the GARCH estimation results is included as an additional feature in the LSTM so that the neural network can recognize when the market is experiencing high volatility or when price movements are relatively stable. The integration of these two models makes predictions more adaptive and responsive to sudden market changes. As a result, the hybrid model is able to follow the actual Bitcoin price patterns more stably and closely than a single model.

Table 12. GARCH Model Evaluation on Bitcoin Data

| ($\alpha$, $\beta$) | AIC |
|---|---|
| (1, 1) | 10833.61 |
| (2, 1) | 10835.44 |
| (1, 2) | 10835.73 |

Based on Table 12, in the Bitcoin data, the GARCH(1, 1) model was selected as the best model based on the lowest AIC value. The estimation results show that the $\alpha + \beta$ parameter is close to 1, which means that Bitcoin volatility is very persistent. The fluctuating volatility information then becomes important input for LSTM. When LSTM receives volatility features from GARCH, the model is able to distinguish between calm and volatile market periods, making predictions more adaptive and closer to actual patterns.


Figure 11. GARCH-LSTM Hybrid Model for Bitcoin

As shown in Figure 11, the orange prediction line closely follows the blue actual data trend with minimal deviation. The model's predictions appear stable but remain responsive to changes in price direction. This indicates that the integration of volatility signals from GARCH successfully enhances the LSTM's ability to recognize both short-term and long-term patterns in Bitcoin prices simultaneously. Overall, the combination of the two models does not reduce prediction errors compared to single models such as ARIMA and pure LSTM. This is in line with the RMSE result of this model, which is 3205.3953.

Gold price movements tend to be more stable than Bitcoin, but still show an up-and-down pattern influenced by global economic conditions, inflation, interest rates, and market sentiment toward safe-haven assets. This relative stability makes gold volatility lower, but it is still important to model because small changes in volatility can provide important signals about the strength of price trends. Therefore, the Hybrid GARCH-LSTM approach is still used to maximize the model's ability to combine short-term volatility dynamics with long-term non-linear patterns.

Table 13. GARCH Model Evaluation on Gold Data

| (α, β) | AIC |
|---|---|
| (1, 1) | 5360.63 |
| (2, 1) | 5362.36 |
| (1, 2) | 5362.38 |

Table 13 shows that the best volatility model was done by comparing several GARCH candidates. Based on the lowest AIC value, GARCH (1, 1) was again the best model for Gold data. Although parameters such as $\omega$, $\alpha$, and $\beta$ are not statistically significant and have wide confidence intervals, this is normal for assets with lower volatility and smoother movements than Bitcoin. Nevertheless, the volatility output from GARCH remains useful as an additional feature in LSTM to distinguish between trend-following and price consolidation periods.

In Figure 12, it can be seen that the predicted line in orange follows the actual data in blue very well, especially during the long trending phase from early 2024 to mid-2025. The predictions appear stable, smooth, and responsive, indicating that LSTM is capable of learning the gradual but consistent price growth patterns of Gold.

When there was a surge in gold prices in mid-2025 triggered by increased demand for hedge assets, the hybrid model successfully captured the acceleration of the trend accurately.


Figure 12. GARCH-LSTM Hybrid Model for Gold

This indicates that even though gold volatility is calmer than Bitcoin, the volatility signals from GARCH still help LSTM in adjusting the sensitivity of predictions to changes in market conditions. Overall, the hybrid model's performance on gold data is consistent with findings on Bitcoin data, producing predictions that are no more stable than non-hybrid models, with an RMSE value of 49.7434.

### 3.9. Interpretation

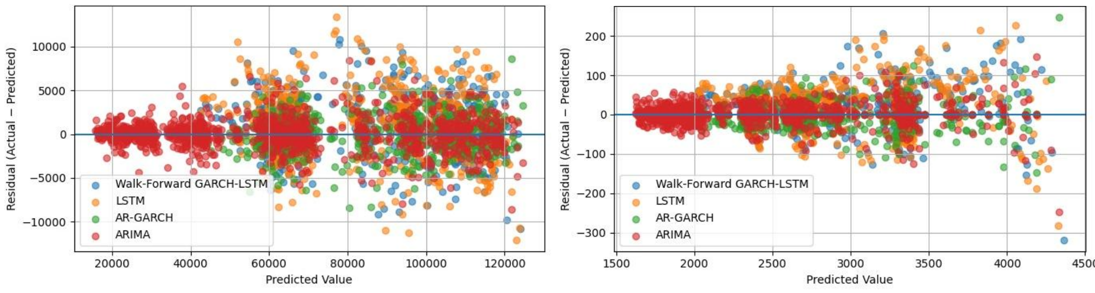
Table 14. Comparative study

| Algoritma Model | Dataset | RMSE |
|---|---|---|
| ARIMA | Bitcoin | 2035.8041 |
| ARIMA | Gold | 30.0767 |
| ARIMA-GARCH | Bitcoin | 2034.55 |
| ARIMA-GARCH | Gold | 30.0556 |
| LSTM | Bitcoin | 3795.9114 |
| LSTM | Gold | 58.0450 |
| GARCH-LSTM | Bitcoin | 3205.3953 |
| GARCH-LSTM | Gold | 49.7434 |

The results of the performance evaluation of the prediction model on Bitcoin data are shown through the RMSE metric as presented in Table 14. In general, the hybrid approach shows better performance than the single model on both types of asset data. In Bitcoin data, which has high volatility, the ARIMA model was able to produce an RMSE of 2035.8041, while the integration of GARCH in ARIMA produced a model with a relatively equivalent RMSE value of 2034.55. Although the difference is not significant, the existence of the GARCH component plays a role in capturing conditional volatility that is not modeled by pure ARIMA. Meanwhile, the LSTM model showed the highest RMSE value of 3795.9114, indicating the limitations of LSTM in modeling extreme fluctuations in crypto data. Integrating GARCH into LSTM reduced the error to 3205.3953, indicating that the hybrid GARCH-LSTM approach provides improved performance compared to pure LSTM.

For Gold data, the ARIMA model produced an RMSE of 30.0767, while the ARIMA-GARCH model produced the lowest RMSE value of 30.0556. These results indicate that even though Gold's volatility is relatively low compared to Bitcoin, GARCH integration still contributes positively to prediction accuracy. The LSTM model again showed the lowest performance with an RMSE value of 58.0450, while the GARCH-LSTM model was able to reduce the RMSE value to 49.7434. Thus, the results on the Gold data show that the hybrid approach is consistently superior to using non-hybrid models.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555




(a) (b)

Figure 13. (a) Bitcoin and (b) Gold Residual Plot

Residual analysis was performed to evaluate the stability and characteristics of prediction errors in both datasets, as shown in Figure 13. In the Bitcoin data, the residuals of the ARIMA and LSTM models showed a fairly clear pattern of heteroscedasticity, marked by an increase in residual variance at higher prediction values. A similar pattern also occurs in the Gold data, albeit with a smaller residual scale. In contrast, the ARIMA-GARCH and GARCH-LSTM hybrid models show a more symmetrical distribution of residuals around zero with relatively more controlled variance in both datasets. This indicates that hybrid models are better able to capture volatility dynamics and produce more statistically stable predictions.

Overall, the evaluation results on Bitcoin and Gold data show that the hybrid model approach provides advantages over single models, especially in terms of prediction stability and residual characteristics. In Bitcoin data, the ARIMA-GARCH model produces a slightly lower RMSE value compared to the ARIMA model. However, GARCH integration still provides advantages in modeling volatility, as reflected in a more controlled and symmetrical residual distribution. Meanwhile, in the Gold data, the ARIMA-GARCH model consistently produced the lowest RMSE value compared to ARIMA. In addition, the GARCH-LSTM model showed a significant improvement in performance compared to LSTM in both datasets, both with high and relatively low volatility. These findings confirm that separating the modeling of the mean and volatility components can improve the stability and reliability of financial data predictions, even though the increase in numerical accuracy is not always dominant across all assets.

## 4. DISCUSSIONS

The analysis results show that the price movements of Bitcoin and Gold during the 2020-2025 period are characterized by unstable volatility changes over time. In some periods, price fluctuations were relatively low, while in other periods, there were significant spikes, especially in the price of Bitcoin. These rapidly changing market conditions meant that the ARIMA approach, which assumes constant variance, was less effective in fully describing the dynamics of risk, even though the model provided good AIC and BIC values.

These limitations indicate that time series modeling of financial data requires an approach that not only focuses on modeling average price values but also represents time-varying volatility dynamics. This aspect becomes increasingly important for assets with high volatility and those experiencing market regime changes, such as those that occurred during the COVID-19 pandemic and the crypto winter phase. Therefore, the use of models that consider volatility dynamics is relevant to support price forecasting and risk analysis.

This study examines two hybrid approaches, namely ARIMA-GARCH and GARCH-LSTM, in an effort to address these issues. The ARIMA-GARCH model combines linear modeling with volatility

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

estimates that change over time, thereby providing a more stable picture of risk in the short term. Meanwhile, the GARCH-LSTM model uses volatility information as additional input for the LSTM, enabling this model to capture non-linear patterns and long-term relationships in price movements. These differences in how they work demonstrate that two hybrid models have complementary functions in modeling market dynamics.

Table 15. Comparative study

| Author | Dataset | Method | RMSE |
|---|---|---|---|
| Adam et al. (2024)[62] | Goto.jk | ARIMA | 3.259 |
| | | LSTM | 3.843 |
| Phung et al. (2023)[38] | Bitcoin | ARIMA | 15380.68 |
| | | ARIMA-GARCH | 14142.99 |
| Mutinda dan Langat (2024)[63] | Stock Airtel | LSTM | 3.9749 |
| | | GARCH-LSTM | 0.2002 |
| Proposed Method | Gold | ARIMA | 30.07 |
| | | ARIMA-GARCH | 30.06 |
| | | LSTM | 58.04 |
| | | GARCH-LSTM | 49.74 |
| Proposed Method | Bitcoin | ARIMA | 2035.80 |
| | | ARIMA-GARCH | 2034.55 |
| | | LSTM | 3795.91 |
| | | GARCH-LSTM | 3205.40 |

To validate this study, a comparison was made with previous studies summarized in Table 15. Previous studies show a consistent pattern that hybrid and deep learning models tend to produce better prediction performance than pure ARIMA models. Phung et al. (2024) showed that ARIMA-GARCH produced lower RMSE values than ARIMA on Bitcoin data. Mutinda and Langat (2024) also reported that GARCH-LSTM provided better performance than single LSTM models. A similar RMSE comparison pattern was also found in this study, where hybrid models tended to produce lower prediction errors than single models. On the other hand, Adam et al. (2024) showed that GOTO.JK stock data, the ARIMA model outperformed LSTM, indicating that model performance still depends on the characteristics of data used.

It should be emphasized that the RMSE values between studies cannot be directly compared due to differences in price scales, data periods, and the features and configurations of the models used. This study uses the 2020-2025 time range, which covers periods of high volatility, so the forecasting challenges faced are different from those in studies that use more stable data periods. Nevertheless, the consistency of the RMSE comparison result shows that the hybrid approach consistently provides relative performance improvements over single models in various data contexts.

Overall, this study shows that a hybrid approach combining statistical models and machine learning can produce a forecasting framework that is more adaptive to changing market conditions. The use of volatility information in the model learning process enables artificial intelligence-based forecasting systems to represent risk dynamics more accurately. From an informatics perspective, this study contributes to the development of a computational framework that integrates statistical modeling and deep learning in complex financial time series analysis. This approach is relevant to support the development of AI forecasting and risk analytics systems, both for assets with high volatility such as Bitcoin and for relatively more stable assets such as Gold.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

## 5. CONCLUSION

This study reveals that the price dynamics of Bitcoin and Gold in the 2020-2025 period have non-stationary patterns with high volatility, both of which have non-linear structures that are difficult to capture by traditional statistical models or single models. The ARIMA model was used as a basic linear approach, but ARIMA is only capable of recognizing short-term dependencies and produces limited prediction performance. This is demonstrated by the RMSE values of 2035.80 for Bitcoin and 30.07 for Gold, as well as the relatively high AIC values of 25265.35 and 12427.80 for the Bitcoin and Gold AR s, respectively. ARIMA tends to flatten prediction patterns, especially when there are extreme fluctuations in the data.

Meanwhile, deep learning-based models, such as LSTM, do not provide better results for long-term patterns and non-linear relationships in the data. The RMSE values of the LSTM model for Bitcoin and Gold are 3795.91 and 58.04, respectively. Based on these RMSE values, it can be seen that even though the model has a long-term memory architecture, this method is still not sufficient to handle data with high noise, such as extreme fluctuations in the data.

Conversely, the findings from this study using a hybrid approach reveal results that effectively improve the model's performance in recognizing the movement patterns of both assets. Hybrid ARIMA-GARCH shows an improvement in recognizing variance changes, with an RMSE of 2034.55 for Bitcoin and 30.05 for Gold, as well as lower AIC values for both assets compared to ARIMA, namely 24003.65 for Bitcoin and 12432.64 for Gold. The most significant improvement in performance was seen in the GARCH-LSTM model, especially for Gold data. The GARCH-LSTM model was able to utilize information related to residual volatility from GARCH as an additional feature that was crucial for LSTM training. The evaluation results indicate that the combination of neural network memory with statistical volatility modeling successfully reduced the RMSE of Gold significantly to 49.7434. Meanwhile, for models with extreme fluctuations such as Bitcoin, the GARCH-LSTM hybrid model was able to reduce the error to 3205.40.

The main impact of this research is to provide a scalable contribution to deep learning-based asset price movement prediction models. This research validates that hybrid architecture is a scalable architecture for various asset characteristics by combining statistical components to understand volatility information and deep learning components to capture complex non-linear patterns. Based on these findings, further research is recommended to integrate the Transformer model and multivariate analysis by adding exogenous variables, such as macroeconomic components and market sentiment, to improve prediction accuracy.

## CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

## ACKNOWLEDGEMENT

The author sincerely extends his highest appreciation to Sebelas Maret University for the research funding provided through the Research Grant Group program, as stated in contract No. 371/UN27.22/PT.01.03/2025. The author also expresses gratitude for the support, facilities, academic atmosphere, encouragement, and essential resources offered by the Faculty of Information Technology and Data Science throughout the completion of this research.

## REFERENCES

[1] N. Maharana, A. K. Panigrahi, and S. K. Chaudhury, “Volatility Persistence and Spillover Effects of Indian Market in the Global Economy: A Pre- and Post-Pandemic Analysis Using VAR-

BEKK-GARCH Model," *J. Risk Financ. Manag.*, vol. 17, no. 7, 2024, doi: 10.3390/jrfm17070294.

[2] N. Fabris, "Monetary Policy Between Stability and Growth," *J. Cent. Bank. Theory Pract.*, vol. 13, no. 1, pp. 27–42, 2024, doi: 10.2478/jcbtp-2024-0002.

[3] J. Zheng, B. Wen, Y. Jiang, X. Wang, and Y. Shen, "Risk spillovers across geopolitical risk and global financial markets," *Energy Econ.*, vol. 127, no. PA, p. 107051, 2023, doi: 10.1016/j.eneco.2023.107051.

[4] J. Chen, "Analysis of Bitcoin Price Prediction Using Machine Learning," *J. Risk Financ. Manag.*, vol. 16, no. 1, 2023, doi: 10.3390/jrfm16010051.

[5] S. Alshammari and H. Obeid, "Analyzing commodity futures and stock market indices: Hedging strategies using asymmetric dynamic conditional correlation models," *Financ. Res. Lett.*, vol. 56, no. May, p. 104081, 2023, doi: 10.1016/j.frl.2023.104081.

[6] N. Azizah, A. Rahmadian, and W. Sirait, "Prediksi Harga Saham Bank BCA, BNI, dan BRI serta Komposisi Portofolio Maksimal Ketiga Saham Berbasis Regresi Linier dan Clustering," vol. 15, no. 1, pp. 32–39, 2025, doi: 10.35200/ex.v15i1.151.

[7] J. Y. Le Chan, S. W. Phoong, S. Y. Phoong, W. K. Cheng, and Y. L. Chen, "The Bitcoin Halving Cycle Volatility Dynamics and Safe Haven-Hedge Properties: A MSGARCH Approach," *Mathematics*, vol. 11, no. 3, 2023, doi: 10.3390/math11030698.

[8] V. Terraza, A. Boru İpek, and M. M. Rounaghi, "The nexus between the volatility of Bitcoin, gold, and American stock markets during the COVID-19 pandemic: evidence from VAR-DCC-EGARCH and ANN models," *Financ. Innov.*, vol. 10, no. 1, 2024, doi: 10.1186/s40854-023-00520-3.

[9] T. Conlon, S. Corbet, and L. Oxley, "Investor Sentiment, Unexpected Inflation, and Bitcoin Basis Risk," *Journal of Futures Markets.*, vol. 44, no. 11, pp. 1807–1831, 2024, doi: 10.1002/fut.22541.

[10] M. R. Maulana, "Bitcoin dan Konsep Uang Digital: Tinjauan Historis dan Teoritis," *Waralaba Journal of Economics and Business.*, vol. 1, no. 2, pp. 69–78, 2024, doi: 10.61590/waralaba.v1i2.144.

[11] U. Q. Bajra and F. Aliu, "Deciphering the cryptocurrency conundrum: Investigating speculative characteristics and volatility," *Financ. Res. Lett.*, vol. 58, no. PC, p. 104589, 2023, doi: 10.1016/j.frl.2023.104589.

[12] M. Anas, E. Bouri, and S. J. H. Shahzad, "A Bibliometric analysis of literature on hedge and safe haven assets," *Journal of Economic Surveys.*, vol. 39, no. 5, pp. 1852–1882, 2025, doi: 10.1111/joes.12677.

[13] M. Ryan, S. Corbet, and L. Oxley, "Is gold always a safe haven?," *Financ. Res. Lett.*, vol. 64, no. March, p. 105438, 2024, doi: 10.1016/j.frl.2024.105438.

[14] A. Wulan, S. Z. Ma'mun, and M. S. Maksar, "ANALISIS ASET SAFE-HAVEN UNTUK PASAR SAHAM DI INDONESIA : STUDI PADA EMAS DAN BITCOIN," *JIMEA | Jurnal Ilmiah MEA (Manajemen, Ekonomi, dan Akuntansi)*, vol. 8, no. 1, pp. 753–769, 2024, doi: 10.31955/mea.v8i1.3737.

[15] L. Rubio, A. Palacio Pinedo, A. Mejía Castaño, and F. Ramos, "Forecasting volatility by using wavelet transform, ARIMA and GARCH models," *Eurasian Econ. Rev.*, vol. 13, no. 3–4, pp. 803–830, 2023, doi: 10.1007/s40822-023-00243-x.

[16] N. W. N. Muhammad Iqbal, "Prediksi Harga Saham Harian PT BTPN Syariah Tbk Menggunakan Model Arima dan Model Garch," *Jurnal Ilmiah Ekonomi Islam*, vol. 7, no. 03, pp. 1573–1580, 2021, doi: dx.doi.org/10.29040/jiei.v7i3.2795 1.

[17] K. Macharia, E. Atitwa, D. Mugo, and M. Kawira, "Modeling stock price trends and volatility in emerging markets using ARIMA and GARCH approaches," *International Journal of Advanced and Applied Sciences.*, vol. 12, no. 7, pp. 134–143, 2025, doi: 10.21833/ijaas.2025.07.013.

[18] C. Cappello, A. Congedi, S. De Iaco, and L. Mariella, "Traditional Prediction Techniques and Machine Learning Approaches for Financial Time Series Analysis," *Mathematics*, vol. 13, no. 3, pp. 1–21, 2025, doi: 10.3390/math13030537.

[19] D. N. Fadhilah, Kankan Parmikanti, and Budi Nurani Ruchjana, "Peramalan Return Saham

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

Subsektor Perbankan Menggunakan Model ARIMA-GARCH," *Jurnal Fourier*, vol. 13, no. 1, pp. 1–19, 2024, doi: 10.14421/fourier.2024.131.1-19.

[20] M. R. Kabir, D. Bhadra, M. Ridoy, and M. Milanova, "LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting," *Sci*, vol. 7, no. 1, 2025, doi: 10.3390/sci7010007.

[21] K. Sahaja and L. Pudukarapu, "Financial Time Series Prediction under COVID-19 Pandemic crisis with Long Short-Term Memory (LSTM) Network," vol. 6, no. 6, pp. 1–4, 2022, doi: https://doi.org/10.1057/s41599-023-02042-w.

[22] N. Hadi, S. Alam, and I. Kurniawan, "Penerapan Model Lstm (Long Short-Term Memory) Dalam Peramalan Inflasi Month on Month Di Indonesia," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 5, pp. 9009–9015, 2025, doi: 10.36040/jati.v9i5.15119.

[23] S. Pan, S. Long, Y. Wang, and Y. Xie, "Nonlinear asset pricing in Chinese stock market: A deep learning approach," *Int. Rev. Financ. Anal.*, vol. 87, no. March, 2023, doi: 10.1016/j.irfa.2023.102627.

[24] R. Liu, Y. Jiang, and J. Lin, "Forecasting the Volatility of Specific Risk for Stocks with LSTM," *Procedia Comput. Sci.*, vol. 202, pp. 111–114, 2022, doi: 10.1016/j.procs.2022.04.015.

[25] D. T. Lim, K. W. Goh, Y. W. Sim, K. Mokhtar, and S. Thinagar, "Estimation of stock market index volatility using the GARCH model: Causality between stock indices," *Asian Econ. Financ. Rev.*, vol. 13, no. 3, pp. 162–179, 2023, doi: 10.55493/5002.v13i3.4738.

[26] A. García-Medina and E. Aguayo-Moreno, *LSTM–GARCH Hybrid Model for the Prediction of Volatility in Cryptocurrency Portfolios*, vol. 63, no. 4. Springer US, 2024. doi: 10.1007/s10614-023-10373-8.

[27] S. Yang, "Pandemic, policy, and markets: insights and learning from COVID-19's impact on global stock behavior," *Empir. Econ.*, vol. 68, no. 2, pp. 555–583, 2025, doi: 10.1007/s00181-024-02648-2.

[28] S. Alam, M. Murshed, and P. Manigandan, "Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools," no. January, 2020, doi: https://doi.org/10.1016/j.resourpol.2023.103342.

[29] A. S. Hasanov, A. U. Burkhanov, B. Usmonov, N. S. Khajimuratov, and M. M. qizi Khurramova, "The role of sudden variance shifts in predicting volatility in bioenergy crop markets under structural breaks," *Energy*, vol. 293, no. July 2023, p. 130535, 2024, doi: 10.1016/j.energy.2024.130535.

[30] D. G. Baur and L. T. Hoang, "A crypto safe haven against Bitcoin," *Financ. Res. Lett.*, vol. 38, no. December 2019, p. 101431, 2021, doi: 10.1016/j.frl.2020.101431.

[31] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, 2020, doi: 10.1007/s00521-020-04867-x.

[32] D. A. D. and W. A. Fuller, "Dickey_Fuller_Distribution of the estimators for autoregressive time series.pdf," *Journal of the American Statistical Association*, vol. Vpl.74, no. No 366 (Jun,1979). pp. 427–341, 1979.

[33] C. Fowler, X. Cai, J. T. Baker, J. P. Onnela, and L. Valeri, "Testing unit root non-stationarity in the presence of missing data in univariate time series of mobile health studies," *J. R. Stat. Soc. Ser. C Appl. Stat.*, vol. 73, no. 3, pp. 755–773, 2024, doi: 10.1093/jrsssc/qlae010.

[34] W. Yiming, L. Xun, M. Umair, and A. Aizhan, "COVID-19 and the transformation of emerging economies: Financialization, green bonds, and stock market volatility," *Resour. Policy*, vol. 92, no. April 2023, p. 104963, 2024, doi: 10.1016/j.resourpol.2024.104963.

[35] R. Cascade-correlation and N. S. Chunking, "Neural Comput.-1997-LSTM-Long short term memory," vol. 9, no. 8, pp. 1–32, 1997.

[36] S. B. Primananda and S. M. Isa, "Forecasting Gold Price in Rupiah using Multivariate Analysis with LSTM and GRU Neural Networks," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 6, no. 2, pp. 245–253, 2021, doi: 10.25046/aj060227.

[37] I. W. R. Pinastawa, M. G. Pradana, D. S. Setiawan, and A. Izzety, "Comparison of ARIMA and GRU Methods in Predicting Cryptocurrency Price Movements," *Sinkron*, vol. 9, no. 1, pp. 96–105, 2025, doi: 10.33395/sinkron.v9i1.14235.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

[38] Q. Phung Duy, O. Nguyen Thi, P. H. Le Thi, H. D. Pham Hoang, K. L. Luong, and K. N. Nguyen Thi, "Estimating and forecasting bitcoin daily prices using ARIMA-GARCH models," *Bus. Anal. J.*, vol. 45, no. 1, pp. 11–23, 2024, doi: 10.1108/baj-05-2024-0027.

[39] G. E. P. Box and D. A. Pierce, "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *J. Am. Stat. Assoc.*, vol. 65, no. 332, p. 1509, 1970, doi: 10.2307/2284333.

[40] G. Dudek, P. Fiszeder, P. Kobus, and W. Orzeszko, "Forecasting cryptocurrencies volatility using statistical and machine learning methods: A comparative study," *Appl. Soft Comput.*, vol. 151, no. December 2023, 2024, doi: 10.1016/j.asoc.2023.111132.

[41] Z. Su, "Gold Price Forecast Based on ARIMA and ETS Models," *Adv. Econ. Manag. Polit. Sci.*, vol. 143, no. 1, pp. 185–189, 2024, doi: 10.54254/2754-1169/2024.ga18972.

[42] T. Pang and Y. Zhao, "On GARCH and Autoregressive Stochastic Volatility Approaches for Market Calibration and Option Pricing," *Risks*, vol. 13, no. 2, pp. 1–24, 2025, doi: 10.3390/risks13020031.

[43] E. S. Nugraha and C. Alvina, "the Application of Standard Generalized Autoregressive Conditional Heteroscedasticity (Sgarch) Model in Forecasting the Stock Price of Barito Pacific," *Barekeng*, vol. 18, no. 2, pp. 0849–0862, 2024, doi: 10.30598/barekengvol18iss2pp0849-0862.

[44] K. Kakade, I. Jain, and A. K. Mishra, "Value-at-Risk forecasting: A hybrid ensemble learning GARCH-LSTM based approach," *Resour. Policy*, vol. 78, no. July, p. 102903, 2022, doi: 10.1016/j.resourpol.2022.102903.

[45] C. Liu, C. Wang, M.-N. Tran, and R. Kohn, "Deep Learning Enhanced Realized GARCH," pp. 1–46, 2023, doi: https://doi.org/10.48550/arXiv.2302.08002.

[46] A. Saini, R. K. Singh, and P. Sinha, "Forecasting gold price using hybrid deep neural network LSTM-autoencoder," *Discov. Artif. Intell.*, vol. 5, no. 1, 2025, doi: 10.1007/s44163-025-00464-w.

[47] S. Setyowibowo, M. As'ad, S. Sujito, and E. Farida, "Forecasting of Daily Gold Price using ARIMA-GARCH Hybrid Model," *J. Ekon. Pembang.*, vol. 19, no. 2, pp. 257–270, 2022, doi: 10.29259/jep.v19i2.13903.

[48] F. Jiang, D. Li, and K. Zhu, "Adaptive inference for a semiparametric generalized autoregressive conditional heteroskedasticity model," *J. Econom.*, vol. 224, no. 2, pp. 306–329, 2021, doi: 10.1016/j.jeconom.2020.10.007.

[49] S. Muneer, C. C. Leal, and B. Oliveira, "Analyzing Volatility Patterns of Bitcoin Using the GARCH Family Models," *Oper. Res. Forum*, vol. 6, no. 2, pp. 1–13, 2025, doi: 10.1007/s43069-025-00482-5.

[50] M. Adnan and M. A. Ahmed, "Predicting stock market indicators using a hybrid model (ARIMA - GARCH): An analytical study of some Gulf financial market indicators," *J. Econ. Adm. Sci.*, vol. 31, no. 148, pp. 130–143, 2025, doi: 10.33095/tkrnay54.

[51] A. Murari, R. Rossi, L. Spolladore, M. Lungaroni, P. Gaudio, and M. Gelfusa, *A practical utility-based but objective approach to model selection for regression in scientific applications*, vol. 56, no. s2. Springer Netherlands, 2023. doi: 10.1007/s10462-023-10591-4.

[52] A. Singh, T. Tripathi, R. Ranjan, and A. K. Tiwari, "Time series forecasting of infant mortality rate in India using Bayesian ARIMA models," *BMC Public Health*, vol. 25, no. 1, 2025, doi: 10.1186/s12889-025-24125-w.

[53] Y. Zhang and G. Meng, "Simulation of an Adaptive Model Based on AIC and BIC ARIMA Predictions," *J. Phys. Conf. Ser.*, vol. 2449, no. 1, 2023, doi: 10.1088/1742-6596/2449/1/012027.

[54] C. Agiakloglou and A. Tsimpanos, *Evaluating the performance of AIC and BIC for selecting spatial econometric models*, vol. 4, no. 1. Springer International Publishing, 2023. doi: 10.1007/s43071-022-00030-x.

[55] S. Albahli and G. N. A. H. Yar, "Defect prediction using Akaike and Bayesian information criterion," *Comput. Syst. Sci. Eng.*, vol. 41, no. 3, pp. 1117–1127, 2022, doi: 10.32604/csse.2022.021750.

[56] L. Zhao, Z. Li, and L. Qu, "Forecasting of Beijing PM2.5 with a hybrid ARIMA model based on integrated AIC and improved GS fixed-order methods and seasonal decomposition," *Heliyon*, vol. 8, no. 12, p. e12239, 2022, doi: 10.1016/j.heliyon.2022.e12239.

Vol. 7, No. 3, June 2026, Page. 2350-2375
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5555

[57] N. A. M. Ikbal, S. A. Halim, and N. Ali, "Estimating Weibull Parameters Using Maximum Likelihood Estimation and Ordinary Least Squares: Simulation Study and Application on Meteorological Data," *Math. Stat.*, vol. 10, no. 2, pp. 269–292, 2022, doi: 10.13189/ms.2022.100201.

[58] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, no. 14, pp. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.

[59] N. L. M. Jailani *et al.*, "Investigating the Power of LSTM-Based Models in Solar Energy Forecasting," *Processes*, vol. 11, no. 5, 2023, doi: 10.3390/pr11051382.

[60] H. Luo and S. G. Paal, "A novel outlier-insensitive local support vector machine for robust data-driven forecasting in engineering," *Eng. Comput.*, vol. 39, no. 5, pp. 3671–3689, 2023, doi: 10.1007/s00366-022-01781-9.

[61] J. Guo, Y. Liu, Q. Zou, L. Ye, S. Zhu, and H. Zhang, "Study on optimization and combination strategy of multiple daily runoff prediction models coupled with physical mechanism and LSTM," *J. Hydrol.*, vol. 624, no. April, p. 129969, 2023, doi: 10.1016/j.jhydrol.2023.129969.

[62] H. A. Adam *et al.*, "Comparison of ARIMA and LSTM Models in Stock Price Forecasting: A Case Study of GOTO.JK," *J. Informatics Telecommun. Eng.*, vol. 7, no. 1, pp. 102–111, 2023, doi: 10.31289/jite.v8i1.11841.

[63] J. K. Mutinda and A. K. Langat, "Stock price prediction using combined GARCH-AI models," *Sci. African*, vol. 26, no. September, p. e02374, 2024, doi: 10.1016/j.sciaf.2024.e02374.