Vol. 12 No.1 Tahun 2026, pp 1-9
Jurnal Teknologi dan Manajemen Informatika
http:// http://jurnal.unmer.ac.id/index.php/jtmi
P-ISSN: 1693-6604 E-ISSN: 2580-8044

Gold Price Prediction Using a Hybrid LSTM and Prophet Model Enhanced with News Sentiment Analysis

Yessica Yamin [1], Robet [2], Hendri [3]

STMIK Time, Medan, Indonesia

Article Info

Article History
Received : 15-10-2025
Revised : 24-11-2025
Accepted : 05-01-2026

Keywords
Gold price forecasting;
LSTM;
Prophet;
Hybrid model;
Sentiment analysis

Corresponding Author
Yessica Yamin
STMIK Time, Medan, Indonesia
yessicayamin01@gmail.com

ABSTRACT

This study aims to improve the accuracy of gold price forecasting by combining statistical and deep learning methods with sentiment analysis. Three models were developed and compared: (1) a pure Long Short-Term Memory (LSTM) model, (2) a hybrid LSTM + Prophet model, and (3) a hybrid LSTM + Prophet + Sentiment model. The datasets consisted of daily gold prices and financial news sentiment from 2013 to 2023. Each model was evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2), and Mean Absolute Percentage Error (MAPE). The pure LSTM model achieved an R^2 of 0.9375, while the hybrid LSTM + Prophet model improved performance to 0.9394 with lower error rates. The integration of sentiment data resulted in stable but not significantly higher accuracy. Overall, the hybrid LSTM + Prophet model produced the best results, confirming that combining statistical trend decomposition with deep learning effectively enhances forecasting stability and interpretability for financial time series data such as gold prices.

INTRODUCTION

Gold has long been recognized as one of the most stable and valuable investment instruments in the world due to its ability to preserve wealth against inflation and economic uncertainty. The price of gold serves as a crucial indicator for investor, reflecting global economic stability. Several factors, such as interest rates, exchange rates, geopolitical dynamics, and financial news, contribute to gold price fluctuations, making prediction a highly challenging task [1][2]. The instability of historical price patterns, influenced by numerous external and nonlinear factors, poses a major problem in forecasting. The ability to accurately predict gold price movements is highly beneficial for both individual investors and financial institutions in the modern data-driven investment era [3].

The volatile nature of gold prices, shaped by multiple external variables, complicates the forecasting process. Traditional statistical models such as Autoregressive Integrated Moving Average (ARIMA) and linear regression often fail to capture complex, nonlinear relationships and long-term temporal dependencies in time series data [4][5] . Therefore, an advanced predictive model capable of learning intricate temporal patterns is required. Among various types of Recurrent Neural Networks (RNNs), the LSTM model is particularly popular due to its strength in handling long-term dependencies in time series and sequential information [6]. LSTM is capable of processing sequential data and retaining long-term information, allowing it to capture historical patterns effectively. Numerous studies have demonstrated the superiority of LSTM in predicting sequential data such as stock prices, cryptocurrencies, and commodities like gold [7]. However, LSTM alone has limitations in capturing long-term trends and seasonal patterns, which are common in economic and financial data.

To address this limitation, the Prophet model, developed by Facebook, provides a complementary approach. Prophet is a time series forecasting model that automatically handles trend, seasonality, and event-based components, such as holidays or global economic events[8]. While LSTM requires extensive training and complex tuning, Prophet is relatively easy to implement and offers interpretable results. When combined into a hybrid framework, Prophet and LSTM can complement each other's strengths—Prophet captures long-term seasonal trends, whereas LSTM focuses on nonlinear short-term dynamics [9]. Previous research has shown that hybrid models combining LSTM with other methods, including Prophet, outperform single-model approaches, particularly in financial forecasting [10].

In addition to historical price data, market sentiment also significantly affects gold price movements. Investor perceptions are shaped by economic news, policy announcements, and media coverage, all of which influence gold demand and pricing behaviour. Recent studies suggest that sentiment analysis from financial text data, such as news articles, can enhance the accuracy of price forecasting models by reflecting the real-time psychological dynamics of the market [11]. Saputra's research demonstrated that integrating sentiment data with historical data improves model responsiveness to market sentiment fluctuations in the context of stocks, cryptocurrencies, and commodities [12]. Similarly, Darvishan's study on gold emphasized that investor sentiment and market volatility indicators strongly influence investment behaviour, leading to more accurate gold price and volatility forecasts [13].

Previous research has successfully applied LSTM to predict various commodities such as crude oil, cryptocurrency, and stock indices. Hybrid models that combine LSTM with methods such as CNN, GRU, or Prophet have also proven to improve predictive accuracy and model stability [14][15]. Moreover, several studies have incorporated sentiment analysis into financial forecasting—especially in stock and Bitcoin prediction—demonstrating enhanced predictive performance [16]. However, few studies have applied the hybrid LSTM + Prophet approach integrated with sentiment variables specifically for gold price forecasting in the Indonesian context. Most existing works remain focused on international financial markets or other asset classes.

Therefore, the objective of this research is to evaluate three predictive modelling approaches for gold price forecasting: (1) pure LSTM, (2) hybrid LSTM + Prophet, and (3) hybrid LSTM + Prophet + News Sentiment. Each model is evaluated using standard performance metrics—MSE, Root Mean Squared Error (RMSE), MAE, R-squared (R^2), and MAPE—on publicly available historical gold price data. This study aims to develop a more accurate and reliable gold price forecasting system that can serve as a valuable decision-support tool for investors, market analysts, and policymakers.

METHODS

Research Framework

This research was conducted through a structured framework consisting of six main stages: data collection, data preprocessing, model construction, model training performance evaluation, and comparative analysis. The complete research workflow is illustrated in Figure 1.

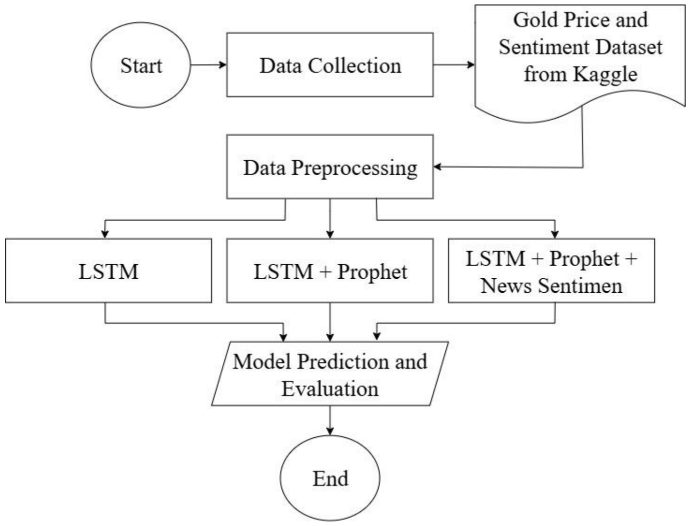

Figure 1. Research Process

The process begins with the collection of two datasets obtained from the Kaggle platform: (1) a historical gold price dataset and (2) a news sentiment dataset related to the gold market. Both datasets were preprocessed to ensure data consistency, followed by the construction and evaluation of three predictive models.

1. Pure LSTM: Focused solely on time-series learning from historical gold prices.
2. Hybrid LSTM + Prophet: Combined the Prophet model for trend and seasonality detection with the LSTM model for short-term dynamics.
3. Hybrid LSTM + Prophet + Sentiment: Integrated sentiment features with Prophet outputs to analyze the impact of investor psychology on price fluctuations.

Each model was evaluated separately using standardized performance metrics to determine accuracy and generalization capability.

Data Collection

Two datasets were used in this study, both retrieved from the Kaggle repository.

1. Gold Price Dataset
   This dataset contains daily gold price data from 2013 to 2023, including attributes such as date, opening price, closing price, highest and lowest price, trading volume, and daily percentage change [17].
2. News Sentiment Dataset
   This second dataset consists of global financial news articles related to gold, each labelled as *positive*, *negative*, or *neutral* according to the sentiment toward price movement [18]. This dataset provides additional psychological and contextual information from market narratives, enabling sentiment-enhanced forecasting.

Together, these datasets form the foundation for constructing both historical and sentiment based hybrid predictive models.

Data Preprocessing

The preprocessing stage aimed to clean, standardize, and integrate both datasets before model training. Several key operations were performed:

1. Data Cleaning and Normalization
   All missing or duplicated entries were removed. Numerical attributes such as price and volume were normalized using the *MinMaxScaler* technique to scale values between 0 and 1.
2. Date Formatting and Sorting
   The *Date* column was converted to the datetime format and ordered chronologically to preserve temporal relationships.
3. Sentiment Transformation

For the sentiment dataset, categorical labels were mapped into numerical scores: *positive = 1*, *neutral = 0*, and *negative = -1*. Daily averages of sentiment scores were computed to synchronize them with gold price data.

4. Dataset Integration
   Both datasets were merged using a left join on the *Date* column, ensuring each date in the gold price dataset was associated with the corresponding sentiment information.
5. Time Series Structuring
   Sliding windows with a *time_step* of 90 were applied to transform the data into sequential format, where each 90-day input predicts the gold price on the 91^{st} day.
6. Data Splitting
   The combined dataset was divided into training and testing subsets using an 80:20 ratio to evaluate generalization performance.

Model Architecture

This study developed three predictive models with distinct structural designs, as described below.

1. Pure LSTM Model
   The first model employed a LSTM neural network, which is a variant of the RNN designed to overcome long-term dependency issues in time-series data [19]. The model architecture consisted of one LSTM layer with 50 hidden units, followed by a dense output layer with one neuron for price regression. The model used the *ReLU* activation function, *Adam* optimizer, and *mean_squared_error* as the loss function. Training was performed for 100 epochs with a batch size of 32.

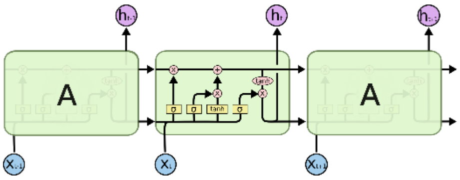

Figure 2. LSTM Model Architecture for Gold Price Prediction

2. Hybrid LSTM + Prophet Model
   The second model combined the Prophet algorithm with LSTM to leverage both methods strengths [20]. Prophet was first used to identify long-term trends and seasonal patterns from the gold price series. The predicted values (*yhat*) from Prophet were subtracted from the actual prices to produce residuals, representing short-term variations not captured by Prophet. These residuals were normalized and then used as additional input features for the LSTM model, allowing LSTM to learn nonlinear dependencies based on Prophet's residual signals. The relationship modelled by Prophet is expressed as follows:

$$y_{(t)} = g_{(t)} + s_{(t)} + h_{(t)} + \epsilon_{(t)} \quad (1)$$

Where:

$g_{(t)}$: represents the main trend component (logistic or piecewise linear)
$s_{(t)}$: denotes periodic seasonal components,
$h_{(t)}$: captures effects of holidays or special events, and
$\epsilon_{(t)}$: indicates random noise or unexplained variations.

This hybrid configuration allowed Prophet to capture trend and seasonality, while LSTM handled complex short-term fluctuations.

3. Hybrid LSTM + Prophet + News Sentiment Model
   The third model extended the hybrid architecture by incorporating sentiment data as an exogenous variable [21]. Sentiment scores were merged with Prophet residuals to form a

multivariate input for LSTM. This combination enabled the model to learn the joint relationship between market sentiment and gold price dynamics. The training configuration was identical to the previous models, with 100 epochs, batch size 32, *Adam* optimizer, and *ReLU* activation. This hybrid multivariate framework provided the ability to analyze how daily sentiment influences short-term deviations from long-term trends.

Evaluation Metrics

To assess the predictive accuracy of each model, five statistical indicators were used. These metrics measure how close the predicted values are to the actual data and provide a quantitative basis for model comparison.

- Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (2)$$

MSE calculates the average squared difference between predicted and observed values. Smaller values indicate higher accuracy.

- Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

RMSE expresses the prediction error in the same unit as the target data, where lower values mean better precision.

- Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (4)$$

MAE represents the average absolute deviation between the predicted and actual values.

- Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

R-squared indicates how much of the variance in the data is explained by the model. A value closer to one reflects stronger performance.

- Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (6)$$

MAPE expresses prediction errors as percentages, with smaller values showing higher forecasting accuracy.

Tools and Experimental Setup

All experiments were conducted using the Python programming language on the Google Collab environment equipped with GPU acceleration. The experiments utilized the following libraries:

- TensorFlow/Keras: for model construction and training,
- Prophet: for time-series trend and seasonality modelling,
- Scikit-learn: for normalization and evaluation metrics,
- Pandas and NumPy: for data manipulation and preprocessing.

Model configurations were standardized as follows:

- Time step: 90 days
- Epochs: 100
- Batch size: 32
- Optimizer: Adam
- Activation function: ReLU
- Training/testing ratio: 80:20

These settings ensured consistent evaluation across all models, allowing fair comparison of predictive performance.

RESULTS AND DISCUSSION

Overview of Experimental Results

This section presents the results obtained from the three predictive models developed in this study:

1. Pure LSTM,
2. Hybrid LSTM + Prophet, and
3. Hybrid LSTM + Prophet + News Sentiment.

Each model was trained and tested using the same dataset split 80% training and 20% testing to ensure a fair comparison. The evaluation was conducted using five quantitative metrics: MSE, RMSE, MAE, R^2 and MAPE. The summary of model performance is shown in Table 1.

Table 1. Model Performance Evaluation Results

Model	Dataset	MSE	RMSE	MAE	R^2	MAPE (%)
LSTM	Train	191.9120	13.8532	9.4709	0.9951	0.70
LSTM	Test	334.4998	18.2893	13.4574	0.9375	0.75
LSTM + Prophet	Train	175.3417	13.2417	9.0014	0.9955	0.66
LSTM + Prophet	Test	324.1363	18.0038	13.0675	0.9394	0.73
LSTM + Prophet + Sentiment	Train	173.3286	13.1654	9.0189	0.9956	0.66
LSTM + Prophet + Sentiment	Test	327.0838	18.0855	13.1536	0.9389	0.73

Performance Comparison

The evaluation results presented in Table 1 show that the pure LSTM model demonstrated strong baseline performance, with an R^2 value of 0.9375 and a MAPE of 75%. After Prophet was integrated as a trend and seasonality detector, the hybrid LSTM + Prophet model achieved better testing performance, with the R^2 increasing to 0.9394 and the MAPE decreasing to 73%. This improvement indicates that Prophet effectively captures long-term seasonal patterns and reduces the burden on LSTM in modelling overall trends.

When sentiment variables were incorporated into the hybrid LSTM + Prophet + Sentiment model, the prediction results remained stable, but no significant improvement was observed. The model achieved an R^2 of 0.9389 with the same MAPE of 73%. Although the difference was minimal, the overall accuracy remained high. The slight variation may be attributed to the limited quality and quantity of sentiment data, as well as possible misalignment between daily news distribution and gold price movements.

Both hybrid models, LSTM + Prophet and LSTM + Prophet + Sentiment, showed excellent performance on training data with R^2 values of 0.9955 and 0.9956, respectively. However, there was no indication of overfitting, since the testing performance remained stable at 0.9394 and 0.9389.

Additionally, the error-based metrics further support these results. The pure LSTM model achieved a MSE of 334.4998, RMSE of 18.2893, and MAE of 13.4574 on the testing dataset. After integrating Prophet, the hybrid LSTM + Prophet model recorded lower errors, with MSE of 324.1363, RMSE of 18.0038, and MAE of 13.0675, confirming improved trend learning and predictive stability. Meanwhile, the hybrid LSTM + Prophet + Sentiment model achieved MSE of 327.0838, RMSE of 18.0855, and MAE of 13.1536. Although the improvement was not substantial, the hybrid models consistently outperformed the pure LSTM model across all evaluation metrics.

In summary, while the pure LSTM model already provided reliable predictions, the hybrid approaches demonstrated better overall accuracy, and stability by combining Prophet's strength in capturing long-term trends with LSTM's ability to model short-term nonlinear dependencies. For comparison, here is a visualization of the comparison between actual prices and the predicted results from the three models in the form of a line graph, which can be seen in the following figure.

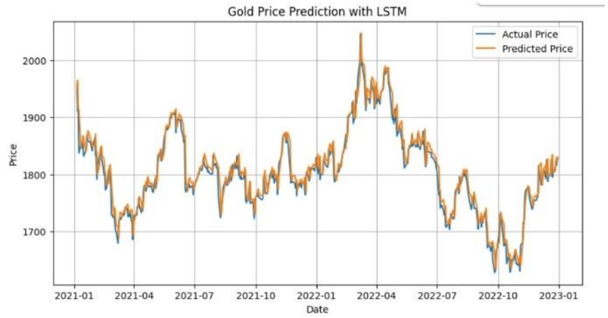

Figure 3. Comparison between actual and predicted gold prices using the pure LSTM model

Figure 4. Comparison between actual and predicted gold prices using the hybrid LSTM + Prophet model

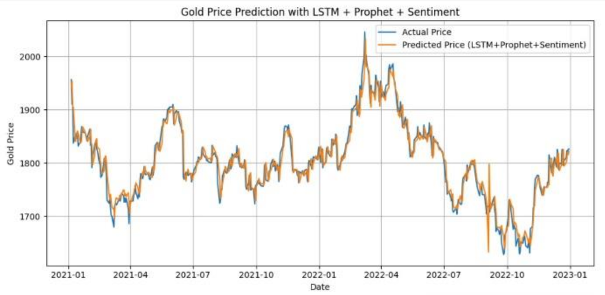

Figure 5. Comparison between actual and predicted gold prices using the hybrid LSTM + Prophet + Sentiment model

Discussion of Findings

The results confirm that integrating deep learning and statistical forecasting methods enhances model robustness for financial time series prediction. Prophet effectively decomposes trend and seasonality components, enabling the hybrid models to focus more on short-term variations rather than long-term drift. This synergy allows the LSTM + Prophet hybrid to outperform the standalone LSTM model, particularly in reducing prediction error and stabilizing output under fluctuating market conditions.

The addition of sentiment features provided marginal improvement, indicating that while sentiment has a psychological impact on market behaviour, its predictive contribution depends heavily on the quality and timing of the textual data. News articles may not always reflect real-time trading decisions, which could explain the minimal accuracy gain observed.

These findings are consistent with previous research [9], [10],[12], and [13], which emphasize the potential of hybrid forecasting approaches and sentiment-based modelling in financial domains. Similar to prior studies on stock and cryptocurrency prediction, the combination of neural networks and statistical decomposition techniques provides a balanced framework capable of modelling both deterministic patterns and nonlinear dynamics effectively.

Overall, the results highlight that hybrid deep learning models offer a more comprehensive understanding of gold price movements, bridging traditional trend analysis with modern sentiment-aware forecasting. This demonstrates the growing importance of integrating multiple data perspectives in developing reliable and interpretable predictive systems for financial applications.

CONCLUSIONS AND RECOMMENDATIONS

This study developed three predictive models for gold price forecasting using pure LSTM, hybrid LSTM + Prophet, and LSTM + Prophet with sentiment. The results show that all models achieved high accuracy, with the hybrid LSTM + Prophet model producing the best overall performance. Combining deep learning with statistical trend decomposition effectively improved forecasting stability and captured both long-term and short-term market patterns.

For future research, improvements can be directed toward enhancing sentiment data using advanced natural language processing techniques and integrating additional economic indicators to strengthen predictive accuracy. Real-time implementation of this hybrid model is also recommended to support decision making in dynamic financial environments.

REFERENCES

[1] D. Makala and Z. Li, "Prediction of gold price with ARIMA and SVM," *J. Phys. Conf. Ser.*, vol. 1767, no. 1, 2021, doi: 10.1088/1742-6596/1767/1/012022.

[2] J. Chai, C. Zhao, Y. Hu, and Z. G. Zhang, "Structural analysis and forecast of gold price returns," *J. Manag. Sci. Eng.*, vol. 6, no. 2, pp. 135–145, 2021, doi: 10.1016/j.jmse.2021.02.011.

[3] N. Tripurana, B. Kar, S. Chakravarty, B. K. Paikaray, and S. Satpathy, "Gold Price Prediction Using Machine Learning Techniques," *CEUR Workshop Proc.*, vol. 3283, pp. 274–281, 2021, doi: 10.54097/gdm0kc53.

[4] 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.

[5] Y. F. Wijaya and A. Triayudi, "Penerapan Data Mining Pada Prediksi Harga Emas dengan Menggunakan Algoritma Regresi Linear Berganda dan ARIMA," *J. Comput. Syst. Informatics*, vol. 5, no. 1, pp. 73–81, 2023, doi: 10.47065/josyc.v5i1.4615.

[6] R. A. Dalimuthe, R. T. Adek, and C. Agusniar, "Prediksi Harga Emas Menggunakan Algoritma Long Short-Term Memory (Lstm)," *SENASTIKA Univ. Malikussaleh PREDIKSI*, pp. 1–10, 2024.

[7] F. Fauzi, S. Aulia, A. R. Syaifullah, and T. W. Utami, "Peramalan Harga Emas Menggunakan Pendekatan Long-Short Term Memory (LSTM)," *J. Edukasi dan Penelit. Inform.*, vol. 10, no. 2, p. 252, 2024, doi: 10.26418/jp.v10i2.78332.

[8] B. Taylor, S. J.Letham, "Forecasting at Scale," *Am. Stat.*, vol. 72, no. 1, pp. 37–45, 2018, doi: 10.1080/00031305.2017.1380080.

[9] A. Primawati, I. S. Sitanggang, A. Annisa, and D. A. Astuti, "Perbandingan Kinerja LSTM dan Prophet untuk Prediksi Deret Waktu (Studi Kasus Produksi Susu Sapi Harian)," *J. Edukasi dan Penelit. Inform.*, vol. 9, no. 3, p. 428, 2023, doi: 10.26418/jp.v9i3.72031.

[10] D. Rizkya, H. Roosaputri, and C. Dewi, "Perbandingan Algoritma ARIMA , Prophet , dan LSTM dalam Prediksi Penjualan Tiket Wisata Taman Hiburan (Studi Kasus : Saloka Theme Park)," vol. 4, no. 3, pp. 507–517, 2023.

[11] M. Costola, O. Hinz, M. Nofer, and L. Pelizzon, "Machine learning sentiment analysis, COVID-19 news and stock market reactions," *Res. Int. Bus. Financ.*, vol. 64, p. 101881, 2023, doi: 10.1016/j.ribaf.2023.101881.

[12] Christophorus Bintang Saputra and D. P. Koesrindartoto, "Pemanfaatan Analisis Sentimen Youtube untuk Prediksi Harga Saham: Studi pada Investor Retail Indonesia," *J. Manaj.*, vol. 21, no. 1, pp. 1–17, 2024, doi: 10.25170/jm.v21i1.5184.

[13] S. Darvishan, S. Heidari, and M. Zarei, "Investigating the Impact of Investor Sentiment

and Volatility on Investment Behavior in the Gold Market," *J. Accounting, Financ. Econ. Sci.*, vol. 4, no. 4, pp. 136–143, 2024, doi: 10.61186/jafes.4.4.136.

[14] G. Jun-gang, "EasyChair Preprint Time-Series Prediction Research Based on Combined Prophet-LSTM Models," 2023, [Online]. Available: https://easychair.org/publications/preprint/42PsG%0A

[15] A. Amini and R. Kalantari, "Gold price prediction by a CNN-Bi-LSTM model along with automatic parameter tuning," *PLoS One*, vol. 19, no. 3 March, pp. 1–17, 2024, doi: 10.1371/journal.pone.0298426.

[16] A. D. Arisandi, L. Atika, Ferdiansyah, E. S. Negara, and K. R. N. Wardani, "Prediksi Mata Uang Bitcoin Menggunakan LSTM dan Analisis Sentimen pada Sosial Media Pendahuluan Metode Penelitian," *J. Ilm. KOMPUTASI*, vol. 19, no. 4, pp. 1–9, 2020.

[17] F. Nekouei, "Gold Price Prediction LSTM 96% Accuracy." Kaggle, 2022. [Online]. Available: https://www.kaggle.com/code/farzadnekouei/gold-price-prediction-lstm-96-accuracy/notebook#Step-12:-Visualizing-Results

[18] A. Sinha, "Sentiment Analysis of Commodity News (Gold)," 2022. [Online]. Available: https://www.kaggle.com/datasets/ankurzing/sentiment-analysis-in-commodity-market-gold/data

[19] M. Owen, V. Vincent, R. Br Ambarita, and E. Indra, "Implementasi Metode Long Short Term Memory Untuk Memprediksi Pergerakan Nilai Harga Emas," *J. Tek. Inf. dan Komput.*, vol. 5, no. 1, p. 96, 2022, doi: 10.37600/tekinkom.v5i1.507.

[20] M. Beck *et al.*, "xLSTM: Extended Long Short-Term Memory," *Adv. Neural Inf. Process. Syst.*, vol. 37, no. NeurIPS, pp. 1–57, 2024.

[21] A. Fernando Pratama *et al.*, "Implementasi Analisis Sentimen dan Model Deep Learning untuk Prediksi Harga Bitcoin," *J. Jupiter*, vol. 15 No. 1, pp. 403–412, 2023.