



Explainable Multimodal Stock Price Forecasting Using a Hybrid LSTM–XGBoost Framework: Evidence from Indonesian LQ45 Companies

Muhamad Sabri Ahmad^{1,2*}, H. Hadiyanto³, Ridwan Sanjaya⁴

¹ Doctoral Program in Information Systems, Universitas Diponegoro, Semarang 50275, Indonesia

² Department of Informatics Engineering, Faculty of Engineering, Universitas Khairun, Ternate 97719, Indonesia

³ Department of Chemical Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang 50275, Indonesia

⁴ Department of Information Systems, Faculty of Computer Science, Soegijapranata Catholic University, Semarang 50234, Indonesia

Corresponding Author Email: msabri@students.undip.ac.id

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310914>

ABSTRACT

Received: 28 April 2026

Revised: 20 July 2026

Accepted: 4 August 2026

Available online: 30 September 2026

Keywords:

stock price prediction, multimodal learning, hybrid LSTM–XGBoost, explainable artificial intelligence, SHapley Additive exPlanations, Indonesian stock market

Accurate stock price forecasting remains challenging due to the nonlinear, noisy, and heterogeneous nature of financial time-series data. This study proposes an explainable multimodal forecasting framework that integrates Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) for next-day stock price prediction. The proposed framework combines historical market prices, technical indicators, daily sentiment information, and company fundamental variables to capture different aspects of market dynamics. The model was evaluated using three companies listed in the Indonesian LQ45 index, representing mining, banking, and telecommunications sectors. A time-based evaluation strategy was adopted, and the proposed framework was compared with standalone LSTM and XGBoost models using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), coefficient of determination (R^2), and Directional Accuracy (DA). Experimental results show that model performance varies across stocks, with XGBoost providing strong and stable performance, while the hybrid model achieved competitive results for selected cases. Ablation analysis indicates that price and technical variables remain the dominant predictors for short-term forecasting, whereas sentiment information contributes mainly to directional prediction. SHapley Additive exPlanations (SHAP) analysis further identifies the influence of individual features on model outputs and provides interpretable insights into prediction behavior. The findings demonstrate that hybrid multimodal learning can provide an interpretable approach for short-term stock price forecasting in emerging markets, while its effectiveness remains dependent on stock characteristics and feature composition.

1. INTRODUCTION

Stock price forecasting is an important research problem in financial modelling because it supports investment analysis, portfolio management, risk assessment, and decision-making in capital markets. Nevertheless, reliable forecasting remains difficult because financial time series are nonlinear, non-stationary, noisy, and sensitive to rapidly changing information. Stock price movements are influenced not only by historical trading patterns but also by market sentiment, company fundamentals, volatility, and sector-specific conditions. These characteristics require forecasting models that can represent heterogeneous information and capture complex nonlinear relationships among variables.

Financial forecasting has been examined using mathematical, statistical, machine learning, and deep learning approaches. Deep learning surveys and decomposition-based hybrid methods have been used to address the nonlinear and non-stationary characteristics of financial time series [1, 2]. More recent studies have shown that artificial intelligence-

based methods can achieve competitive forecasting performance by learning complex relationships from historical market data [3-5]. However, model effectiveness varies according to the structure of the dataset, the forecasting horizon, and the characteristics of the predicted asset.

Long Short-Term Memory (LSTM) networks are widely used in financial time-series forecasting because their recurrent structure enables them to learn temporal dependencies from sequential observations. LSTM-based models and their optimized variants have been applied to stock and cryptocurrency forecasting with promising results [6-9]. However, their performance is sensitive to sequence length, model configuration, data volume, and market noise. Comparative studies have also shown that LSTM does not consistently outperform conventional machine learning models, particularly when the prediction task relies heavily on engineered tabular features [10]. Consequently, the superiority of LSTM should be established empirically rather than assumed for every financial forecasting problem.

Extreme Gradient Boosting (XGBoost) provides an

alternative approach for structured financial data. It can model nonlinear relationships and feature interactions while controlling model complexity through regularization. In short-term stock forecasting, XGBoost is particularly suitable for technical indicators, lagged prices, rolling statistics, sentiment measures, and fundamental ratios. Previous comparisons have demonstrated that XGBoost can produce stable and competitive results relative to LSTM in short-term financial prediction [11, 12]. However, XGBoost does not inherently learn sequential dependencies. Temporal information must instead be represented explicitly through lag features, rolling indicators, or other forms of feature engineering.

The complementary characteristics of LSTM and XGBoost motivate their integration within a hybrid architecture. LSTM can learn temporal representations from sequential market observations, whereas XGBoost can model nonlinear interactions among engineered tabular features. Hybrid and ensemble methods have been increasingly applied in financial forecasting because they can combine different modelling capabilities and improve robustness under complex data conditions [13-15]. Nevertheless, increased architectural complexity does not automatically improve predictive performance. A hybrid model should therefore be evaluated against strong standalone baselines and tested across assets with different market characteristics.

Data representation is another important consideration in stock price forecasting. Many existing models rely primarily on historical prices and technical indicators, although stock prices may also respond to public sentiment and company-specific financial information. Multi-source heterogeneous data can provide a broader representation of market behavior by combining price-based information with external signals [11]. Sentiment derived from social media and other textual sources has also been used as an additional predictor of stock movements [16, 17]. However, the contribution of sentiment and fundamental information may vary across stocks and prediction horizons. Their usefulness should therefore be evaluated empirically rather than presumed.

Model interpretability is equally important in financial applications. Machine learning and deep learning models are often regarded as black boxes because their predictions are not accompanied by clear explanations of feature influence. This limitation is important when model outputs are used to support investment or risk-related decisions. Explainable artificial intelligence (XAI) methods, including Local Interpretable Model-Agnostic Explanations and SHapley Additive exPlanations (SHAP), have been developed to improve the transparency of predictive models [18, 19]. SHAP is particularly relevant because it provides consistent global and local explanations based on Shapley values. Its application in financial prediction enables researchers to identify the variables that influence overall model behavior and individual predictions [20].

Despite these developments, four research gaps remain. First, many studies evaluate LSTM and XGBoost separately, while fewer studies integrate their complementary capabilities through a stacking architecture for next-day stock price forecasting. Second, studies that simultaneously combine price, technical, sentiment, and fundamental features remain limited, particularly in the Indonesian stock market. Third, many empirical studies evaluate only one stock, index, or asset, which restricts the assessment of model consistency across different sectors. Fourth, forecasting accuracy is often emphasized without a systematic examination of feature-

group contributions and model interpretability.

To address these gaps, this study develops an explainable hybrid LSTM-XGBoost framework for multimodal next-day stock price forecasting. The framework integrates historical prices, technical indicators, daily sentiment features, and company fundamental indicators. It is evaluated on three stocks included in the Indonesian LQ45 index: PT Aneka Tambang Tbk (ANTM), PT Bank Central Asia Tbk (BBCA), and PT Telkom Indonesia Tbk (TLKM). These stocks represent the mining, banking, and telecommunications sectors, respectively. The prediction target is the next-day closing price, denoted as $Close(t+1)$.

The main contributions of this study are fourfold. First, it develops a stacking-based hybrid LSTM-XGBoost architecture that combines temporal learning with nonlinear tabular regression. Second, it evaluates the framework using multimodal features derived from market prices, technical indicators, daily sentiment, and company fundamentals. Third, it compares the proposed model with standalone LSTM and XGBoost baselines across three stocks from different sectors. Fourth, it incorporates an ablation study and SHAP analysis to examine feature-group contributions and provide global and local explanations of model predictions.

2. RELATED WORKS

Stock price forecasting has been investigated using statistical methods, machine learning, deep learning, hybrid forecasting, multimodal data integration, and XAI. These research streams differ in the type of information they process and, in the mechanisms, used to represent temporal, nonlinear, and external market influences. This section reviews the literature most closely related to the present study, including financial time-series forecasting, LSTM-based prediction, XGBoost-based prediction, hybrid and ensemble models, multimodal stock forecasting, and XAI.

2.1 Financial time-series and stock price forecasting

Financial time-series forecasting aims to estimate future market behavior using historical observations and relevant explanatory variables. Unlike many conventional prediction problems, financial forecasting is complicated by nonlinearity, non-stationarity, volatility clustering, noise, and uncertainty. These properties limit the effectiveness of models that assume stable statistical relationships over time.

Comprehensive reviews show that deep learning models have been widely applied to financial time-series forecasting because they can represent nonlinear and dynamic market behavior. Sezer et al. [1] reviewed deep learning applications in financial time-series forecasting and showed that different architectures are used for price, trend, and volatility prediction. Other studies have combined signal decomposition, statistical forecasting, and artificial intelligence to address the complex structure of financial data. Lu et al. [2], for example, proposed an attention-based EMD-LSTM model in which empirical mode decomposition was used before temporal learning. Such an approach is intended to reduce signal complexity and improve the representation of non-stationary patterns.

Artificial intelligence-based forecasting has become increasingly prominent because machine learning and deep learning models can learn nonlinear relationships directly from market data. Xu et al. [3] proposed a hybrid artificial

intelligence model for stock price forecasting, while Xiao et al. [4] examined hybrid modelling for REIT volatility forecasting. Although these studies addressed different financial targets, both demonstrated that combined modelling strategies can be useful when a single model is insufficient to represent complex market behavior.

Comparative evidence also suggests that model performance is highly dependent on the dataset and prediction setting. Notaria et al. [21] compared several time-series models for stock price prediction and showed that no model consistently dominates across all conditions. Therefore, financial forecasting models should be evaluated empirically under the specific characteristics of the asset, feature structure, and prediction horizon.

2.2 Long Short-Term Memory-based stock prediction

LSTM networks are widely used in stock and financial forecasting because they are designed to capture temporal dependencies in sequential data. Their memory-cell architecture allows information from previous observations to influence current predictions, making LSTM suitable for ordered financial series.

Nelson et al. [6] applied LSTM neural networks to stock-market price-movement prediction, demonstrating the relevance of recurrent architectures for financial sequence modelling. Yuliyono and Girsang [7] applied an artificial bee colony-optimized LSTM model to Bitcoin price prediction and showed that metaheuristic optimization can improve recurrent-model configuration. Gülmez [8] also developed an optimized deep LSTM architecture using the Artificial Rabbits Optimization algorithm and reported improved stock price prediction performance. These studies indicate that LSTM performance can be enhanced through parameter optimization and architectural refinement.

Despite these advantages, LSTM is sensitive to sequence length, model architecture, training data size, feature scaling, and market noise. Its performance may decline when the available information is dominated by engineered tabular variables rather than pure sequential patterns. Kumar et al. [9] compared LSTM and XGBoost for short-term Bitcoin price prediction and showed that tree-based models can remain competitive in short-horizon forecasting. Hossain and Kaur [22] similarly compared XGBoost and LSTM for stock market prediction and found that relative performance depends on the structure of the data and the experimental design.

These findings indicate that LSTM should not be assumed to provide the best performance in every financial forecasting task. Its effectiveness must be assessed in relation to the temporal structure of the data, the prediction horizon, and the characteristics of the input features.

2.3 Extreme Gradient Boosting-based stock prediction

XGBoost is a tree-based ensemble algorithm designed for structured and tabular prediction problems. It is effective in capturing nonlinear relationships, modelling feature interactions, handling heterogeneous variables, and controlling overfitting through regularization.

In stock price prediction, XGBoost is particularly suitable for engineered financial features, including moving averages, volatility measures, technical indicators, lagged prices, sentiment scores, and financial ratios. Hossain and Kaur [22] demonstrated that XGBoost can provide competitive

performance relative to LSTM in stock market prediction. Kumar et al. [9] also found that XGBoost remained effective for short-term cryptocurrency price forecasting.

The strength of XGBoost lies in its ability to learn complex relationships from tabular features without requiring a sequential input structure. However, this advantage is accompanied by an important limitation. XGBoost does not inherently model temporal dependencies in the same manner as recurrent neural networks. Temporal information must be explicitly represented through lag features, rolling statistics, trend indicators, or other forms of feature engineering.

This limitation motivates the integration of XGBoost with sequence-based learning. A hybrid architecture can use LSTM to capture temporal patterns and XGBoost to model nonlinear relationships among engineered multimodal features.

2.4 Hybrid and ensemble models for financial forecasting

Hybrid and ensemble models have been increasingly used in financial forecasting because they combine the capabilities of multiple algorithms. The central assumption is that different models may capture different aspects of market behavior, and that their integration can improve robustness or predictive accuracy.

Hossain and Kaur [22] compared XGBoost and LSTM for stock-market prediction and showed that model performance can vary according to the experimental setting. Lv [13] developed a hybrid transfer-learning framework for stock index forecasting, demonstrating the potential of knowledge transfer and model integration in financial prediction. Aravindhan and Manikavelan [15] proposed a decomposition-based Transformer-LSTM framework for stock index forecasting, illustrating how decomposition and deep sequence modelling can be combined to represent complex temporal patterns.

Liu et al. [23] developed an ensemble stock price prediction model using multi-source heterogeneous data, an LSTM attention mechanism, and a multidimensional grey model. Their study is particularly relevant because it integrates heterogeneous financial information and multiple modelling mechanisms within one forecasting framework.

Hybrid approaches have also been applied in other forecasting domains. Baimel et al. [5] used SHAP-based error compensation in electricity price forecasting and demonstrated that explainability can be integrated with predictive correction mechanisms. Although the application domain differs, the study reinforces the broader relevance of hybrid and explainable forecasting architectures.

Nevertheless, hybridization does not automatically guarantee better performance. Its effectiveness depends on the compatibility of the component models, the alignment of their outputs, the amount and quality of the data, and the characteristics of the prediction task. Therefore, hybrid models must be evaluated against strong standalone baselines rather than assumed to be superior because of their greater complexity.

2.5 Multimodal data in stock price prediction

Stock price movement is influenced by multiple types of information. Historical prices and technical indicators reflect recent market behavior, sentiment represents investor or public perception, and fundamental indicators represent the financial condition of the company. Integrating these data

sources can provide a broader representation of market dynamics.

Liu et al. [14] showed that multi-source heterogeneous data can improve stock price prediction by combining information that captures different dimensions of market behavior. Gandhudi et al. [16] examined the influence of tweets on stock price prediction using an explainable hybrid quantum neural network and demonstrated that textual sentiment can provide additional predictive information. Arif et al. [17] also combined historical stock data and social media sentiment from X to predict digital bank stock prices, providing evidence that sentiment can be incorporated into stock forecasting in the Indonesian context.

However, the contribution of each data modality is not necessarily uniform. Price and technical indicators are directly linked to short-term market movement, while sentiment may be more useful for identifying directional changes or temporary shifts in investor expectations. Fundamental variables are reported at lower frequencies and may have stronger relevance for medium- or long-term valuation than for next-day price movement.

For this reason, multimodal features should not be assumed to improve forecasting performance automatically. Their contribution should be evaluated using controlled comparisons, such as an ablation study that measures the incremental value of sentiment and fundamental information beyond price-based features.

2.6 Explainable artificial intelligence in financial prediction

Interpretability is an important requirement in financial prediction because model outputs may influence investment, portfolio, and risk-related decisions. Although black-box models can achieve high predictive performance, their use is limited when users cannot determine which variables drive the predictions.

Ribeiro et al. [18] introduced Local Interpretable Model-Agnostic Explanations as a method for explaining individual predictions from complex classifiers. The method approximates local model behavior using interpretable representations. Lundberg and Lee [19] later introduced SHAP, a unified framework based on cooperative game theory. SHAP assigns an importance value to each feature by estimating its contribution to the model output.

SHAP is widely used because it supports both global and local interpretation. Global explanations identify the variables that most strongly influence predictions across a dataset, while local explanations show how individual feature values increase or decrease a specific prediction. This dual capability is particularly valuable in financial forecasting, where both overall model behavior and case-specific reasoning may be required.

Muhammad et al. [20] applied an explainable deep learning approach to stock market trend prediction and demonstrated the relevance of model transparency in financial applications. Gandhudi et al. [16] also incorporated explainability into a tweet-based stock prediction framework. These studies indicate that explainability can complement predictive performance by showing whether a model relies on financially meaningful signals.

In the present study, SHAP is applied to the XGBoost-based component of the forecasting framework. This choice is appropriate because TreeSHAP provides an efficient and

theoretically consistent explanation method for tree-based models.

2.7 Research gap

The reviewed literature shows substantial progress in financial time-series forecasting, deep learning, gradient boosting, hybrid modelling, multimodal data integration, and XAI. Nevertheless, several gaps remain.

First, many studies compare LSTM and XGBoost as separate models, whereas fewer studies integrate their complementary capabilities through a stacking architecture for next-day stock price forecasting. Second, multimodal stock forecasting studies that simultaneously combine historical prices, technical indicators, sentiment features, and company fundamentals remain limited, particularly in the Indonesian stock market. Third, many studies focus on a single stock, index, or asset, which limits the evaluation of model consistency across sectors with different market characteristics. Fourth, additional data modalities are often included without a controlled assessment of whether they improve numerical accuracy or directional prediction. Fifth, many forecasting studies emphasize performance metrics but provide limited explanation of feature contributions.

To address these gaps, this study develops an explainable hybrid LSTM–XGBoost framework for multimodal next-day stock price forecasting. The framework combines sequential learning from LSTM with nonlinear tabular regression from XGBoost. It is evaluated using three Indonesian LQ45 stocks from the mining, banking, and telecommunications sectors. The study also conducts an ablation analysis to evaluate the contribution of price, sentiment, and fundamental features and applies SHAP to provide global and local explanations of model predictions.

3. METHODOLOGY

This study employed a quantitative experimental design to develop and evaluate an explainable hybrid LSTM–XGBoost framework for multimodal stock price forecasting. The methodology comprised data collection, preprocessing, multimodal data integration, feature engineering, model development, model training, performance evaluation, ablation analysis, and SHAP-based explainability analysis.

The forecasting objective was to estimate the next-day closing price, denoted as $\text{Close}(t+1)$, using information available up to trading day t . The overall experimental workflow is presented in Figure 1.

3.1 Research objects and data sources

The empirical evaluation was conducted on three stocks listed in the Indonesian LQ45 index: ANTM, BBCA, and TLKM. These stocks were selected because they represent three distinct sectors: mining, banking, and telecommunications. Table 1 summarizes the selected companies, sectors, data types, main variables, and observation periods.

Three categories of data were used in this study. The first category consisted of historical stock-market data, including open, high, low, close, and volume variables. The second category consisted of daily sentiment indicators derived from textual information. The third category consisted of company

fundamental indicators obtained from financial reports.

The prediction target was defined as the closing price on the next trading day:

$$y_t = \text{Close}(t + 1)$$

(1)

where, y_t denotes the target predicted using the information available at time t , and $\text{Close}(t+1)$ denotes the closing price on the following trading day.

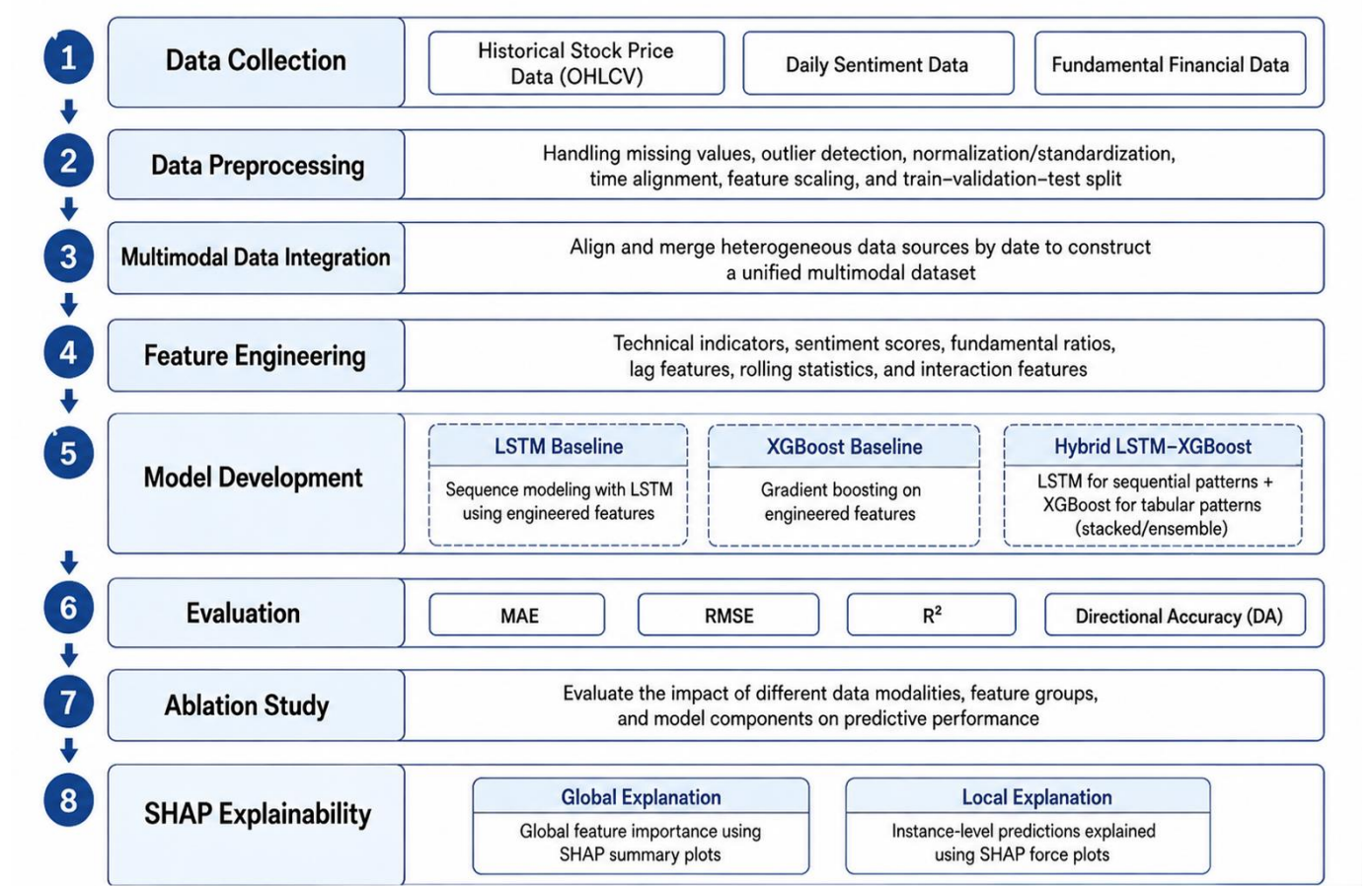


Figure 1. Overall research framework of the proposed study

Table 1. Research objects and data sources

Stock Code	Company	Sector	Data Types	Main Variables	Observation Period
ANTM	PT Aneka Tambang Tbk	Mining	Price, sentiment, fundamental	OHLCV, technical indicators, sentiment indicators, financial ratios	2018-2025
BBCA	PT Bank Central Asia Tbk	Banking	Price, sentiment, fundamental	OHLCV, technical indicators, sentiment indicators, financial ratios	2018-2025
TLKM	PT Telkom Indonesia Tbk	Telecommunications	Price, sentiment, fundamental	OHLCV, technical indicators, sentiment indicators, financial ratios	2018-2025

3.2 Data preprocessing and multimodal integration

Before model development, all datasets were preprocessed to ensure temporal and structural consistency. Stock-price records were sorted chronologically, converted into numerical format, and checked for duplicate and missing observations. Sentiment records were aggregated into daily indicators to align them with the trading calendar. Fundamental data, which were reported periodically, were aligned with the corresponding reporting periods and mapped to daily observations.

Because the three data modalities had different temporal frequencies, a multimodal integration process was required. Price and sentiment variables were treated as daily observations, whereas fundamental variables were aligned

according to their reporting availability and then associated with the corresponding trading dates. This procedure produced a daily integrated record containing market, technical, sentiment, and fundamental information. The multimodal data integration process is illustrated in Figure 2.

The observations were maintained in chronological order throughout preprocessing and model development. All predictive variables were constructed using information available at or before time t . This constraint was applied to prevent look-ahead bias and data leakage.

3.3 Feature engineering

Feature engineering transformed the raw observations into predictive variables. The final feature set consisted of four

main groups: raw market features, technical indicators, sentiment features, and fundamental features. Table 2 summarizes the feature groups and representative variables used in the final modelling dataset.

Raw market features included open, high, low, close, and volume. Technical indicators were generated from historical price data to represent returns, trends, volatility, momentum, and short-term temporal dependence. These variables included log return, moving averages, rolling volatility, Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), MACD signal, and lagged closing prices.

The logarithmic return was computed as:

$$r_t = \ln\left(\frac{\text{Close}(t)}{\text{Close}(t-1)}\right) \tag{2}$$

where, r_t denotes the log return at time t .

The moving average over a window of k trading days was defined as:

$$MA_{k,t} = \frac{1}{k} \sum_{j=0}^{k-1} \text{Close}(t-j) \tag{3}$$

Rolling volatility was calculated as the standard deviation of log returns over a specified window.

Sentiment features consisted of the total number of textual observations, positive sentiment count, neutral sentiment count, negative sentiment count, and mean sentiment score. These variables represented daily public or market perception.

Fundamental features included revenue, net income, total assets, total liabilities, total equity, earnings per share, return on equity, return on assets, net profit margin, debt-to-equity ratio, current ratio, asset turnover, and year-on-year growth indicators.

Rows containing incomplete values resulting from rolling-window calculations, lag construction, or target shifting were removed before model training.

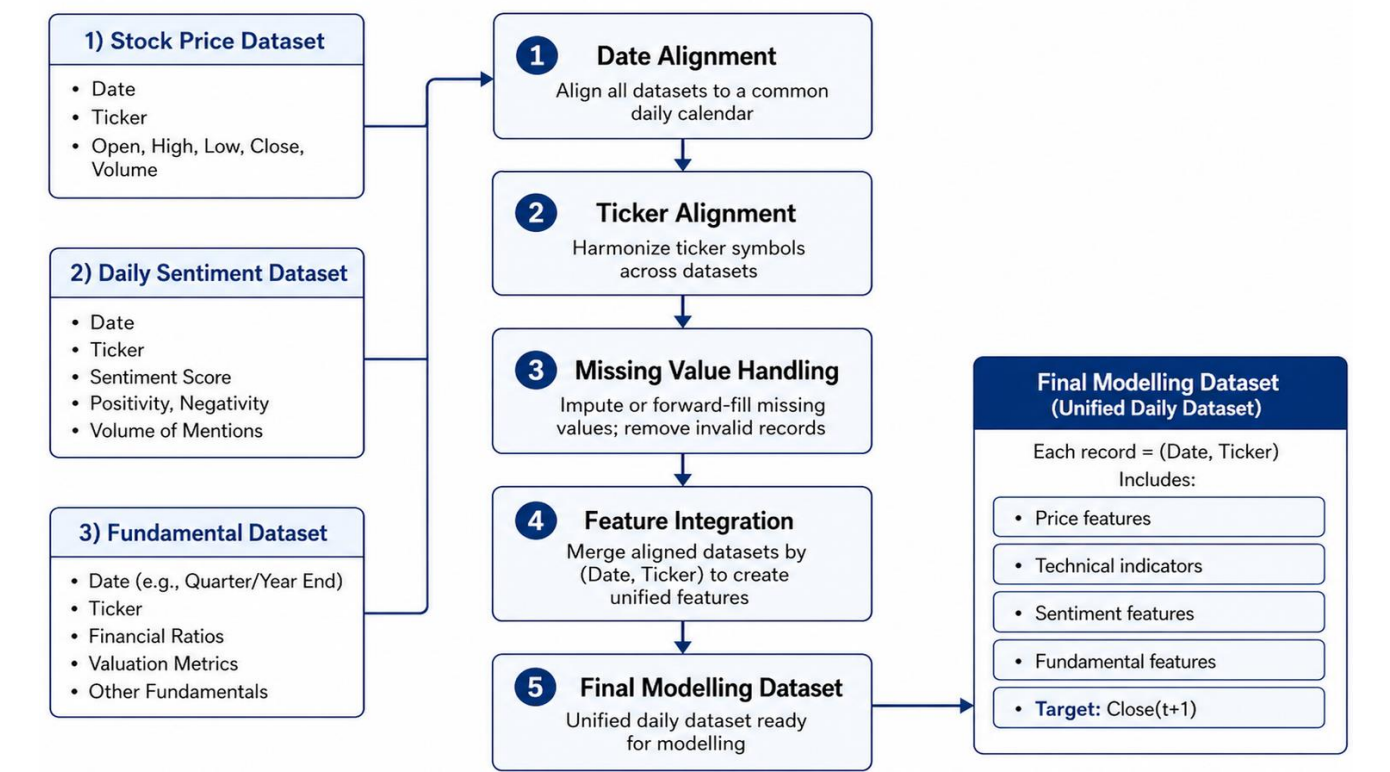


Figure 2. Multimodal data integration process

Table 2. Feature groups and representative variables

Feature Group	Variables	Description
Price data	open, high, low, close, volume	Raw daily market information
Technical indicators	log_return, ma_5, ma_10, ma_20, volatility_5, volatility_10, rsi_14, macd, macd_signal	Trend, return, volatility, and momentum information
Lag features	close_lag1, close_lag2, close_lag3	Previous closing-price information
Sentiment features	total_tweets, positive, neutral, negative, mean_score	Daily aggregated sentiment indicators
Fundamental features	revenue, net_income, total_assets, total_liabilities, total_equity, eps, roe, roa, net_margin, de_ratio, current_ratio, asset_turnover, yoy_revenue_growth, yoy_net_income_growth	Company financial condition and ratios
Target	Close($t+1$)	Next-day closing price

3.4 Model development

Three forecasting models were developed and compared: LSTM, XGBoost, and hybrid LSTM–XGBoost. LSTM served as the temporal deep-learning baseline, XGBoost served as the structured-data baseline, and the hybrid model combined temporal representation with nonlinear tabular regression.

3.4.1 Long Short-Term Memory baseline model

The LSTM model was used to learn temporal relationships from sequential observations. The input data were transformed into fixed-length sequences using a 20-day rolling window. The LSTM input gate was defined as:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (4)$$

The forget gate was defined as:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (5)$$

The candidate cell state was computed as:

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (6)$$

The cell state was updated as:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

The output gate and hidden state were calculated as:

$$\begin{aligned} o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (8)$$

where, x_t denotes the input vector at time t ; h_{t-1} denotes the

previous hidden state; c_{t-1} denotes the previous cell state; i_t , f_t , and o_t denote the input, forget, and output gates, respectively; σ denotes the sigmoid activation function; and $\odot$ denotes element-wise multiplication.

3.4.2 Extreme Gradient Boosting baseline model

XGBoost was used to model nonlinear relationships among engineered tabular features. The model generates predictions by combining multiple regression trees. The predicted output at iteration K is expressed as:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F} \quad (9)$$

where, f_k denotes the k -th regression tree and $\mathcal{F}$ denotes the space of possible regression trees.

The XGBoost objective function was defined as:

$$\mathcal{L}^{(K)} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (10)$$

where, $l(y_i, \hat{y}_i)$ denotes the loss between the actual and predicted values, and $\Omega(f_k)$ denotes the regularization term used to control tree complexity and reduce overfitting.

3.4.3 Proposed hybrid Long Short-Term Memory-Extreme Gradient Boosting model

The proposed hybrid model was constructed using a stacking strategy. First, the LSTM model learned temporal patterns from 20-day input sequences. The resulting LSTM prediction was then appended to the engineered tabular feature vector. XGBoost used the augmented feature vector to produce the final forecast.

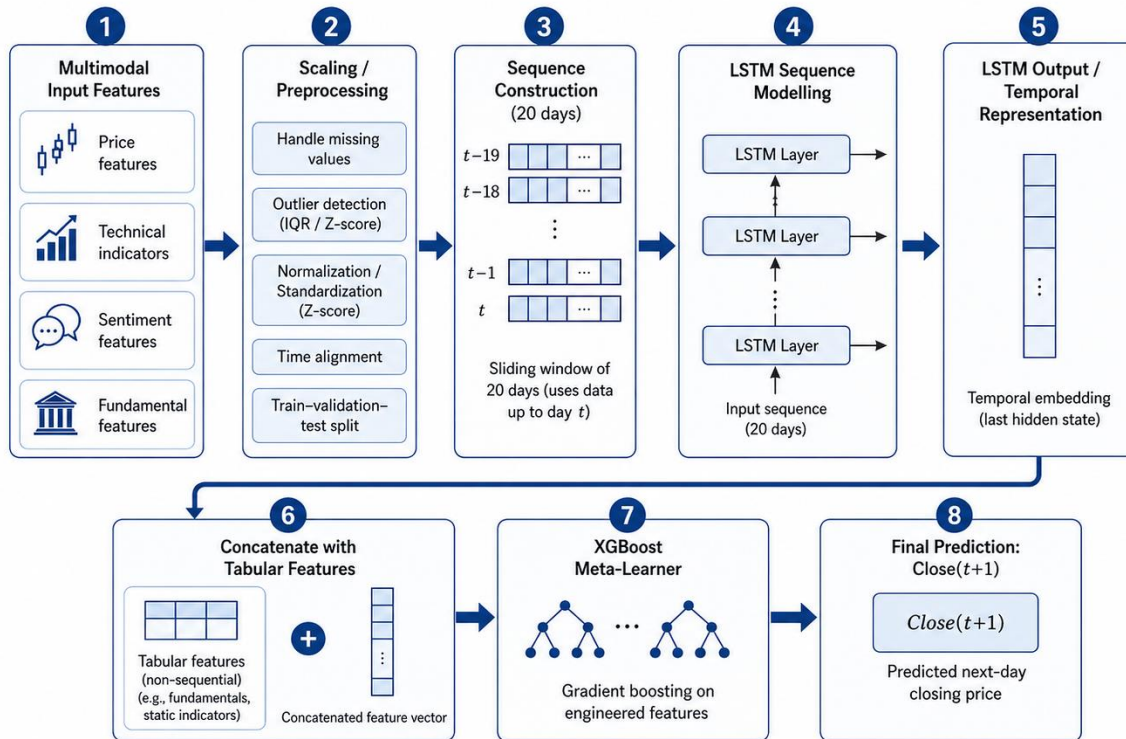


Figure 3. Architecture of the proposed Hybrid Long Short-Term Memory (LSTM)-Extreme Gradient Boosting (XGBoost) model

Table 3. Model configuration and experimental settings

Component	Configuration
Prediction target	Close($t + 1$)
Data splitting	Time-based split
Split proportion	66% training, 13% validation, 21% testing
LSTM sequence length	20 trading days
LSTM optimizer	Adam
LSTM loss function	Mean Squared Error
LSTM epochs	100
LSTM batch size	32
XGBoost n_estimators	300
XGBoost max_depth	6
XGBoost learning_rate	0.05
XGBoost subsample	0.8
XGBoost colsample_bytree	0.8
XGBoost objective	reg:squarederror
Random state	42
Explainability method	SHAP applied to the XGBoost-based component

Note: LSTM = Long Short-Term Memory, XGBoost = Extreme Gradient Boosting, SHAP = SHapley Additive exPlanations.

Table 4. Experimental scenarios used in the ablation study

Scenario	Feature Composition	Evaluation Purpose
Model A	Price and technical features	Evaluate the predictive strength of historical market information
Model B	Price, technical, and sentiment features	Assess the incremental contribution of daily sentiment
Model C	Price, technical, sentiment, and fundamental features	Evaluate the complete multimodal framework

The hybrid input representation was formulated as:

$$z_t = [x_t^{\text{tabular}}, \hat{y}_t^{\text{LSTM}}] \quad (11)$$

where, x_t^{tabular} denotes the engineered multimodal tabular features and $\hat{y}_t^{\text{LSTM}}$ denotes the prediction generated by LSTM.

The final hybrid prediction was expressed as:

$$\hat{y}_t^{\text{Hybrid}} = f_{\text{XGB}}(z_t) \quad (12)$$

where, f_{XGB} denotes the XGBoost meta-learner and $\hat{y}_t^{\text{Hybrid}}$ denotes the final next-day closing-price prediction.

The architecture shown in Figure 3 enabled the model to combine temporal information learned by LSTM with nonlinear interactions among price, technical, sentiment, and fundamental variables. Table 3 summarizes the model configuration and experimental settings.

3.5 Experimental scenarios and ablation study

An ablation study was conducted to evaluate the incremental contribution of each data modality. Three feature scenarios were tested. Model A used price and technical features only. Model B combined price, technical, and sentiment features. Model C used the complete multimodal feature set consisting of price, technical, sentiment, and fundamental variables.

The ablation analysis was designed to determine whether sentiment and fundamental variables improved forecasting performance beyond price-based information. This comparison was necessary because additional data sources do not necessarily improve next-day forecasting accuracy. Table 4 defines the feature composition and evaluation purpose for each ablation scenario.

3.6 Data splitting and model training procedure

A chronological data-splitting strategy was applied to

preserve the temporal order of observations. The integrated dataset was divided into training, validation, and testing subsets using proportions of 66%, 13%, and 21%, respectively. Random splitting was not used because it could introduce future information into the training process and result in data leakage.

Feature scaling was applied separately for each modelling approach. Min–max normalization was used for the LSTM input because recurrent neural networks are sensitive to input scale. Standardization was used for the XGBoost input to maintain consistency among heterogeneous tabular variables. All scalers were fitted exclusively on the training subset and were subsequently applied to the validation and testing subsets.

The LSTM model was trained using sequences of 20 consecutive trading days. The LSTM predictions were subsequently aligned with the corresponding XGBoost observations. For the hybrid model, the aligned LSTM output was appended as an additional feature before the XGBoost meta-learner was trained.

The three models were evaluated using the same aligned testing period. This procedure ensured that performance differences were attributable to the modelling approaches rather than differences in the test observations.

3.7 Evaluation metrics

Model performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), coefficient of determination (R^2), and Directional Accuracy (DA). These metrics measured both numerical forecasting accuracy and the ability to predict the direction of stock-price movement.

MAE was calculated as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

RMSE was calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

The R^2 was calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (15)$$

DA was calculated as:

$$DA = \frac{1}{n-1} \sum_{t=2}^n \mathbb{I}[\text{sign}(y_t - y_{t-1}) = \text{sign}(\hat{y}_t - \hat{y}_{t-1})] \quad (16)$$

where, y_i denotes the actual value, $\hat{y}_i$ denotes the predicted value, $\bar{y}$ denotes the mean of the actual values, n denotes the number of test observations, and $\mathbb{I}[\cdot]$ denotes the indicator function.

Lower MAE and RMSE values indicate smaller prediction errors. A higher R^2 value indicates that the model explains a larger proportion of the variance in the observed stock prices. A higher DA value indicates greater accuracy in predicting the direction of price movement.

3.8 Explainability analysis using SHapley Additive exPlanations

SHAP was applied to improve the interpretability of the XGBoost-based prediction component. TreeSHAP was

selected because it is computationally efficient for gradient-boosting models and provides consistent feature-attribution values.

Two levels of interpretation were used. Global interpretation was performed using a SHAP summary plot and a feature-importance bar plot. These visualizations identified the variables with the largest contributions across the test observations. Local interpretation was performed using a SHAP waterfall plot, which showed how individual feature values increased or decreased a selected prediction.

The explainability analysis focused on ANTM as the primary case study. This analysis was used to evaluate whether the prediction model relied on financially meaningful features and whether its output could be explained at both global and individual levels. Figure 4 presents the SHAP explainability workflow used in the study.

3.9 Implementation environment

The experiments were implemented in Python using a cloud-based computational environment. Data preparation and feature engineering were performed using standard numerical and data-processing libraries. The LSTM model was developed using a deep-learning framework, whereas XGBoost was used for tree-based regression. Additional libraries were used for scaling, evaluation, visualization, and SHAP-based explainability.

The same preprocessing, feature-engineering, splitting, training, and evaluation procedures were applied to ANTM, BBKA, and TLKM to maintain experimental consistency across the three stocks.

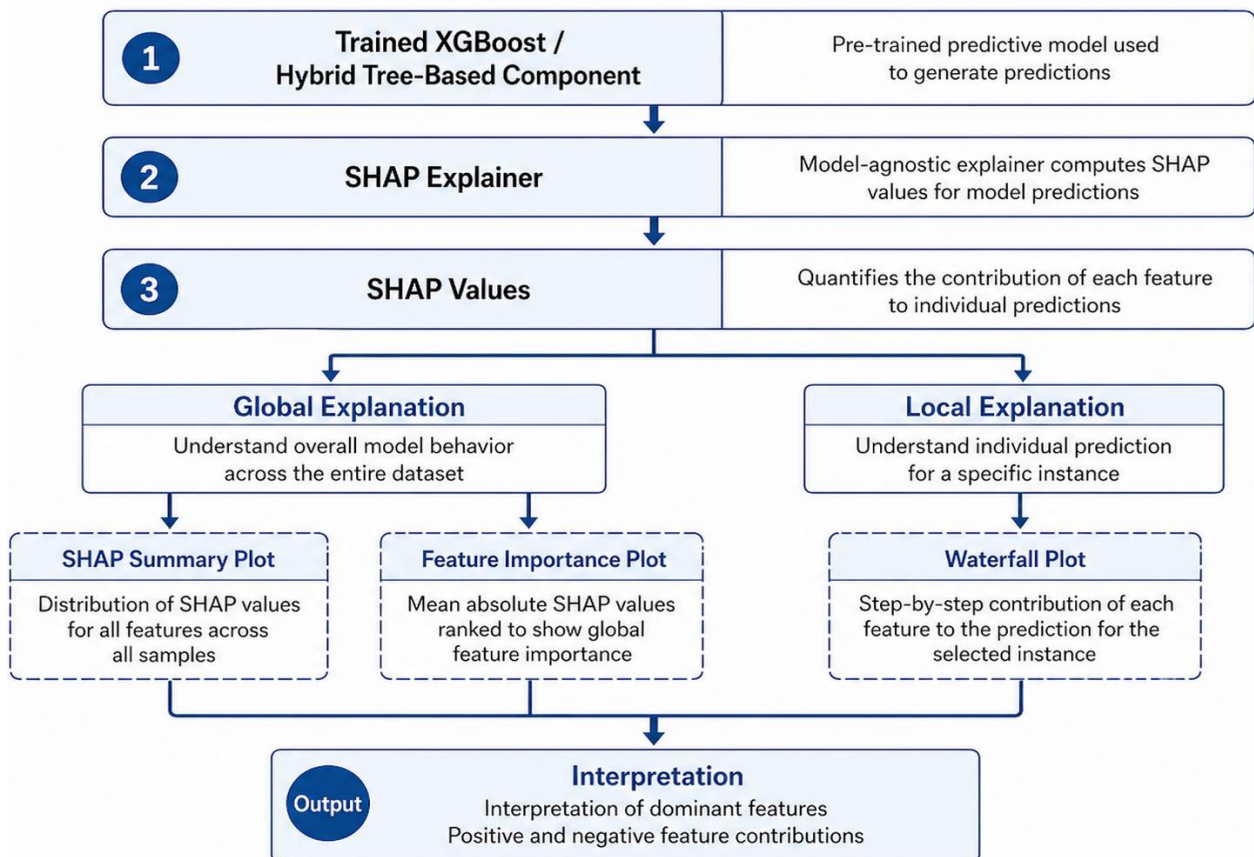


Figure 4. SHapley Additive exPlanations (SHAP) explainability workflow

4. RESULTS AND DISCUSSION

This section presents the empirical results of the proposed multimodal stock-price forecasting framework. The analysis covers the integrated dataset, the comparative performance of LSTM, XGBoost, and hybrid LSTM–XGBoost, cross-stock performance differences, the ablation study, SHAP-based explainability, and the principal implications of the findings. Model performance was evaluated using MAE, RMSE, R^2 , and DA.

4.1 Dataset summary

The final dataset integrated historical stock prices, technical indicators, daily sentiment features, and company fundamental variables. Historical price data constituted the primary daily input and were enriched with technical features derived from price, return, volatility, momentum, and lagged observations. Sentiment information was aggregated at the daily level, while fundamental variables were aligned with trading dates according to their reporting availability.

The empirical analysis covered three LQ45 stocks from different sectors: ANTM, BBCA, and TLKM. The integrated dataset contained 5,844 price observations, 2,479 sentiment-related articles covering 1,832 sentiment days, 108 fundamental records, and 49 final variables.

Table 5 summarizes the scope of the integrated dataset. The inclusion of multiple feature groups enabled the models to learn from short-term market behavior, aggregated sentiment, and company-level financial information. However, the contribution of each feature group was evaluated separately through the ablation study rather than assumed to be uniformly beneficial.

4.2 Baseline and hybrid model performance

Three forecasting models were evaluated: LSTM, XGBoost, and hybrid LSTM–XGBoost. LSTM served as the temporal deep-learning baseline, XGBoost served as the tabular machine-learning baseline, and the hybrid model combined the LSTM prediction with engineered tabular

features through a stacking mechanism.

As shown in Table 6, XGBoost produced the most stable overall performance across the three stocks. For ANTM, XGBoost achieved the lowest MAE and RMSE, the highest R^2 , and the highest DA. Its MAE was 345.92, RMSE was 529.24, R^2 was 0.708, and DA was 0.539. These results indicate that the engineered tabular features provided substantial predictive information for ANTM.

For BBCA, the hybrid LSTM–XGBoost model achieved the best numerical forecasting performance. It produced an MAE of 435.15, an RMSE of 676.63, and an R^2 of 0.635. These values were better than those obtained by the standalone XGBoost model. However, XGBoost achieved a slightly higher DA of 0.415 compared with 0.404 for the hybrid model. Therefore, the hybrid architecture improved price-level accuracy for BBCA but did not improve directional prediction.

For TLKM, the hybrid LSTM–XGBoost model produced the lowest MAE and RMSE, with values of 66.94 and 88.52, respectively. XGBoost, however, achieved the highest R^2 of 0.937, compared with 0.921 for the hybrid model. Both models obtained the same DA of 0.488. These findings indicate that XGBoost and the hybrid model were both competitive for TLKM, although their relative advantages differed across evaluation metrics.

The standalone LSTM model produced the weakest results for all three stocks. Its R^2 values were negative for BBCA and TLKM, indicating that its predictions explained less variance than a mean-based reference model over the evaluated test period. The LSTM predictions also produced low DA values, particularly for BBCA and TLKM. Under the present experimental configuration, the sequential model alone was therefore less effective than the tree-based and hybrid alternatives.

Figures 5–7 compare the actual and predicted stock prices. The LSTM predictions were generally smoother and less responsive to short-term price changes. In contrast, the XGBoost and hybrid LSTM–XGBoost predictions followed the observed price patterns more closely. The visual results were consistent with the numerical performance reported in Table 6.

Table 5. Summary of the research dataset

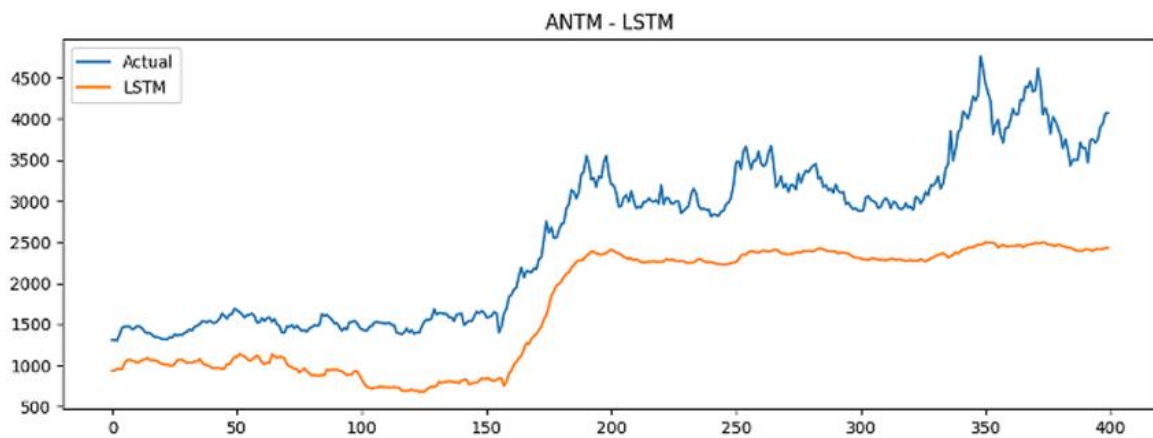
Data Component	Description	Total or Scope
Stocks	ANTM, BBCA, TLKM	3 stocks
Sectors	Mining, banking, telecommunications	3 sectors
Price observations	Daily OHLCV records	5,844 observations
Sentiment data	Daily aggregated sentiment indicators	2,479 articles; 1,832 sentiment days
Fundamental records	Financial statement-based indicators	108 records
Final feature set	Price, technical, sentiment, and fundamental variables	49 columns
Prediction target	Next-day closing price	$Close_{t+1}$

Note: ANTM = PT Aneka Tambang Tbk, BBCA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk.

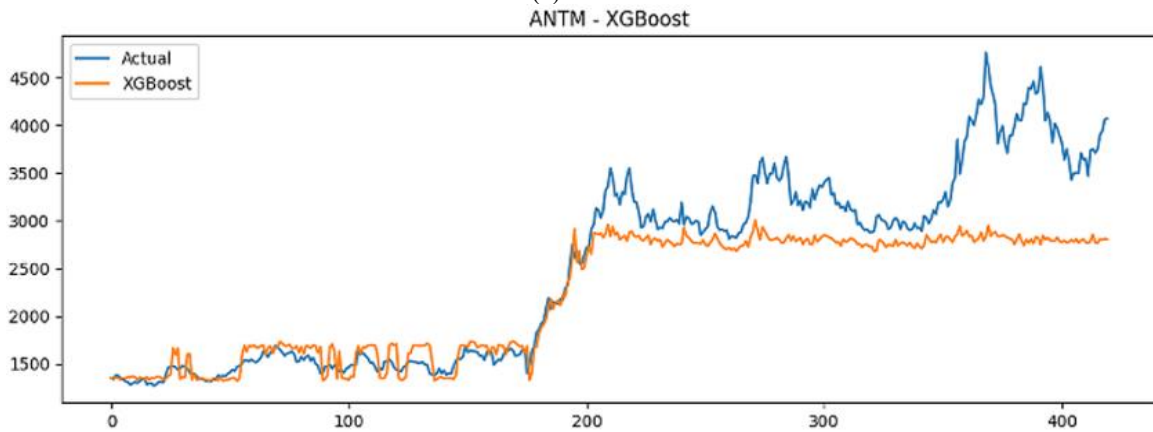
Table 6. Prediction performance of Long Short-Term Memory (LSTM), Extreme Gradient Boosting (XGBoost), and hybrid LSTM–XGBoost

Stock	Model	MAE	RMSE	R^2	Directional Accuracy (DA)
ANTM	LSTM	777.19	861.73	0.205	0.481
ANTM	XGBoost	345.92	529.24	0.708	0.539
ANTM	Hybrid LSTM–XGBoost	423.61	606.44	0.606	0.489
BBCA	LSTM	1960.24	2239.41	-3.001	0.145
BBCA	XGBoost	466.28	703.80	0.613	0.415
BBCA	Hybrid LSTM–XGBoost	435.15	676.63	0.635	0.404
TLKM	LSTM	558.89	625.77	-2.956	0.057
TLKM	XGBoost	69.41	90.94	0.937	0.488
TLKM	Hybrid LSTM–XGBoost	66.94	88.52	0.921	0.488

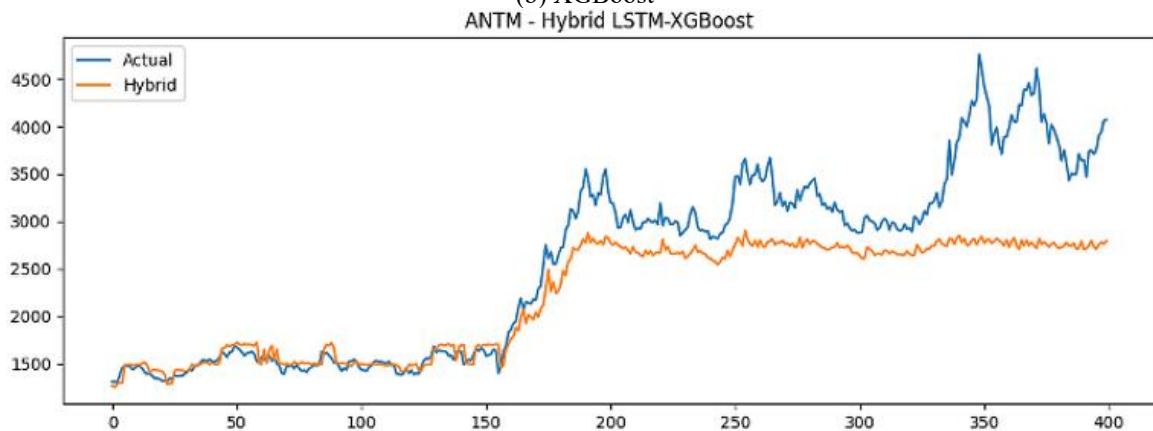
Note: ANTM = PT Aneka Tambang Tbk, BBCA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk, RMSE = Root Mean Squared Error.



(a) LSTM



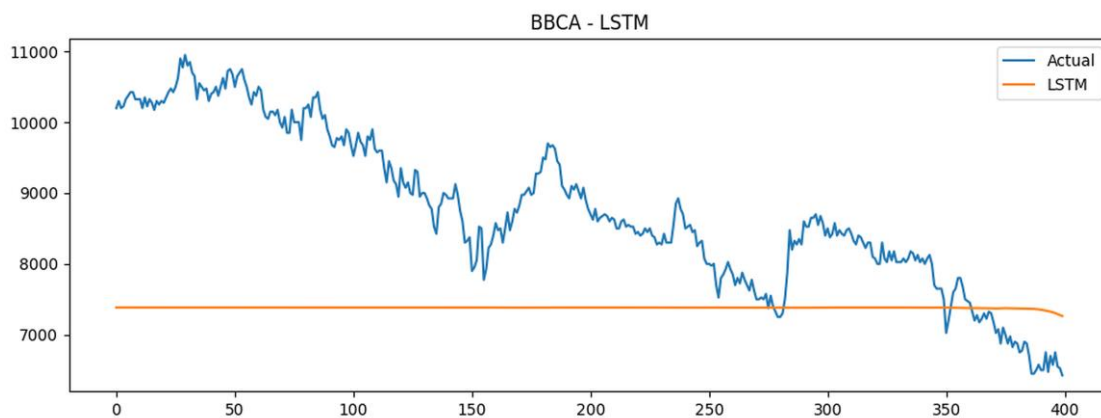
(b) XGBoost



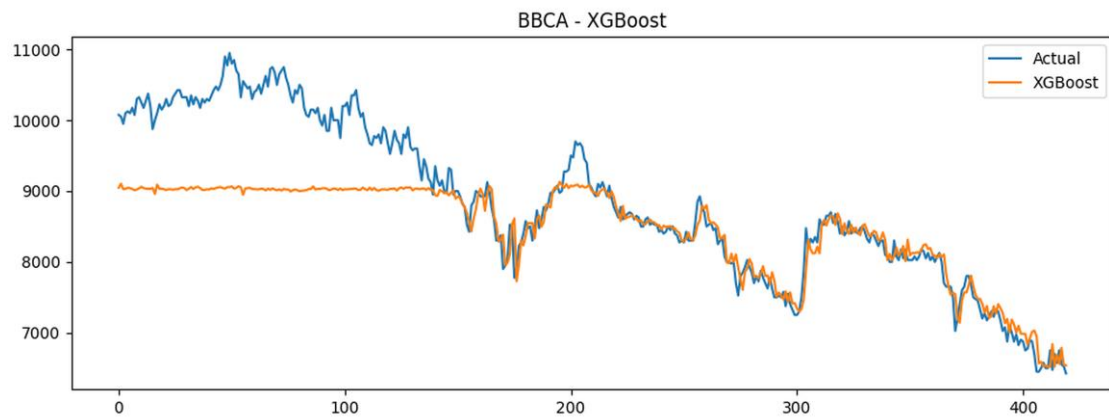
(c) Hybrid LSTM-XGBoost

Figure 5. Actual and predicted stock prices for ANTM

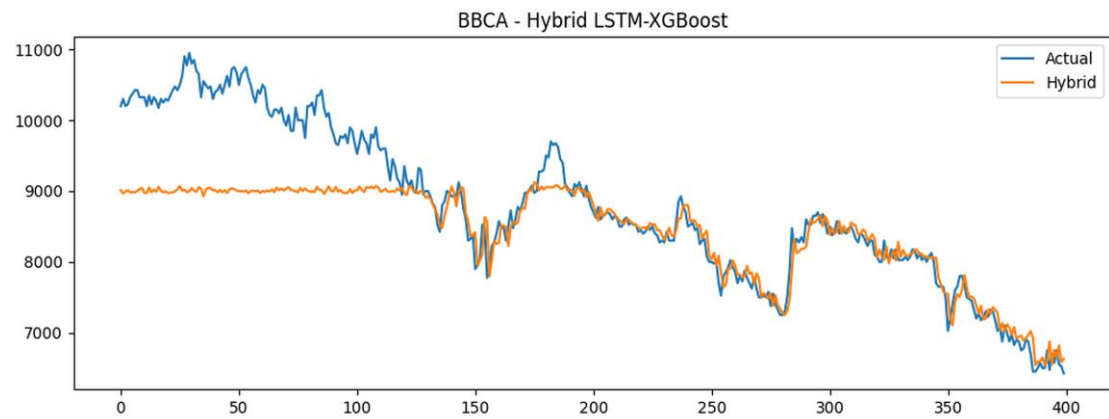
Note: ANTM = PT Aneka Tambang Tbk, RMSE = Root Mean Squared Error, DA = Directional Accuracy.



(a) LSTM



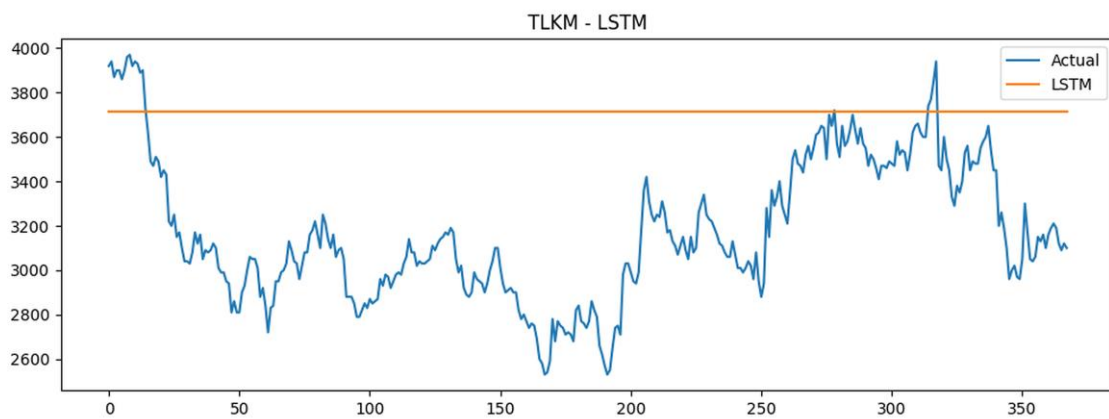
(b) XGBoost



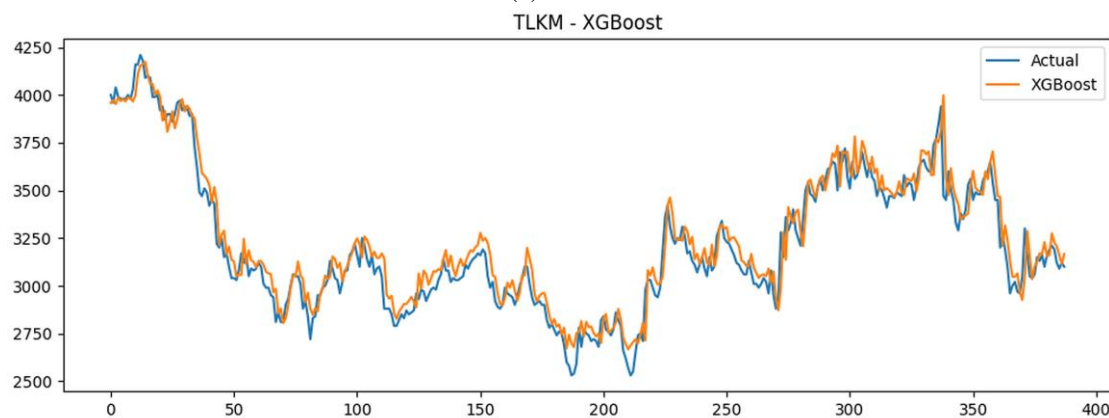
(c) Hybrid LSTM-XGBoost

Figure 6. Actual and predicted stock prices for BBCA

Note: BBCA = PT Bank Central Asia Tbk, RMSE = Root Mean Squared Error, DA = Directional Accuracy.



(a) LSTM



(b) XGBoost

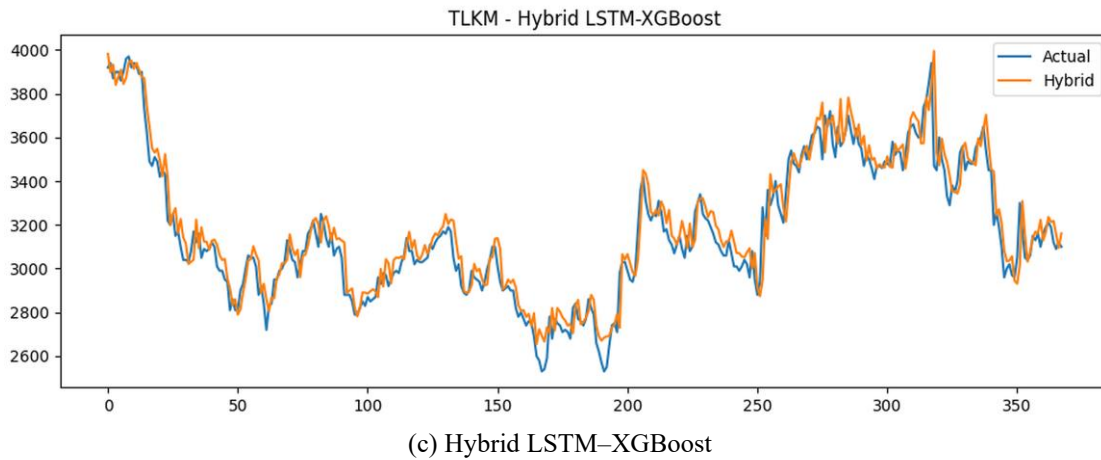


Figure 7. Actual and predicted stock prices for TLKM

Note: TLKM = PT Telkom Indonesia Tbk, RMSE = Root Mean Squared Error, DA = Directional Accuracy.

4.3 Cross-stock performance analysis

The comparative results demonstrate that model performance varied across stocks. No single model achieved the best result for every stock and metric. XGBoost was the strongest model for ANTM and achieved the highest R^2 for TLKM, whereas the hybrid model provided the best error-based performance for BBKA and the lowest MAE and RMSE for TLKM.

These differences indicate that the effectiveness of the hybrid architecture was stock-dependent. For BBKA, the LSTM-derived temporal prediction contributed useful information when combined with the XGBoost feature set. For ANTM, however, the standalone XGBoost model remained superior. The additional LSTM prediction did not improve the forecasting result and may have introduced information that was less useful than the engineered tabular features.

The results also show that greater architectural complexity did not guarantee superior performance. The hybrid LSTM–XGBoost model improved performance in selected cases, but it did not consistently outperform XGBoost. This finding supports a context-dependent interpretation of hybrid modelling: the value of model integration depends on the interaction between stock characteristics, temporal patterns, feature structure, and the forecasting horizon.

The cross-stock results further suggest that engineered tabular variables were highly informative for next-day forecasting. XGBoost remained competitive across all three stocks because the feature set explicitly represented recent prices, lagged values, technical indicators, sentiment measures, and fundamental information.

4.4 Ablation study results

The ablation study evaluated the incremental contribution of sentiment and fundamental features. Three feature scenarios were examined:

Model A: price and technical features;

Model B: price, technical, and sentiment features;

Model C: price, technical, sentiment, and fundamental features.

Table 7 shows that Model A, which used price and technical features only, achieved the best MAE, RMSE, and R^2 for all three stocks. For ANTM, Model A produced an MAE of 312.56, an RMSE of 490.76, and an R^2 of 0.749. For BBKA,

it achieved an MAE of 449.54, an RMSE of 679.77, and an R^2 of 0.639. For TLKM, it produced an MAE of 60.99, an RMSE of 83.70, and an R^2 of 0.947.

These results indicate that recent prices and technical indicators were the strongest predictors of next-day price levels in the evaluated dataset. The finding is consistent with the short forecasting horizon because price, trend, volatility, momentum, and lagged information are directly aligned with daily market dynamics.

Sentiment features contributed more clearly to directional prediction than to price-level accuracy. For ANTM, DA increased from 0.518 in Model A to 0.535 in Model B and 0.539 in Model C. For BBKA, the highest DA of 0.430 was achieved by Model B. For TLKM, Model B also achieved the highest DA, with a value of 0.506.

These findings suggest that sentiment indicators contained information related to the direction of price movement, although they did not consistently reduce MAE or RMSE. The contribution of sentiment was therefore metric-dependent: it was more evident in directional classification than in continuous price prediction.

Fundamental features did not consistently improve next-day forecasting performance. Model C produced higher MAE and RMSE and lower R^2 than Model A for all three stocks. A plausible explanation is the mismatch between the daily prediction horizon and the lower reporting frequency of fundamental information. However, the present experiment does not establish that fundamental features are ineffective for all forecasting horizons. Their contribution may differ in weekly, monthly, or longer-term prediction settings.

Figures 8–10 illustrate the differences among the three feature scenarios. The MAE and R^2 comparisons confirm the advantage of price and technical features for price-level forecasting. The DA comparison shows that sentiment information provided incremental directional value for selected stocks.

4.5 SHapley Additive exPlanations-based explainability analysis

SHAP analysis was used to interpret the XGBoost-based prediction component. The analysis focused on ANTM as the primary interpretability case. Global interpretation was performed using the SHAP summary and feature-importance plots, while local interpretation was performed using a

waterfall plot for one selected prediction.

As shown in Figure 11, the SHAP summary plot showed that recent price-based features had the largest influence on ANTM predictions. The most influential variables included

open, high, close, the five-day moving average, and low. Their high SHAP magnitudes indicate that the model relied primarily on recent market-price information when generating next-day forecasts.

Table 7. Ablation study results for ANTM, BBKA, and TLKM

Stock	Scenario	MAE	RMSE	R ²	DA
ANTM	Model A: Price	312.56	490.76	0.749	0.518
ANTM	Model B: Price + Sentiment	334.09	518.39	0.720	0.535
ANTM	Model C: Price + Sentiment + Fundamental	345.92	529.24	0.708	0.539
BBKA	Model A: Price	449.54	679.77	0.639	0.418
BBKA	Model B: Price + Sentiment	455.29	689.31	0.629	0.430
BBKA	Model C: Price + Sentiment + Fundamental	466.28	703.80	0.613	0.415
TLKM	Model A: Price	60.99	83.70	0.947	0.478
TLKM	Model B: Price + Sentiment	63.37	85.60	0.944	0.506
TLKM	Model C: Price + Sentiment + Fundamental	69.41	90.94	0.937	0.488

Note: ANTM = PT Aneka Tambang Tbk, BBKA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk, RMSE = Root Mean Squared Error, R² = coefficient of determination, DA = Directional Accuracy.

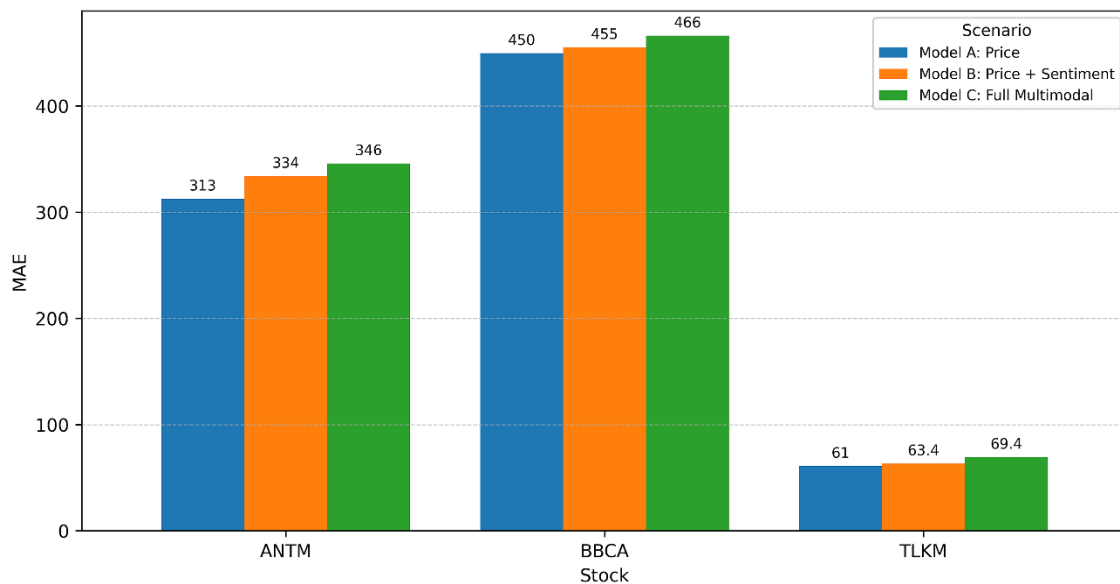


Figure 8. Comparison of Mean Absolute Error (MAE) values in the ablation study

Note: ANTM = PT Aneka Tambang Tbk, BBKA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk.

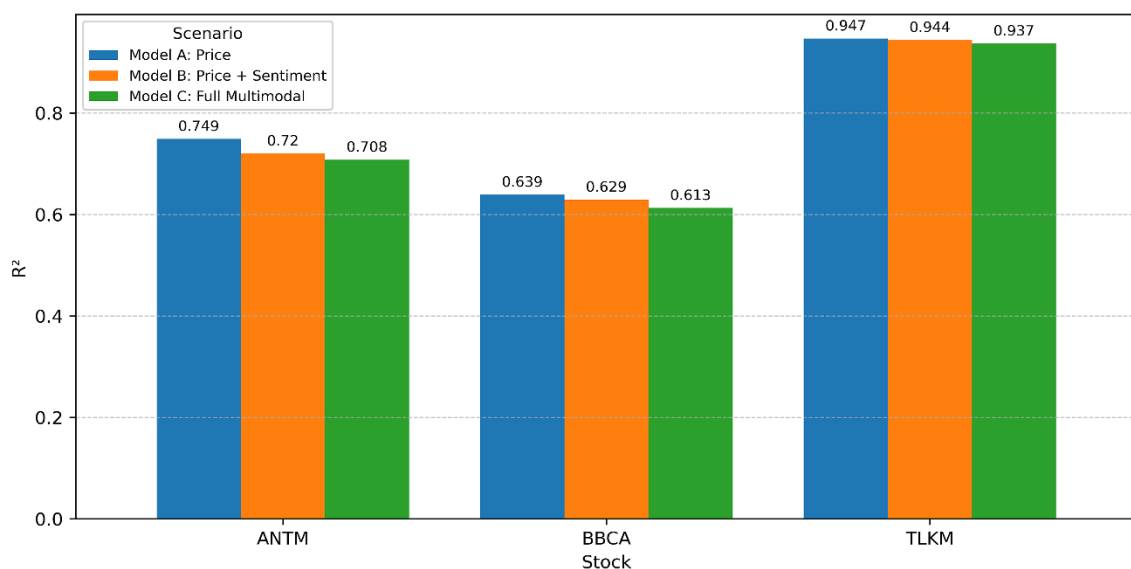


Figure 9. Comparison of coefficient of determination (R²) values in the ablation study

Note: ANTM = PT Aneka Tambang Tbk, BBKA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk.

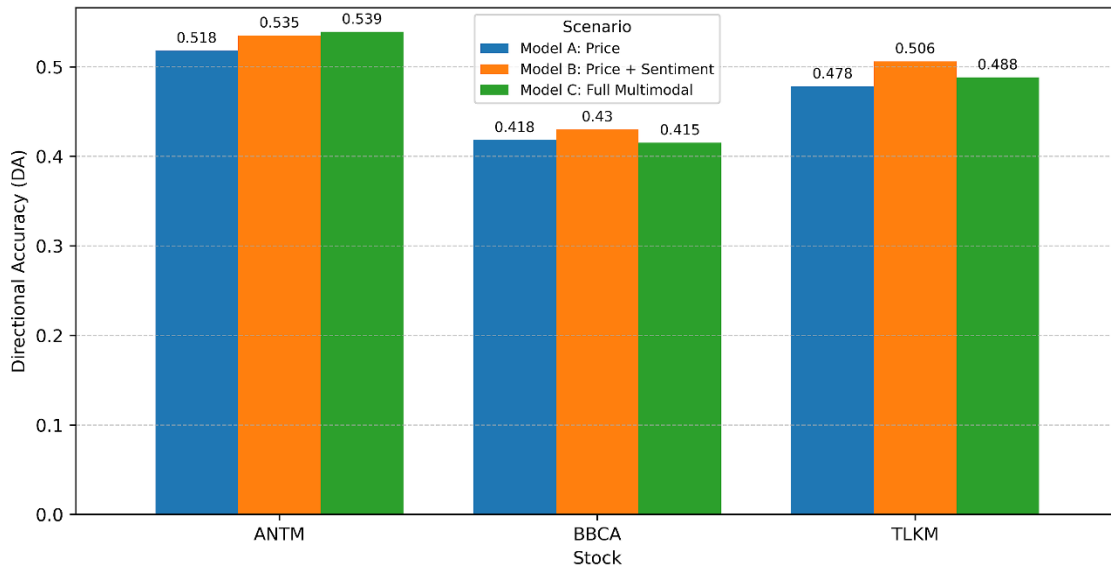


Figure 10. Comparison of Directional Accuracy (DA) values in the ablation study
 Note: ANTM = PT Aneka Tambang Tbk, BBCA = PT Bank Central Asia Tbk, TLKM = PT Telkom Indonesia Tbk.

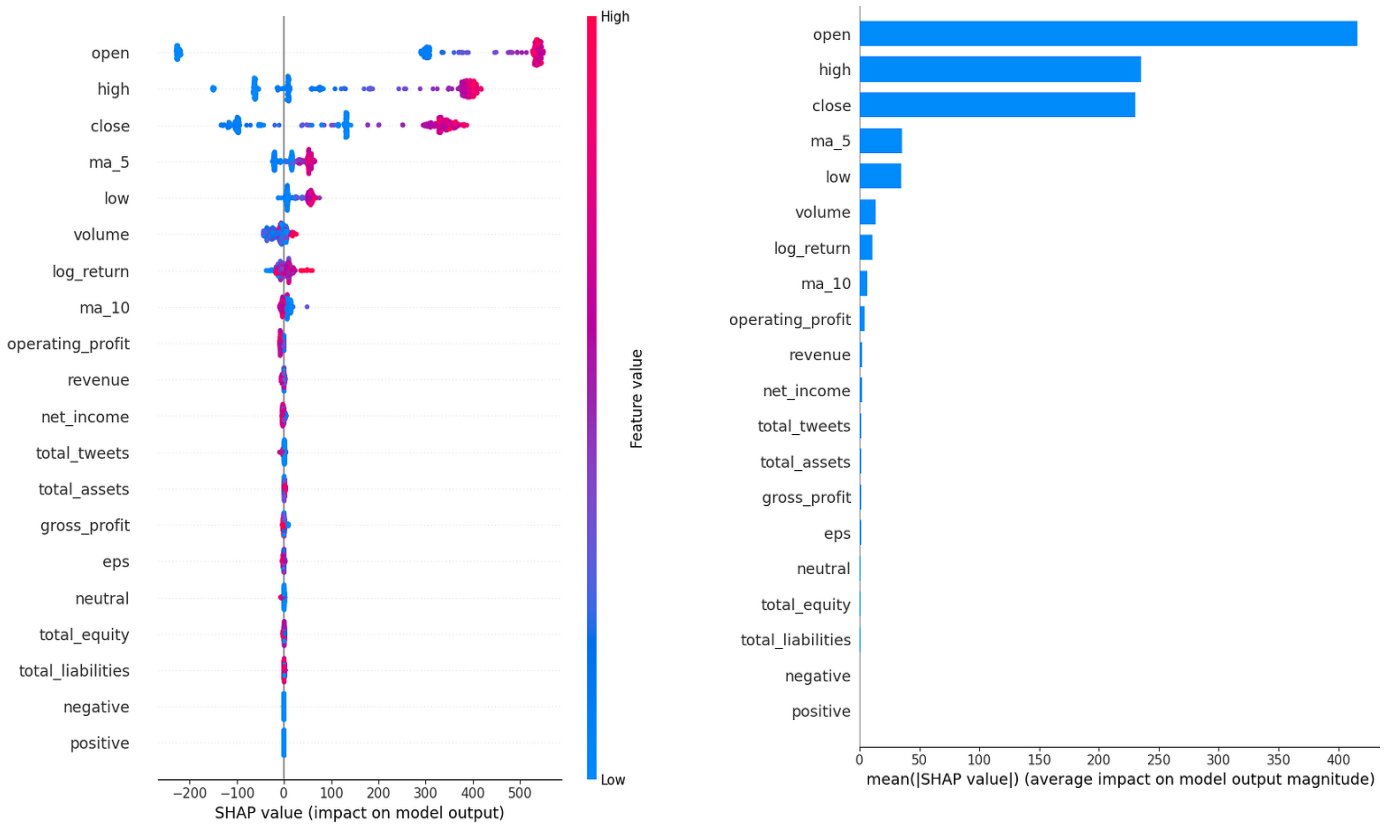


Figure 11. SHapley Additive exPlanations (SHAP) summary plot for PT Aneka Tambang Tbk (ANTM)

The direction and magnitude of the SHAP values also showed that the effect of each variable depended on its observed value and its interaction with the fitted model. Positive SHAP values increased the predicted closing price, whereas negative SHAP values reduced its relative to the model baseline.

As shown in Figure 12, the SHAP feature-importance bar plot ranked the predictors according to their mean absolute SHAP values. Open, high, close, ma_5, and low were among the highest-ranked variables. This ranking confirms that short-term price and technical information dominated the model's global prediction behavior.

Figure 12. SHapley Additive exPlanations (SHAP) feature-importance bar plot for PT Aneka Tambang Tbk (ANTM)

The explainability results were consistent with the ablation analysis. The ablation study showed that price and technical features produced the strongest price-level forecasting performance, while SHAP independently identified variables from these feature groups as the most influential predictors. This consistency strengthens the interpretation that the model learned economically interpretable short-term market relationships rather than relying primarily on lower-frequency fundamental variables.

As shown in Figure 13, the SHAP waterfall plot explained an individual prediction by showing how each feature shifted the model output from the expected value to the final predicted value. Features with positive contributions increased the

prediction, while features with negative contributions reduced it. This local explanation complements the global feature rankings by showing that the contribution of a feature can differ across individual observations.

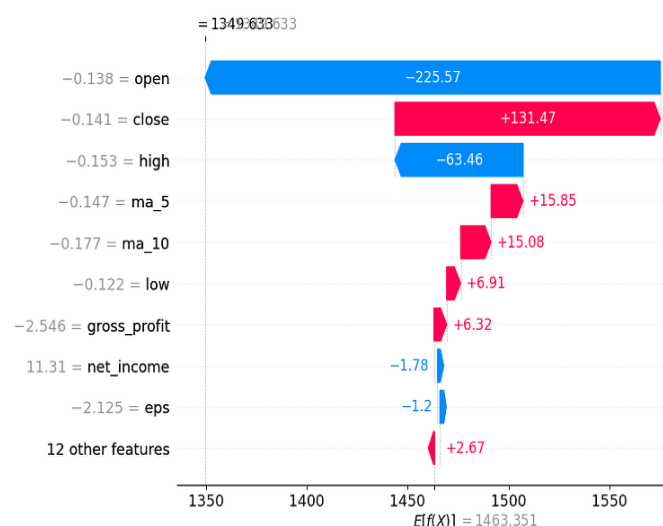


Figure 13. SHapley Additive exPlanations (SHAP) waterfall plot for an individual PT Aneka Tambang Tbk (ANTM) prediction

Overall, the SHAP analysis demonstrated that the XGBoost-based component was interpretable at both global and local levels. The model relied mainly on recent price-based and technical variables, while sentiment and fundamental variables had smaller overall contributions in the ANTM case.

4.6 Discussion

The empirical findings provide five principal insights.

First, XGBoost was the most stable model across the three stocks. It achieved the best complete set of metrics for ANTM and the highest R^2 for TLKM. This result indicates that tree-based boosting was well suited to the engineered multimodal tabular dataset used in this study.

Second, the hybrid LSTM–XGBoost model provided contextual rather than universal improvement. It achieved the best MAE, RMSE, and R^2 for BBKA and the lowest MAE and RMSE for TLKM. However, it did not outperform XGBoost for ANTM. The value of hybridization therefore depended on the specific stock and evaluation metric.

Third, the standalone LSTM model produced the weakest results. This outcome suggests that, under the present sequence length, feature structure, training configuration, and prediction horizon, the recurrent architecture did not capture the daily price dynamics as effectively as XGBoost. The result should not be interpreted as evidence that LSTM is generally unsuitable for financial forecasting. Instead, it shows that its effectiveness is sensitive to the experimental context.

Fourth, the ablation study demonstrated that the inclusion of additional data modalities did not automatically improve next-day price forecasting. Price and technical features produced the lowest errors and highest R^2 values, while sentiment features provided more visible improvements in DA. Fundamental features did not produce consistent gains for the daily prediction horizon.

Fifth, SHAP improved the transparency of the prediction

framework. The feature-attribution results identified recent price and technical variables as the dominant predictors and provided a local explanation for an individual prediction. This interpretability is important because it enables users to evaluate whether the model relies on plausible financial signals.

Taken together, the results indicate that forecasting performance depends on the alignment among model architecture, feature representation, stock characteristics, prediction horizon, and evaluation objective. A complex hybrid model may improve price-level accuracy for one stock without improving directional performance or generalizing uniformly across other stocks.

4.7 Limitations and challenges

This study has several limitations.

First, the forecasting horizon was restricted to the next-day closing price. The usefulness of sentiment and fundamental information may differ for weekly, monthly, or quarterly horizons.

Second, sentiment data were aggregated into daily indicators. This aggregation reduced intraday temporal detail and may have obscured short-lived market reactions.

Third, the SHAP analysis focused on ANTM as the primary interpretability case. Although this approach provided a detailed case-level explanation, equivalent analyses for BBKA and TLKM would provide a broader assessment of cross-stock feature importance.

Fourth, the empirical evaluation covered three LQ45 stocks. These stocks represented different sectors, but the sample was not sufficient to establish generalization across the entire Indonesian market.

Fifth, the hybrid stacking procedure used the LSTM prediction as an additional XGBoost input. Future work should examine more rigorous out-of-fold stacking procedures, alternative meta-learners, and repeated temporal validation to evaluate the robustness of the hybrid architecture.

Despite these limitations, the study provides a consistent framework for comparing standalone and hybrid models, evaluating multimodal feature contributions, and explaining model predictions in the Indonesian stock-market context.

5. CONCLUSIONS

This study developed an explainable hybrid LSTM–XGBoost framework for next-day stock price forecasting using price, technical, sentiment, and fundamental features. The model was evaluated on ANTM, BBKA, and TLKM to examine its performance across different stock characteristics.

The results showed that XGBoost provided the most stable performance, while the hybrid model produced improvements for selected stocks, particularly BBKA. However, the hybrid architecture did not consistently outperform standalone XGBoost, indicating that model effectiveness depends on the interaction between stock characteristics, feature structure, and forecasting horizon.

The ablation study demonstrated that price and technical features were the strongest predictors of next-day prices. Sentiment features contributed more clearly to directional prediction, whereas fundamental features did not consistently improve short-term forecasting accuracy. SHAP analysis confirmed that recent price-based variables were the dominant

contributors to the model predictions.

Overall, the findings show that increasing model complexity does not automatically produce better forecasting performance. Effective stock prediction requires an appropriate alignment between the model architecture, data representation, prediction objective, and interpretability requirements. Future research should evaluate the framework using additional stocks, longer forecasting horizons, and more robust temporal validation strategies.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Universitas Khairun for the institutional support provided during the completion of this research.

REFERENCES

- [1] Sezer, O.B., Gudelek, M.U., Ozbayoglu, A.M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90: 106181. <https://doi.org/10.1016/j.asoc.2020.106181>
- [2] Lu, C., Chi, Y.G., Guan, Y.Y., Fan, J.L. (2019). A hybrid attention-based EMD-LSTM model for financial time series prediction. In 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 113-118. <https://doi.org/10.1109/ICAIBD.2019.8837038>
- [3] Xu, W., Chen, J.L., Xiao, J. (2024). A hybrid price forecasting model for the stock trading market based on AI technique. In 2024 Ninth International Conference on Informatics and Computing (ICIC), Medan, Indonesia, pp. 1-5. <https://doi.org/10.1109/ICIC64337.2024.10957173>
- [4] Xiao, Z.Y., Li, S.M., Lin, Z.X. (2012). A hybrid modeling approach for forecasting the volatility of REITs index in US market. In 2012 International Conference on Management Science & Engineering 19th Annual Conference Proceedings, Dallas, TX, USA, pp. 1861-1867. <https://doi.org/10.1109/ICMSE.2012.6414425>
- [5] Baimel, D., Belikov, J., Heistrene, L. et al. (2025). An improved and explainable electricity price forecasting model via SHAP-based error compensation approach. *IEEE Transactions on Artificial Intelligence*, 6(1): 159-168. <https://doi.org/10.1109/TAI.2024.3455313>
- [6] Nelson, D.M.Q., Pereira, A.C.M., de Oliveira, R.A. (2017). Stock market's price movement prediction with LSTM neural networks. In 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, pp. 1419-1426. <https://doi.org/10.1109/IJCNN.2017.7966019>
- [7] Yuliyono, A.D., Girsang, A.S. (2019). Artificial bee colony-optimized LSTM for bitcoin price prediction. *Advances in Science, Technology and Engineering Systems*, 4(5): 375-383. <https://doi.org/10.25046/aj040549>
- [8] Gülmez, B. (2023). Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. *Expert Systems with Applications*, 227: 120346. <https://doi.org/10.1016/j.eswa.2023.120346>
- [9] Kumar, K.S., Sree, D.I., Devi, P.Y., Pujitha, M.V. (2024). Comparative analysis of LSTM and XGBoost models for short-term Bitcoin price prediction. In 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, pp. 932-939. <https://doi.org/10.1109/ICAAIC60222.2024.10575848>
- [10] Yadav, A.K., Vishwakarma, V.P. (2024). An integrated blockchain based real time stock price prediction model by CNN, Bi LSTM and AM. *Procedia Computer Science*, 235: 2630-2640. <https://doi.org/10.1016/j.procs.2024.04.248>
- [11] Chen, X.Q., Gupta, L. (2024). Training LSTMs with circular-shift epochs for accurate event forecasting in imbalanced time series. *Expert Systems with Applications*, 238(Part C): 121701. <https://doi.org/10.1016/j.eswa.2023.121701>
- [12] Ahmad, M.S., Hadiyanto, H., Sanjaya, R. (2026). An explainable hybrid long short-term memory and extreme gradient boosting framework for daily stock price forecasting with contextual latent features. *Mathematical Modelling of Engineering Problems*, 13(3): 513-520. <https://doi.org/10.18280/mmep.130306>
- [13] Lv, C.K., Gao, B.Y., Yu, C. (2021). A hybrid transfer learning framework for stock price index forecasting. In 2021 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress, AB, Canada, pp. 355-359. <https://doi.org/10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00067>
- [14] Liu, Q.Y., Hu, Y.R., Liu, H.J. (2024). Enhanced stock price prediction with optimized ensemble modeling using multi-source heterogeneous data: Integrating LSTM attention mechanism and multidimensional gray model. *Journal of Industrial Information Integration*, 42: 100711. <https://doi.org/10.1016/j.jii.2024.100711>
- [15] Aravindhan, M., Manikavelan, D. (2025). Decomposition-based transformer-LSTM framework for stock index forecasting. In 2025 3rd World Conference on Communication & Computing (WCONF), Raipur, India, pp. 1-6. <https://doi.org/10.1109/WCONF64849.2025.11233358>
- [16] Gandhudi, M., Alphonse, P.J.A., Fiore, U., Gangadharan, G.R. (2024). Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction. *Computers and Electrical Engineering*, 118(PA): 109302. <https://doi.org/10.1016/j.compeleceng.2024.109302>
- [17] Arif, E., Suherman, S., Widodo, A.P. (2025). Predicting stock prices of digital banks: A machine learning approach combining historical data and social media sentiment from X. *Ingénierie des Systèmes d'Information*, 30(3): 687-701. <https://doi.org/10.18280/isi.300313>
- [18] Ribeiro, M.T., Singh, S., Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, San Diego, California, pp. 97-101. <https://doi.org/10.18653/v1/n16-3020>

[19] Lundberg, S.M., Lee, S.I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30: 4765-4774. https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

[20] Muhammad, D., Ahmed, I., Naveed, K., Bendeache, M. (2024). An explainable deep learning approach for stock market trend prediction. *Heliyon*, 10(21): e40095. <https://doi.org/10.1016/j.heliyon.2024.e40095>

[21] Notaria, H., Shah, S., Thopte, D., Soneji, H., Bari, P., Deulkar, K. (2024). Comparative analysis of stock price prediction using time series models. In *2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, India, pp. 1-6. <https://doi.org/10.1109/ICCUBEA61740.2024.10775112>

[22] Hossain, S., Kaur, G. (2024). Stock market prediction: XGBoost and LSTM comparative analysis. In *2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT)*, Vellore, India, pp. 1-6. <https://doi.org/10.1109/AIIoT58432.2024.10574794>

[23] Liu, Z.Z., Su, Z.H., Shang, L., Sun, H.X., Zhao, B. (2025). An approach to stock price prediction based on improved EEMD and attention-enhanced BiLSTM. *Expert Systems with Applications*, 284: 127802. <https://doi.org/10.1016/j.eswa.2025.127802>

NOMENCLATURE

Close(t) Closing stock price at trading day (t)

c_t	LSTM cell state at time (t)
$\tilde{c}_t$	Candidate cell state at time (t)
DA	Directional Accuracy
f_k	The k -th regression tree in XGBoost
f_i	Forget-gate vector at time (t)
f_{XGB}	XGBoost prediction function or meta-learner
h_t	LSTM hidden state at time (t)
i_t	Input-gate vector at time (t)
$\mathbb{I}[\cdot]$	Indicator function
K	Number of regression trees
$l(y_i, \hat{y}_i)$	Loss function between the actual and predicted values
MAE	Mean Absolute Error
$MA_{k,t}$	Moving average over (k) trading days at time (t)
n	Number of observations
o_t	Output-gate vector at time (t)
R^2	Coefficient of determination
r_t	Logarithmic return at time (t)
RMSE	Root Mean Squared Error
t	Trading-day index
x_i	Input feature vector for observation (i)
x_t	LSTM input vector at time (t)
x_t^{tabular}	Engineered multimodal tabular feature vector at time (t)
y_i	Actual target value for observation (i)
y_t	Target value predicted using information available at time (t)
$\hat{y}_i$	Predicted value for observation (i)
$\hat{y}_t^{\text{LSTM}}$	Prediction generated by the LSTM model
$\hat{y}_t^{\text{Hybrid}}$	Final prediction generated by the Hybrid LSTM–XGBoost model
$\bar{y}$	Mean of the actual target values
z_t	Augmented hybrid feature vector at time (t)