Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

Stock Price Prediction Using LSTM and XGBoost with Social Media Sentiment

Nisa Hanum Harani[1*], Marismati Marismati[2]

[1]Vocational, Computer Science, International Business Logistics University, Bandung, Indonesia
[2]Vocational, Accounting, International Business Logistics University, Bandung, Indonesia
Email: [1]nisa@ulbi.ac.id, [2]marismati@ulbi.ac.id
Correspondence Author Email: nisa@ulbi.ac.id
Submitted: **24/08/2025**; Accepted: **20/09/2025**; Published**: 21/09/2025**

Abstract−The influence of social media on financial markets is growing and motivates research on the predictive role of sentiment in stock price movements. Bank Negara Indonesia (BBNI) is part of the Danantara holding company, and BBNI's strategic position is an important indicator for measuring the performance of the broader financial ecosystem in Indonesia. This study analyzes the influence of social media sentiment on the stock price prediction of Bank Negara Indonesia (BBNI), which is part of the state-owned holding company Danantara. Historical market data is combined with sentiment indicators obtained from public conversations on X/Twitter. Daily sentiment features are then integrated with market variables, including OHLCV data, to form a combined dataset. Two machine learning approaches were employed: Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost). The results revealed contrasting patterns between the two models. The LSTM Baseline consistently produced RMSE around (≈46–65) across all scenarios. However, XGBoost-Extended is the best-performing and recommended model for sentiment-integrated prediction with RMSE (≈30–40).

Keywords: Danantara; LSTM; Price Prediction; Sentiment Analysis; XGBoost

1. INTRODUCTION

The stock market is a key component of a country's financial system that reflects its economic well-being [1]. Stock price fluctuations are typically influenced by various factors, such as company performance, economic policies, and investor psychology. In practice, decisions to buy or sell stocks are not solely based on rational considerations but also on the perceptions and emotions of market participants, which can change at any time . This situation means that public information and opinions can be the primary cause of stock price changes, especially in the short term [2]. However, recent empirical evidence increasingly challenges this view, suggesting that certain types of alternative data may have predictive power. In particular, public sentiment expressed through digital platforms has emerged as a significant variable influencing stock price behaviour [3]. This perspective is particularly relevant in the context of Indonesian banking stocks, such as Bank Negara Indonesia (BBNI), where market perceptions and investor confidence are heavily influenced by information flows outside conventional financial reports. Bank Negara Indonesia (BNI) is part of Danantara, which consolidates key national-economy players, including Bank Mandiri, BNI, BRI, Pertamina, PLN, MIND ID, and Telkom Indonesia [4]. Danantara aims to strengthen financial resilience, optimize asset management, and accelerate digital transformation across multiple sectors .

The development of information technology, particularly social media such as Twitter, has changed the way individuals give and receive information. Twitter is currently a highly active forum for sharing comments, news, and reactions to various economic and corporate issues. The speed and volume of information disseminated on Twitter have the potential to provide a direct snapshot of public opinion regarding a company or business area [5] . Therefore, tweet sentiment analysis has the potential to become a valuable secondary source of data for market prediction [6] . Previous studies have shown that sentiment indicators derived from Twitter can enhance predictive models, with Bollen et al [7] reporting accuracy rates above 80% for major stock indices. These findings highlight the value of integrating sentiment analysis into stock price prediction frameworks, particularly in emerging markets where investor behaviour is often influenced by sentiment. Although historical stock prices have been the basis for prediction models built over many years, this method is incomplete. Traditional prediction models cannot capture the psychological forces and rapidly changing mindset of the market. During crises or the announcement of important news in unstable situations, social media data can provide earlier warnings than technical indicators.

Therefore, the integration of sentiment data with historical data holds great potential for improving the accuracy of stock prediction models.

Previous studies have indicated that social media sentiment is related to stock price movements[8]. Several previous studies have been conducted to integrate sentiment data with historical data in stock price prediction. A study by Maharani et al. conducted post-training on IndoBERT using an Indonesian financial corpus to improve the accuracy of sentiment and topic analysis in the financial domain, demonstrating the potential for developing domain-specific RoBERTa models for financial sentiment analysis [9]. Another study combined tweet and news data in a stock price prediction model using MLP and LSTM, where the results showed that integrating both data sources could improve model accuracy compared to using historical data alone [10].

At the methodological level, stock price forecasting remains a challenging problem due to the non-linear and non-stationary characteristics of financial time series. Traditional linear regression models are often unable to capture this complexity. Recent advancements in machine learning, particularly in the application of Long Short-Term

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

Memory (LSTM) networks, have demonstrated superior performance in handling sequential dependencies and volatility patterns inherent in financial data . Empirical results consistently show that LSTMs outperform traditional machine learning approaches in predicting stock prices across various markets and sectors [11], [12] , [13]. These results indicate that positive public sentiment leads to upward stock price movements, while negative sentiment can serve as a precursor to downward price movements.

While many existing studies have show the potential of social media sentiment to influence stock price movements, most prediction models still rely solely on historical or technical data. These traditional models have limitations because they are unable to capture the social dynamics and real-time information disseminated through digital platforms. Thus, there is a research gap, such as the need for an informatics-based computational approach that can combine historical data with social media sentiment using machine learning methods. . A more formal approach is needed to combine both types of data under a predictive model framework. Based on this knowledge, this study combines financial indicators with Twitter-based sentiment features to predict BBNI stock price movements. By combining conventional data such as closing prices and trading volumes with sentiment variables including polarity, subjectivity, and the number of positive or negative tweets, this model aims to provide a more comprehensive understanding of the factors influencing BBNI stock performance. This study is aimed at filling this gap by applying two different algorithms, namely Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost), to systematically predict BBNI stock price movements.

2. RESEARCH METHODOLOGY

Previous research on stock price prediction based on machine learning can be grouped into three main directions. First, time-series-based deep learning approaches emphasise the ability of LSTM and BiLSTM in predicting stock price patterns. Roondiwala et al. [14] [15] proved that the LSTM architecture is capable of capturing non-linear relationships and improving accuracy compared to traditional regression. In Indonesia, Afrianto et al. [16]used BiLSTM with public sentiment factors on BBCA stocks and demonstrated that integrating price and public sentiment data resulted in lower prediction errors (RMSE). Second, sentiment analysis-based approaches that focus on transforming public opinion into quantitative features. Rezaeinia et al. [17] introduced the Improved Word Vectors (IWV) method to improve the accuracy of word embedding in sentiment classification. This study is relevant because the quality of text representation greatly affects the performance of Twitter-based price prediction. Similar research also emphasises the role of pre-trained embeddings such as Word2Vec and GloVe for sentiment analysis [18] Third, hybrid approaches combine statistical models with machine learning. Al Amrani et al. [19]proposed a hybrid Random Forest–SVM (RFSVM) for sentiment classification, which proved to be more accurate than single algorithms. In the financial domain, hybrid approaches are also adopted by combining stock price prediction models with transformer-based sentiment indices such as BERT. For example, research by Hiew et al. [20] shows that combining BERT-based sentiment indices with LSTM models can improve stock return prediction capabilities.

2.1 Object and Scope

The object of this study is the stock of PT Bank Negara Indonesia (Persero) Tbk (BBNI), one of the main holding companies of Danantara. BBNI was selected based on its strategic position in the national banking sector and its relevance to market perceptions of Danantara as a whole. This study applies an integrative approach that combines historical stock price data with sentiment data from Twitter to build a stock price movement prediction model for BBNI .The social media as an alternative data source is based on findings that public opinion spread through digital platforms can influence market dynamics, particularly in highly exposed stocks such as BBNI. Historical data was obtained from investing.com, which provides market indicators such as closing prices, transaction volumes, and daily price changes. Meanwhile, sentiment data was obtained from Twitter by collecting tweets containing the keywords "BBNI", "Bank Negara Indonesia", and other relevant terms as shown in Table I. The tweets obtained were classified into positive, negative, and neutral sentiments using the RoBERTa language model adapted for Indonesian. This sentiment classification aims to capture the nuances of public opinion towards the issuer on specific days. The initial stage of the research began with data pre-processing. For text data, cleaning was performed by removing URLs, punctuation marks, symbols, and meaningless words. The cleaned tweets are then analysed using a deep learning-based sentiment classification model. The RoBERTa model was chosen for its ability to understand complex linguistic contexts in Indonesian, as demonstrated by Maharani et al. in the development of a domain-specific model for local social media. After obtaining daily sentiment data, the integration process is carried out by combining the sentiment dataset and historical data based on stock trading dates.

2.2 Research Stage

The research activities adopted the CRISP-DM framework [21], [22] consisting of: (1) business understanding (identifying BBNI's role in Danantara), (2) market and sentiment data collection, (3) text preprocessing (URL cleaning, punctuation, slang normalization), (4) sentiment classification with a transformer model, (5) integration of market and sentiment datasets, (6) model development using LSTM and XGBoost, and (7) model performance evaluation using RMSE, MAE, and MAPE metrics , as illustrated in Figure 1. Sentiment classification is performed

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

using a customised model to generate indicators including 15 sentiment features. To facilitate collaboration between the two datasets, both have been time-normalised at the daily level. For trading days with no sentiment data (e.g., non-trading days or missing observations), sentiment variables are either forward-filled or marked as missing, depending on the completeness threshold. After normalisation, the datasets are merged into a single daily panel indexed by date, ensuring that for each trading day, the stock's OHLCV values and combined sentiment features are available in the same row. To avoid lookahead bias, all predictor variables are lagged by one day, meaning that features used to predict the closing price on day t only come from day *t–1*.

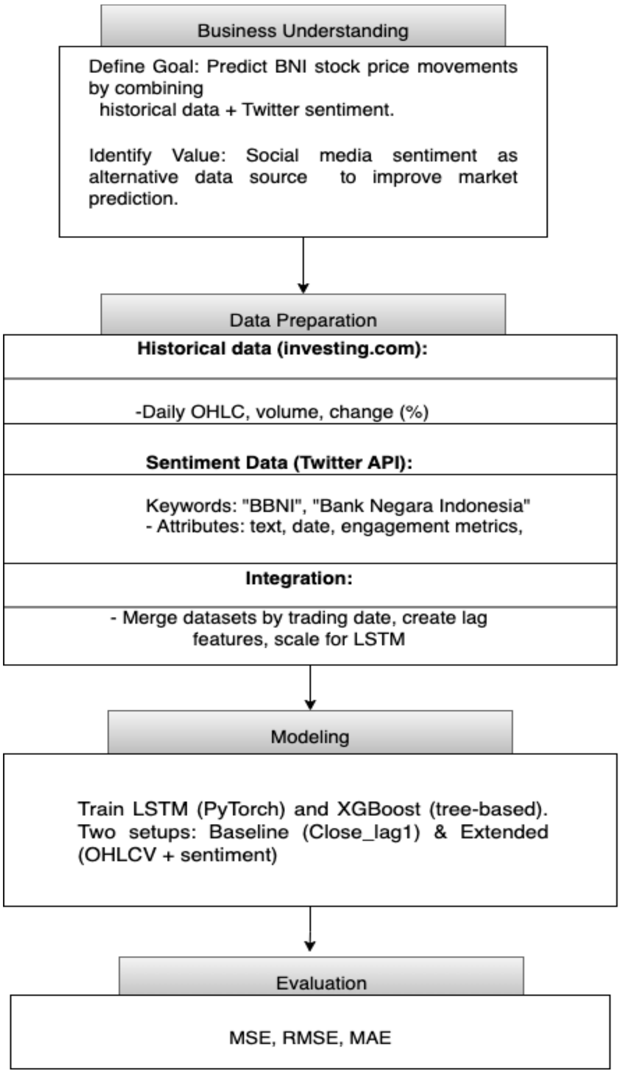

Figure 1. Methodology

This is implemented by creating a lag-1 version of each feature (e.g., Close_lag1, Volume_lag1, avg_signed_sentiment_lag1). The target variable is defined as Close_t, which represents the actual closing price on day t. Missing values due to the lag (typically in the first observation) are removed. The final dataset reflects a combination of fundamental factors (market-based) and alternative factors (public sentiment-based) that are expected to improve the accuracy of stock price predictions. The stock market dataset consists of OHLCV (Open, High, Low, Close, Volume) records comprising 7 daily features for Bank Negara Indonesia (BNI), collected from official market sources. Twitter sentiment data is generated through a separate sentiment analysis pipeline applied to raw tweet text containing keywords related to BNI shown in Table 1.

Table 1. Dataset Description

Type	Source	Size	Attributes
sentiment	x/twitter	5222	15 (conversation_id_str, created_at, favourite_count, full_text, id_str, image_url, in_reply_to_screen_name, lang, location, quote_count, reply_count, retweet_count, tweet_url, user_id_str, username)
historical	website (investing.com)	164 / day	7 (Date, Last, Open, High, Low, Volume, Change%)

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

3. RESULT AND DISCUSSION

This section describes the results of the experiment and provides an in-depth analysis of their implications. Each section describes the process performed, explains the methodological choices made, and presents the patterns observed in the dataset.

3.1 Business Understanding

As one of Indonesia's largest state-owned banks, PT Bank Negara Indonesia (Persero) Tbk (BNI) plays a central role in the newly formed Danantara group, a strategic holding entity that consolidates key state-owned companies across various sectors. This consolidation has significant implications for market perception, as developments within Danantara—whether related to corporate governance, investment strategies, or cross-sector synergies—can influence investor sentiment not only toward the group as a whole but also toward individual member companies such as BNI. While traditional stock price forecasting models in Indonesia focus on macroeconomic indicators, company fundamentals, and historical price patterns, these models often overlook real-time public sentiment signals that can change rapidly in response to news, policy announcements, or market rumours. In particular, Twitter and other social media platforms act as amplifiers of investor sentiment, with sentiment changes potentially preceding or reinforcing price movements.

While there is growing international evidence of the predictive value of social media sentiment, research in the Indonesian market—particularly regarding the interrelated impact of strategic ownership structures such as Danantara—remains limited. The Efficient Market Hypothesis (EMH) states that stock prices fully reflect all available information, meaning that predicting price movements based on publicly available data should be impossible. However, recent empirical evidence challenges this assumption, suggesting that non-traditional and real-time sources of information—such as social media sentiment—may contain predictive signals that are overlooked by conventional models.

3.2 Data Preparation

The empirical analysis in this study is based on two main data sources: (1) historical stock market data of PT Bank Negara Indonesia (Persero) Tbk (BNI), and (2) text data obtained from Twitter containing public discussions related to BNI and its affiliation with the Danantara group. The integration of these two datasets facilitates the analysis of quantitative market indicators and qualitative sentiment signals, enabling a more holistic understanding of stock price movements.

3.2.1 Stock Market Data

Historical price data for BNI, as shown in Figure 2, including opening, high, low, and closing prices (OHLC) as well as trading volume, was obtained from a reliable financial market repository, ensuring data completeness and accuracy. This data set covers several months to capture typical trading patterns and periods of high volatility. The closing price (close price) is used as the target variable in the forecasting model, while the previous OHLC values and trading volume serve as base predictors in accordance with established financial time series modelling practices. To maintain consistency with sentiment data, stock prices are resampled to a consistent 30-minute interval, reflecting the short-term forecasting horizon that is the focus of this study.

Out[247…

Date	Open	High	Low	Close	Volume	Dividends	Stock Splits
2025-02-28 00:00:00+07:00	3918.961387	3918.961387	3680.336103	3698.691895	156574800	0.0	0.0
2025-03-03 00:00:00+07:00	3836.360523	3955.673171	3781.293148	3909.783691	156616300	0.0	0.0
2025-03-04 00:00:00+07:00	3891.427813	3937.317292	3827.182543	3873.072021	71340000	0.0	0.0
2025-03-05 00:00:00+07:00	3900.605740	4166.764720	3900.605740	4093.341553	101995600	0.0	0.0
2025-03-06 00:00:00+07:00	4175.942713	4313.611154	4166.764817	4212.654297	104226000	0.0	0.0
...	...	...	...	...	...	...	...
2025-07-10 00:00:00+07:00	3980.000000	4140.000000	3980.000000	4110.000000	44060900	0.0	0.0
2025-07-11 00:00:00+07:00	4130.000000	4220.000000	4120.000000	4180.000000	48939300	0.0	0.0
2025-07-14 00:00:00+07:00	4190.000000	4190.000000	4000.000000	4040.000000	68377100	0.0	0.0
2025-07-15 00:00:00+07:00	4040.000000	4070.000000	4010.000000	4050.000000	30497400	0.0	0.0
2025-07-16 00:00:00+07:00	4130.000000	4150.000000	4080.000000	4130.000000	66349800	0.0	0.0

Figure 2 . Stock Dataset

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

3.2.2 Sentiment data

Twitter data was collected via the Twitter API, using keyword-based queries and cashtags targeting explicit references to BNI (e.g., "Bank Negara Indonesia," "BNI") and its association with the Danantara group during the time period from the formation of Danantara on 28 February 2025 to 15 July 2025, as shown in Figure 3. This approach ensures coverage of direct corporate mentions as well as broader market discussions that may indirectly influence investor perceptions. The collected tweets include metadata such as timestamps, user locations (if available), and text content.

	roberta	finbert	vader	text	agree_all	agree_majority	majority_label
0	neutral	negative	positive	kegiatan ini dilaksanakan dengan menggandeng n...	False	False	no agreement
1	negative	neutral	positive	rakyat butuh tentara beneran buat menjaga keda...	False	False	no agreement
2	neutral	negative	positive	hai kak , dengan senang hati mimin bantu infoi...	False	False	no agreement
3	negative	neutral	positive	aku kira presiden prabowo pintar, nyatanya ema...	False	False	no agreement
4	neutral	negative	positive	turnamen ini akan diikuti oleh negara, dan tim...	False	False	no agreement
...	...	...	...	...	...	...	...
153	positive	neutral	positive	industri halal di berbagai sektor terus dikemb...	False	True	positive
154	positive	neutral	positive	sudah aman, terima kasih atas respon cepat nya	False	True	positive
155	positive	negative	positive	tidak terasa, sudah di pertengahan tahun . gim...	False	True	positive
156	positive	neutral	positive	promo hemat tiap sabtu di hartono! dapatkan di...	False	True	positive
157	positive	neutral	positive	Pelopor Indonesia menerima kartu BNI Bank BNI ...	False	True	positive

Figure 3. Sentiment Dataset

3.2.3 Preprocessing and Feature Extraction

After data collection, the Twitter dataset underwent a step-by-step pre-processing process. Non-text elements such as URLs, mentions, hashtags, and retweet markers were removed to reduce noise. A domain-specific normalisation dictionary was used to standardise Indonesian financial slang, abbreviations, and informal expressions commonly used in online discussions. Language detection ensures that only Indonesian-language tweets are retained for analysis, given the composition of local investors in the market. Using the TextBlob sentiment analysis framework—which has been adapted for Indonesian—two main sentiment measures are generated: polarity, which indicates the degree of positive or negative tone, and subjectivity, which indicates the extent to which the content is opinion versus fact. Additional variables, including the number of positive tweets, the number of negative tweets, and the frequency of tweets per 30-minute interval, are calculated to capture the intensity and direction of sentiment dynamics, as shown in Table 2.

Table 2 . Dataset

Column Name	Description
date	The 30-minute trading period in which stock market data and sentiment data are stored
Open	The opening price of BNI shares at the start of the specified 30-minute interval, representing the first price traded during that period.
High	The highest transaction value of BNI shares in the specified 30-minute interval
Low	The lowest transaction price of BNI shares during the specified 30-minute interval.
Close	The final transaction price of BNI shares at the end of the specified 30-minute interval
Volume	The total number of shares traded during the specified interval
Dividends	The amount of dividends per share distributed during a specific period,
Stock Splits	The ratio of the number of shares divided (stock splits) that occurred during a specific period
avg_signed_sentiment	The average sentiment polarity value for all tweets related to BNI in that interval ranges from -1 (very negative) to +1 (very positive), providing a measure of the direction of public opinion.
count_positive	The total number of tweets classified with positive sentiment towards BNI during the observation period, indicating a positive market perception.
count_negative	Total number of tweets classified with negative sentiment towards BNI during the observation period
count_neutral	Total number of tweets classified with neutral sentiment towards BNI during the observation period

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

Column Name	Description
total_tweets	Total volume of tweets related to BNI during the observation period, covering all sentiment categories, and serving as an indicator of overall public attention

3.2.4 Data Integration

To align sentiment indicators with stock market variables, a time-window matching procedure was applied. For each 30-minute trading interval, sentiment metrics were calculated based on tweets posted in the previous interval, ensuring temporal consistency and preventing look-ahead bias. The resulting dataset consists of a combination of financial market features and sentiment features, enabling direct comparison between the basic LSTM model (only financial indicators) and the developed model incorporating sentiment features, as shown in Figure 4.

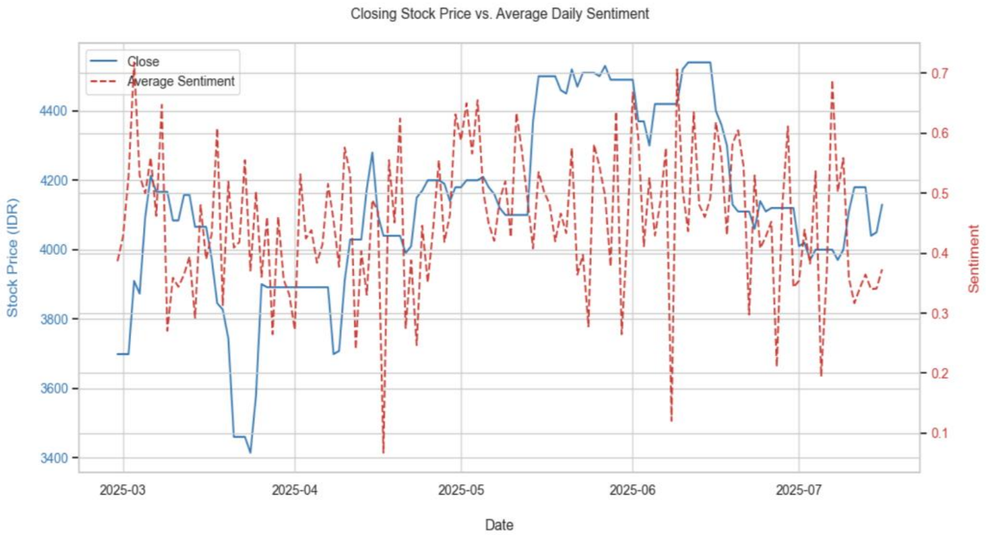

Figure 4 . Closing Stock Price vs Average Daily Sentiment

3.3 Modelling

In this study, a model was developed to evaluate sentiment features for stock price forecasting. Two models commonly used in the sentiment-stock literature represent different yet complementary paradigms. First, Long Short-Term Memory (LSTM) represents a sequential architecture designed to capture nonlinear temporal dependencies and short-term patterns in financial time series. Second, Extreme Gradient Boosting (XGBoost) represents a decision tree-based boosting approach known for its robustness on tabular data, capable of modelling nonlinear interactions between features. To compare the models, this study created two comparison scenarios. Baseline uses only financial information, such as closing prices (e.g., Close_lag1).

Extended adds input signals, such as lagged OHLCV and sentiment features related to BNI (e.g., average polarity, number of positive/negative tweets, and conversation intensity). By directly comparing Baseline and Extended on LSTM and XGBoost, we can estimate the marginal contribution of sentiment features over stock information while testing the sensitivity of each learning paradigm to feature dimension expansion.

3.3.1 Data Split

In financial time series modelling, an important aspect is how data is split for training, validation, and testing. Unlike tabular data, stock price data has time-dependent properties, so the split must be done in chronological order. Therefore, the split and validation strategies must be adjusted to the temporal dynamics so that the evaluation results are more representative of real conditions. In this study, several data splitting schemes were used, as shown in Table 3.

Table 3 . Split Scenario

Scheme	Train%	Val%	Test
Split 80/10/10	80	10	10
Split 70/15/15	70	15	15
Split 60/20/20	60	20	20
Walk-forward (3f)	70	10	20

3.3.2 Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM)

In this study, a model was developed to evaluate sentiment features for stock price forecasting. Two models commonly used in the sentiment-stock literature represent different but complementary paradigms.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

a. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM**)** represents a sequential architecture designed to capture non-linear temporal dependencies and short-term patterns in financial time series. LSTM is a type of recurrent neural network (RNN) designed to capture long-term and short-term dependencies in time series data. This architecture utilises *gate* mechanisms (input, forget, and output) to regulate information flow, making it more effective than conventional RNNs, which are prone to vanishing gradient problems. LSTM is often used in previous research because, in the context of stock price prediction, daily prices have strong temporal patterns and are influenced by short-term dynamics. In this study, two configurations were tested: LSTM Baseline, which only uses simple lags of the closing price (Close_lag1), and LSTM Extended, which adds OHLCV lags as well as sentiment and Twitter behaviour variables.

The comparison of these two configurations aims to assess the extent to which additional information from sentiment can improve prediction accuracy in sequential models, while identifying potential overfitting risks when feature complexity is increased on a limited dataset.

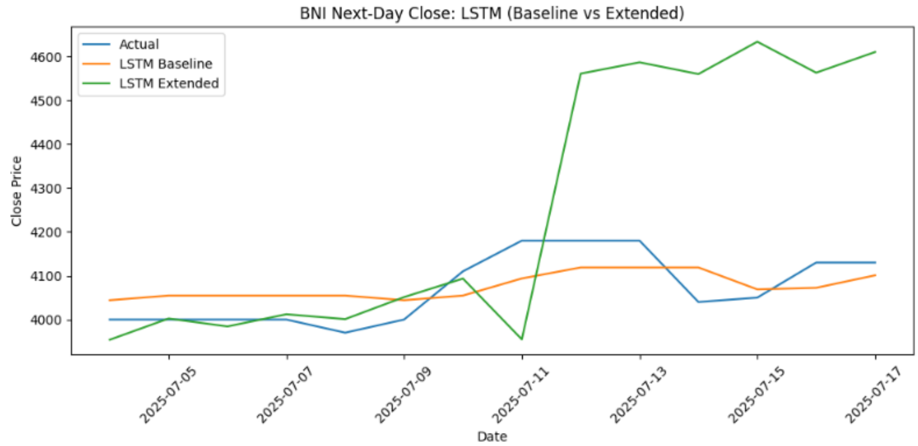

Figure 5 . BNI Next-Day Close : LSTM

Figure 5 shows a comparison between the LSTM Baseline and LSTM Extended predictions of BNI's actual closing price with a horizon of one day ahead. It can be seen that LSTM Baseline (orange line) is able to follow the actual trend relatively stably, although its predictions tend *to underfit* with flatter fluctuations than the actual data. In contrast, LSTM Extended (green line) exhibits a prediction pattern that deviates drastically after the midpoint of the test period. Extended predictions surge far above actual values, resulting in significant systematic bias. This phenomenon reflects issues of overfitting and scale mismatch when sentiment features and Twitter behaviour are added to the LSTM model on a relatively limited dataset.

b. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a decision tree-based boosting approach known for its robustness on tabular data, capable of modelling non-linear interactions between features.

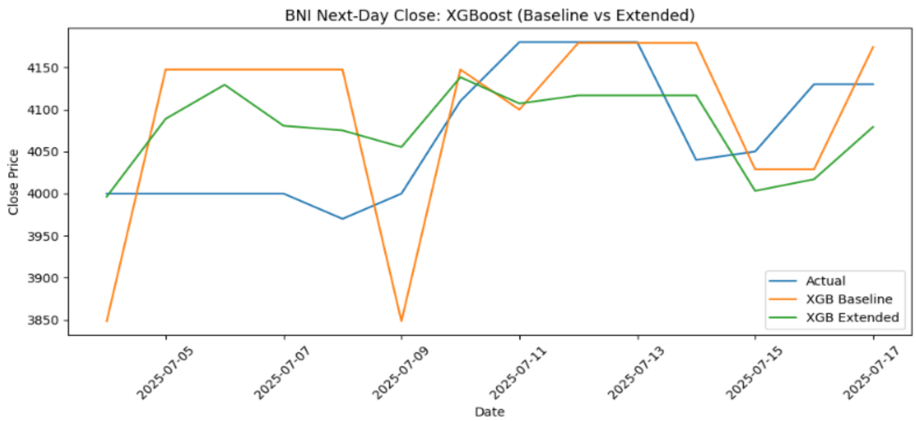

Figure 6 . BNI Next-Day Close : XGBoost

The main advantage of XGBoost is its ability to combine the results of multiple decision trees iteratively, thereby capturing non-linear relationships between variables with efficient computational performance. In this experiment, two configurations were tested. XGBoost Baseline only used one lag feature of the closing price (Close_lag1), while XGBoost Extended added OHLCV lag as well as sentiment and Twitter behaviour features. The purpose of this comparison was to evaluate whether the addition of non-financial variables could improve the performance of

ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

tree-based models. Figure 6 shows the prediction results of XGBoost Baseline and Extended against the actual closing price of BNI. XGBoost Baseline (orange line) shows a more volatile pattern, with some extreme predictions that deviate significantly from the actual values, such as a sharp decline around 9 July.
This is consistent with quantitative metrics, where the Baseline RMSE is above 110 in most split schemes. In contrast, XGBoost Extended (green line) shows a smoother and more stable prediction pattern, more closely following the actual price trend despite a slight *lag*. The addition of sentiment and behaviour features has proven to help XGBoost capture price movement dynamics better, resulting in an RMSE reduction of around 20–40% compared to the Baseline in various scenarios.

3.4 Evalution

The model evaluation was conducted using two main approaches, namely static split (80/10/10, 70/15/15, and 60/20/20) and walk-forward validation (3 fold). The LSTM model was trained under the conditions : 2 hidden layers, 50 neurons per layer, *batch size* 32, *learning rate* 0.001, *loss function* MSE, and *optimizer* Adam. The XGBoost model uses the following parameters: n_estimators = 500, max_depth = 6, and learning_rate = 0.1. All experiments are tracked using *mlflow* and *Weights & Biases (wandb)* to ensure reproducibility of results. The problem was formulated as univariate regression with the target $yt + 1$, which is the daily closing price on the following day. The prediction error function used is Mean Squared Error (MSE), with reporting metrics Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). For additional interpretation, Mean Absolute Percentage Error (MAPE) and Normalised RMSE (NRMSE_mean) are calculated as relative measures against the price scale.

Table 4 . Evaluation Result on Spliting Scenario

Scheme	Model	Features	MSE	RMSE	MAE
Split 80/10/10	XGBoost	Baseline	13035.22	114.17	96.29
Split 80/10/10	XGBoost	Extended	5912.78	76.89	69.85
Split 80/10/10	LSTM	Baseline	3347.01	57.85	54.58
Split 80/10/10	LSTM	Extended	36957.63	192.24	143.98
Split 70/15/15	XGBoost	Baseline	10,482.03	102.38	82.76
Split 70/15/15	XGBoost	Extended	3959.74	62.93	51
Split 70/15/15	LSTM	Baseline	2622.68	51.21	35.83
Split 70/15/15	LSTM	Extended	68305.18	261.35	166.54
Split 60/20/20	XGBoost	Baseline	8974.27	94.73	74.82
Split 60/20/20	XGBoost	Extended	3612.71	60.11	48.61
Split 60/20/20	LSTM	Baseline	2162.89	46.51	33
Split 60/20/20	LSTM	Extended	146,726.55	383.05	208.1

Table 4 shows the evaluation results of the model on three static split scenarios (80/10/10, 70/15/15, and 60/20/20). There are noticeable differences in performance patterns between LSTM and XGBoost. LSTM Baseline consistently achieves the lowest RMSE across all static scenarios, with values of 46.51 in the 60/20/20 split and 51.21 in the 70/15/15 split. This indicates that on a limited dataset, sequential models are quite effective in capturing short-term dependencies in stock prices. Conversely, LSTM Extended experiences significant degradation across all splits, for example, RMSE of 192.24 (80/10/10) and 383.05 (60/20/20), indicating overfitting due to the addition of feature dimensions. Meanwhile, XGBoost Extended consistently outperformed the Baseline in every split. Relative improvements ranged from 32% to 37% (e.g., from RMSE 114.17 to 76.89 in the 80/10/10 split).

This consistency confirms that integrating sentiment and Twitter behaviour features enriches the tree boosting model without sacrificing stability. Thus, the static split results show a contrast: LSTM is effective with minimal features, while XGBoost is optimal when features are expanded. This is important given the relatively limited dataset (~140 observations), making robustness to split changes an indicator of model reliability.

Table 5 . Validation Result (3-Fold)

Scheme	Model	Features	RMSE	MAE
WF (k=1)	LSTM	Baseline	51.57	36.25
WF (k=1)	XGBoost	Extended	106.96	69.77
WF (k=2)	LSTM	Extended	49.49	37.70
WF (k=2)	XGBoost	Extended	62.02	55.60
WF (k=3)	LSTM	Baseline	64.70	59.22
WF (k=3)	XGBoost	Extended	63.18	46.49

Table 5 presents the evaluation using the walk-forward validation approach (3 folds), which better reflects real-world implementation conditions when the model is updated over time. The results show that the LSTM Baseline returns to stability with an RMSE of around 51–65 across all folds, while the LSTM Extended fluctuates sharply: performing very well on the second fold (RMSE 49.49) but poorly on the third fold (RMSE 325.92). This fluctuation indicates that LSTM Extended is highly sensitive to variations in data distribution across different folds. In contrast, XGBoost Extended remained consistently superior to Baseline on every fold. The most significant improvement occurred on fold 2 (RMSE 62.02 vs 106.08), with a relative improvement of approximately 41%.

Overall, with minimal features, LSTM-Baseline achieves the lowest RMSE among all baseline configurations; however, when sentiment features are included, XGBoost-Extended delivers the best performance and stability, outperforming LSTM-Extended across all splits. Therefore, the preferred model depends on the feature regime: LSTM-Baseline for sparse/price-only setting, XGBoost-Extended for sentiment-integrated forecasting

4. CONCLUSION

This study aims to evaluate the two machine learning algorithms, LSTM and XGBoost, in predicting the daily closing price of Bank Negara Indonesia (BBNI) stock with a one-day horizon. The two feature configurations tested were: *Baseline* (only closing price lag) and *Extended* (combination of OHLCV and Twitter sentiment).The baseline LSTM consistently produced the lowest error across various data splitting schemes, confirming the effectiveness of simple sequential models in capturing short-term autoregressive patterns. However, LSTM performance declined sharply in the Extended configuration due to overfitting and dataset size limitations . Conversely, XGBoost shows the opposite pattern. The Baseline configuration tends to be less accurate, but performance improves significantly in the Extended configuration. The addition of OHLCV and public sentiment features has proven to enrich tree-based models, resulting in a 30–40% reduction in RMSE compared to the Baseline, as well as greater stability in walk-forward evaluation. Thus, XGBoost Extended emerges as the most reliable model in leveraging the combination of financial and non-financial signals. These findings have two main implications. First, architectural complexity does not always correlate with accuracy, especially on limited datasets; simple models such as LSTM Baseline can outperform more complex versions. Second, tree-based algorithms such as XGBoost are relatively more resilient to feature addition and data distribution variations, making them a more stable choice for short-term stock price prediction based on market data and sentiment integration.

ACKNOWLEDGMENT

The author would like to express his gratitude to the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology for the funding support provided. This research was funded through research grants from the Directorate, and therefore, every publication, paper, or dissemination of the research results must acknowledge the funding agency as a form of acknowledgment and academic responsibility.

REFERENCES

[1] A. Bagheffar and C. Saous, "The Impact of Investor Sentiment on Stock Returns in the Indonesian Stock Market: An Econometric Study," *Finance and Business Economics Review*, vol. 7, pp. 166–183, Dec. 2023.

[2] G. Liu, Y. Yang, W. Mo, W. Gu, and R. Wang, "Private Placement, Investor Sentiment, and Stock Price Anomaly," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 27, no. 5, pp. 771–779, Sep. 2023, doi: 10.20965/jaciii.2023.p0771.

[3] D. Choi, C. J. Shallue, Z. Nado, J. Lee, C. J. Maddison, and G. E. Dahl, "On Empirical Comparisons of Optimizers for Deep Learning," Jun. 2020, [Online]. Available: http://arxiv.org/abs/1910.05446

[4] J. M. Simanjuntak and K. N. Widyadhana, "Danantara, the SOE Superholding, and the Pillar of Indonesia's Economic Future?," *International Journal of Economics Development Research*, vol. 6, no. 4, pp. 1777–1796, Jun. 2025.

[5] P. Patel, "Real-Time Sentiment Analysis of Twitter Streams for Stock Forecasting," *International Journal of Computer Trends and Technology*, vol. 72, no. 5, pp. 204–209, May 2024, doi: 10.14445/22312803/ijctt-v72i5p125.

[6] M. Mokhtari, A. Seraj, N. Saeedi, and A. Karshenas, "The Impact of Twitter Sentiments on Stock Market Trends," Isfahan University of Technology, Isfahan, Iran, 2023.

[7] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *J Comput Sci*, vol. 2, no. 1, pp. 1–8, Mar. 2011, doi: 10.1016/j.jocs.2010.12.007.

[8] Z. Li, "The Impact of Social Media Sentiment on Stock Price Changes," *Advances in Economics, Management and Political Sciences*, vol. 170, no. 1, pp. 49–59, Jun. 2025, doi: 10.54254/2754-1169/2025.lh23972.

[9] N. P. I. Maharani, Y. Yustiawan, F. C. Rochim, and A. Purwarianti, "Domain-Specific Language Model Post-Training for Indonesian Financial NLP," Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.09736

[10] H. Zolfagharinia, M. Najafi, S. Rizvi, and A. Haghighi, "Unleashing the Power of Tweets and News in Stock-Price Prediction Using Machine-Learning Techniques," *Algorithms*, vol. 17, no. 6, Jun. 2024, doi: 10.3390/a17060234.

[11] T. Fischer and C. Krauss, "Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions," *Eur J Oper Res*, vol. 270, no. 2, pp. 654–669, Oct. 2018, doi: 10.1016/j.ejor.2017.11.054.

[12] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. C. Khatri, "Predicting stock market index using LSTM," *Machine Learning with Applications*, vol. 9, p. 100320, Sep. 2022, doi: 10.1016/j.mlwa.2022.100320.

[13] E. Arif, S. Suherman, and A. P. Widodo, "Predicting Stock Prices of Digital Banks: A Machine Learning Approach Combining Historical Data and Social Media Sentiment from X," *Ingenierie des Systemes d'Information*, vol. 30, no. 3, pp. 687–701, Mar. 2025, doi: 10.18280/isi.300313.

[14] M. Roondiwala, H. Patel, and S. Varma, "Predicting Stock Prices Using LSTM," *International Journal of Science and Research*, vol. 6, Apr. 2017, [Online]. Available: https://www.quandl.com/data/NSE

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1380−1389
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8284

[15] M. L. Thormann, J. Farchmin, C. Weisser, R. M. Kruse, B. Safken, and A. Silbersdorff, "Stock Price Predictions with LSTM Neural Networks and Twitter Sentiment," *Statistics, Optimization and Information Computing*, vol. 9, no. 2, pp. 268–287, 2021, doi: 10.19139/soic-2310-5070-1202.

[16] N. Afrianto, D. H. Fudholi, and S. Rani, "Prediksi Harga Saham Menggunakan BiLSTM dengan Faktor Sentimen Publik," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 41–46, Feb. 2022, doi: 10.29207/resti.v6i1.3676.

[17] S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings," *Expert Syst Appl*, vol. 117, pp. 139–147, Mar. 2019, doi: 10.1016/j.eswa.2018.08.044.

[18] E. Ria Devina, "Sentiment Analysis between VADER and EDA for the US Presidential Election 2020 on Twitter Datasets," *Journal of Applied Data Sciences*, vol. 2, no. 1, pp. 8–18, 2021.

[19] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, "Random forest and support vector machine based hybrid approach to sentiment analysis," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 511–520. doi: 10.1016/j.procs.2018.01.150.

[20] J. Z. G. Hiew, X. Huang, H. Mou, D. Li, Q. Wu, and Y. Xu, "BERT-based Financial Sentiment Index and LSTM-based Stock Return Predictability," *arXiv preprint arXiv:1906.09024*, Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.09024

[21] J. S. Saltz, "CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps," in *IEEE International Conference on Big Data*, 2021, pp. 2337–2344. doi: 10.1109/BigData52589.2021.9671634.

[22] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 526–534. doi: 10.1016/j.procs.2021.01.199.