

# Bitcoin Trend Prediction Using Bitcoin Influence Index Weighted Sentiment and Hybrid Ensemble Learning

Akhlis Munazilin<sup>1,\*</sup>, Mochamad Agung Wibowo<sup>2</sup>, Rizky Parluka<sup>3</sup>

<sup>1</sup>Doctoral Program of Information Systems, Postgraduate School, Universitas Diponegoro, 50241, Semarang, Indonesia

<sup>1</sup>Department of Computer Science, Faculty of Science and Technology, Universitas Ibrahimy, 68374, Situbondo, Indonesia

<sup>2</sup>Postgraduate School, Universitas Diponegoro, 50241, Semarang, Indonesia

<sup>3</sup>Department of Informatics, Faculty of Computer Science, University of Pembangunan Nasional "Veteran" Jawa Timur, 60294, Surabaya, Indonesia

(Received: March 11, 2026; Revised: May 15, 2026; Accepted: September 1, 2026; Available online: October 14, 2026)

## Abstract

Predicting Bitcoin trends is a complex problem because market movements are influenced not only by historical data and technical indicators but also by sentiment and rapidly evolving information disseminated through social media. Previous studies have shown that sentiment information can improve cryptocurrency prediction; however, most approaches aggregate opinions without explicitly distinguishing the characteristics of their information sources. This study develops the Bitcoin Influence Index (BII) as a multidimensional representation of social media account influence based on four dimensions: Reach Influence, Credibility Influence, Activity Influence, and Content Influence. BII is integrated with FinBERT based sentiment analysis to construct BII weighted sentiment, which is subsequently transformed into temporal features through lag, rolling average, momentum, interaction, and shock measures. Next day Bitcoin direction is predicted using a hybrid ensemble that combines XGBoost, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) as base learners, with Logistic Regression as the meta learner. The experiment compares two scenarios, Without BII and with BII, using 96 daily observations obtained from the integration of market and social media data. The results show that The with BII scenario increases accuracy from 40.00% to 75.00%, recall from 14.29% to 78.57%, F1-score from 0.2500 to 0.8148, and area under the curve from 0.3809 to 0.6667, while precision changes from 100.00% to 84.62%. Paired t-test yields  $p = 0.030771$ , and Wilcoxon signed rank produces  $p = 0.034808$ , while McNemar produces  $p = 0.065430$ . These findings suggest that opinion source characteristics can provide additional predictive information when integrated with sentiment, technical indicators, and temporal models. The main contribution of this study lies in extending sentiment aware forecasting toward influence aware Bitcoin trend prediction, an approach that considers not only opinion polarity but also the multidimensional characteristics of the information sources conveying those opinions.

**Keywords:** Bitcoin Trend Prediction, Bitcoin Influence Index, Influence Aware Sentiment, FinBERT, Hybrid Ensemble Learning

## 1. Introduction

Bitcoin has developed into one of the most valuable digital assets and has become extensively traded and studied in financial data science [1], [2], [3], [4]. Unlike many traditional assets, Bitcoin is traded continuously in a highly information-intensive digital environment and exhibits relatively high volatility [5], [6], [7], [8]. Cryptocurrency price changes can be associated with market dynamics, sentiment, regulation, technological developments, and other external factors that influence investor expectations [9], [10], [11], [12]. Sarmini et al. [13] demonstrated that cryptocurrency volatility is an important consideration in risk analysis, whereas Quang et al. [14] showed that market volatility and dynamic price movements require Bitcoin forecasting models and feature representations capable of capturing temporal dependencies.

Bitcoin prediction methods have evolved from statistical forecasting toward machine learning, deep learning, and multimodal forecasting. Historical prices and trading volumes remain important baseline information, while technical indicators such as Moving Average, Relative Strength Index, Moving Average Convergence Divergence, and Bollinger Bands are widely used to transform price movements into trend and momentum information [5], [15], [16], [17]. Ye et

\*Corresponding author: Akhlis Munazilin (akhlimunazilin@students.undip.ac.id)

DOI: <https://doi.org/10.47738/jads.v7i4.1322>

This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

© Authors retain all copyrights

al. [18] combined price information, technical indicators, and sentiment indices in a stacking ensemble based on Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks. Similarly, Lee [19] employed attention-based recurrent models with technical indicators to predict Bitcoin trends. These studies indicate that the quality of feature representation can be as important as the choice of prediction algorithm.

Another increasingly prominent development is the integration of sentiment information. Bitcoin has a large digital discussion ecosystem, making Twitter/X, Reddit, news outlets, and financial platforms valuable sources of alternative data. Arslan [20] identified Twitter as a source of real-time information that can represent users' emotions and perceptions of cryptocurrency, whereas Saraswathi et al. [21] combined historical prices, social media sentiment, and ensemble deep learning. El Abaji and Haraty [22] found that sentiment integration improved LSTM performance but produced different effects on Prophet and Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors (SARIMAX). These findings indicate that sentiment information can be relevant to Bitcoin forecasting; however, its inclusion does not automatically improve the performance of every forecasting model.

A remaining limitation concerns how sentiment information is constructed. In many previous studies, tweets are classified as positive, neutral, or negative and subsequently aggregated into daily sentiment indicators. Although practical, this approach implicitly treats the contributions of individual observations as relatively equivalent. For example, a positive tweet from an account with substantial audience reach, an established account history, consistent activity, and relevant Bitcoin-related content may receive the same contribution during aggregation as a positive tweet from an account with a limited audience or little relevance to Bitcoin. This condition can be described as an unweighted sentiment assumption, in which sentiment polarity is aggregated without explicitly accounting for the characteristics and influence of the information source.

The literature has begun to demonstrate the importance of source-level information. Feizian and Amiri [23] assigned weights to sentiment polarity scores based on normalized follower counts and showed that influence-weighted sentiment could improve cryptocurrency prediction performance. They also emphasized that user characteristics, such as follower and friend counts, had not been extensively utilized in previous studies. Jahanbin and Chahooki [24] specifically incorporated influencer data into cryptocurrency trend prediction. These findings provide an empirical basis for the argument that opinions originating from different sources may carry different levels of informational value.

Nevertheless, follower count represents only one dimension of influence. Popularity does not necessarily correspond to credibility, credible accounts are not always active, and active accounts are not necessarily relevant to Bitcoin. From a communication perspective, the two-step flow theory explains how information can be mediated by opinion leaders before reaching broader audiences, although digital media have made information flows more complex than a simple two-stage process [25]. In the context of market prediction, Akdogan and Anbar [26] further demonstrated that social media information encompasses not only sentiment polarity but also social, cognitive, and behavioral aspects.

To address this research gap, this study develops the Bitcoin Influence Index (BII) as a composite representation of account influence characteristics. The BII comprises four dimensions: Reach Influence (RI), Credibility Influence (CI), Activity Influence (AI), and Content Influence (CoI). RI represents an account's audience reach, CI provides an operational proxy for credibility, AI captures account activity, and CoI measures the relevance of account content to Bitcoin. The BII is subsequently integrated with FinBERT outputs so that each tweet is represented not only by its sentiment direction but also by its prediction confidence and source influence weight.

This study also considers the temporal characteristics of social media information. Sentiment signals do not necessarily correspond to market movements on the same day. Anbaee Farimani et al. [27] constructed financial mood time series while accounting for lag windows between sentiment information and market data, whereas Han et al. [28] incorporated time-lagged sentiment into multimodal Bitcoin forecasting. Building on these findings, this study develops BII-weighted sentiment features with time lags of one to three days, rolling-window statistics, BII momentum, interaction features, and sentiment shock features. These representations are designed to capture the persistence, evolution, and potential delayed effects of influence-weighted social signals on Bitcoin market trends.

Based on the identified research gap, this study addresses three research questions (RQs): RQ1: How can social media influence be represented multidimensionally through the BII?; RQ2: How can the BII be integrated with FinBERT to construct influence-aware temporal sentiment features?; RQ3: To what extent is the addition of BII-derived features associated with changes in next-day Bitcoin trend prediction performance compared with a scenario without BII-derived features?

The contributions of this study are fourfold. First, it develops a multidimensional BII that extends beyond follower count as the sole measure of influence. Second, it constructs BII-weighted sentiment representations by integrating sentiment polarity, prediction confidence, and source influence. Third, it develops temporal BII-derived features to capture the persistence and dynamics of social signals. Fourth, it evaluates the predictive utility of these representations using a hybrid ensemble comprising XGBoost, LSTM, and GRU, with Logistic Regression serving as the meta-learner.

Accordingly, the novelty of this study does not lie in the isolated use of FinBERT or ensemble learning, but rather in the integration of multidimensional source influence, influence-weighted sentiment, temporal social signal representations, and hybrid Bitcoin trend prediction within a unified framework. This approach provides a systematic means of investigating whether accounting for the characteristics of information sources can improve the representation of social media sentiment and contribute to next-day Bitcoin trend prediction.

## 2. Literature Review

### 2.1. Evolution of Bitcoin Predictions

Bitcoin forecasting is growing from statistical time series models to machine learning and deep learning. ARIMA and GARCH models are widely used. used for catch structure historical and volatility, while machine learning provides ability more flexible in studying nonlinear relationships. Literature Then develop towards recurrent networks such as LSTM and GRU because financial data has temporal sequence that is not may treated as observation independent [22].

Technical indicators remain become source information important on modern models. Ye et al. [18] uses MACD, SMA, RSI, and other indicators along with price and sentiment data. Quang et al. [14] show that integration moving averages with LSTM can produce different performance depends frequency and configuration of indicators. Lee [19] also combining attention based recurrent models with OHLCV and technical indicators for trend prediction. With Thus, more technical information appropriate understood as representation condition market that can combined with source information other, not as the only one predictor.

### 2.2. Sentiment and Multimodal Forecasting

Utilization of sentiment analysis pushed by limitations historical data in catch change perception investors. Twitter/X becomes Wrong One source popular Because conversation cryptocurrency ongoing in a way fast and continuous. Shang [29] combines 124 technical indicators with Twitter RoBERTa deep sentiment prediction next day Bitcoin price range and report improvement average accuracy from 0.56 to 0.62 after sentiment feature added. Although however, the benefits of sentiment data are not uniform. El Abaji and Haraty [22] found improvement on LSTM after sentiment entered, but Prophet experience decreased and SARIMAX shows results no consistent. This finding is important Because show that More sentiment data The same with automatically better prediction. Quality sentiment representation, temporal alignment, and model architecture determine to what extent are the information social can utilize.

Transformer based sentiment analysis offers more representation contextual than lexical scoring. Anbae Farimani et al. [27] use FinBERT which has trained and adapted for domain financial, then combine mood time series with informative market feature. Research the show that integration technical information and financial sentiment can give more information strong than use Wrong One source only. FinBERT chosen in this research because suitability domain and its capabilities produce positive, neutral, and negative probabilities. However, the research No claim that FinBERT always superior to RoBERTa or domain adapted large language models; comparison comprehensive NLP intermodels are beyond the goal main study.

### 2.3. From Unweighted to Influence Aware Sentiment

The most relevant research gaps lie in the source heterogeneity. Feizian and Amiri [23] show that tweets can give weight based on the publisher's influence factor and gain results more goods compared to polarity score without weight on some cryptocurrency. They use normalized follower count as representation of influence.

Contribution the important, but follower count only describes reach. Influence social may also be related with credibility, activity, and relevance. Jahanbin and Chahooki [24] strengthen attention towards influencers with using aspect based sentiment towards cryptocurrency influencer. Meanwhile, Sundarasan and Saleem [30] further highlighted the growing role of social media information in cryptocurrency research, reinforcing the importance of considering contextual characteristics beyond conventional market data. Building on these findings, this study advances the conceptual progression from conventional sentiment analysis to influence weighted sentiment and, ultimately, to

multidimensional influence aware sentiment. BII is not intended as a causal measure of social influence, but rather as an operational representation of source characteristics that can be incorporated as predictive information.

## 2.4. Hybrid and Ensemble Learning

Ensemble learning used when different learners capture different patterns. Ye et al. [18] developed a stacking LSTM and GRU with a metamodel and used out of fold prediction to building second level features. Saraswathi et al. [21] also show that the BiLSTM and GRU ensemble with market and sentiment information can produce performance more goods compared to a number of individual models in their experiment.

This research uses XGBoost, LSTM, and GRU for utilizes two types of data structures. XGBoost handle tabular nonlinear interactions, while LSTM and GRU are used for temporal sequence. Hybrid architecture No claimed as novelty main. Model functions as an experimental engine for test whether BII derived information provides additional predictive utility.

## 3. Methodology

### 3.1. Research Framework and Data

This study uses two primary data sources: historical Bitcoin market data and Twitter/X data. Market data consist of Date, Open, High, Low, Close, and Volume. Twitter/X data contains text and account metadata such as user followers, friends, favorites, verification status, account creation date, hashtags, and retweet information. The market data were cleaned, converted into numerical format, and arranged chronologically. The technical features constructed from the market data include Return, Volatility, RSI, MACD, MACD signal, SMA-7, SMA-14, EMA-7, EMA-14, and Bollinger Bands. The use of technical indicators is consistent with previous cryptocurrency forecasting studies that combine historical prices with derived market information [14], [18], [31]. The overall research process is summarized in figure 1, which illustrates the integration of Bitcoin historical data and Twitter/X data through preprocessing, technical indicator construction, trend labeling, BII computation, FinBERT sentiment analysis, temporal feature engineering, hybrid ensemble learning, performance evaluation, and statistical testing.

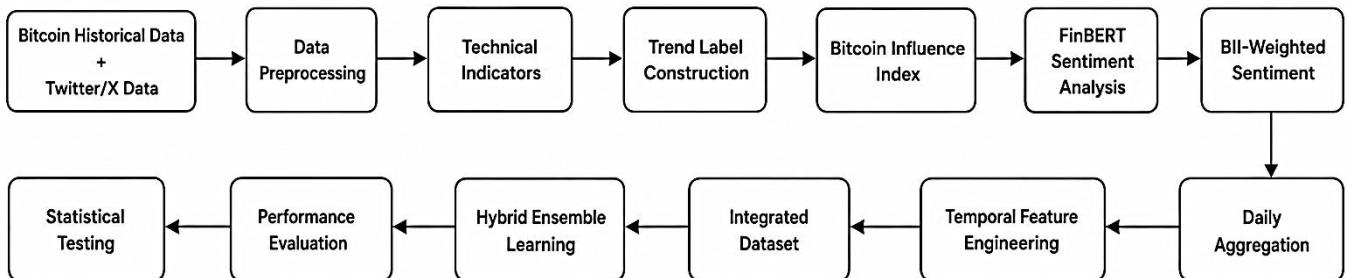


**Figure 1.** Proposed Influence Aware Bitcoin Trend Prediction Framework

These multi sources approach is consistent with the growing trend in Bitcoin forecasting toward integrating price information, technical indicators, sentiment signals, and other external information sources. Ye et al. [18] for example, used price data, technical indicators, and Twitter derived sentiment in a stacking ensemble LSTM–GRU, while Han et al. [28] showed trend towards multimodal fusion that combines textual sentiment and numerical market informations. The research target is next day direction:

$$Trend_t = \begin{cases} 1, & Close_{t+1} > Close_t \\ 0, & Close_{t+1} \leq Close_t \end{cases} \quad (1)$$

Thus, the research is a binary classification, not a continuous price regression. Twitter/X text is processed through lowercasing, URL removal, mention removal, deletion hashtag symbols, non alphabeticals character filtering, and whitespace normalization. Metadata related to followers, friends, favorites, verification status, and account age are converted into numerical features. Retweet status is used in activity penalties. A cleaning approach is required because Twitter data on previous cryptocurrency research was also found contains links, mentions, duplication, advertising, and various form of noise [18], [29].

### 3.2. Bitcoin Influence Index Construction

BII was formed from four raw dimensions. Reach Influence is calculated from log followers and follower to friend ratio:

$$RI_i^{\text{raw}} = \log(1 + \text{Followers}_i) + \log\left(1 + \frac{\text{Followers}_i}{\text{Friends}_i + 1}\right) \quad (2)$$

Credibility Influence uses verification status, log account age, and domain relevance:

$$CI_i^{\text{raw}} = \text{Verified}_i + \log(1 + \text{Age}_i) + \text{DomainRel}_i \quad (3)$$

Activity Influence uses favorites and retweet penalty:

$$AI_i^{\text{raw}} = \log(1 + \text{Favourites}_i) \times RP_i \quad (4)$$

Content Influence gained through content relevance and hashtag relevance:

$$CoI_i^{\text{raw}} = \text{ContentRel}_i + \text{HashtagRel}_i \quad (5)$$

These formulas are consistent with the final experimental implementation. Because each raw dimension has different scales, normalization is done using Min–Max:

$$RX_i = \frac{X_i^{\text{raw}} - X_{\min}^{\text{raw}}}{X_{\max}^{\text{raw}} - X_{\min}^{\text{raw}}}, \quad X \in \{RI, CI, AI, CoI\} \quad (6)$$

BII next calculated:

$$BII_i = w_R RI_i + w_C CI_i + w_A AI_i + w_{Co} CoI_i \quad (7)$$

with:

$$w_R + w_C + w_A + w_{Co} = 1 \quad (8)$$

Equal Weight is used as a baseline:

$$w_R = w_C = w_A = w_{Co} = 0.25 \quad (9)$$

Weight sensitivity was also tested through RID, CID, AID, and COID scenarios. In each dominant-weight scenario, one dimension is assigned a weight of 0.40, while the remaining three dimensions are assigned weights of 0.20 each. The purpose of this test is not to identify a “true” causal weight, but to examine whether BII scores and rankings change substantially when the relative importance of the dimensions is adjusted. The Credibility Influence dimension should be interpreted cautiously. CI is not a factual reliability score or a measure of an account's ability to predict Bitcoin prices; rather, it is an operational credibility proxy based on the available metadata.

### 3.3. FinBERT and BII Weighted Sentiment

Sentiment analysis was performed using FinBERT. Positive labels are converted to +1, neutral labels to 0, and negative labels to −1. In addition to the sentiment label, FinBERT produces a confidence score  $C_i$ . BII weighted sentiment at the tweet level is calculated:

$$BWS_i = S_i \times C_i \times BII_i \quad (10)$$

With  $S_i$  as a numeric sentiment label,  $C_i$  as confidence, and  $BII_i$  as a source influence score. This formula is consistent with the implementation used in the hybrid FinBERT experiment. Under this approach, two tweets with the same positive sentiment label may contribute differently when their confidence scores or BII values differ. This distinguishes conventional sentiment representation from influence aware sentiment.

### 3.4. Daily Aggregation and Temporal Features

Tweet level output then aggregated per day into sentiment label mean, confidence mean, BII mean, BII max, BII weighted sentiment mean, BII weighted sentiment sum, and tweet count. Implementation experiment uses arithmetic aggregation of tweet level BII weighted sentiment, not a normalized weighted average formula.

For catch delayed information, used:

$$BWSLag_k(t) = BWSMean_{t-k}, \quad k = 1, 2, 3 \quad (11)$$

Rolling features are calculated use windows three and seven day:

$$BWSRoll_n(t) = \frac{1}{n} \sum_{j=0}^{n-1} BWSMean_{t-j}, \quad n \in \{3, 7\} \quad (12)$$

In addition, it was formed:

$$BII Momentum_t = BII Mean_t - BII Mean_{t-1} \quad (13)$$

$$BII Interaction_t = BII Mean_t \times TweetCount_t \quad (14)$$

Shock feature obtained through deviation against the rolling mean of seven day, then used for forming BII sentiment shock and BII volume shock. All feature the is part from the final pipeline.

### 3.5. Experimental Scenarios and Data Split

Experiment comparing two scenarios. Without BII using market, technical, ordinary sentiment, and tweet volume features. With BII using all over baseline features plus BII mean, BII max, weighted sentiment, lag, rolling, momentum, interaction, and shock features. Experiment code shows 19 baseline features and feature space expansion in the scenario with BII.

After daily aggregation, lag construction, integration with market data, and the removal of missing values, the modeling dataset consists of 96 daily observations, with 58 labels Trend = 1 and 38 labels Trend = 0. Observations shared Chronologically, approximately 70 % training, 10% validation, and 20% testing. Sequence model used seven days previously as input. Standardization done with parameters that obtained from the training set so validation and testing no used for estimate scaler.

### 3.6. Hybrid Ensemble

The base learners consist of XGBoost, LSTM, and GRU. LSTM and GRU use 64 recurrent units, 32 dense layers, dropout of 0.30 and 0.20, Adam optimizer, learning rate of 0.001, binary cross entropy, batch size of 16, maximum of 100 epochs, and early stopping with patience of 15. Each model produce probability  $p_{XGB,t}$ ,  $p_{LSTM,t}$ ,  $p_{GRU,t}$ . Probability the used as input to the Logistic Regression meta learner together mean and standard deviation probability:

$$M_t = [p_{XGB,t}, p_{LSTM,t}, p_{GRU,t}, \mu_{p,t}, \sigma_{p,t}] \quad (15)$$

Final probability:

$$P(y_t = 1 | M_t) = \frac{1}{1 + \exp[-(\beta_0 + \beta^T M_t)]} \quad (16)$$

The decision threshold was not automatically fixed at 0.50; instead, it was selected using the validation data based on the F1-score. Evaluation using Accuracy, Precision, Recall, F1-score, ROC-AUC, and confusion matrix. Statistical comparison was performed using paired t-test, Wilcoxon signed-rank, and McNemar's exact test.

It should be noted that the XGBoost configurations used in the two scenarios were not fully identical. Therefore, the main results should be interpreted as a scenario comparison rather than a pure feature ablation experiment. In addition, the meta-training procedure does not yet fully employ out-of-fold predictions as used in more rigorous stacking approaches, such as that of Ye et al. [18]. Both issues are therefore treated as methodological limitations.

## 4. Results and Discussion

### 4.1. Characteristics of BII

The construction of BII from nearly two million tweets resulted in mean RI, CI, AI, CoI, and composite BII values of approximately 0.2265, 0.4384, 0.4468, 0.1933, and 0.3263, respectively. To examine the overall distribution of influence scores across the analyzed social media data, the distribution of the composite BII values is presented in figure 2. In addition to the overall distribution, figure 3 presents the top 20 accounts with the highest BII values, providing an overview of the accounts that exhibit the strongest multidimensional influence characteristics in the analyzed dataset.

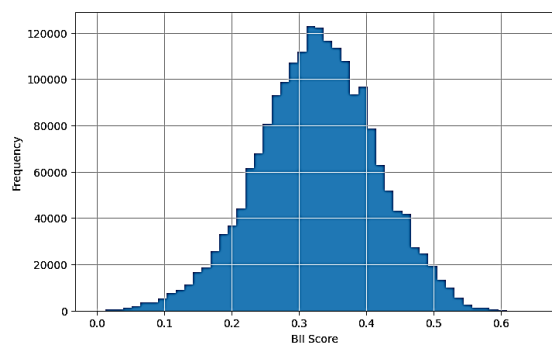


**Figure 2.** Visualization of BII Distribution

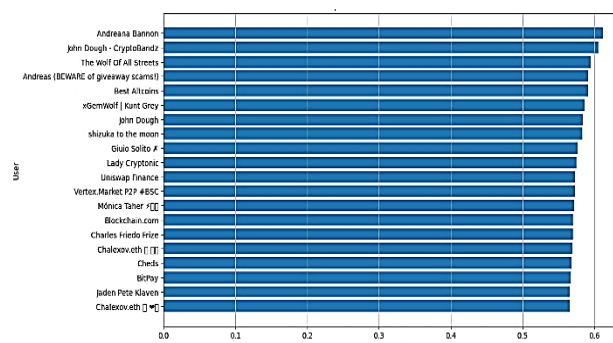


**Figure 3.** Top 20 Accounts with highest BII value

The results show that no single observation achieved the maximum score across all four dimensions simultaneously. This finding highlights the multidimensional nature of social influence. An account with high reach does not necessarily exhibit high activity or content relevance; conversely, an account with a smaller audience may achieve relatively high credibility or relevance scores. These findings extend the work of Feizian and Amiri [23], who demonstrated the benefits of influence weighting based on follower information. In the present study, follower derived reach represents only one component of the multidimensional influence construct. Thus, the model does not equate popularity with overall social influence.

## 4.2. Influence Aware Sentiment Representation

Daily aggregation shows that ordinary sentiment and BII weighted sentiment do not always exhibit similar magnitudes. This is expected because weighted sentiment incorporates three components: polarity, confidence, and BII. These results indicate that sentiment polarity alone does not fully represent the structure of social signals. Days with similar mean sentiment may produce different weighted signals when the confidence scores and influence characteristics of active accounts differ.

Development of temporal features also give context additional. Lagged sentiment represents possibility delayed market response, rolling features capture persistence, whereas shock features capture deviation from local patterns. This approach is consistent with financial sentiment studies that use delay windows and time lagged sentiment to align sentiment information with market dynamics [27], [28], [32].

## 4.3. Prediction Performance

The predictive performance of the Hybrid Ensemble under the Without BII and With BII scenarios is compared in Table 1 and figure 4 using Accuracy, Precision, Recall, F1-score, and ROC-AUC.

**Table 1.** Performance Comparison Between Without BII and With BII Models

| Metric    | Without BII | With BII | Change |
|-----------|-------------|----------|--------|
| Accuracy  | 40.00%      | 75.00%   | +35.00 |
| Precision | 100.00%     | 84.62%   | -15.38 |
| Recall    | 14.29%      | 78.57%   | +64.28 |
| F1-score  | 0.25        | 0.8148   | 0.5648 |
| ROC-AUC   | 0.3809      | 0.6667   | 0.2858 |

The improved performance in the With BII scenario can be attributed to the additional information provided by influence-aware sentiment features. BII differentiates social signals according to source characteristics and integrates them with temporal features, providing complementary information for trend classification. The 100% precision of the Without BII model, despite only 40.00% accuracy, does not necessarily indicate data leakage; its low recall (14.29%) shows that the model rarely predicted the positive class, resulting in few but correct positive predictions while missing most actual upward movements. With BII, precision decreased to 84.62%, while accuracy increased to 75.00%, recall to 78.57%, F1-score to 0.8148, and ROC-AUC to 0.6667, indicating a more balanced classification performance. To minimize temporal data leakage, observations were split chronologically into approximately 70% training, 10% validation, and 20% testing; standardization parameters were estimated only from the training set, and the prediction threshold was determined using validation rather than test data.

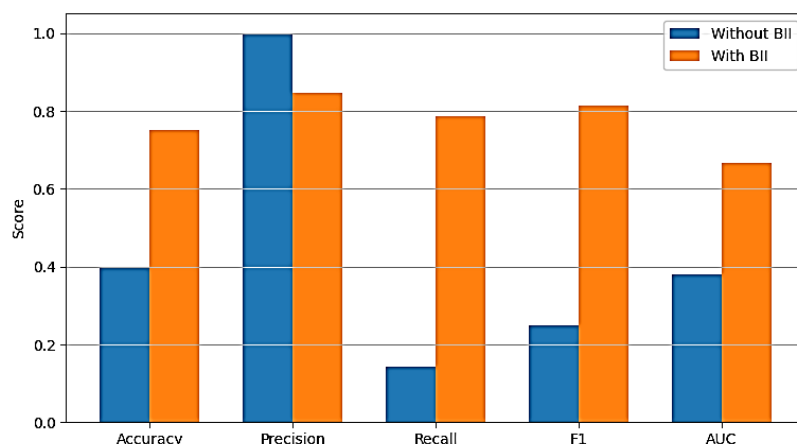


**Figure 4.** Performance Comparison Between Without BII and With BII Models

This finding supports literature pattern that social information can repair prediction if its representation appropriate, but the result No always shaped improvement every metric. El Abaji and Haraty [22] too shows different responses between models after sentiment integrations. Shang [29] shows improvement accuracy after Twitter sentiment was added on high dimensional technical features, whereas Feizian and Amiri [23] found weighted sentiment can increase cryptocurrency prediction.

#### 4.4. Statistical Significance

Paired correctness comparison produces a paired t-test:  $p = 0.030771$  and Wilcoxon:  $p = 0.034808$ . Both p-values are below the 0.05 significance threshold. In contrast, McNemar's exact test yielded  $p = 0.065430$  and therefore did not reach conventional statistical significance. These findings should be interpreted conservatively. The paired t-test and Wilcoxon signed-rank test provide evidence of differences in the paired correctness vectors, whereas McNemar's test does not provide sufficient evidence at the 5% significance level. To determine whether the performance differences between the Without BII and With BII scenarios are statistically supported, table 2 summarizes the results of the paired t-test, Wilcoxon signed-rank test, and McNemar's exact test.

**Table 2.** Statistical Testing Results

| Statistical Test     | p-value  | Decision at $\alpha = 0.05$ |
|----------------------|----------|-----------------------------|
| Paired t-test        | 0.030771 | Significant                 |
| Wilcoxon signed-rank | 0.034808 | Significant                 |
| McNemar              | 0.065430 | Not significant             |

Thus, these findings should not be interpreted as definitive evidence that BII consistently improves prediction across all periods. A more appropriate conclusion is that the addition of BII derived features was associated with improved predictive performance in the evaluated temporal sample, with partial statistical support from the paired tests.

#### 4.5. Feature Importance and Interpretation

Feature importance analysis for the XGBoost branch in the With BII scenario shows that SMA-14, SMA-7, and EMA-14 remain among the most important features. BII weighted sentiment lag-1 also appears among the relatively important predictors. To examine the relative contribution of the input features to the XGBoost branch of the With BII scenario, figure 5 presents the 15 features with the highest feature-importance scores.

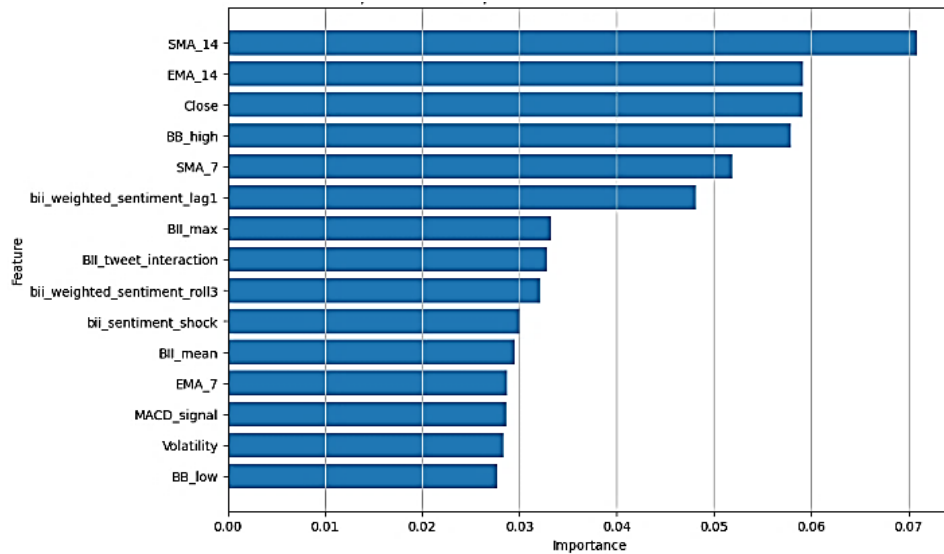


**Figure 5.** Top 15 most important features

These results indicate that social information does not replace technical information. Technical indicators continue to play a major role in representing market conditions, while BII-derived signals function as a complementary information layer. The appearance of lag-1 BII-weighted sentiment among the important features also provides support for the temporal hypothesis that previous social information may have predictive relevance for subsequent market direction.

The appearance of lag-1 weighted sentiment too give support against the temporal hypothesis that information social previously can own predictive relevance to direction next. This is consistent with use delayed sentiment on study Anbaee Farimani et al. [27] and Han et al. [28]. Feature importance should not be interpreted as causal importance. SMA-14 or weighted sentiment lag-1 does not can stated to “cause” the movement of Bitcoin only Because used model. Interpretation limited to relative predictive utility in the fitted XGBoost branch.

#### 4.6. Theoretical Implications

This research provides implications on method market sentiment conceptualized. Traditional sentiment-based forecasting especially modeling content opinion:

$$\text{MarketSentiment}_t = f(\text{Sentiment}_t) \quad (17)$$

Influence aware representation expands it become:

$$\text{MarketSignal}_t = f(\text{Sentiment}_t, \text{SourceInfluence}_t, \text{TemporalDynamics}_t) \quad (18)$$

In a way theory communication, Katz [25] shows that primary groups and opinion leaders can become part important in flow information and formation decisions, although modern communication processes are more complex than the two stages model classic. On modern digital context, Akdogan and Anbar [26] also emphasize that social media information can load dimensions social and behavioral besides sentiment content. Therefore, treating every post as unit identical information can remove part context source. Contribution conceptual study with thus No state that influencers cause change price. BII operates difference character source become a predictor. The relationship obtained is a predictive association, not a causal market influence.

#### 4.7. Practical Implications and Economic Significance

In a way practically, BII can used as a supplementary signal in cryptocurrency analytics or decision support systems. Increased tweet volume can interpret in a way different when happen together Low BII\_mean compared to when Lots account with relative BII tall follow involved. BII momentum, interaction, and shock Also allows system No only observe sentiment level, but change composition the actor who involved in discussion. Approach this kind of thing can useful for market monitoring systems that want differentiate popular chatter of conversations involving account with character influence more stronges. However, research Not yet Proving economic profitability. There are no trading backtests, transaction costs, slippage, Sharpe ratio, maximum drawdown, or cumulative returns in experiment this. Accuracy 75% no can translate in a way direct as a 75% trading return. El Abaji and Haraty [22] also emphasize that integration of predictive analytics into algorithmic trading faces challenge addition such as volatility, overfitting, and

real time processing. Therefore, the contribution practical study restricted on predictive information value. Economic significance is necessary tested in research continued.

## 5. Conclusion

This study develops an influence-aware Bitcoin trend prediction framework that integrates the multidimensional Bitcoin Influence Index (BII), FinBERT based sentiment analysis, temporal feature engineering, and hybrid ensemble learning. The BII comprises four dimensions: Reach Influence, Credibility Influence, Activity Influence, and Content Influence to represent source heterogeneity that is not fully captured when sentiment scores are aggregated without considering differences in source characteristics.

BII is combined with the FinBERT sentiment label and confidence score to generate BII weighted sentiment. This social signal is further transformed into lag, rolling, momentum, interaction, and shock features, enabling the model to capture not only same-day sentiment conditions but also their temporal dynamics. On the evaluated temporal sample, the With BII scenario yielded an accuracy of 75.00%, a recall of 78.57%, an F1-score of 0.8148, and an AUC of 0.6667, compared to 40.00%, 14.29%, 0.2500, and 0.3809 for the Without BII scenario. Precision decreased from 100.00% to 84.62%; therefore, the benefit associated with BII is more appropriately interpreted as an improvement in overall classification balance rather than a uniform improvement across all performance metrics.

The paired t-test ( $p = 0.030771$ ) and Wilcoxon signed-rank test ( $p = 0.034808$ ) indicated statistically significant differences at  $\alpha = 0.05$ , whereas McNemar's exact test ( $p = 0.065430$ ) did not reach the conventional significance threshold. These results provide preliminary evidence that BII derived information has predictive utility, although they do not support universal or causal claims. Theoretically, this study extends sentiment-aware forecasting toward influence aware forecasting by conceptualizing social market information as a function of sentiment content, source characteristics, and temporal dynamics. Practically, BII has the potential to serve as a supplementary signal in cryptocurrency monitoring and decision-support systems.

The generalizability of the findings remains limited by the relatively small daily sample size, FinBERT sampling, reliance on a single social-media platform, operational proxies used in the BII, a single temporal holdout, differences in baseline configurations, and a stacking procedure that does not yet employ out-of-fold meta-training. Future research should evaluate BII over longer periods and across different market regimes, develop adaptive or data driven weighting schemes, integrate multi source social and on chain data, employ walk-forward evaluation, implement out of fold stacking, and assess economic significance through trading backtesting and risk adjusted performance measures. Thus, the primary contribution of this study is not to claim that influencers cause Bitcoin price movements, but to demonstrate that heterogeneity among opinion sources can be translated into structured information with the potential to enrich Bitcoin trend prediction.

## 6. Declarations

### 6.1. Author Contributions

Conceptualization: A.M., M.A.W.; Methodology: A.M., M.A.W., R.P.; Software: A.M.; Validation: A.M., M.A.W., R.P.; Formal Analysis: A.M., R.P.; Investigation: A.M.; Data Curation: A.M.; Visualization: A.M.; Writing – Original Draft Preparation: A.M.; Writing – Review and Editing: M.A.W., R.P.; Supervision: M.A.W., R.P.; Project Administration: M.A.W.; All authors have read and agreed to the published version of the manuscript.

### 6.2. Data Availability Statement

The data used in this study consists of historical Bitcoin market data and publicly available social media data. Processed datasets and supporting experimental outputs may be made available by the corresponding author upon reasonable request, subject to the terms and conditions of the original data providers.

### 6.3. Funding

The authors would like to thank the Indonesian Education Scholarship Institute (BPI), the Center for Higher Education Funding and Assessment (PPAPT), the Ministry of Higher Education, Science and Technology of the Republic of Indonesia, the Education Endowment Fund Agency (LPDP), and the Ministry of Finance of the Republic of Indonesia for their support of this research. They also extend their gratitude to Universitas Diponegoro and Universitas Ibrahimy, as well as the Institute for Research and Community Service (LPPM) for their support.

#### 6.4. Institutional Review Board Statement

Not applicable.

#### 6.5. Informed Consent Statement

Not applicable.

#### 6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

#### References

- [1] J. V. Critien, A. Gatt, and J. Ellul, "Bitcoin price change and trend prediction through Twitter sentiment and data volume," *Financ. Innov.*, vol. 8, no. 1, pp. 1–20, 2022, doi: 10.1186/s40854-022-00352-7.
- [2] M. Nair, M. I. Marie, and L. A. Abd-Elmegid, "Prediction of cryptocurrency price using time series data and deep learning algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 338–347, 2023, doi: 10.14569/IJACSA.2023.0140837.
- [3] M. Z. Maliha, A. S. Trisha, A. M. T. Khan, P. Das, M. I. Hossain, and R. Rahman, "A deep learning approach to predict cryptocurrency price by evaluating sentiment and stock market correlations," *Lect. Notes Networks Syst.*, vol. 867, no. 1, pp. 3–17, 2024, doi: 10.1007/978-981-99-8937-9\_1.
- [4] R. U. Din, S. Ahmed, S. H. Khan, A. Albanyan, J. Hoxha, and B. F. Alkhamees, "A novel decision ensemble framework: Attention-customized BiLSTM and XGBoost for speculative stock price forecasting," *PLoS One*, vol. 20, no. 4, pp. 1–31, Apr. 2025, doi: 10.1371/journal.pone.0320089.
- [5] H. S. Jung, S. H. Lee, H. Lee, and J. H. Kim, "Predicting Bitcoin trends through machine learning using sentiment analysis with technical indicators," *Comput. Syst. Sci. Eng.*, vol. 46, no. 2, pp. 2231–2246, 2023, doi: 10.32604/csse.2023.034466.
- [6] G. Wang, C. Liu, E. Sbair, M. S. Sheng, J. Hu, and M. Tao, "Interrelations between Bitcoin market sentiment, crude oil, gold, and the stock market with Bitcoin prices: Vision from the hedging market," *Stud. Econ. Financ.*, vol. 41, no. 5, pp. 1166–1190, 2024, doi: 10.1108/SEF-03-2024-0137.
- [7] A. I. Mubarak, "Neural networks in financial engineering: Forecasting using RNN, LSTM and hybrid models in varying volatilities," *Int. J. Intell. Eng. Syst.*, vol. 18, no. 7, pp. 68–81, 2025, doi: 10.22266/ijies2025.0831.06.
- [8] H. S. Htay, M. Ghahremani, and S. Shiales, "Enhancing Bitcoin price prediction with deep learning: Integrating social media sentiment and historical data," *Appl. Sci.*, vol. 15, no. 3, pp. 1–21, Feb. 2025, doi: 10.3390/app15031554.
- [9] S. Sun and Y. Qin, "Fuzzy clustering algorithm for trend prediction of the digital currency market," *Salud Cienc. Tecnol. Ser. Conf.*, vol. 3, no. 1, pp. 1–14, 2024, doi: 10.56294/sctconf20241094.
- [10] N. P. Patel et al., "Fusion in cryptocurrency price prediction: A decade survey on recent advancements, architecture, and potential future directions," *IEEE Access*, vol. 10, no. 1, pp. 34511–34538, 2022, doi: 10.1109/ACCESS.2022.3163023.
- [11] J. W. Choi and Y. K. Choi, "The prediction of Bitcoin price through gold price using long short-term memory model," *IAES Int. J. Artif. Intell.*, vol. 13, no. 1, pp. 909–916, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp909-916.
- [12] D. Ganesh and V. C. Harshavardhan, "Cryptocurrency market dynamics: A machine learning-based approach for price prediction," *Int. Conf. Smart Syst. Appl. Electr. Sci.*, vol. 2024, no. May, pp. 1–6, May 2024, doi: 10.1109/ICSSSES62373.2024.10561446.
- [13] Sarmini, C. R. A. Widiawati, D. R. Febrianti, and D. Yuliana, "Volatility analysis of cryptocurrencies using statistical approach and GARCH model: A case study on daily percentage change," *J. Appl. Data Sci.*, vol. 5, no. 3, pp. 838–848, 2024, doi: 10.47738/jads.v5i3.261.
- [14] P. D. Quang, N. H. Duy, P. Q. Khoai, and B. D. Duong, "Integrating moving average indicators with long short-term memory model in Bitcoin price forecasting," *J. Appl. Data Sci.*, vol. 6, no. 3, pp. 1850–1864, 2025, doi: 10.47738/jads.v6i3.766.
- [15] N. Latif, J. D. J. D. Selvam, M. Kapse, V. Sharma, and V. C. Mahajan, "Comparative performance of LSTM and ARIMA for the short-term prediction of Bitcoin prices," *Australas. Account. Bus. Financ. J.*, vol. 17, no. 1, pp. 256–276, 2023, doi: 10.14453/aabfj.v17i1.15.
- [16] M. Ortu, N. Uras, C. Conversano, S. Bartolucci, and G. Destefanis, "On technical trading and social media indicators for cryptocurrency price classification through deep learning," *Expert Syst. Appl.*, vol. 198, no. 1, pp. 1–12, 2022, doi: 10.1016/j.eswa.2022.116804.

- 
- [17] N. N. AlMadany, O. Hujran, G. Al Naymat, and A. Maghyereh, "Forecasting cryptocurrency returns using classical statistical and deep learning techniques," *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, pp. 1–23, 2024, doi: 10.1016/j.jjime.2024.100251.
- [18] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A stacking ensemble deep learning model for Bitcoin price prediction using Twitter comments on Bitcoin," *Mathematics*, vol. 10, no. 8, pp. 1–21, Apr. 2022, doi: 10.3390/math10081307.
- [19] M. M.-C. Lee, "Bitcoin trend prediction with attention-based deep learning models and technical indicators," *Systems*, vol. 12, no. 11, pp. 1–19, Nov. 2024, doi: 10.3390/systems12110498.
- [20] S. Arslan, "Bitcoin price prediction using sentiment analysis and empirical mode decomposition," *Comput. Econ.*, vol. 65, no. 4, pp. 2227–2248, 2025, doi: 10.1007/s10614-024-10588-3.
- [21] S. Saraswathi et al., "An ensemble machine learning technique for Bitcoin price prediction," *J. Trends Comput. Sci. Smart Technol.*, vol. 6, no. 2, pp. 153–167, 2024, doi: 10.36548/jtcsst.2024.2.005.
- [22] M. El Abaji and R. A. Haraty, "Enhancing Bitcoin forecast accuracy by integrating AI, sentiment analysis, and financial models," *Int. J. Comput. Digit. Syst.*, vol. 17, no. 1, pp. 1–15, 2025, doi: 10.12785/ijcds/1571026011.
- [23] F. Feizian and B. Amiri, "Cryptocurrency price prediction model based on sentiment analysis and social influence," *IEEE Access*, vol. 11, no. 1, pp. 142177–142195, 2023, doi: 10.1109/ACCESS.2023.3342688.
- [24] K. Jahanbin and M. A. Z. Chahooki, "Aspect-based sentiment analysis of Twitter influencers to predict the trend of cryptocurrencies based on hybrid deep transfer learning models," *IEEE Access*, vol. 11, no. 1, pp. 121656–121670, 2023, doi: 10.1109/ACCESS.2023.3327060.
- [25] E. Katz, "Communications research since Lazarsfeld," *Public Opin. Q.*, vol. 51, no. 4, pp. S25–S45, 1987, doi: 10.1093/poq/51.4\_PART\_2.S25.
- [26] Y. E. Akdogan and A. Anbar, "More than just sentiment: Using social, cognitive, and behavioral information of social media to predict stock markets with artificial intelligence and big data," *Borsa Istanbul Rev.*, vol. 25, no. 1, pp. 1–12, 2025, doi: 10.1016/j.bir.2024.12.003.
- [27] S. Anbae Farimani, M. Vafaei Jahan, A. Milani Fard, and S. R. K. Tabbakh, "Investigating the informativeness of technical indicators and news sentiment in financial market price prediction," *Knowl.-Based Syst.*, vol. 247, no. 1, pp. 1–12, 2022, doi: 10.1016/j.knosys.2022.108742.
- [28] P. Han, H. Chen, A. Rasool, Q. Jiang, and M. Yang, "MFB: A generalized multimodal fusion approach for Bitcoin price prediction using time-lagged sentiment and indicator features," *Expert Syst. Appl.*, vol. 261, no. 1, pp. 1–15, 2025, doi: 10.1016/j.eswa.2024.125515.
- [29] L. Shang, "Sentiment-driven Bitcoin price range forecasting: Enhancing CART decision trees with high-dimensional indicators and Twitter dynamics," *IEEE Access*, vol. 13, no. 1, pp. 60508–60518, 2025, doi: 10.1109/ACCESS.2025.3557186.
- [30] S. Sundarasan and F. Saleem, "From tweets to trades: A bibliometric and systematic review of social media's influence on cryptocurrency," *Int. J. Financ. Stud.*, vol. 13, no. 2, pp. 1–18, Mar. 2025, doi: 10.3390/ijfs13020087.
- [31] Q. Phung Duy, O. Nguyen Thi, P. H. Le Thi, H. D. Pham Hoang, K. L. Luong, and K. N. Nguyen Thi, "Estimating and forecasting Bitcoin daily prices using ARIMA-GARCH models," *Bus. Anal. J.*, vol. 45, no. 1, pp. 11–23, 2024, doi: 10.1108/BAJ-05-2024-0027.
- [32] M. A. Naeem, I. Mbarki, and S. J. H. Shahzad, "Predictive role of online investor sentiment for cryptocurrency market: Evidence from happiness and fears," *Int. Rev. Econ. Financ.*, vol. 73, no. 1, pp. 496–514, 2021, doi: 10.1016/j.iref.2021.01.008.