

Technema: Journal of Intelligent Engineering and Computing
Vol 1 No 3 September 2026, Hal 428-440
ISSN: 3163-6020 (Print) ISSN: 3163-5997 (Electronic)
Open Access: https://sovereignresearch.org/technema

Forecasting the IDX Composite Index Through the End of 2026 Using the XGBoost Algorithm

Dede Prabowo Wiguna[1*], Sakaria Efrata Ginting[2], Izba Zaheer[3], Abeysekara S. S. K.[4]
[1-2] Mandiri Bina Prestasi University, Indonesia
[3] Muslim Youth (MY) University, Pakistan
[4] Nāgānanda International Institute for Buddhist Studies, Sri Lanka
email: gedewiguna71@gmail.com[1]

Article Info :
Received:
28-09-2026
Revised:
12-10-2026
Accepted:
20-10-2026

Abstract

This study examined the ability of Extreme Gradient Boosting (XGBoost) to predict the direction of the IDX Composite Index (IHSG) across 20-, 40-, and 60-trading-day horizons associated with October–December 2026. Daily IHSG observations and supporting financial indicators, including energy prices, the USD/IDR exchange rate, major United States equity indices, and the VIX, were obtained from Yahoo Finance for December 17, 2010, through October 5, 2026, covering 3,828 trading days. Predictor variables were constructed using 1-, 5-, and 20-day return ratios, and binary targets distinguished upward from downward index movements. XGBoost classifiers were implemented in Python using Google Colaboratory, with chronological training and testing partitions of 80% and 20%. Testing accuracies were 55.25%, 52.23%, and 50.13% across the respective horizons. All three forecasts indicated UP/BULLISH movements, with estimated probabilities of 52.86%, 65.47%, and 60.11%. Confusion matrices revealed frequent false-positive bullish classifications and limited detection of declining markets. Relative to retrospective majority-class baselines, performance was marginally better at 20 days, weaker at 40 days, and equivalent at 60 days. Forecast probabilities should not be interpreted as expected percentage changes in IHSG. These findings indicate conditional upward signals but insufficient classification reliability to establish an impending market recovery or support independent investment decisions.

Keywords: *Classification Accuracy, Financial Forecasting, IDX Composite Index, Machine Learning, XGBoost.*

©2026 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.
(https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUCTION

Movements in the IDX Composite Index (IHSG) reflect changes in valuations across Indonesia's equity market, although the resulting fluctuations are shaped by interactions extending beyond individual securities and domestic trading conditions. Evidence from the Indonesian banking sector indicates that systemic disturbances can affect the composite index and banking share prices simultaneously, complicating forecasts based on isolated market movements (Ulyah et al., 2024). Recent research has also approached IHSG prediction through multivariate and multi-horizon forecasting, recognizing the importance of temporal dependencies and information from multiple explanatory variables (Deni & Amalita, 2026). Analyses of Indonesian equities have examined momentum, volatility, trend, and trading-volume indicators as potential predictors, illustrating the importance of information selection in modeling financial market behavior (Indra et al., 2025). The broader forecasting literature identifies technical, macroeconomic, and fundamental variables as relevant inputs, although their predictive contributions vary across markets and methodological specifications (Kumbure et al., 2022). Such characteristics make forecasting the composite index particularly challenging when historical price relationships are exposed to changing economic conditions, international capital movements, exchange-rate fluctuations, and commodity-market developments. For the remaining months of 2026, the central analytical concern involves estimating prospective IHSG movements while distinguishing potential market recovery from continued downward pressure and recognizing the uncertainty associated with longer forecasting horizons.

Research on stock index forecasting has progressed from conventional statistical time-series specifications toward machine learning architectures capable of representing more complex

relationships among financial observations. Applications combining autoregressive modeling, conditional variance estimation, and Fourier-series analysis demonstrate how alternative assumptions about temporal structure influence forecasting accuracy for the Indonesian Composite Index (Mardianto et al., 2026). Comparisons involving ARIMA, Long Short-Term Memory (LSTM), and hybrid ARIMA–LSTM models further establish that architectural configuration, historical coverage, and parameter selection influence the accuracy of IHSG opening and closing price predictions (Nensi et al., 2025). International forecasting experiments have extended these comparisons to XGBoost, LSTM, and their hybrid configurations, incorporating backtesting procedures to examine model performance across different equity markets (Oukhouya et al., 2023). A related methodological development employs explainable XGBoost with investor-sentiment information to classify movements in Chinese stock indices, emphasizing directional prediction rather than the exclusive estimation of numerical index values (Deng et al., 2023). These approaches demonstrate substantial methodological development, but their forecasting results cannot be considered directly interchangeable because they differ in prediction targets, explanatory variables, market characteristics, and evaluation horizons.

Existing studies of financial-market prediction leave a specific methodological and contextual question concerning the application of XGBoost to multi-horizon forecasting of the Indonesian Composite Index toward the end of 2026. Reviews of financial forecasting research reveal considerable variation in the selection of predictive variables, investigated markets, and modeling procedures, limiting the extent to which performance reported for one setting can be generalized to another (Kumbure et al., 2022). Existing Indonesian studies comparing autoregressive and recurrent neural network models provide historical forecasting benchmarks, but they do not directly establish the predictive characteristics of XGBoost for IHSG movements across the proposed forecasting horizons (Nensi et al., 2025). Research applying XGBoost to longer-term equity-market prediction addresses the classification of prospective market returns, whereas forecasting the numerical level of a particular composite index presents a different predictive objective (Sürek & Lau, 2025). Similarly, evidence supporting sentiment-based XGBoost models for Chinese stock indices concerns market-direction classification under an information structure that differs from the Indonesian financial environment (Deng et al., 2023). The resulting research gap concerns the empirical assessment of whether a consistently specified XGBoost framework can generate informative IHSG forecasts across 20-, 40-, and 60-trading-day horizons while accounting for historical price movements and relevant external market indicators. This question also requires distinguishing numerical forecasting accuracy from directional performance because a model that approximates index values adequately may not demonstrate equivalent reliability in identifying bearish or bullish movements.

Forecasting the composite index across different temporal horizons presents methodological challenges for financial computing and practical consequences for market participants who depend on projections with explicitly defined temporal horizons. Comparative investigations of the Indonesian Composite Index indicate that alternative representations of conditional variance and temporal patterns can produce different forecasting errors, reinforcing the importance of evaluating model suitability against the intended prediction task (Mardianto et al., 2026). Evidence from selected LQ45 equities also requires careful interpretation because predictive relationships established for individual stock returns cannot automatically be generalized to the broader composite index, which represents a different level of market aggregation (Indra et al., 2025). International comparisons involving machine learning and hybrid forecasting architectures demonstrate the relevance of backtesting in assessing whether predictive performance extends beyond the observations used for model development (Oukhouya et al., 2023). Research on longer-term stock market forecasting with XGBoost further connects predictive outcomes to investment exposure decisions, indicating that forecasting utility depends on the relevance of model outputs to the decisions they are intended to inform (Sürek & Lau, 2025). Without horizon-specific evaluation, traders, investors, and market analysts may interpret model-generated trajectories as reliable indicators of future conditions even when forecasting uncertainty increases with the projection interval. A forecasting framework that distinguishes historical fit, forward predictive performance, and directional interpretation would consequently provide a more defensible analytical basis for evaluating prospective IHSG movements without presenting conditional estimates as certain market outcomes.

To examine these forecasting challenges, XGBoost is employed as a computational approach for forecasting the IDX Composite Index across 20-, 40-, and 60-trading-day horizons using historical

market observations extending from December 2010 to October 2026. XGBoost employs an efficient gradient-boosted decision-tree architecture that can represent nonlinear relationships and interactions among predictive features without requiring a predetermined linear functional specification (Chen & Guestrin, 2016). This approach differs from Indonesian composite-index forecasting methods based on autoregressive conditional-variance structures and Fourier-series representations, which model temporal behavior through explicitly specified statistical components (Mardianto et al., 2026). It also differs from multivariate Temporal Fusion Transformer applications for the Indonesian Composite Index, while retaining a shared interest in incorporating multiple information sources and evaluating forecasts across different temporal horizons (Deni & Amalita, 2026). International evidence showing that hybrid LSTM–XGBoost configurations can outperform standalone alternatives under particular experimental conditions indicates that the selection of XGBoost does not independently establish superior predictive accuracy (Oukhouya et al., 2023). The proposed contribution instead concerns the integration of historical IHSG information, selected domestic and international financial indicators, and explicitly differentiated forecasting horizons within a consistent machine learning framework directed toward the remaining period of 2026. Its intended novelty is contextual and evaluative rather than algorithmic, emphasizing the assessment of index-level projections and associated market-direction indications under a common forecasting design for the Indonesian equity market.

This study aims to forecast the IDX Composite Index through the end of 2026 by developing an XGBoost-based prediction model using historical observations from December 2010 through the available portion of October 2026. The investigation evaluates forecasting behavior across 20-, 40-, and 60-trading-day horizons to determine how predictive performance changes as the projection interval increases and to examine potential indications of continued bearish conditions or emerging market recovery. Predictor development considers historical index information alongside selected domestic and international financial variables, reflecting the established relevance of predictor composition in stock-market forecasting applications (Kumbure et al., 2022). Model evaluation is directed toward numerical forecasting accuracy and the consistency of predicted upward or downward index movements, with particular attention to separating historical predictive evidence from projections of future market conditions. This distinction accommodates the methodological difference between index-level forecasting and directional classification previously examined through XGBoost-based financial prediction models (Deng et al., 2023). The expected contribution is a reproducible, horizon-specific forecasting assessment that extends the application of boosted-tree modeling to the Indonesian Composite Index, complements existing research on longer-term market prediction, and provides conditional analytical information for investors, financial analysts, and policymakers assessing prospective market developments through December 2026 (Sürek & Lau, 2025).

RESEARCH METHODS

This research employed a quantitative predictive modeling design based on financial time-series data to classify future movements of the IDX Composite Index (IHSG) across multiple forecasting horizons. Historical IHSG observations and supporting financial-market variables were obtained from Yahoo Finance, covering December 17, 2010, to October 5, 2026, with a total of 3,828 trading days (Yahoo Finance, 2026). The trading day served as the unit of analysis, with IHSG values representing the reference index and associated financial indicators providing explanatory variables for predictive modeling. Feature engineering involved calculating return ratios over 1-day, 5-day, and 20-day intervals for each explanatory variable to represent market fluctuations across different temporal windows, consistent with the application of temporally derived features in financial forecasting (Kumbure et al., 2022). Binary classification targets were constructed separately for the 20-day, 40-day, and 60-day forecasting horizons, assigning a value of 1 when the future IHSG exceeded its current reference value and 0 when the future value was lower. This target specification distinguished the prediction of market direction from numerical index-level forecasting, reflecting the methodological characteristics of stock index direction classification using machine learning (Deng et al., 2023).

The dataset was divided chronologically into training and testing subsets at a ratio of 80% to 20%, preserving the temporal sequence of market observations to reduce the risk of future-information bias. The earlier 80% of observations constituted the training dataset, while the remaining 20% was reserved for evaluating predictions against subsequent market observations. The Extreme Gradient Boosting (XGBoost) algorithm was applied to model relationships between the engineered return

features and horizon-specific binary targets through an ensemble of gradient-boosted decision trees (Chen & Guestrin, 2016). The analytical procedure comprised data preparation, feature construction, target assignment, chronological partitioning, model training, and directional prediction across the three forecasting horizons, following the general computational framework applied in machine learning-based stock market forecasting (Oukhouya et al., 2023). Data processing, model implementation, and prediction generation were performed using Python programming in the Google Colaboratory computational environment (Google, n.d.). The resulting predictions were interpreted as classifications of potential upward or downward IHSG movements over 20, 40, and 60 trading days, providing a basis for evaluating directional forecasting performance without treating binary outputs as numerical estimates of future index values.

RESULTS AND DISCUSSION

Multi-Horizon Forecasting Results of the IDX Composite Index

The XGBoost classification framework generated directional forecasts for three IDX Composite Index (IHSG) horizons using financial market observations available through October 5, 2026. The historical dataset covered December 17, 2010, to October 5, 2026, comprising 3,828 trading days and providing a common observational basis for each forecasting scenario. Each forecast classified prospective index movements according to whether the future IHSG would exceed or fall below its corresponding reference value. The resulting outputs represented directional classifications rather than numerical estimates of future index levels, distinguishing the analytical objective from conventional index-value forecasting. Previous applications of machine learning to Indonesian financial markets have addressed different prediction targets, including composite index values and stock returns, demonstrating the relevance of target specification to model evaluation (Deni & Amalita, 2026). The distinction is particularly relevant when interpreting bullish predictions because an anticipated increase does not independently establish the magnitude, duration, or persistence of subsequent price movements. Forecasting outcomes were consequently examined as conditional market-direction indicators across three temporal horizons rather than definitive estimates of future market conditions.

Predictor construction transformed historical financial observations into return ratios calculated over 1-day, 5-day, and 20-day intervals to represent market fluctuations across different temporal windows. These transformations incorporated information about recent price changes without requiring the original observations to follow a predetermined linear relationship. Financial forecasting research identifies the selection and transformation of explanatory variables as consequential components of predictive modeling, particularly when market relationships vary across observation periods (Kumbure et al., 2022). The constructed features allowed XGBoost to evaluate associations between historical return patterns and subsequent index movements using information from the available financial series. Applications involving Indonesian equities have likewise examined technical indicators derived from historical observations, supporting the relevance of temporally structured financial information for market prediction (Indra et al., 2025). Nevertheless, the presence of several return intervals does not establish the individual predictive contribution of each financial indicator, since the reported results do not provide variable-level importance estimates or independent feature assessments.

Three forecasting horizons were specified at 20, 40, and 60 trading days to examine directional predictions over different periods following the reference observation. Positive classifications indicated that the future IHSG was expected to exceed its current value, whereas negative classifications represented an anticipated decline relative to the same reference. Although the forecasting intervals differed, the underlying classification objective remained consistent, allowing directional predictions to be examined under a common analytical framework. XGBoost constructs predictive functions through sequentially optimized decision trees, permitting nonlinear relationships and interactions among explanatory variables to contribute to classification decisions (Chen & Guestrin, 2016). Related research on stock index direction forecasting has demonstrated the application of gradient-boosting algorithms to financial classification problems involving market movements and additional explanatory information (Deng et al., 2023). The three horizons were associated with the October, November, and December 2026 forecasting scenarios reported in the original analysis, preserving the temporal designations assigned to the model outputs. These designations identify the reported forecasting scenarios and do not imply that the model independently generated continuous daily index trajectories throughout each corresponding calendar month.

Table 1. Forecast Direction of IHSG Movements, October–December 2026

Target Period	Horizon (Days)	Model Accuracy	Forecast Direction	Probability (%)
October 2026	20	55.25%	UP/BULLISH	52,86%
November 2026	40	52.23%	UP/BULLISH	65,47%
December 2026	60	50.13%	UP/BULLISH	60,11%

Source: Processed data, 2026.

Table 1 presents the XGBoost forecasting results, including the target period, forecasting horizon, classification accuracy, predicted market direction, and probability associated with the predicted classification. The model assigned an UP/BULLISH direction to all three forecasting scenarios, indicating a consistent expectation of upward index movements across the reported horizons. Classification accuracy and forecast probability nevertheless differed, suggesting that agreement in predicted direction should not be interpreted as equivalent predictive reliability. The reported accuracy values describe the proportion of correctly classified observations, whereas the probability values express the model-assigned likelihood of the predicted upward movement. Comparative research involving XGBoost and hybrid stock-market forecasting models emphasizes the need to assess predictive classifications using appropriate evaluation measures rather than relying exclusively on directional outputs (Oukhouya et al., 2023). The numerical results provide the empirical basis for examining how the model's directional predictions and classification performance vary across the three forecasting horizons.

For October 2026, the 20-trading-day forecasting model recorded a classification accuracy of 55.25%, accompanied by an UP/BULLISH probability of 52,86%. The predicted probability was only slightly above an equal division between the two directional categories, indicating a relatively weak preference for upward movement. Although the accuracy exceeded 50%, the observed classification performance still indicated substantial uncertainty in distinguishing positive and negative market outcomes. The October prediction should consequently be interpreted as a conditional indication of potential recovery rather than conclusive evidence that an upward market transition would occur. Research comparing gradient-boosting variants for stock-price direction prediction on the Indonesia Stock Exchange reinforces the importance of evaluating classification performance independently of the predicted directional category (Kho et al., 2025). The relatively limited probability assigned to the bullish class also restricts the strength of any interpretation concerning the persistence or magnitude of a potential market rebound. Accordingly, the shortest forecasting horizon provided a positive directional signal, but its reported performance does not establish sufficiently strong predictive reliability for definitive investment judgments.

The November 2026 scenario extended the forecasting interval to 40 trading days and produced an accuracy of 52.23%, with a predicted UP/BULLISH probability of 65,47%. This probability represented the highest bullish indication among the three forecasting scenarios, despite the reduction in classification accuracy relative to October. The discrepancy illustrates that greater probability assigned to a particular forecast does not necessarily correspond to higher accuracy across the evaluated observations. A model can express greater confidence in an individual directional outcome while continuing to exhibit considerable classification errors across the testing dataset. Investigations involving ensemble learning and XGBoost-based financial forecasting similarly distinguish the strength of predictive outputs from the broader question of model performance under financial market uncertainty (Din et al., 2025). The November prediction therefore indicated a stronger model-assigned preference for upward movement, although the reported accuracy provided limited support for interpreting that preference as an improvement in forecasting reliability.

The December 2026 forecast employed the longest observation horizon of 60 trading days, generating an UP/BULLISH probability of 60,11% and classification accuracy of 50.13%. Although the model continued to favor an upward movement, its accuracy was lower than the corresponding values obtained for October and November. Classification performance near 50% requires particular interpretive caution because the reported accuracy alone does not establish whether the model performs better than a simple prediction rule based on the distribution of actual market directions. The relatively favorable bullish probability also cannot be interpreted as evidence of a proportionate expected increase

in the IHSG, since the output refers to directional classification rather than the magnitude of an index change. Research examining longer-term stock market forecasting with XGBoost identifies prediction-horizon specification and evaluation procedures as relevant considerations when assessing the practical significance of forecasting outputs (Sürek & Lau, 2025). Broader assessments of machine learning and deep learning models for stock-market prediction similarly indicate that predictive effectiveness depends on the characteristics of the forecasting task and the evaluation approach employed (Sonkavde et al., 2023). The December scenario consequently suggested a possible upward direction toward the end of 2026, but the comparatively low classification accuracy restricted the strength of the empirical evidence supporting that prediction.

Comparative examination of the three scenarios revealed bullish probabilities of 52,86%, 65,47%, and 60,11% for the 20-day, 40-day, and 60-day horizons, respectively. Classification accuracy decreased progressively from 55.25% to 52.23% and subsequently to 50.13%, indicating weaker observed classification performance at longer forecasting horizons. The probabilities did not follow the same pattern, increasing between the first two scenarios before declining for the final forecasting interval. This divergence demonstrates that predicted directional probabilities and aggregate classification accuracy represent distinct dimensions of forecasting performance and should not be interpreted interchangeably. Comparative studies involving XGBoost and recurrent neural network models have also addressed forecasting performance as an empirical characteristic requiring assessment under the relevant prediction settings rather than an inherent property of the selected algorithm (Vuong et al., 2022). The consistent UP/BULLISH classifications identify a conditional expectation of market recovery, while the declining accuracy values establish the need for a more detailed examination of classification errors before assessing the reliability and practical implications of these forecasts.

XGBoost Classification Performance and Confusion Matrix Analysis

The evaluation of the XGBoost forecasts was based on three confusion matrices constructed from the chronologically reserved testing observations. Each matrix cross-tabulated observed and predicted IHSG directions using downward movement as class 0 and upward movement as class 1. Correct classifications occupied the true-negative and true-positive cells, whereas false-positive and false-negative cells represented the two types of directional error. This distinction matters because an upward prediction can appear frequent even when actual downward movements are inadequately recognized. The two-class framework also makes it possible to distinguish general accuracy from the model's ability to detect either market direction. Research on XGBoost-based equity classification has likewise treated prediction categories as operational outcomes whose accuracy must be assessed against observed market movements (Han et al., 2023). The supplied confusion-matrix counts provide the basis for examining these properties without introducing additional model outputs.

Every horizon-specific matrix contains 762 test observations, calculated by summing its four reported classification cells. The equal evaluation counts permit direct comparison of error frequencies, although the associated future-return labels are horizon-specific. This comparability does not mean that observations at the three horizons are independent or that their target windows are mutually exclusive. Under the study's chronological split, model performance was assessed on observations later than those assigned to training. A methodological review of stock-market prediction highlights the importance of clear data partitioning and evaluation criteria when comparing financial classifiers (Nti et al., 2020). Consequently, the confusion matrices were interpreted as test-set outcomes rather than evidence of verified market conditions after October 5, 2026.

The error analysis concentrates on the relative balance between upward and downward classifications rather than treating the largest diagonal cell as sufficient evidence of model effectiveness. True-positive counts describe accurately recognized upward movements, while true-negative counts represent correctly identified downward movements. False positives are particularly consequential for bullish interpretations because they occur when a decline is incorrectly identified as an increase. False negatives correspond to upward movements that are instead classified as downward. The distinction is consistent with research on feature-engineered stock-market classifiers, where directional outcomes have different consequences for evaluating forecasting signals (Yun et al., 2021). Class-specific performance must be examined alongside aggregate accuracy because uneven detection rates may be obscured by a single summary statistic. This evaluation is especially pertinent when the forecast classifications favor one market direction more often than the observed data justify.

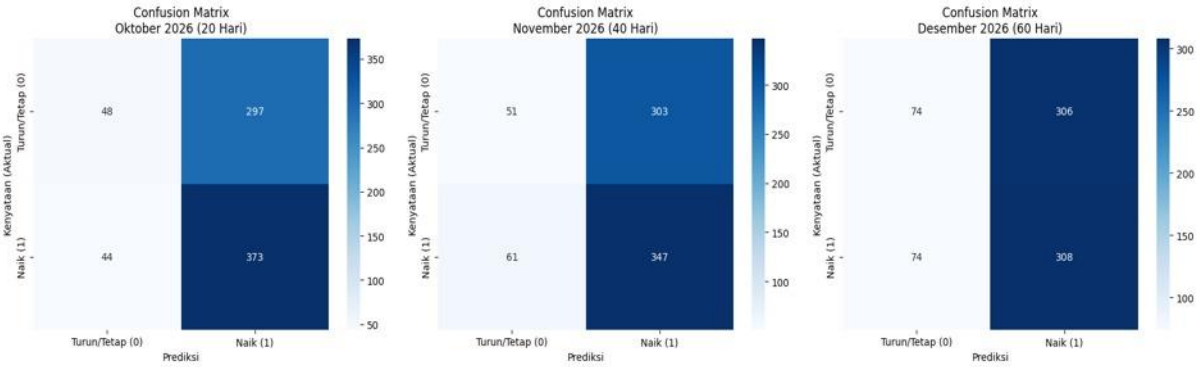


Figure 1. Confusion Matrix Results

Figure 1 shows that the 20-day classifier produced 48 true negatives, 297 false positives, 44 false negatives, and 373 true positives. The diagonal cells comprise 421 correct predictions out of 762 observations, corresponding to the reported accuracy of 55.25%. The true-positive count substantially exceeds the true-negative count, indicating a pronounced difference in the model's capacity to identify upward and downward movements. Among 345 actual downward observations, 297 were incorrectly classified as upward, making bullish false alarms the principal source of classification error. This pattern limits the practical interpretation of the positive forecast because many observed declines were not recognized by the classifier. Recent research on XGBoost-based stock-market prediction illustrates the relevance of evaluating class-specific behavior rather than equating a positive forecast with dependable market recovery (Gao, 2025).

Table 2. Classification Performance Metrics Across Forecasting Horizons

Horizon (Days)	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)
20	55.25	55.67	89.45	13.91	68.63
40	52.23	53.38	85.05	14.41	65.60
60	50.13	50.16	80.63	19.47	61.85

Source: Calculated from the original confusion matrix data, 2026

Table 2 summarizes classification metrics calculated directly from the four cells of each horizon-specific confusion matrix. The 20-day model recorded bullish precision of 55.67%, recall of 89.45%, specificity of 13.91%, and an F1-score of 68.63%. At the 40-day horizon, the corresponding values were 53.38%, 85.05%, 14.41%, and 65.60%. At 60 days, precision declined to 50.16%, recall to 80.63%, and the F1-score to 61.85%, while specificity increased to 19.47%. The comparatively high recall figures show that the model identified many actual increases, but its consistently low specificity demonstrates weak identification of actual declines. The contrast is important because high recall can coexist with numerous false-positive classifications, particularly when the predicted positive class predominates. Research on hyperparameter optimization for stock-market forecasting illustrates why model assessment should consider several complementary performance measures instead of relying exclusively on accuracy (Pebrianti et al., 2024).

The 40-day confusion matrix records 51 true negatives, 303 false positives, 61 false negatives, and 347 true positives. Its 398 correctly classified observations correspond to 52.23% accuracy, while the majority of its 364 errors arise from predicting increases that did not occur. At 60 days, the model produced 74 true negatives, 306 false positives, 74 false negatives, and 308 true positives. These counts yield 382 correct classifications and an accuracy of 50.13%, matching the reported result for the longest horizon. Between the shortest and longest horizons, correct bearish detections increased from 48 to 74, whereas correct bullish detections decreased from 373 to 308. Research applying XGBoost to financial markets emphasizes the importance of evaluating predictive outputs in relation to the particular task specification and market data used (Liu, 2024).

The observed classification asymmetry also appears in the distributions of predicted and actual upward movements. At 20 trading days, the classifier predicted 670 increases although only 417 increases occurred among the 762 reference outcomes. The respective counts were 650 predicted increases versus 408 actual increases for 40 days and 614 versus 382 for 60 days. A tendency to produce upward classifications more frequently than upward outcomes can increase recall while generating substantial false-positive errors. This combination helps explain why the models achieved relatively high bullish recall without corresponding strength in bearish detection. Research evaluating ensemble techniques for financial return prediction similarly supports examining how modeling decisions influence directional classification rather than interpreting one performance metric in isolation (Mohapatra et al., 2022). The reported probabilities of bullish movement in the forecasting summary must consequently be considered alongside this testing-set asymmetry, since a positive directional prediction does not independently establish the reliability of the underlying classification.

The classification results indicate that the three XGBoost models have limited ability to distinguish downward IHSG movements under the reported evaluation design. Their aggregate accuracies decrease with forecasting horizon, but the error counts reveal that performance limitations are more specifically associated with frequent bullish false positives. The available results do not establish probability calibration, feature contribution, or performance stability under alternative temporal testing periods. Research on financial time-series modeling identifies sensitivity to model configurations and changing data characteristics as important challenges for applying predictive algorithms to financial markets (Bao et al., 2025). The confusion matrices consequently provide a stronger empirical basis for describing observed classification behavior than for asserting that the index will recover during each projected period. These findings establish the need for a benchmark-based appraisal of whether the directional signals provide incremental forecasting information beyond the prevailing class distribution, particularly when the results are considered for financial decision-making.

Forecast Reliability, Market Implications, and Predictive Limitations

The bullish classifications reported for the three forecasting horizons should be evaluated as conditional outputs of the XGBoost model rather than verified recoveries in the IDX Composite Index (IHSG). The analysis used market observations through October 5, 2026, meaning that the October–December designations represented prospective forecasting scenarios rather than completed observations of subsequent market conditions. Contemporary reports documenting substantial IHSG weakness during the first half of 2026 established the financial-market circumstances motivating the forecasts, but these developments did not independently validate the predicted recovery (IDXChannel, 2026). Accounts of the same market correction also indicated that the observed decline reflected prevailing market conditions rather than evidence concerning the timing or probability of a subsequent reversal (Lidyana, 2026). An UP/BULLISH classification indicates that the model predicts a higher index level at the specified future horizon without estimating the magnitude of the anticipated increase. The reported bullish probabilities of 52.86%, 65.47%, and 60.11% describe the model's directional signals, although the available evaluation does not establish whether these probability estimates are calibrated against observed outcome frequencies. Their practical significance consequently depends on the historical classification errors and the extent to which the model provides predictive information beyond elementary directional expectations.

Classification accuracy decreased from 55.25% at the 20-day horizon to 52.23% at 40 days and 50.13% at 60 days, indicating weaker observed predictive performance as the evaluated forecasting interval increased. These results describe differences among three horizon-specific classification tasks rather than establishing that every extension of the forecasting period necessarily produces a corresponding reduction in accuracy. Longer forecasting intervals can involve different distributions of market movements and overlapping future-return observations, making the interpretation of aggregate accuracy dependent on the characteristics of each prediction target. Research examining autoregressive conditional-variance models and Fourier-series forecasting for the Indonesian Composite Index illustrates that model assessment depends on the representation of financial dynamics and the predictive objective being investigated (Mardianto et al., 2026). Comparisons involving ARIMA, LSTM, and hybrid forecasting approaches similarly demonstrate the relevance of evaluating predictive performance under clearly specified temporal structures and modeling configurations (Nensi et al., 2025). The

observed accuracy pattern therefore requires an additional benchmark to determine whether XGBoost classifications offer meaningful improvement over a simple directional prediction rule.

Classification accuracy exceeding 50% does not independently demonstrate meaningful predictive discrimination when upward and downward market movements are unequally represented within the testing observations. A simple classifier that consistently selects the more frequent actual class can achieve an accuracy equivalent to that class's proportion without distinguishing individual observations through explanatory features. Each reported confusion matrix contains 762 testing observations, allowing the frequency of actual upward and downward market movements to be calculated directly from the supplied results. The 20-day evaluation contains 417 upward and 345 downward observations, while the 40-day evaluation contains 408 upward and 354 downward observations. At the 60-day horizon, the corresponding distribution comprises 382 upward and 380 downward observations, indicating an almost balanced test population. Empirical financial research using machine learning emphasizes the importance of identifying predictive relationships that extend beyond simple patterns and distributions within historical market observations (Gu et al., 2020). These class frequencies provide the basis for a retrospective majority-class benchmark, although they cannot substitute for a baseline model developed exclusively from training information.

Table 3. Comparison of XGBoost Accuracy and Majority-Class Baseline

Forecast Horizon (Days)	XGBoost Accuracy (%)	Majority-Class Baseline (%)	Accuracy Difference (Percentage Points)
20	55.25	54.72	+0.52
40	52.23	53.54	−1.31
60	50.13	50.13	0.00

Source: Calculated from the original confusion matrix data, 2026

Table 3 demonstrates that the 20-day XGBoost classifier achieved an accuracy of 55.25%, exceeding the retrospective majority-class baseline of 54.72% by approximately 0.52 percentage points. At the 40-day horizon, the model's accuracy of 52.23% was 1.31 percentage points below the corresponding baseline of 53.54%, indicating that the reported classification performance did not exceed this elementary reference. The 60-day model and the retrospective majority-class benchmark produced identical rounded accuracies of 50.13%, providing no observed improvement in aggregate accuracy under this comparison. These differences indicate limited discriminatory advantages, although the retrospective baseline cannot establish how XGBoost would perform against a competing classifier trained and evaluated through the same chronological procedure. The high false-positive frequencies identified in the confusion matrices reinforce this limitation because the models repeatedly classified actual market declines as upward movements. Research reviewing technical and fundamental approaches to stock-market prediction emphasizes that the suitability of evaluation measures depends on the prediction objective and the characteristics of classification errors (Nti et al., 2020).

Comparisons with earlier IHSG forecasting research require careful distinction between directional classification and the prediction of numerical index values. Statistical approaches integrating autoregressive structures, conditional-variance models, and Fourier-series representations investigate time-series characteristics through predictive specifications that differ from the binary targets evaluated in the present analysis (Mardianto et al., 2026). Studies comparing ARIMA, LSTM, and hybrid architectures also examine the forecasting properties of Indonesian composite-index observations, although their results cannot be ranked directly against classification accuracy without equivalent target definitions and evaluation measures (Nensi et al., 2025). Multivariate regression incorporating time-series errors represents another approach to modeling IHSG fluctuations, particularly when relationships among financial variables are considered alongside changing market conditions (Ulyah et al., 2024). Research comparing Prophet, XGBoost, and ridge regression for Indonesian banking stocks provides evidence of alternative forecasting architectures, but conclusions derived from individual banking securities cannot automatically be transferred to a broad composite index (Tony et al., 2025). The present results do not establish the superiority of XGBoost over these methods because no competing algorithms were evaluated using identical training periods, testing observations, and prediction targets. Their empirical contribution instead concerns the documented behavior of a common

binary classification framework across three IHSG forecasting horizons, particularly the divergence between positive directional probabilities and limited observed classification accuracy.

Market conditions occurring after the observation cutoff constitute an additional source of uncertainty that cannot be fully represented by historical model performance. Financial-market volatility indicators such as the VIX provide information about expectations of market uncertainty, but they do not independently determine subsequent movements in a particular national equity index (Whaley, 2000). Reports identifying individual stocks that supported or suppressed IHSG performance during the first half of 2026 further demonstrate that changes in the aggregate index can reflect heterogeneous movements among its constituent securities (Al Ishaqi, 2026). The study incorporated historical financial information related to domestic and international market indicators, although the reported results do not quantify the separate contributions of these variables to classification decisions. Relationships estimated from historical observations may also change following shifts in monetary conditions, international capital flows, commodity prices, or investor expectations. Research applying deep learning to financial-market prediction has examined the challenges of extracting predictive information from historical market data, reinforcing the need to distinguish observed model performance from the certainty of future outcomes (Fischer & Krauss, 2018).

For investors and portfolio managers, the reported bullish classifications provide conditional directional information but do not establish a sufficiently reliable basis for independent trading decisions. A positive prediction does not quantify the expected index increase, downside exposure, transaction costs, or the financial consequences of an incorrectly classified market decline. Research combining XGBoost with financial signal-processing techniques for portfolio forecasting illustrates the importance of connecting predictive outputs with the requirements of investment decision-making rather than treating forecasting signals as complete investment strategies (Dezhkam & Manzuri, 2023). Financial analysts may find the classification results useful for examining the relative frequency of upward and downward predictions, particularly because the confusion matrices reveal a substantial tendency toward bullish false positives. For policymakers and institutional observers, the three forecast scenarios offer a limited computational perspective on prospective market developments but cannot independently demonstrate a sustained improvement in Indonesia's economic or financial conditions. Recent applications of optimized XGBoost models to stock-index prediction also illustrate the continuing methodological interest in improving predictive architectures through financial feature selection and model optimization (Anam et al., 2026). Responsible interpretation of the present forecasts requires retaining the original observation cutoff, the three trading-day horizons, and the documented limitations of the testing results when communicating their potential implications.

The empirical findings identify declining classification accuracy, a pronounced tendency toward upward predictions, and minimal improvement relative to a retrospective majority-class benchmark across the three forecasting horizons. The results establish that XGBoost generated UP/BULLISH classifications for the October, November, and December 2026 scenarios, but they do not demonstrate that the corresponding market recoveries had already occurred or would necessarily materialize. The 20-, 40-, and 60-day intervals refer to trading sessions rather than fixed calendar-month boundaries, requiring careful interpretation of the reported target periods when relating model outputs to future market dates. Although the documented analytical sequence includes feature engineering, chronological data partitioning, model training, and prediction generation, consistent with a standard machine learning workflow, the available results do not establish probability calibration, feature-level contributions, or repeated temporal validation (Daqiqil, 2021). Research examining longer-term stock-market prediction with XGBoost underscores the relevance of forecasting design and evaluation procedures when assessing the significance of predictive outputs (Sürek & Lau, 2025). Subsequent research could evaluate prospectively implemented baseline classifiers, probability calibration, alternative temporal validation schemes, and competing predictive architectures before assigning stronger practical significance to the reported bullish forecasting signals.

CONCLUSION

This study concludes that the XGBoost algorithm generated directional forecasts for the IDX Composite Index (IHSG) across three forecasting horizons extending from October through December 2026. The model consistently projected upward movements, indicating a potential bullish trend, with predicted probabilities of 52.86%, 65.47%, and 60.11% for the 20-, 40-, and 60-trading-day horizons,

respectively. The corresponding classification accuracies were 55.25%, 52.23%, and 50.13%, demonstrating a progressive decline in predictive performance as the forecasting horizon increased. Confusion matrix analysis revealed a pronounced tendency toward bullish classifications, with frequent false-positive predictions indicating limited capability to identify actual market declines. Comparison with the retrospective majority-class baseline showed only marginal improvement at the 20-day horizon, inferior performance at 40 days, and equivalent accuracy at 60 days. These findings indicate that XGBoost can generate consistent directional projections, although the observed classification performance does not provide sufficient evidence to confirm a sustained IHSG recovery toward the end of 2026.

The forecasting results may serve as supplementary analytical information for individual investors, investment managers, financial analysts, and government authorities when assessing potential market developments ahead of the year-end financial reporting period. However, the relatively modest classification accuracy and substantial bullish prediction errors require cautious interpretation, particularly when the projections are considered for asset allocation and financial decision-making. The study contributes empirical evidence concerning the application of XGBoost to Indonesian stock-market direction classification using historical financial observations and differentiated forecasting horizons. Its principal methodological contribution lies in distinguishing directional probabilities from actual predictive accuracy while demonstrating the importance of evaluating classification performance against a basic benchmark. Further research should incorporate alternative machine learning algorithms, probability calibration, additional temporal validation procedures, and systematic evaluation of class-specific prediction errors to establish more reliable forecasting performance. The findings provide a basis for refining computational approaches to IHSG forecasting while emphasizing that predictions derived from historical market patterns remain subject to uncertainty and should not be interpreted as definitive investment signals.

REFERENCES

Al Ishaqi, A. M. (2026, July 1). Daftar 10 saham penopang dan penekan IHSG semester I/2026. *Bisnis.com*. https://market.bisnis.com/read/20260701/7/1984660/daftar-10-saham-penopang-dan-penekan-ihsg-semester-i2026/2

Anam, S., Shamsuddin, M. R., Kustiawan, E., Mahanani, D. M., & Yusuf, F. I. (2026). Grey wolf-optimized XGBoost regressor for stock index prediction with financial features. *BAREKENG: Jurnal Ilmu Matematika dan Terapan, 20*(3), 2131–2150. https://doi.org/10.30598/barekengvol20iss3pp2131-2150

Bao, W., Cao, Y., Yang, Y., Che, H., Huang, J., & Wen, S. (2025). Data-driven stock forecasting models based on neural networks: A review. *Information Fusion, 113*, 102616. https://doi.org/10.1016/j.inffus.2024.102616

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785

Daqiqil, I. (2021). *Machine learning: Teori, studi kasus, dan implementasi menggunakan Python*. UR Press.

Deng, S., Huang, X., Zhu, Y., Su, Z., Fu, Z., & Shimada, T. (2023). Stock index direction forecasting using an explainable eXtreme Gradient Boosting and investor sentiments. *The North American Journal of Economics and Finance, 64*, 101848. https://doi.org/10.1016/j.najef.2022.101848

Deni, D. R., & Amalita, N. (2026). Application of temporal fusion transformer for multivariate forecasting of Indonesian composite stock price index. *Journal of Multidisciplinary Science: MIKAILALSYS, 4*(3), 6481–6508. https://doi.org/10.58578/mikailalsys.v4i3.12032

Dezhkam, A., & Manzuri, M. T. (2023). Forecasting stock market for an efficient portfolio by combining XGBoost and Hilbert–Huang transform. *Engineering Applications of Artificial Intelligence, 118*, 105626. https://doi.org/10.1016/j.engappai.2022.105626

Din, R. U., Ahmed, S., Khan, S. H., Albanyan, A., Hoxha, J., & Alkhamees, B. (2025). A novel decision ensemble framework: Attention-customized BiLSTM and XGBoost for speculative stock price forecasting. *PLOS ONE, 20*(4), e0320089. https://doi.org/10.1371/journal.pone.0320089

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research, 270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Gao, J. (2025). Application of XGBoost in the A-shares stock market forecasting. *In Proceedings of the 2024 5th International Conference on Big Data Economy and Information Management* (pp. 503–511). Association for Computing Machinery. https://doi.org/10.1145/3724154.3724237

Google. (n.d.). *Google Colaboratory notebook* [Notebook]. https://colab.research.google.com/drive/1ETWilEnIu_vE5cx_S8AGZ8NV8tWQ7P_U

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Han, Y., Kim, J., & Enke, D. (2023). A machine learning trading system for the stock market based on N-period Min-Max labeling using XGBoost. *Expert Systems with Applications, 211*, 118581. https://doi.org/10.1016/j.eswa.2022.118581

IDXChannel. (2026, July 7). *IHSG koreksi 34,74 persen sepanjang semester I-2026, OJK ungkap penyebabnya*. RCTI+. https://www.rctiplus.com/news/detail/idxchannel/5422806/ihsg-koreksi-34-74-persen-sepanjang-semester-i-2026--ojk-ungkap-penyebabnya

Indra, Supian, S., Sukono, Riaman, Saputra, M. P. A., Azahra, A. S., & Pirdaus, D. I. (2025). Stock return prediction on the LQ45 market index in the Indonesia Stock Exchange using a machine learning algorithm based on technical indicators. *Journal of Risk and Financial Management, 18*(12), 714. https://doi.org/10.3390/jrfm18120714

Kho, D. C., Purnomo, H. D., & Hendry, H. (2025). Performance analysis of gradient boosting models variants in predicting the direction of stock closing prices on the Indonesia Stock Exchange. *BAREKENG: Jurnal Ilmu Matematika dan Terapan, 19*(2), 1393–1408. https://doi.org/10.30598/barekengvol19iss2pp1393-1408

Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications, 197*, 116659. https://doi.org/10.1016/j.eswa.2022.116659

Lidyana, V. (2026). *IHSG babak belur di akhir semester I 2026, OJK ungkap penyebabnya*. IDN Times. https://www.idntimes.com/business/economy/ihsg-babak-belur-di-akhir-semester-i-2026-ojk-ungkap-penyebabnya-00-3m8tp-ztyn6y

Liu, J. (2024). Predicting Chinese stock market using XGBoost multi-objective optimization with optimal weighting. *PeerJ Computer Science, 10*, e1931. https://doi.org/10.7717/peerj-cs.1931

Mardianto, M. F. F., Valida, H., Putri, F. F., Fauzi, D. M., & Pusporani, E. (2026). Prediction of the Indonesia Composite Index (ICI) using the ARCH GARCH approach and the Fourier series. *BAREKENG: Journal of Mathematics and Its Applications, 20*(1), 271–286. https://doi.org/10.30598/barekengvol20iss1pp0271-0286

Mohapatra, S., Mukherjee, R., Roy, A., Sengupta, A., & Puniyani, A. (2022). Can ensemble machine learning methods predict stock returns for Indian banks using technical indicators? *Journal of Risk and Financial Management, 15*(8), 350. https://doi.org/10.3390/jrfm15080350

Nensi, A. I. E., Al Maida, M., Notodiputro, K. A., Angraini, Y., & Mualifah, L. N. A. (2025). Performance analysis of ARIMA, LSTM, and hybrid ARIMA-LSTM in forecasting the composite stock price index. *CAUCHY: Jurnal Matematika Murni dan Aplikasi, 10*(2), 588–604. https://doi.org/10.18860/cauchy.v10i2.33379

Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review, 53*(4), 3007–3057. https://doi.org/10.1007/s10462-019-09754-z

Oukhouya, H., Kadiri, H., El Himdi, K., & Guerbaz, R. (2023). Forecasting international stock market trends: XGBoost, LSTM, LSTM-XGBoost, and backtesting XGBoost models. *Statistics, Optimization & Information Computing, 12*(1), 200–209. https://doi.org/10.19139/soic-2310-5070-1822

Pebrianti, D., Kurniawan, H., Bayuaji, L., & Rusdah, R. (2024). XGBoost hyper-parameter tuning using particle swarm optimization for stock price forecasting. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika, 9*(4), 1179–1195. https://doi.org/10.26555/jiteki.v9i4.27712

Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic

review, performance analysis and discussion of implications. *International Journal of Financial Studies, 11*(3), 94. https://doi.org/10.3390/ijfs11030094

Sürek, R. E., & Lau, W.-Y. (2025). A refined methodological approach: Long-term stock market forecasting with XGBoost. *Journal of Intelligent Systems, 34*(1), 823–846. https://doi.org/10.1515/jisys-2025-0027

Tony, Ratchagit, M., & Hiryanto, L. (2025). Forecasting Indonesian banking stock prices using Prophet, XGBoost, and ridge regression: A comparative analysis. *Jurnal Teknik Informatika, 6*(4), 2890–2901. https://doi.org/10.52436/1.jutif.2025.6.4.4973

Ulyah, S. M., Susanti, R., Andreas, C., Rahmayanti, I. A., Rifada, M., Fitriyani, N. L., & Ana, E. (2024). A multivariate regression with time series error in forecasting Jakarta Composite Index and stock prices of banking industry in Indonesia by considering COVID-19 effect. *International Journal of Technology, 15*(6), 1839–1850. https://doi.org/10.14716/ijtech.v15i6.5469

Vuong, P. H., Dat, T. T., Mai, T. K., Uyen, P. H., & The Bao, P. (2022). Stock-price forecasting based on XGBoost and LSTM. *Computer Systems Science and Engineering, 40*(1), 237–246. https://doi.org/10.32604/csse.2022.017685

Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management, 26*(3), 12–17. https://doi.org/10.3905/jpm.2000.319728

Yahoo Finance. (2026). *Historical market data: IHSG, S&P 500, NASDAQ, Nikkei, Hang Seng, STI, VIX, UST 10Y, USD/IDR, DXY, Brent, WTI, gold, and coal* [Data set]. Retrieved October 5, 2026, from https://finance.yahoo.com

Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications, 186*, 115716. https://doi.org/10.1016/j.eswa.2021.115716