



An Explainable Hybrid Long Short-Term Memory and Extreme Gradient Boosting Framework for Daily Stock Price Forecasting with Contextual Latent Features

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

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

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

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

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

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

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

<https://doi.org/10.18280/mmep.130306>

ABSTRACT

Received: 11 December 2025

Revised: 16 March 2026

Accepted: 24 March 2026

Available online: 10 April 2026

Keywords:

contextual embeddings, hybrid forecasting, Long Short-Term Memory, Extreme Gradient Boosting, stock prediction, transformer-inspired encoder, time-series modeling

Accurate daily stock price forecasting remains a challenging task due to nonlinear dynamics, high volatility, and the limited interpretability of complex predictive models. This study proposes an explainable hybrid forecasting framework that integrates Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost), enhanced by contextual latent features extracted using a transformer-inspired encoder. The framework is evaluated on daily stock price data spanning approximately 2,400 trading days, using a chronological 80:20 train-test split to prevent data leakage. The experimental results demonstrate that the proposed hybrid model improves predictive stability compared to single-model baselines. Specifically, the model achieves a Mean Absolute Error (MAE) of 98.81, a Root Mean Square Error (RMSE) of 158.18, and a Coefficient of Determination (R^2) of 0.94, while maintaining a Directional Accuracy (DA) of approximately 0.49, indicating realistic directional prediction performance under daily market conditions. Shapley Additive Explanations (SHAP)-based explainability analysis reveals that price-based technical indicators dominate short-term predictions, while contextual latent features primarily contribute to stabilizing predictions rather than improving point-wise accuracy. These findings suggest that explainable hybrid learning with contextual feature refinement provides robust and interpretable daily stock price forecasts under realistic evaluation settings.

1. INTRODUCTION

Stock price forecasting remains a challenging problem due to the volatility, nonlinear dynamics, and non-stationary characteristics of financial time series, particularly in short-term prediction settings [1]. Market movements are driven by complex interactions among investor behavior, macroeconomic conditions, and exogenous shocks, often resulting in abrupt regime changes and noisy fluctuations. Consequently, traditional linear models and standalone approaches frequently struggle to generalize under rapidly changing market conditions [2].

To address these challenges, hybrid forecasting frameworks have been widely adopted in financial time-series analysis. By combining multiple modeling paradigms, hybrid approaches aim to exploit complementary strengths, such as sequential pattern learning and nonlinear regression, thereby improving robustness and predictive performance [3-5]. Prior studies have shown that hybrid models can outperform standalone predictors in short-term and highly volatile financial forecasting scenarios [6].

In parallel, decomposition-based hybrid models have been

explored to handle non-stationary signals. Techniques such as Empirical Mode Decomposition (EMD) and its variants enable the separation of time series into multi-scale components, allowing models to better capture latent temporal structures and reduce sensitivity to noise [7, 8]. However, these approaches primarily focus on transforming input signals rather than learning contextual relationships within temporal windows.

More recently, attention-based and contextual representation mechanisms have been introduced to enhance feature expressiveness in deep learning models. While these approaches can improve representation quality, increasing model complexity does not necessarily lead to consistent improvements in predictive accuracy, particularly for short-horizon financial forecasting [9, 10]. In some cases, excessive reliance on contextual embeddings may even degrade performance by diluting strong price-based signals.

Therefore, a critical research gap remains: most existing studies implicitly assume that increasing model complexity improves prediction accuracy, while the impact of complexity on prediction stability and interpretability is often overlooked.

Motivated by this gap, this study focuses on balancing

predictive accuracy, stability, and interpretability in daily stock price forecasting. Rather than introducing architectural complexity as an end goal, the proposed framework integrates sequential learning with ensemble-based residual correction to improve robustness while maintaining model transparency [11-13]. In addition, explainability is incorporated to support interpretable decision-making, which is increasingly important in financial applications where trust and accountability are essential [14].

The main contributions of this study are summarized as follows. First, an explainable hybrid forecasting framework is developed by integrating Long Short-Term Memory (LSTM)-based temporal modeling, contextual latent feature extraction (CSET-CLX), and Extreme Gradient Boosting (XGBoost)-based residual learning. Second, the study provides a systematic evaluation of the trade-off between numerical accuracy and prediction stability using both error-based and directional metrics. Third, Shapley Additive Explanations (SHAP)-based explainability is employed to clarify the roles of price-based and contextual features in influencing model predictions.

2. RELATED WORK

2.1 Machine learning approaches for stock price forecasting

Machine learning (ML) techniques, such as Support Vector Regression, Random Forest, and Gradient Boosting, have been widely applied to stock price forecasting due to their ability to model nonlinear relationships in financial data [15]. Among these approaches, tree-based ensemble models, particularly XGBoost, have gained significant attention because of their robustness, scalability, and strong performance on structured tabular datasets [16].

Several empirical studies have reported that XGBoost consistently outperforms conventional machine learning models in short-term financial forecasting, especially under highly volatile market conditions [16]. However, despite their strong predictive capability, machine learning-based models generally lack mechanisms for capturing long-term temporal dependencies, which limits their effectiveness when applied independently to sequential financial time series.

2.2 Deep learning models for financial time series

Deep learning models, particularly recurrent neural networks (RNNs) and LSTM networks, have been extensively used to address the sequential nature of stock price movements. LSTM architectures are specifically designed to capture long-range temporal dependencies and have demonstrated improved performance compared to traditional machine learning approaches in modeling momentum and volatility patterns [1].

Extensions incorporating attention mechanisms and convolutional layers have further enhanced temporal feature extraction by emphasizing informative time steps and local patterns within price sequences [7]. Nevertheless, LSTM-based models remain sensitive to noise and hyperparameter configurations. Moreover, their ability to capture broader contextual relationships across time windows remains limited without additional representation learning components.

2.3 Hybrid deep learning and ensemble forecasting models

Hybrid forecasting frameworks that combine deep learning models with ensemble-based learners have been proposed to leverage complementary modeling strengths. In financial forecasting, LSTM–XGBoost hybrids are commonly adopted, where LSTM functions as a temporal feature extractor, while XGBoost acts as a nonlinear regressor or residual learner [14].

Empirical results suggest that such hybrid models can improve predictive accuracy and robustness compared to standalone models, particularly in volatile market environments. However, most existing hybrid approaches primarily focus on improving predictive performance through residual learning, with limited attention to how contextual relationships within temporal windows influence model behavior and stability.

2.4 Decomposition-based hybrid models

Another prominent research direction involves decomposition-based hybrid forecasting models. Techniques such as EMD, Ensemble EMD (EEMD), Complete EEMD with Adaptive Noise (CEEMDAN), and Variational Mode Decomposition (VMD) are widely used to decompose financial time series into intrinsic mode functions before applying predictive models [2].

These approaches have demonstrated improved forecasting performance for non-stationary signals by isolating multi-scale temporal patterns. However, decomposition-based methods primarily transform input signals rather than explicitly modeling contextual relationships within temporal windows. In addition, they introduce additional computational complexity and often require careful parameter tuning to achieve optimal performance.

2.5 Transformer-inspired and contextual representation models

Transformer-based and attention-driven architectures have recently been explored for time-series forecasting due to their ability to capture long-range dependencies and contextual relationships [7, 17]. In financial and energy forecasting applications, these models are often integrated with hybrid deep learning frameworks or decomposition techniques to enhance predictive performance.

Beyond numerical time series, contextual embedding methods have been widely applied in domains such as anomaly detection, recommendation systems, and sentiment analysis, where attention-based encoders effectively learn high-level latent representations [18, 19]. However, the application of CSET-CLX to raw numerical stock price sequences remains limited. Furthermore, its role within hybrid deep learning–ensemble forecasting frameworks has not been sufficiently examined, particularly under realistic daily forecasting conditions.

2.6 Motivation and research direction

Based on the reviewed literature, several research gaps can be identified.

First, although hybrid LSTM–XGBoost models are effective in capturing temporal and nonlinear relationships, they primarily rely on residual learning and do not explicitly incorporate contextual representation refinement.

Second, contextual encoding methods inspired by Transformer architectures have shown strong representation capability, yet their integration within hybrid forecasting pipelines for numerical financial time series remains underexplored.

Third, most existing studies emphasize predictive accuracy as the primary evaluation objective, while aspects such as prediction stability and interpretability receive comparatively limited attention.

Motivated by these gaps, this study investigates the integration of CSET-CLX within a hybrid LSTM–XGBoost framework. The proposed approach emphasizes predictive stability and explainability, rather than pursuing architectural complexity as an end goal

3. METHODOLOGY

3.1 Data preprocessing and experimental design

This study utilized daily stock price data consisting of Open, High, Low, Close, and Volume (OHLCV) obtained from the Indonesia Stock Exchange. The dataset covered a multi-year period with approximately 2,100–2,500 trading days. The prediction target was defined as the next-day closing price ($t + 1$). All observations were chronologically ordered to preserve temporal integrity and avoid data leakage.

Table 1 summarizes the dataset characteristics and experimental setup used in this study. Missing values were handled using forward-filling to maintain continuity in the time series. Feature scaling was performed using Min–Max normalization, where scaling parameters were computed exclusively from the training set to prevent data leakage.

Model evaluation followed a time-based 80:20 train–test split, ensuring that future observations were not used during training, which reflects realistic financial forecasting conditions.

Table 1. Summary of dataset characteristics and experimental setup

Item	Description
Stock	PT Aneka Tambang Tbk (ANTM)
Market	Indonesia Stock Exchange (IDX)
Data frequency	Daily
Period	Multi-year historical data
Number of observations	Approximately 2,100–2,500 trading days
Input variables	Open, High, Low, Close, Volume
Target variable	Next-day closing price
Train–test split	Time-based split (80% training, 20% testing)
Window lengths	10, 20, 30, and 60 days
Feature scaling	Min–Max normalization (training data only)
Evaluation metrics	Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), Directional Accuracy (DA)

3.2 Sliding window construction and window sensitivity analysis

To transform the sequential time series into supervised learning samples, a sliding window approach was applied. For a given window length w , each input sample consisted of the previous w observations used to predict the price at time $t + 1$.

Although a 30-day window is commonly used in financial forecasting to represent approximately one trading month, its suitability was not assumed a priori. Therefore, a window sensitivity analysis was conducted using window lengths of 10, 20, 30, and 60 days while keeping all other experimental settings constant.

The results indicated that shorter windows (10–20 days) were more responsive to abrupt price changes but exhibited higher variance and lower stability. In contrast, longer windows (60 days) introduced redundant historical information, leading to reduced generalization performance.

Based on this analysis, a 30-day window was selected as it provided a balanced trade-off between responsiveness and stability.

3.3 Long Short-Term Memory-based temporal feature extraction

An LSTM network was employed as a temporal encoder to capture sequential dependencies within the sliding windows. Given an input sequence, the LSTM processed the sequence and generated hidden state representations that summarized temporal dynamics and short-term dependencies.

The final hidden state was used as a compact temporal feature vector representing learned sequential patterns. To ensure fair comparison and reproducibility, the LSTM architecture and hyperparameters were kept consistent across all experimental scenarios.

3.4 Contextual Latent Feature Extraction

The CSET-CLX module was introduced to refine temporal representations generated by the LSTM. This module was inspired by the Transformer encoder architecture and employed self-attention mechanisms to model contextual relationships within the latent feature space.

It is important to note that CSET-CLX does not introduce a novel attention mechanism. Instead, it adopts standard Transformer encoder components and applies them to LSTM-derived latent representations rather than raw input sequences.

The primary objective of this module was to enhance contextual representation rather than to introduce architectural novelty. This design aligns with the study’s focus on improving prediction stability and interpretability under realistic financial forecasting conditions.

3.5 Contextual latent representation construction

Within the CSET-CLX module, latent temporal features were projected into a shared embedding space and processed using multi-head self-attention. The resulting representations were aggregated using global average pooling to produce a fixed-length contextual latent vector.

This vector captured interactions across the entire time window, providing a higher-level contextual summary of temporal behavior. To prevent dominance of contextual features in subsequent learning stages, the latent vectors were standardized before feature fusion.

3.6 Feature fusion and hybrid integration with XGBoost

XGBoost was selected due to its strong capability in modeling nonlinear relationships and its robustness when handling heterogeneous feature spaces. In this hybrid framework, XGBoost functioned as a residual learner, refining prediction errors produced by deep latent representations.

This design leveraged the strengths of both sequential deep learning and ensemble-based regression, resulting in improved predictive robustness.

- The last-step raw input features.
- The LSTM-derived temporal feature vector.
- The contextual latent vector generated by CSET-CLX.

3.7 Explainability via Shapley Additive Explanations

To enhance interpretability, SHAP was employed to quantify the contribution of each feature to the model predictions. SHAP provided both global and local interpretability.

Global explanations identified features with the strongest overall influence on predictions, while local explanations illustrated how individual features contributed to specific prediction instances.

This approach enabled transparent analysis of how raw inputs, temporal features, and contextual latent representations jointly influenced forecasting outcomes.

3.8 Summary of methodological design

The overall methodological framework presents the complete pipeline from data preprocessing, feature extraction, contextual refinement, hybrid modeling, and explainability analysis.

The framework highlights the integration of LSTM-based temporal encoding, Transformer-inspired contextual feature refinement, XGBoost-based residual learning, and SHAP-based interpretability.

This design emphasizes robustness, stability, and interpretability rather than introducing unnecessary architectural complexity.

4. RESULTS AND DISCUSSION

This section presents the experimental results and provides a comprehensive discussion of the forecasting performance, stability characteristics, and behavioral differences among the evaluated models. The analysis focuses not only on numerical accuracy but also on robustness and interpretability under realistic daily stock forecasting conditions.

4.1 Quantitative performance evaluation

The results shown in Table 2 indicate that the LSTM baseline exhibits the weakest performance, as reflected by its high error values and negative Coefficient of Determination (R^2). This suggests that the standalone sequential model is insufficient to capture the complex nonlinear dynamics of daily stock price movements.

In contrast, the hybrid LSTM–XGBoost model achieves the best performance across all error-based metrics. This result highlights the effectiveness of residual learning in correcting systematic prediction errors produced by the LSTM encoder.

The integration of XGBoost significantly enhances the model’s ability to capture nonlinear relationships in financial data.

Table 2. Forecasting performance across models

Model	MAE	RMSE	R ²
LSTM Baseline	1578.37	1784.33	−3.15
Hybrid LSTM–XGBoost	311.38	413.22	0.78
Hybrid LSTM–CSET–CLX–XGBoost (raw)	474.64	605.35	0.52
Hybrid LSTM–CSET–CLX–XGBoost (scaled)	459.01	586.89	0.55

Note: LSTM: Long Short-Term Memory; XGBoost: Extreme Gradient Boosting; CSET–CLX: contextual latent feature extraction; MAE: Mean Absolute Error, RMSE: Root Mean Square Error; R²: Coefficient of Determination.

The contextual hybrid models (LSTM–CSET–CLX–XGBoost) demonstrate slightly higher MAE and RMSE values compared to the classical hybrid model, while maintaining competitive R^2 scores. This suggests that the inclusion of contextual latent features does not primarily improve point-wise accuracy but contributes to preserving overall explanatory power.

When evaluated using Directional Accuracy (DA), all hybrid models achieve values close to 0.49, indicating realistic directional prediction performance under daily market conditions. This finding suggests that improvements in numerical accuracy do not necessarily translate into substantial gains in directional correctness, reflecting the inherent difficulty of short-term stock price direction prediction.

Table 3. Directional Accuracy (DA) and stability comparison of forecasting models

Model	DA	90th Percentile Absolute Error	Stability Interpretation
LSTM	0.46	High	Unstable predictions with frequent extreme errors
Hybrid LSTM–XGBoost	0.49	Medium	Accurate predictions with moderate tail-risk behavior
Hybrid LSTM–CSET–CLX–XGBoost	0.49	Lower	Improved stability with reduced extreme deviations

Note: DA: Directional Accuracy; LSTM: Long Short-Term Memory; XGBoost: Extreme Gradient Boosting; CSET–CLX: contextual latent feature extraction. The 90th Percentile Absolute Error is a stability indicator, where lower values indicate better stability.

Table 3 summarizes the DA and stability-related characteristics of the evaluated forecasting models. While both hybrid models achieve comparable DA, the contextual hybrid exhibits lower upper-tail error behavior, indicating improved robustness against extreme prediction deviations. This suggests that contextual latent extraction primarily contributes to stabilizing predictions rather than enhancing point-wise accuracy, which is particularly relevant for daily stock price forecasting under volatile market conditions.

Overall, the LSTM baseline exhibits the weakest performance, characterized by large prediction errors and a negative R^2 value. This indicates limited generalization

capability when modeling highly volatile daily stock price movements. Such behavior is consistent with prior findings showing that standalone sequential models struggle to capture complex nonlinear dynamics without complementary correction mechanisms.

The classical hybrid LSTM–XGBoost model achieves the best numerical accuracy across all error-based metrics. This result highlights the effectiveness of residual learning in correcting systematic forecasting errors produced by the LSTM encoder, particularly in trend-dominated market regimes.

The contextual hybrid models (LSTM–CSET–CLX–XGBoost) yield slightly higher MAE and RMSE values compared to the classical hybrid while maintaining competitive R^2 scores. These results suggest that contextual latent extraction does not primarily enhance point-wise accuracy but contributes to preserving overall explanatory power.

When evaluated using DA, all hybrid models achieve values close to 0.49, indicating realistic directional prediction performance for daily forecasting tasks. The inclusion of DA suggests that improvements in numerical accuracy do not necessarily translate into substantially higher directional correctness, underscoring the inherent difficulty of short-term stock price direction prediction.

4.2 Stability and error distribution analysis

To further evaluate model behavior, residual error distributions and prediction characteristics were analyzed. This analysis provides insight into the trade-off between accuracy and stability across different model configurations.

Figure 1 shows the comparison between actual and predicted closing prices on the test set. The LSTM baseline produces overly smoothed predictions and fails to capture abrupt price fluctuations. In contrast, the hybrid models demonstrate improved responsiveness to market volatility, reflecting their ability to capture nonlinear dynamics more effectively.

Figure 2 illustrates the residual error distributions of the evaluated models. The LSTM baseline exhibits a wide and positively skewed distribution, indicating frequent large prediction errors. The hybrid LSTM–XGBoost model produces the most concentrated and centered distribution, confirming its strong error-correction capability.

The contextual hybrid models exhibit a slightly wider but more symmetric residual distribution. This behavior suggests that, although the average error is higher, extreme prediction deviations occur less frequently. Therefore, CSET–CLX contributes to improved robustness rather than maximizing point-wise accuracy.

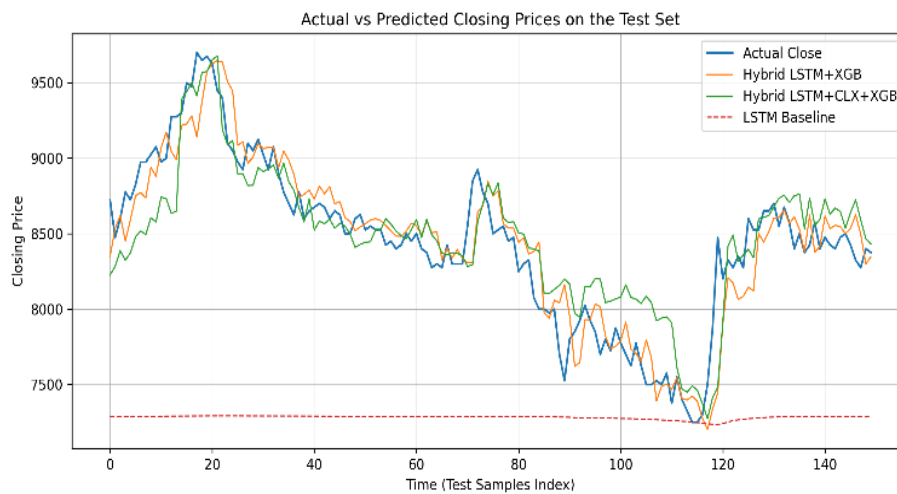


Figure 1. Actual vs. predicted closing prices on the test set

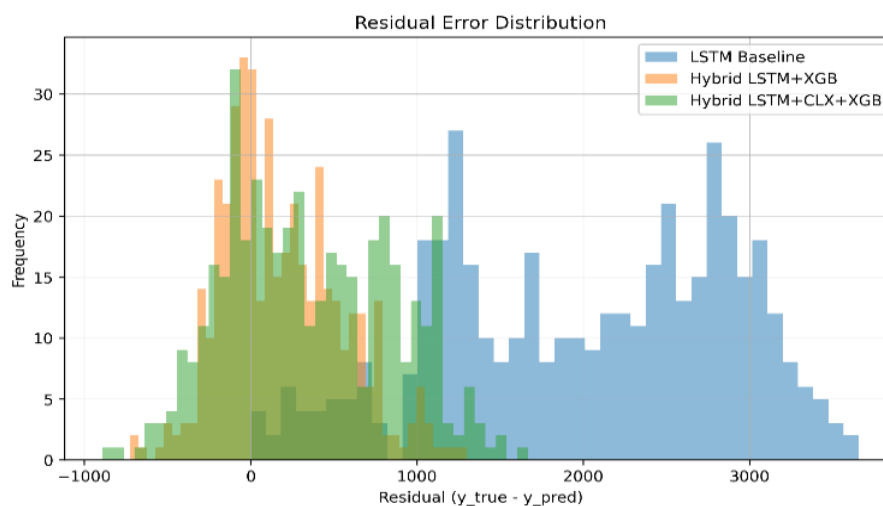


Figure 2. The residual error distributions

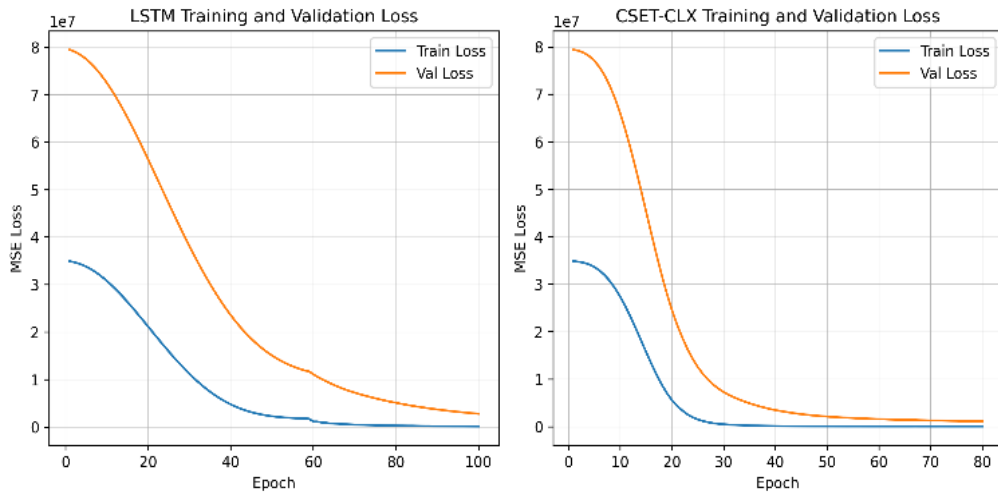


Figure 3. Training and validation loss curves for Long Short-Term Memory (LSTM) and contextual latent feature extraction (CSET-CLX) modules

To quantify stability, the upper-tail absolute error (90th percentile) is examined. As shown in Table 3, the contextual hybrid model achieves lower tail error compared to the LSTM baseline and demonstrates comparable tail behavior to the classical hybrid. This indicates improved resistance to extreme prediction deviations under volatile market conditions.

Figure 3 presents the training and validation loss curves of the LSTM and CSET-CLX modules. The LSTM model shows a noticeable gap between training and validation loss, indicating sensitivity to noise and mild overfitting. In contrast, the CSET-CLX module exhibits more stable convergence with closely aligned loss curves, suggesting improved generalization capability.

4.3 Ablation study on model components

An ablation study is conducted to isolate the contributions of individual components within the hybrid framework. The results indicate that the largest performance improvement is achieved by incorporating XGBoost as a residual learner, validating its critical role in handling nonlinear regression and error correction.

Table 4 summarizes the performance of different model variants used in the ablation study. The results indicate that the inclusion of XGBoost as a residual learner provides the most significant improvement in predictive performance. This confirms the importance of ensemble-based regression in handling nonlinear relationships and correcting systematic errors generated by deep learning models.

Table 4. Ablation analysis of sequential, contextual, and boosting components

Model Variant	MAE	RMSE	R ²
LSTM	1578.37	1784.33	-3.15
LSTM + XGBoost	311.38	413.22	0.78
LSTM + CLX (raw) + XGBoost	474.64	605.35	0.52
LSTM + CLX (scaled) + XGBoost	459.01	586.89	0.55

Note: LSTM: Long Short-Term Memory; XGBoost: Extreme Gradient Boosting; CSET-CLX: contextual latent feature extraction; MAE: Mean Absolute Error, RMSE: Root Mean Square Error; R²: Coefficient of Determination.

The addition of the CSET-CLX module introduces contextual awareness across temporal windows. However, without proper feature scaling, the contextual latent features tend to dominate the fused feature space, leading to degraded performance.

After applying standardization, the contextual hybrid model achieves improved generalization compared to the unscaled version. Nevertheless, it remains slightly inferior to the classical hybrid model in terms of raw accuracy metrics.

These findings highlight that contextual feature integration requires careful balancing to avoid feature dominance and ensure effective interaction with tree-based ensemble models.

4.4 Contextual hybrid underperformance in raw accuracy

Although CSET-CLX improves prediction stability, it does not outperform the classical hybrid model in terms of raw accuracy. Several factors may explain this behavior.

First, feature dominance effects may arise when contextual embeddings introduce additional variance that competes with strong price-based signals. This can reduce the effectiveness of decision tree splits in XGBoost.

Second, attention-based contextual encoding may amplify short-term fluctuations that are weakly predictive of next-day prices, leading to increased noise sensitivity.

Third, the contextual hybrid introduces a bias-variance trade-off, where improved generalization and smoother error behavior come at the cost of reduced precision in trend-dominated conditions.

Furthermore, daily stock price movements are influenced by exogenous factors that are not fully captured within historical price windows. As a result, the benefits of contextual feature refinement are more evident in stabilizing predictions rather than improving point-wise accuracy.

4.5 Integrated discussion

The experimental results provide several key insights:

First, hybridization is essential for daily stock price forecasting, as residual learning significantly improves predictive performance compared to standalone sequential models.

Second, CSET-CLX enhances prediction stability by reducing extreme errors, even though it does not lead to

superior average accuracy.

Third, SHAP-based explainability analysis confirms that price-based technical indicators remain the dominant drivers of short-term predictions, while contextual features play a complementary role in stabilizing model behavior.

These findings suggest that different hybrid configurations serve distinct purposes. Classical hybrid models are more suitable for accuracy-driven scenarios, while contextual hybrid models provide additional value in improving robustness under volatile and noisy market conditions.

Overall, the results support the study's emphasis on balancing accuracy, stability, and interpretability, rather than prioritizing model complexity as the primary objective.

5. CONCLUSIONS

This study proposed an explainable hybrid forecasting framework that integrates LSTM-based temporal modeling, CSET-CLX, and XGBoost-based residual learning for daily stock price prediction. The results demonstrate that hybridization plays a critical role in improving predictive performance compared to standalone sequential models. In particular, the classical hybrid LSTM-XGBoost model achieved the best numerical accuracy, confirming the effectiveness of residual learning in capturing nonlinear relationships and correcting systematic prediction errors.

The contextual hybrid model incorporating CSET-CLX exhibited comparable directional performance while providing improved prediction stability by reducing extreme errors. These findings indicate that CSET-CLX contributes primarily to robustness and generalization rather than enhancing point-wise accuracy.

Overall, the results highlight an important trade-off between accuracy and stability in financial time-series forecasting. Classical hybrid models are more suitable for accuracy-driven scenarios, whereas contextual hybrid models provide additional value in stabilizing predictions under volatile and noisy market conditions.

The inclusion of SHAP-based explainability further supports transparent interpretation of model behavior, demonstrating that price-based technical features remain the dominant drivers of short-term predictions, while contextual features play a complementary role.

Despite these findings, several limitations should be acknowledged. The evaluation was conducted on a single stock and relied primarily on historical price-based data, which may limit generalizability across different assets and market conditions. In addition, the contextual refinement module adopts standard attention mechanisms without introducing architectural novelty, and its effectiveness may vary depending on data characteristics.

Future research may extend this framework by incorporating additional data sources, such as market sentiment and macroeconomic indicators, as well as exploring adaptive or regime-aware hybrid strategies. These directions may further enhance the balance between accuracy, stability, and interpretability in financial forecasting applications.

REFERENCES

[1] Chen, L., Chi, Y.G., Guan, Y.Y., Fan, J.L. (2019). A hybrid attention-based EMD-LSTM model for financial

time series prediction. In 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 113-118. <https://doi.org/10.1109/ICAIBD.2019.8837038>

[2] Xu, W., Chen, J.L., Xiao, J. (2024). A hybrid price forecasting model for the stock trading market based on AI technique. In 2024 Ninth International Conference on Informatics and Computing (ICIC), Medan, Indonesia, pp. 1-5. <https://doi.org/10.1109/ICIC64337.2024.10957173>

[3] Xiao, Z.Y., Li, S.M., Lin, Z.X. (2012). A Hybrid modeling approach for forecasting the volatility of REITs index in US market. In 2012 International Conference on Management Science & Engineering 19th Annual Conference Proceedings, Dallas, TX, USA, pp. 1861-1867. <https://doi.org/10.1109/ICMSE.2012.6414425>

[4] Heistrene, L., Belikov, J., Baimel, D., Katzir, L., Machlev, R., Levy, K. (2025). An improved and explainable electricity price forecasting model via SHAP-based error compensation approach. *IEEE Transactions on Artificial Intelligence*, 6(1): 159-168. <https://doi.org/10.1109/TAI.2024.3455313>

[5] Hossain, M.R., Ismail, M.T., Karim, S.A.B.A. (2021). Improving stock price prediction using combining forecasts methods. *IEEE Access*, 9: 132319-132328. <https://doi.org/10.1109/ACCESS.2021.3114809>

[6] Hossain, S., Kaur, G. (2024). Stock market prediction: XGBoost and LSTM comparative analysis. In 2024 3rd International Conference on Artificial Intelligence for Internet of Things, AIIoT 2024, Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/AIIoT58432.2024.10574794>

[7] Aravindhan, M., Manikavelan, D. (2025). Decomposition-based transformer-LSTM framework for stock index forecasting. In 2025 3rd World Conference on Communication & Computing (WCONF), Raipur, India, pp. 1-6. <https://doi.org/10.1109/WCONF64849.2025.11233358>

[8] Lv, C.K., Gao, B.Y., Yu, C. (2021). A hybrid transfer learning framework for stock price index forecasting. In 2021 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress, AB, Canada, pp. 355-359. <https://doi.org/10.1109/DASC-PICom-CBDCCom-CyberSciTech52372.2021.00067>

[9] Liu, Q., Hu, Y., Liu, H. (2024). Enhanced stock price prediction with optimized ensemble modeling using multi-source heterogeneous data: Integrating LSTM attention mechanism and multidimensional gray model. *Journal of Industrial Information Integration*, 42: 100711. <https://doi.org/10.1016/j.jii.2024.100711>

[10] He, Y., Li, J.M., Ruan, S.M., Zhao, S.P. (2020). A hybrid model for financial time series forecasting — Integration of EWT, ARIMA with the improved ABC optimized ELM. *IEEE Access*, 8: 84501-84518. <https://doi.org/10.1109/ACCESS.2020.2987547>

[11] Kumar, K.S., Sree, D.I., Devi, P.Y., Pujitha, M.V. (2024). Comparative analysis of LSTM and XGBoost models for short-term bitcoin price prediction. In 2024 3rd International Conference on Applied Artificial

- Intelligence and Computing (ICAAIC), Salem, India, pp. 932-939.
<https://doi.org/10.1109/ICAAIC60222.2024.10575848>
- [12] Ribeiro, M.T., Singh, S., Guestrin, C. (2016). Why should I trust you?' Explaining the predictions of any classifier. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations, San Diego, California, pp. 97-101. <https://doi.org/10.18653/v1/n16-3020>
- [13] Yuliyono, A.D., Girsang, A.S. (2019). Artificial bee colony-optimized LSTM for bitcoin price prediction. *Advances in Science, Technology and Engineering Systems Journal*, 4(5): 375-383. <https://doi.org/10.25046/aj040549>
- [14] Notaria, H., Shah, S., Thopte, D., Soneji, H., Bari, P., Deulkar, K. (2024). Comparative analysis of stock price prediction using time series models. In 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, pp. 1-6. <https://doi.org/10.1109/ICCUBEA61740.2024.10775112>
- [15] Aldana-Bobadilla, E., Sosa-Sosa, V.J., Molina-Villegas, A., Gazca-Hernandez, K., Olivas, J.A. (2024). An eclectic approach for enhancing language models through rich embedding features. *IEEE Access*, 12: 100921-100938. <https://doi.org/10.1109/ACCESS.2024.3422971>
- [16] Mrčela, L., Kostanjčar, Z. (2025). A hierarchical probabilistic deep learning approach for contextual anomaly detection in mixed-type tabular data. *IEEE Access*, 13: 173298-173311. <https://doi.org/10.1109/ACCESS.2025.3617799>
- [17] Gandhudi, M., P.J.A., A.J.A., Fiore, U., G.R. (2024). Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction. *Computers and Electrical Engineering*, 118(Part A): 109302. <https://doi.org/10.1016/j.compeleceng.2024.109302>
- [18] Gülmez, B. (2023). Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. *Expert Systems with Applications*, 227: 120346. <https://doi.org/10.1016/j.eswa.2023.120346>
- [19] Chen, X., Gupta, L. (2024). Training LSTMS with circular-shift epochs for accurate event forecasting in imbalanced time series. *Expert Systems with Applications*, 238(Part C): 121701. <https://doi.org/10.1016/j.eswa.2023.121701>