JOURNAL OF INNOVATION AND TECHNOLOGY
eISSN:2805-5179 Vol. 2025, Issue 2, No.11

Integrating Explainable AI, Ensemble Technique and Sentiment Analysis for Stock Market Forecasting

Vaishnavi M. Gokhale*, Aruna M.

Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka, India.

***Email:** gokhalevaishnavi56@gmail.com

Abstract

Financial forecasting that is both accurate and comprehensible, is crucial for stock market risk management. Many of the innovative approaches that have surfaced in recent years make use of deep learning, ensemble methods, and ensemble techniques that combine many approaches on financial data. Although these techniques frequently produce impressive results, they frequently function as "black boxes," making it challenging for people to examine how predictions are made because of a lack of transparency that breeds mistrust. This review examines studies that combine a variety of models. It makes things more readable and trustworthy by utilizing XAI, deep learning, and machine learning. LSTM, BiLSTM, CNN, XGBoost, ARIMA, and Prophet are a few of these. These are used to identify important trends in financial data and to monitor patterns over time. In addition to SHAP, LIME, and Layer-wise Relevance Propagation, there are more XAI technologies that can be useful. It displays the factors that assist models in making predictions. These justifications encourage more individuals to have faith in the outcome. The review discusses major issues, such as making models easy to comprehend, how well these models hold up in new marketplaces, and leveraging people's emotions when forecasting.

Keywords

Explainable AI, Stock Market Forecasting, Deep Learning, Hybrid Models, Interpretability.

Introduction

Markets are subject to sudden fluctuations. Prices fluctuate so rapidly that it is nearly impossible for anyone to forecast the future simply by looking at the charts. It is just impossible for human eyes to detect every minute movement or subtle pattern. This explains why machine learning has been so valuable in the modern era. Deep learning models like CNN, LSTM, and BiLSTM are

Submission: 29 September 2025; **Acceptance:** 24 October 2025; **Available online:** December 2025

now used by people to identify patterns that they might otherwise miss. In order to capture both short-term movements and long-term correlations in the data, some researchers even integrate different models. Investors can now forecast stocks more accurately and easily thanks to these contemporary technologies.

The impact of social media is another significant shift in today's industry. Information, thoughts, and reactions about financial events are continuously generated on platforms such as Reddit and Twitter. This deluge of text can be read by new algorithms to determine people's emotions, such as fear, excitement, confidence, or perplexity. The model gains a better understanding of why prices may rise or fall when it integrates these feelings with market data. This makes it much easier for investors to understand what is influencing the market at any given time. However, despite all these developments, there remains a significant obstacle. Models for deep learning are frequently "black boxes." In addition to wanting accurate forecasts, traders often want to know why the model arrived at a specific conclusion. Explainable AI (XAI) can help with this. People can examine the model and determine which elements—such as technical indicators, trading volume, or social sentiment had the most effects on the prediction by using tools like SHAP, LIME, and LRP. This degree of transparency fosters confidence and facilitates people's responsible reliance on these mechanisms.

According to recent study, more people are attempting to integrate explainable AI with deep learning to make predictions that are both precise and intelligible. For instance, Valensky and Mohaghegh (2023) emphasised the requirement for precise, uniform tools by contrasting SHAP, LIME, and counterfactual explanations. FinXABSA, developed by Ong, Prasad, and Verma (2023), divides financial language into straightforward categories to facilitate the understanding of sentiment and risk. Tanveer and Latif suggested a technique in 2024 that combines news emotions and numbers to enhance stock projections. In order to investigate how online discussions impact the market, Choi and Mezouar (2024) used Reddit posts. SHAP revealed which comments were more important. By creating systems that offer explanations at both a broad and specific level, Kalra and Mittal (2024) went one step further and increased transparency.

By 2024, explainable AI was being used in fields other than stock movement, such as financial dashboards, risk management, and cryptocurrency prediction. An XAI framework created for situations with stringent compliance was created by Singh, Thomas, and Banerjee (2024). In order to assist traders, Patel, Krishnan, and Roy (2024) created a dashboard that combines LSTM and Prophet models with straightforward visual explanations. According to Khan and Sharma (2024), projections that include SHAP or LIME explanations are more likely to be trusted by investors. Reddy, Agarwal, and Das (2024) developed a Bitcoin prediction algorithm that adapts to sentiment in real time. And most recently, by integrating attention mechanisms with SHAP, Sharma, Gupta, and Banerjee (2025) further enhanced interpretability, producing models that are not only precise but also transparent about their decision-making process.

Methodology

Machine learning (ML), deep learning (DL), hybrid ensembles, sentiment analysis, and Explainable AI (XAI) frameworks are all included in the methods. Despite variations in breadth, the majority employ a tiered system that includes evaluation, interpretability integration, predictive modelling, data collecting, and preprocessing. Figure 1 shows the system architecture and explained in detail below.

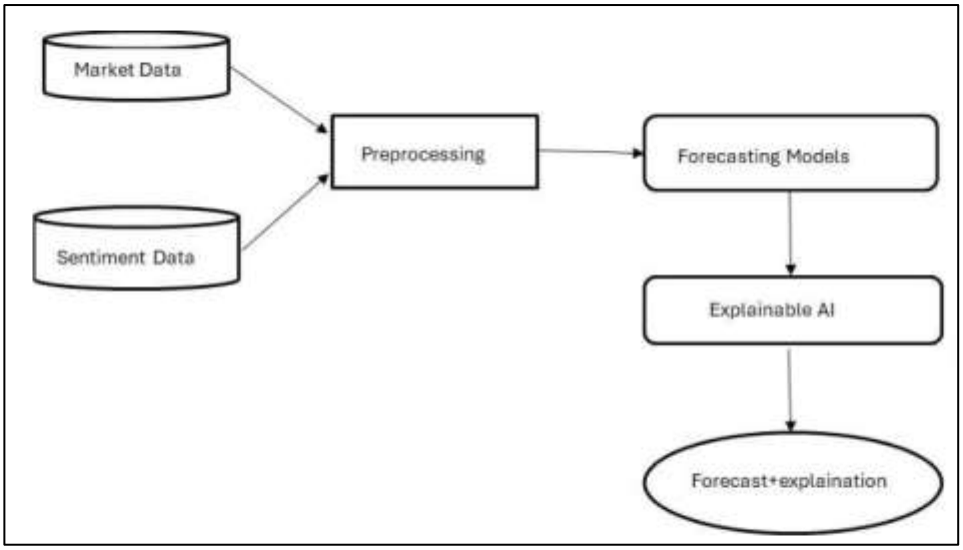

Fig. 1: System architecture

(i). Data Acquisition and Preprocessing

Data acquisition strategies varied across the studies:

- **Market indices and exchange data:** Global indices such as S&P500, FTSE100, DAX30, Nikkei225, and PSX were used for price/time-series prediction (Valensky et al., 2023; Tanveer & Latif, 2024; Kalra & Mittal, 2024).
- **Cryptocurrency data:** Bitcoin streaming data from Yahoo Finance was employed for dynamic forecasting (Reddy et al., 2024).
- **Social media sentiment:** Reddit forums (WallStreetBets, r/Stocks) and Twitter feeds were incorporated to capture investor psychology (Choi & Mezouar, 2024; Ong et al., 2023).
- **Feature engineering:** Technical indicators (moving averages, RSI, open interest, exchange rates) were extracted, while some studies applied feature selection via LASSO regression or Random Forest importance ranking (Kalra & Mittal, 2024; Ong et al., 2023).
- **Preprocessing:** Common techniques included normalization (Min-Max scaling), to handle outliers, tokenization for textual data, and bootstrapping for imbalance correction.

(ii). Predictive Modeling Approaches

The studies employed a variety of predictive models tailored to different financial tasks:

- **Classical ML models: '**Decision Trees', ' Random Forest', and Support Vector Machines were applied as benchmarks (Kalra & Mittal, 2024).

- **Ensemble methods:** XGBoost and bagging ensembles integrated with sentiment analysis consistently outperformed baselines (Choi & Mezouar, 2024).
- **Hybrid deep learning:** CNN-BiLSTM models were proposed for capturing both spatial and temporal dependencies in market data (Tanveer & Latif, 2024).
- **Stacked and hybrid architectures:** Prophet combined with LSTM in multi-layer dashboards enabled better trend capture (Patel et al., 2024).
- **Aspect-based NLP:** FinXABSA used Sentic GCN for aspect-level sentiment modeling linked to market drivers (Ong et al., 2023).
- **Comparative DL frameworks:** Studies tested CNN, LSTM, GRU, and DNNs for trend prediction and stability analysis (Valensky et al., 2023; Sharma et al., 2025).
- **Classical statistical forecasting:** ARIMA and Prophet were compared alongside DL models for Bitcoin forecasting (Reddy et al., 2024).

(iii). Integration of Explainability

Explainability was central to all works, ensuring trust and interpretability:

- **Post-hoc interpretability:** SHAP and LIME were the most widely used to rank feature importance and provide local/global explanations (Valensky et al., 2023; Sharma et al., 2025).
- **Layer-wise analysis:** LRP was employed in hybrid CNN-BiLSTM architectures to highlight the contribution of technical and macroeconomic indicators (Tanveer & Latif, 2024).
- **Aspect-level explanations:** FinXABSA provided token- and aspect-specific sentiment attribution, linking financial commentary directly to price shifts (Ong et al., 2023).
- **Dashboard-driven transparency:** Patel et al. (2024) developed interactive visualizations of LSTM/Prophet forecasts with SHAP/LIME overlays.
- **Governance frameworks:** Some papers proposed systematic XAI frameworks for risk management and regulatory compliance, emphasizing stability and fidelity of explanations (Singh et al., 2024).

(iv). Evaluation Metrics

Performance was evaluated through classification and regression metrics:

- **Classification metrics: '** Accuracy', ' Precision', ' Recall', and F1-score were used in trend classification tasks (Sharma et al., 2025; Tanveer & Latif, 2024).
- **Regression metrics:** RMSE, MAE, and R^2 were applied for price forecasting, particularly in Bitcoin and index predictions (Reddy et al., 2024; Patel et al., 2024).
- **Interpretability evaluation:** Fidelity, stability, and completeness of SHAP/LIME explanations were occasionally measured (Valensky et al., 2023).
- **Benchmark comparisons:** Traditional ML models (RF, SVM, LR) were used as baselines against deep learning or hybrid architectures (Kalra & Mittal, 2024; Sharma et al., 2025).

(v). Emerging Methodological Trends

The review highlights several methodological trends:

- **Hybrid and stacked modeling:** Combining DL with statistical methods (LSTM + Prophet, CNN-BiLSTM) shows growing adoption.
- **Sentiment-aware forecasting:** Integration of Reddit, Twitter, and ABSA pipelines reflects rising importance of investor psychology.

- **Dashboard-driven decision support:** Interactive XAI dashboards demonstrate practical applications in trading.
- **Risk management orientation:** Frameworks stress compliance, interpretability, and governance alongside predictive accuracy.
- **Multi-class forecasting:** Moving beyond binary up/down to multi-pattern trend classification improves realism in financial modelling.

Results and Discussion

More adaptable deep learning, ensemble approaches, and explainable AI are clearly replacing older, conventional models in financial forecasting, according to recent research. Because they are straightforward, clear, and offer solid baselines, traditional methods like ARIMA, Prophet, LASSO, and Decision Trees are still helpful (Reddy et al., 2024; Kalra & Mittal, 2024). However, hybrid and ensemble models have seized the lead as markets become more sentiment-driven and sophisticated. Combining numerical and emotional data results in considerably more accurate forecasts, according to studies utilizing CNN-BiLSTM networks (Tanveer & Latif, 2024), XGBoost with Reddit sentiment (Choi & Mezouar, 2024), and LSTM Prophet dashboards (Patel et al., 2024). Additionally, tools such as FinXABSA (Ong et al., 2023) show how separating financial text into specific sentiment categories might enhance forecasts in rapidly fluctuating markets.

Simultaneously, explainability has gained equal importance with accuracy. Regulators and investors want to know why a model predicts something, not just the outcome. These days, it's common practice to highlight which indicators or sentiment elements have the greatest influence on a model's decisions using techniques like SHAP, LIME, and LRP. Scholars have also highlighted how transparent models promote safer decision-making and foster trust (Singh et al., 2024; Valensky & Mohaghegh, 2023). According to Sharma et al. (2025), more sophisticated deep learning algorithms may even recognize several market patterns, including upward, downward, and rounded-top trends, while still providing an explanation for their actions. In general, modern financial AI is shifting towards hybrid models that give real-time, sentiment-aware insights that people can comprehend and trust by fusing accuracy and clarity.

Table 1: Comparison of Reviewed Studies (2023–2025)

Ref	Author(s), Year	Method(s) Used	Key Innovation / Focus
[1]	Tanveer & Latif, 2024	CNN-BiLSTM + LRP	Hybrid DL with feature-level explanations for PSX trends
[2]	Choi & Mezouar, 2024	XGBoost + Reddit Sentiment + SHAP	Ensemble learning with sentiment integration and interpretability
[3]	Valensky & Mohaghegh, 2023	LSTM, CNN, Decision Trees + SHAP, LIME	Cross-model comparison of transparency and explanation stability
[4]	Kalra & Mittal, 2024	LASSO + CDNN	Decision-support pipeline using feature selection along with

			interpretability
[5]	Ong et al., 2023	FinXABSA (Sentic GCN)	Aspect-based sentiment classification for financial forecasting
[6]	Singh et al., 2024	Integrated XAI Framework	Systematic XAI for financial risk management and compliance
[7]	Patel et al., 2024	Hybrid LSTM + Prophet + Dashboard	Dashboard-driven forecasting with SHAP/LIME explanations
[8]	Tanveer & Latif, 2024	Conv1D-BiLSTM + SHAP, LIME	Stability classification for PSX with interpretable pipeline
[9]	Reddy et al., 2024	LSTM, ARIMA, Prophet, XGBoost + Sentiment	Streaming Bitcoin forecasting with hybrid DL and statistical models
[10]	Sharma et al., 2025	Deep Neural Network + SHAP, LIME	Multi-class trend prediction (5 patterns) with explainable DL

Conclusion

Advanced models like LSTM, BiLSTM, CNN, and XGBoost offer improved predictive performance by capturing nonlinear dependencies and temporal dynamics in stock and cryptocurrency markets, but more conventional techniques like ARIMA, Random Forest, and SVM remain dependable baselines. By combining both long-term trends and short-term variations, hybrid frameworks like CNN BiLSTM pipelines and LSTM Prophet dashboards significantly increase predictive capabilities. Interpretability and decision assistance are becoming more and more important in the studies, in addition to pure prediction accuracy. By revealing important indicators like moving averages, trade volume, and sentiment polarity, explainable AI (XAI) techniques specifically, SHAP, LIME, and Layer-wise Relevance Propagation—have been used to close the gap between model outputs and financial reasoning. Aspect-based NLP models and sentiment analysis from Reddit are used in several works to broaden this focus and emphasize how market psychology influences price movements. Frameworks for dashboard-driven visualization and governance-focused XAI further demonstrate the shift from experimental models to useful instruments for making decisions.

The two foundations of contemporary financial forecasting are hybridization and explainability, as evidenced by common findings throughout the studied publications. While XAI improves trust, accountability, and regulatory compliance, deep learning guarantees flexibility in turbulent and complicated markets. Significantly, new developments that tackle the challenges of transparency, robustness, and deployment readiness include sentiment-aware forecasting, multi-class trend prediction, and real-time dashboard integration. In conclusion, the integration of explainability frameworks and ML/DL algorithms signifies a departure from forecasting that is focused on accuracy and a move towards transparent, intelligent, and adaptive financial decision-support systems that can help analysts, investors, and regulators navigate a more unpredictable market environment.

Acknowledgements

There is no grant or funding bodies to be acknowledged for preparing this paper

References

Choi, J., & Mezouar, Y. (2024**).** Enhancing equity trading through ensemble learning with Reddit sentiment analysis and explainable artificial intelligence. In *Proceedings of the IEEEInternational Conference on Big Data and Smart Computing (BigComp)* (pp. 1–6). https://doi.org/10.1109/ICCIMS61672.2024.10690490

Jagannathan, J., & Goditi, V. V. S. K. (2023). Integration of a hybrid model with XAI for stock price forecasting: A dashboard-driven approach. *IEEE Access*, *11*, 1–12. https://doi.org/10.1109/ITIKD63574.2025.11005353

Kalra, A., & Mittal, R. (2024). Explainable AI for improved financial decision support in trading. In *Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)* (pp. 1–6). https://doi.org/10.1109/ICRITO61523.2024.10522366

Khan, H., & Sharma, V. (2024). Interpretability in financial forecasting: The role of explainable AI in stock market prediction. In *Proceedings of the IEEE International Conference on Data Science and Machine Learning (ICDSML)* (pp. 1–6). https://doi.org/10.1109/ICSTE63875.2024.00039

Mallik, M. A., Reddi, S. K., Nirmala, G., Banoth, S., Reddy, P. C. S., & Sahni, N. (2023)**.** Stock market trend prediction using an explainable deep learning approach. *IEEE Access*, *11*, 1-12. https://doi.org/10.1109/ICKECS65700.2025.11035281

Parashiva Murthy, S., Gaur, S., Jolly, T., R, S., Sharma, G., & Rathore, R. (2023). Integratedexplainable AI for financial risk management: A systematic approach. *IEEE Access*, *11*, 1–12. https://doi.org/10.1109/IATMSI64286.2025.10984539

Reddy, S., Agarwal, P., & Das, K. (2024). Prediction of Bitcoin price through LSTM, ARIMA, XGBoost, Prophet, and sentiment analysis on dynamic streaming data. In *Proceedings of the International Conference on Emerging Technologies in Computing (iCETiC)* (pp.1-10). https://doi.org/10.1109/ICACCS57279.2023.10113014

Salaheldeen, K., Saad, A. A., Elagamy, M. N., & Fathy, C. (2023). A hybrid sentiment analysismodel for enhanced stock market movement predictions. *IEEE Access*, *11*, 1–12. https://doi.org/10.1109/ICEENG64546.2025.11031261

Thugu Rajesh Kumar, Reddy and Chitra, and Jeyarani, Periasamy (2024) *Online Product Evaluation System Based on Ratings and Review.* Journal of Data Science, 2024 (07). pp. 1-6. ISSN 2805-5160, https://doi.org/10.61453/jods.v2024no07

Valensky, D., & Mohaghegh, M. (2023). Evaluating transparency: A cross-model exploration of explainable AI in financial forecasting. *IEEE Access*, *11*, 1–12. https://doi.org/10.1109/CSDE59766.2023.10487732