Journal of Engineering and Technology Management 79 (2026)

Stock Market Prediction using AI Techniques: A Review

Moksh Tekam
Sagar Institute of Research & Technology, Bhopal, M.P., India

Vandana Dubey
Sagar Institute of Research & Technology, Bhopal, M.P., India

Mohit Singh Tomar
Sagar Institute of Research & Technology, Bhopal, M.P., India

***Abstract*— This review explores the application of Artificial Intelligence (AI) techniques in stock market prediction, emphasizing their growing importance in financial forecasting and decision-making. Traditional statistical methods often struggle to capture the nonlinear and volatile nature of financial markets, whereas AI models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests, and Deep Learning architectures like LSTM and CNN offer improved predictive accuracy by learning complex patterns from large datasets. The paper highlights various AI approaches, comparative performances, and the integration of sentiment analysis and hybrid models to enhance forecasting reliability. It also discusses current challenges, such as data quality, model interpretability, and market uncertainty, while identifying future research directions toward more robust and explainable AI-driven prediction systems.**

Keywords*— Stock Market Prediction, Artificial Intelligence, Machine Learning, Deep Learning, Financial Forecasting, LSTM Neural Networks.

I. INTRODUCTION

The stock market is a dynamic and complex system that plays a crucial role in the economic development of any country. It serves as a platform where investors buy and sell shares of publicly traded companies, influencing the overall flow of capital in the economy. Predicting stock market trends has always been a challenging task due to its inherent volatility, uncertainty, and dependency on a wide range of economic, political, psychological, and global factors. Despite these challenges, the ability to accurately forecast stock prices or market movements holds immense value for investors, financial institutions, and policymakers. Accurate predictions can lead to more informed investment decisions, better portfolio management, reduced risks, and optimized returns.

Traditional stock market analysis methods, such as fundamental analysis and technical analysis, have been widely used for decades. Fundamental analysis focuses on evaluating a company's intrinsic value by examining financial statements, market position, and macroeconomic conditions. On the other hand, technical analysis relies on studying historical price patterns, trading volumes, and statistical indicators to anticipate future movements. While both methods have provided valuable insights, they often fall short in addressing the high-dimensional, nonlinear, and rapidly changing nature of modern financial markets. As a result, the quest for more precise and adaptive prediction models has led to the emergence of Artificial Intelligence (AI) as a transformative approach in financial forecasting.

AI techniques, particularly those under the umbrella of Machine Learning (ML) and Deep Learning (DL), have demonstrated remarkable potential in uncovering hidden relationships within large and complex datasets. Unlike traditional statistical models that depend on pre-defined assumptions, AI systems can automatically learn from data, recognize intricate patterns, and make predictions with improved accuracy. Methods such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests, Decision Trees, and more recently, advanced architectures like Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), have shown exceptional promise in modeling temporal dependencies and nonlinear relationships in financial data. These techniques have enabled analysts to predict stock price movements, volatility trends, and even investor sentiment based on diverse data sources.

Figure 1: Stock Market

Moreover, the integration of AI with natural language processing (NLP) and sentiment analysis has further revolutionized stock market prediction. By analyzing news articles, social media trends, and financial reports, AI models can gauge public mood and assess how market sentiment affects price movements. This holistic approach allows for a more comprehensive understanding of the forces driving the stock market, thereby improving prediction reliability. In addition, hybrid models that combine multiple AI and statistical techniques have been proposed to enhance robustness and minimize prediction errors, especially in volatile market conditions.

Despite these advancements, stock market prediction remains a highly challenging endeavor. Factors such as sudden geopolitical events, unexpected policy changes, and behavioral biases can cause unpredictable fluctuations that even the most advanced AI models struggle to capture. Furthermore, the interpretability of AI-based predictions is a growing concern, as financial experts often require transparent models to justify investment decisions. Nevertheless, continuous improvements in AI algorithms, computational power, and access to real-time data are progressively addressing these limitations.

Stock market prediction has evolved from traditional analytical techniques to sophisticated AI-driven approaches capable of handling the complexity of financial systems. With ongoing research and innovation, AI is expected to play an increasingly vital role in building intelligent, adaptive, and explainable models for financial forecasting. The growing intersection of finance and artificial intelligence marks a new era of data-driven decision-making, offering promising opportunities for investors, analysts, and researchers in navigating the uncertain landscape of global stock markets.

II. LITERATURE SURVEY

C. Raj Kumar et al., [1] proposed SEMP-TA, a stacking ensemble-based machine learning framework designed to improve the accuracy of stock market trend prediction. Their model integrates multiple base learners, including Random Forest, Gradient Boosting, and Neural Networks, within a meta-learning architecture to effectively capture nonlinear relationships in financial data. The study emphasizes trend analysis by combining technical indicators and historical stock prices, achieving higher accuracy and stability compared to single-model approaches. The experimental results on benchmark datasets demonstrated that the ensemble model significantly reduced prediction variance and improved robustness under varying market conditions. This research highlights the importance of ensemble learning in addressing overfitting and enhancing generalization performance in financial forecasting tasks.

Y. Yang et al., [2] conducted an extensive survey on ensemble learning methodologies, focusing on their evolution and performance enhancement in the deep learning era. The study outlines how combining multiple models, such as bagging, boosting, and stacking, contributes to higher predictive reliability in complex domains like stock market forecasting. The authors provided a detailed taxonomy of ensemble architectures and discussed their applicability in financial time series analysis. Moreover, they emphasized the integration of ensemble frameworks with neural architectures such as CNNs and LSTMs for improved trend recognition. Their work provides a foundation for understanding the strengths and limitations of ensemble learning as a mechanism for achieving high-performance AI-driven stock market prediction.

Y. Soun et al., [3] developed a self-supervised learning model that utilizes sparse and noisy Twitter data to predict stock price movements with remarkable accuracy. By applying graph-based representation learning and transformer architectures, the model effectively captures the latent relationship between public sentiment and stock volatility. The research demonstrated that even with incomplete or unstructured text data, self-supervised models could outperform supervised baselines in short-term trend forecasting. The authors showed that incorporating real-time social media data provides valuable contextual insights that enhance market reaction prediction. This approach bridges the gap between financial modeling and sentiment analysis, paving the way for data-driven, emotion-aware trading strategies.

S. Mohapatra, et al., [4] investigated the predictive capabilities of ensemble machine learning algorithms in forecasting stock returns for Indian banking firms. Using technical indicators such as moving averages, momentum oscillators, and relative strength indices, the study implemented models like Random Forest, AdaBoost, and Gradient Boosting. Results showed that ensemble models consistently achieved superior accuracy and reduced error rates compared to single learners. The research also discussed the economic relevance of accurate stock return prediction for portfolio management and risk mitigation. Overall, this study highlighted how ensemble methods can enhance forecasting robustness, particularly in emerging financial markets like India.

E. Leung et al., [5] examined the promises and pitfalls of machine learning in predicting stock returns through an extensive empirical evaluation. Their research compared traditional regression-based techniques with modern AI approaches, including deep neural networks and boosted tree algorithms. The authors observed that while machine learning provides significant improvements in predictive accuracy, it is also prone to overfitting when trained on limited or biased datasets. The study stressed the importance of data preprocessing, feature selection, and cross-validation to maintain generalization ability. Furthermore, they discussed interpretability issues and proposed transparency frameworks for better trust and adoption of AI in financial modeling.

L. Prokhorenkova et al. [6] introduced CatBoost, a novel gradient boosting algorithm that efficiently handles categorical features without introducing bias. In stock market prediction, CatBoost demonstrated robust performance by managing complex nonlinear dependencies and high-dimensional financial datasets. The algorithm employs ordered boosting and minimal data leakage techniques, reducing overfitting and improving model reliability. The study illustrated that CatBoost outperformed traditional tree-based models such as XGBoost and LightGBM when applied to financial time series forecasting. Its ability to process categorical stock indicators without extensive data encoding makes it particularly valuable for real-world trading systems.

T. Chen et al., [7] presented XGBoost, a scalable tree boosting framework that has become a benchmark for predictive modeling, including in financial analytics. Their study focused on optimizing computational efficiency while maintaining high prediction accuracy, making it suitable for large-scale stock market datasets. XGBoost incorporates regularization, parallel computation, and sparsity-aware algorithms, which enhance both speed and performance. In financial prediction tasks, it has shown strong capability in handling heterogeneous data, detecting patterns in price movements, and managing missing values effectively. The framework's adaptability and scalability have made it a preferred choice for hybrid AI systems in financial forecasting.

Y. Zhang et al., [8] explored the integration of machine learning models with fundamental economic indicators to predict foreign exchange rates, offering valuable insights applicable to stock markets as well. Their approach combined regression-based learning with time-series analysis to capture the nonlinear dynamics of financial systems. The authors found that blending fundamental analysis with machine learning improved predictive power over traditional econometric models. Their results suggest that hybrid frameworks can effectively incorporate both quantitative and qualitative factors influencing market behavior. This study underscores the value of combining economic theory with data-driven intelligence for more comprehensive financial predictions.

I. K. Nti et al., [9] performed a comprehensive evaluation of ensemble learning techniques for stock market prediction using large-scale data from multiple exchanges. They analyzed the comparative performance of bagging, boosting, and stacking algorithms and concluded that ensemble-based models outperform individual classifiers across different financial contexts. The authors emphasized the importance of feature engineering and parameter optimization to enhance model performance. Additionally, they highlighted challenges in maintaining model generalization when exposed to market anomalies. The study established ensemble learning as a robust and adaptable strategy for financial forecasting in dynamic environments.

J. Patel et al., [10] investigated trend-based deterministic data preparation and machine learning techniques for predicting stock price movements. Their research utilized algorithms such as SVM, Random Forest, and Neural Networks, incorporating trend indicators to improve learning precision. The results indicated that models trained with well-structured feature engineering outperformed those using raw market data. They also stressed the significance of preprocessing, normalization, and data transformation in enhancing predictive stability. This study set a foundational framework for integrating technical and statistical perspectives in developing more accurate AI-driven stock forecasting systems.

Table 1: Summary of literature review

Sr. No.	Author with year	Work	Outcome
1	C. Raj Kumar (2025)	SEMP-TA: A Novel Stock Market Prediction Approach Based on Stacking Ensemble Machine Learning for Effective Trend Analysis	Achieved higher prediction accuracy and robustness against market volatility.
2	Y. Yang (2024)	A Survey on Ensemble Learning Under the Era of Deep Learning	Improved prediction reliability and performance in complex financial data.
3	Y. Soun (2022)	Accurate Stock Movement Prediction with Self-Supervised Learning from Sparse Noisy Tweets	Improved short-term prediction accuracy.
4	S. Mohapatra (2022)	Can Ensemble Machine Learning Methods Predict Stock Returns for Indian Banks Using Technical Indicators?	Accurate than single classifiers in financial return forecasting.
5	E. Leung (2021)	The Promises and Pitfalls of Machine Learning for Predicting Stock Returns	High potential but cautioned against overfitting, emphasizing the need for explainable AI in finance.
6	L. Prokhorenkova (2018)	CatBoost: Unbiased Boosting with Categorical Features	Improved performance and reduced overfitting in financial time-series modeling.
7	T. Chen (2016)	XGBoost: A Scalable Tree Boosting System	Widely used for large-scale financial prediction

			tasks due to its scalability and precision.
8	Y. Zhang (2020)	The Predictability of the Exchange Rate When Combining Machine Learning and Fundamental Models	Enhanced forecasting accuracy over traditional econometric methods.
9	I. K. Nti (2020)	A Comprehensive Evaluation of Ensemble Learning for Stock-Market Prediction	Boosting outperform individual models in diverse market conditions.
10	J. Patel (2015)	Predicting Stock and Stock Price Index Movement Using Trend Deterministic Data Preparation and Machine Learning Techniques	Demonstrated that well-structured feature engineering significantly boosts prediction performance.

III. CHALLENGES

Challenges in Stock Market Prediction Using AI Techniques-

1. **Market Volatility and Uncertainty:** Stock markets are highly dynamic and influenced by numerous unpredictable factors such as political events, economic shifts, and investor sentiment. This volatility makes it difficult for AI models to maintain consistent predictive accuracy across different market conditions.

2. **Data Quality and Noise:** Financial datasets often contain missing values, noise, or inaccurate entries. Poor data quality can significantly affect model training, leading to biased or misleading predictions. Data cleaning and preprocessing remain major challenges for reliable AI forecasting.

3. **Non-Stationary Nature of Data:** Stock market data is non-stationary, meaning statistical properties like mean and variance change over time. This causes AI models trained on past data to underperform when market patterns shift, reducing long-term generalization.

4. **Overfitting and Generalization Issues:** AI models, particularly deep learning architectures, are prone to overfitting—performing well on training data but failing on unseen test data. Overfitting leads to unreliable predictions in real-world market scenarios.

5. **Interpretability and Transparency:** Many AI techniques, such as deep neural networks, operate as “black boxes,” offering limited explanation for their predictions. Lack of interpretability hinders trust and adoption by financial analysts and investors who require transparent decision-making processes.

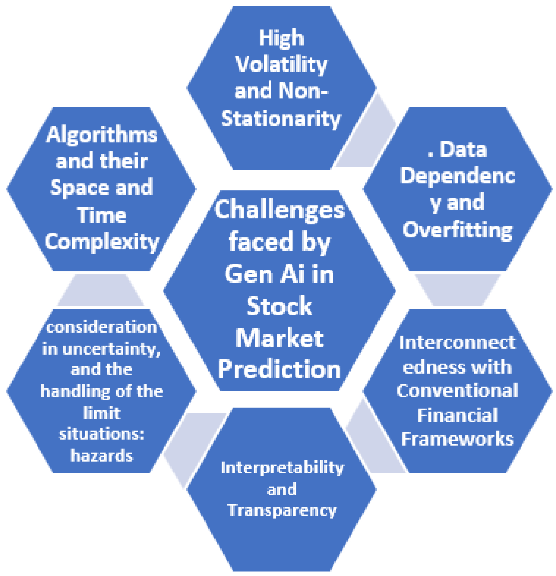

Figure 2: Challenges

6. **Integration of Multi-Source Data:** Combining heterogeneous data sources—like historical prices, company reports, social media sentiment, and macroeconomic indicators—is complex. Synchronizing and weighting such diverse inputs remains a major challenge in developing holistic AI models.

7. **Real-Time Prediction and Computational Complexity:** Stock trading requires near-instant decision-making, but AI models—especially deep learning systems—often demand high computational power and time for processing. Achieving real-time prediction with low latency remains difficult.

8. **Ethical and Regulatory Concerns:** The use of AI in stock prediction raises concerns about algorithmic bias, insider information misuse, and regulatory compliance. Ensuring fairness, accountability, and transparency in AI-driven financial systems is essential but challenging to implement effectively.

IV. PROPOSED PLAN

Proposed Plan for Stock Market Prediction Using AI Techniques

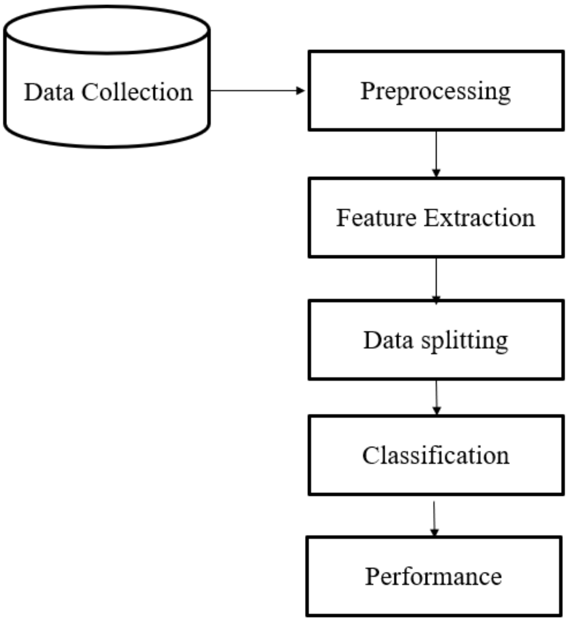

Figure 3: Proposed Plan

The proposed plan focuses on developing an intelligent, data-driven framework for accurate and reliable stock market prediction using advanced Artificial Intelligence (AI) techniques. The plan involves a structured approach divided into several stages to ensure systematic data handling, model optimization, and performance evaluation.

1. **Problem Definition and Objective Setting:** Clearly define the scope of the study, such as predicting stock prices, price movement direction (up/down), or market trend analysis. Establish measurable objectives—like achieving a specific prediction accuracy or minimizing forecasting error.
2. **Data Collection:** Gather relevant datasets from reliable sources, including historical stock prices, trading volumes, financial news, economic indicators, and social media sentiment. Use platforms such as Yahoo Finance, Bloomberg, and Twitter APIs to ensure comprehensive data coverage.
3. **Data Preprocessing and Feature Engineering:** Clean the collected data by handling missing values, removing noise, and normalizing numerical features. Extract meaningful technical indicators such as Moving Averages (MA), Relative Strength Index (RSI), and MACD. Additionally, use sentiment analysis techniques on textual data to include market psychology as an input feature.
4. **Model Selection and Design:** Choose suitable AI models based on the nature of data and prediction goals. Machine Learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting can be combined with Deep Learning models like LSTM and CNN for sequential and pattern-based prediction. A hybrid ensemble model can be developed to leverage the strengths of multiple approaches.
5. **Model Training and Optimization:** Train the selected models on historical data using supervised and semi-supervised learning techniques. Optimize hyperparameters using methods like Grid Search and Bayesian Optimization to enhance model accuracy and generalization. Use cross-validation techniques to avoid overfitting and ensure robustness.
6. **Performance Evaluation:** Evaluate the models using key performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), accuracy, and F1-score. Compare results across different AI models and hybrid frameworks to identify the most effective approach.

V. CONCLUSION

The stock market prediction using Artificial Intelligence represents a transformative approach to understanding and forecasting complex financial behaviors. By leveraging the capabilities of machine learning and deep learning techniques, investors and researchers can uncover hidden patterns, model nonlinear relationships, and make data-driven decisions with greater confidence. Although challenges such as data volatility, overfitting, interpretability, and integration of diverse data sources persist, continuous advancements in AI algorithms and computational power are steadily addressing these limitations. The integration of sentiment analysis, ensemble learning, and real-time adaptive models has further enhanced prediction reliability. Ultimately, the adoption of AI-based frameworks in financial forecasting promises a more intelligent, efficient, and responsive system for navigating the uncertainties of global stock markets.

REFERENCES

1. C. Raj Kumar and S. Manikandan, "SEMP-TA: A Novel Stock Market Prediction Approach Based on Stacking Ensemble Machine Learning for Effective Trend Analysis," in IEEE Access, vol. 13, pp. 117479-117490, 2025, doi: 10.1109/ACCESS.2025.3586233.

2. Y. Yang, H. Lv, and N. Chen, "A survey on ensemble learning under the era of deep learning," Artif. Intell. Rev., vol. 56, no. 6, pp. 5545–5589, Jun. 2024, doi: 10.1007/s10462-022-10283-5.

3. Y. Soun, J. Yoo, M. Cho, J. Jeon, and U. Kang, "Accurate stock movement prediction with self-supervised learning from sparse noisy tweets," in Proc. IEEE Int. Conf. Big Data (Big Data), Osaka, Japan, Dec. 2022, pp. 1691–1700, doi: 10.1109/BIGDATA55660.2022.10020720.

4. S. Mohapatra, R. Mukherjee, A. Roy, A. Sengupta, and A. Puniyani, "Can ensemble machine learning methods predict stock returns for Indian banks using technical indicators?," J. Risk Financial Manage., vol. 15, no. 8, p. 350, Aug. 2022, doi: 10.3390/jrfm15080350.

5. E. Leung, H. Lohre, D. Mischlich, Y. Shea, and M. Stroh, "The promises and pitfalls of machine learning for predicting stock returns," J. Financial Data Sci., vol. 3, no. 2, pp. 21–50, Apr. 2021. [Online]. Available: https://www.PM-research.com/content/iijjfds/3/2/21.

6. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 31, Dec. 2018, pp. 6639–6649.

7. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

8. Y. Zhang and S. Hamori, "The predictability of the exchange rate when combining machine learning and fundamental models," *J. Risk Financial Manage.*, vol. 13, no. 3, p. 48, Mar. 2020, doi: 10.3390/jrfm13030048.

9. I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A comprehensive evaluation of ensemble learning for stock-market prediction," *J. Big Data*, vol. 7, no. 1, p. 20, Dec. 2020, doi: 10.1186/s40537-020-00299.

10. J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," Expert Syst. Appl., vol. 42, no. 1, pp. 259–268, Jan. 2015, doi: 10.1016/j.eswa.2014.07.040.

11. https://www.kaggle.com/competitions/kaggle-stock-market-prediction/data.