International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 2, Mar-Apr 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

SafeGrowth: AI-Powered Portfolio Management for Smarter Investment Decisions

Sujal Ambatkar*, Omkar Survase**, Ankit Nair***, Mohammed Razeen****, Jaymala Chavan*****

*(Department of Computer Engineering, Pillai College of Engineering, Panvel.
Email: ambatkar22comp@student.mes.ac.in)
**(Department of Computer Engineering, Pillai College of Engineering, Panvel.
Email: osurve22comp@student.mes.ac.in)
***(Department of Computer Engineering, Pillai College of Engineering, Panvel.
Email: ankit22comp@student.mes.ac.in)
****(Department of Computer Engineering, Pillai College of Engineering, Panvel.
Email: morazeen23comp@student.mes.ac.in)
*****(Department of Computer Engineering, Pillai College of Engineering, Panvel.
Email: jaymalachavan@mes.ac.in)

--************************----------------------------------

Abstract:

SafeGrowth framework uses artificial intelligence in assisting investment decisions in a data-driven portfolio optimization environment. Unlike other methods, which are mostly based on historical prices are often unable to adapt well to changing market conditions, the proposed framework combines time-series forecasting with sentiment analysis of financial news articles. In the proposed framework, a Bi-LSTM model is used for predicting stock prices, a FinBERT-based sentiment analysis module is used for sentiment analysis of financial news articles. This approach helps in understanding the market behaviour by combining quantitative as well as qualitative analysis. The Experimental results show a significant improvement in performance over the baseline LSTM model. The prediction error is improved by 27.0%, from 312.47 to 144.19, and the Sharpe ratio is improved by 122.0%, from 0.556 to 1.237. In order to increase the robustness of the proposed framework, Conditional Value at Risk (CVaR) is used for risk-based portfolio optimization, which takes into account the risk in a highly uncertain market environment. Additionally, a chatbot-based AI model is also used for better user interaction, which can help in better understanding and decision-making.

Keywords — ***Portfolio Optimization, Bi-LSTM, Sentiment Analysis, FinBERT, Stock Price Prediction, Financial News Analysis, Conditional Value at Risk, Artificial Intelligence in Finance.***

--************************----------------------------------

I. INTRODUCTION

The global financial system is highly volatile because of economic, geopolitical issues, and weak handling of digital information. Modern Portfolio Theory (MPT) and the Capital Asset Pricing Model (CAPM) are based on certain principles of market efficiency and rational behaviour of the investor. However, the theories mentioned above are based in impractical assumptions which are not applicable during the times of economic crisis.

Another even more important drawback of existing investment systems is that they tend to ignore the ability to integrate the quantitatively expressed financial data and qualitative information like financial news. Time series models only look at past prices, so they miss out on other things happening in the market. Meanwhile, models built around sentiment aim to capture these outside influences but rarely take center stage in forecasting frameworks.

Recent advances in deep learning and natural language processing (NLP) have been maturing so that there are opportunities to apply deep neural networks to enhance financial forecasting models. Newer models connect time-series patterns with structures like Bi-LSTM, rather than using separate parts to manage each task. In addition, risk-aware optimization methods, such as CVaR, were developed to account for the influences of extreme risks. Still, current approaches tend to keep forecasting, emotion reading, and safety measures apart. So far, nobody has built a unified system capable of live decisions combining all three pieces together.

Instead of treating forecasting, sentiment analysis, and portfolio optimization separately, SafeGrowth merges them into one system. The SafeGrowth framework, however, combines Bi-LSTM based timeseries predictions, Fin BERT based sentiment, and CVaR based risk management into a single, real-time platform. The framework incorporates a streaming data pipeline and a conversational interface based on AI, leading to fast predictions as well as more interpretable explanations. The ability to incorporate multi-modal data fusion and user-centric explanations into an efficient, real-time platform distinguishes this approach from its predecessors in AI-guided portfolio management.

The major contributions of this work are summarized as follows:

- Development of a hybrid framework combining Bi-LSTM and FinBERT for improved stock prediction.
- Integration of financial news-based sentiment analysis with numerical time-series data.
- Application of CVaR for risk-aware portfolio optimization.
- Design of an AI-powered conversational chatbot for explainable insights, predictions, and real-time investment alerts.

II. RELATED WORK

The state-of-the-art advancements of Artificial Intelligence has massively enhanced financial forecasting system and portfolios management system. Time-series modelling via deep learning models is one of the research interests among many. Long Short Term Memory (LSTM) and Bidirectional Long Short Term Memory (Bi-LSTM) are both prevalent approaches to forecast stock price [2], [3], [10], since both of them are robust in learning the dependence between adjacent financial time-series. Previous existing researches show that Bi-LSTM is better than other statistical methods such as ARIMA and GARCH in modelling non-linear and variance market behaviour [3]. These models only consider comparing historic prices without considering external market-based effects.

In order to address this issue, NLP models have been used in event-driven prediction systems and extracting sentiment from text data. State-of-the-art NLP models that are domain-specific (e.g., FinBERT) have proven effective in classifying financial news and regressing market sentiment scores [1], [11]. Research revealed that the combination of sentiment signals with the numerical time-series data increased the prediction accuracy and added a new dimension of market interpretation. Nevertheless, certain issues such as sparsity, domain adaptation and the noise of textual data still limit the model performances.

Apart from prediction, risk management is also a fundamental task for the portfolio optimization problem. Risk measures such as variance do not always give good small and large loss projections in down market, new measures of risk have also been suggested, for example, the Conditional Value at Risk (CVaR), which can more accurately measure the tail risk of large losses, helps prevent portfolios from fragile conditions under uncertainty [8]. Research on CVaR has demonstrated the successful use of it in risk minimization, but a real interesting research topic is combining CVaR with the AI prediction techniques.

While there have been these advances, none have combined all three that we implement in our system: a framework for time-series prediction; sentiment analysis of financial news; and risk sensitive portfolio optimization. One aspect that has not been focused on enough is enhancing user friendliness and human interaction and understanding.

In order to address the above failures, the integrated approach SafeGrowth integrates Bidirectional LSTM price prediction, financial news sentiment analysis with FinBERT and portfolio optimization based CVaR within a single system. And even a new AI-based chatting platform system is implemented to improve users' understanding.

III. EXISTING SYSTEM

There are typically three main layers that make up the current framework of the portfolio management systems. These include the Input Layer, the Processing Layer and the Output Layer.

Input Layer: In this module, the stock time-series data (Including historical data and real-time data of the stock price) as well as financial news is collected. These data sources are providing necessary quantitative as well as qualitative data needed for a complete prediction of the different factors and events influencing the stock market. But, algorithms existing today are either depending only on the numerical data or are not including the news content for prediction purpose.

Processing Layer: In this layer, we would be applying different machine learning and deep learning based algorithms for analysis of data. For stock price prediction, time-series forecasting models such as Long Short Term Memory (LSTM) based neural networks provide prediction of stock changes based on its financial information over time. For extracting financial news sentiment, NLP methods such as BERT Base Model are used. However, these approaches are still independent of one another and there is very less interaction between the two representations.

In addition, most systems of that time have not still incorporated sound risk management schemes. Notably, traditional methods tend to measure variance, which cannot address market extreme risks and tails. The unreliability of portfolio optimization scheme under uncertain and volatile markets arises from such lack of risk scrutiny techniques.

The Output Layer enables the outputted results visually via visualization dashboards, such as predicted prices, sentiment trends, and some findings. While these dashboards display important information, they are often not very interactive to facilitate the user‘s understanding of the model results.

However, there still remains many limitations of current systems, such as heavy computation of deep learning, the lag of implementing real-time data, no unified framework that combines prediction, sentiment analysis and risk optimization, and limited user interaction that makes the system less accessible to the public to guide practical trading.

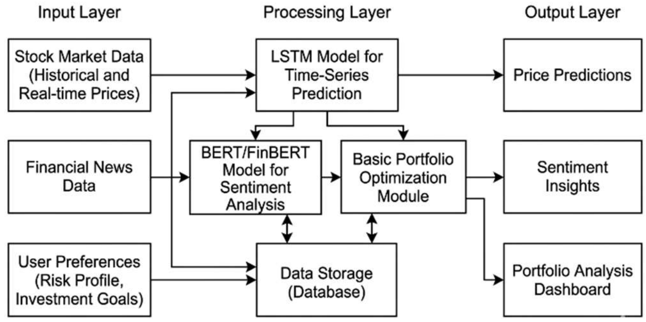

Fig. 1. Existing System Architecture

IV. PROPOSED SYSTEM

The proposed SafeGrowth architecture is a portfolio management system that uses real-time data and tries to fix some of the problems with current architectural methods. This framework combines time-series forecasting, sentiment analysis based on news, and risk-aware portfolio optimization to make decisions more accurate, improve performance, and make the system more scalable.

A. Data Sources Layer

The Data Sources segment gets both quantitative and qualitative financial data. The Twelve API provided the study with both historical and real-time stock market data. This information is the basis for analysis. This information lets the system pick up on both quantitative market trends and signals based on people's feelings. User-defined preferences, such as risk tolerance and investment goals, are included to make the portfolio recommendations more specific to the user.

B. Data Pipeline and Feature Engineering

Data Pipeline is in charge of getting data, cleaning it up, and changing it. To keep latency low and real-time responsiveness, a streaming processing method was chosen. The pipeline gets data from the outside world, makes it all the same, and then gets it ready for the model to use.

A centralized Feature Store is used to make sure that the training and inference processes are always the same. This method makes sure that all machine learning models always use the same feature sets, which makes the system more efficient, scalable, and reliable.

C. AI Engine

The AI Engine is the most important part of the SafeGrowth system. It brings together different smart parts that can make predictions, analyze feelings, and manage risks.

A Bidirectional Long Short Term Memory (Bi-LSTM) model is used to predict stock prices. The main difference between Bi-LSTM models and regular LSTM models is that Bi-LSTM models can capture both past and future temporal dependencies. This skill makes it easier to predict what will happen in financial markets that are known for being unstable.

A model based on FinBERT looks at financial news to figure out what people are feeling and what signals they are sending. This lets the system use qualitative data, which changes how markets work. This is an important link to understanding these systems.

Combining the outputs of the Bi-LSTM model and FinBERT gives you a better idea of how the market is doing. We use a Conditional Value at Risk (CVaR)-based method to make sure that portfolio allocation lowers losses, especially when the market is doing poorly. All of this minimizes risk.

D. Output Layer

The Output Layer shows the processed results in a clear way. A portfolio analytics dashboard shows stock predictions, sentiment analysis, and how to divide up your portfolio. To make things better for users, an AI Chatbot connects to the system. The Chatbot helps people understand model outputs and answers questions directly to help them make smart investment choices. This makes the system easier to use, which helps people understand complicated analytical results.

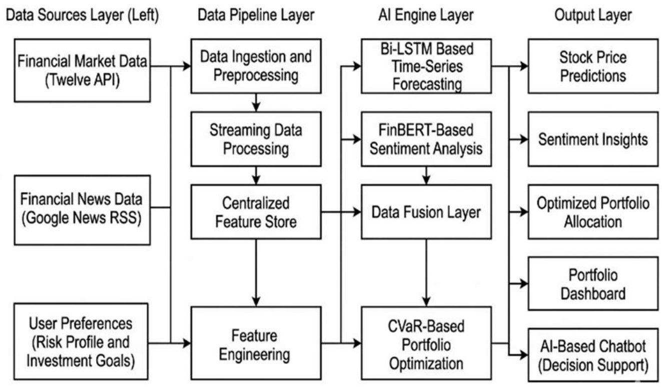

Fig. 2. Proposed System Architecture

V. IMPLEMENTATION

The SafeGrowth system functions as a modular and scalable portfolio management framework. It incorporates data processing, prediction models, sentiment analysis, and risk optimization strategies. The system was developed using Python 3.9, employing TensorFlow and PyTorch for deep learning applications, and FastAPI for deploying RESTful services in various contexts.

A. Data Collection and Preprocessing

Financial data is collected using the Twelve Data API, providing current and past stock information, including Open, High, Low, Close, and Volume features. Financial news is obtained from Google News RSS feeds. Missing data points were addressed through forward filling. The numerical features underwent normalization using a Min-Max scaling approach. Time-series sequences are produced using a sliding window approach, which helps in the generation process. This method allows for a methodical examination of data, making sure each resulting series retains its core features. The application of this technique contributes notably to the coherent structuring of temporal data and permits a more detailed analysis of patterns that might otherwise be overlooked. This approach ensures precision in the formation of each time-series segment. Textual data undergoes tokenization, and these tokens are then converted into numerical embeddings for sentiment analysis.

B. Prediction Model (Bi-LSTM)

Bi-directional Long Short-Term Memory networks are used to account for temporal dependencies [10], considering both past and future contexts within a sequence. The model underwent training with Mean Squared Error (MSE) loss, with its parameters refined through the Adam optimizer.

C. Sentiment Analysis (FinBERT)

This research uses a FinBERT-based model [11] to classify financial news as positive, negative, or neutral sentiments. These sentiment scores are another form of data that enters the system, which contributes to better predictions of outputs.

D. Risk-Aware Portfolio Optimization (CVaR)

Portfolio optimization techniques frequently involve Conditional Value at Risk (CVaR) to reduce extreme losses by addressing tail risk. The computation of asset allocation aims to achieve equilibrium between anticipated returns and potential adverse risk.

E. System Interface and Decision Support

The architecture provides RESTful Application Programming Interfaces for both prediction generation and the allocation of portfolios. The output data are presented on an interactive dashboard. An AI-powered chatbot allows users to make sense of predictions and secure actionable insights.

F. Real-Time Processing Capability

The system functions with near real-time operation, experiencing an average response time of about one to two seconds for each request. Efficient data preprocessing, streamlined model inference, and a streaming architectural design lead to this low-latency performance in the system. Such responsiveness helps keep portfolio recommendations current and appropriate within changing financial markets.

VI. RESULTS AND EVALUATION

A. Dataset Description

The SafeGrowth system underwent an experimental evaluation using historical daily market data. This data pertains to the NIFTY 50 index and was sourced from the National Stock Exchange of India (NSE). The dataset spans a one-year period, from March 2025 to March 2026. This time frame was supplemented by the collection of additional historical data to ensure an adequate number of samples for model training. Each record contains key financial characteristics such as Open, High, Low, Close, trading volume, and turnover. Sentiment signals were extracted from financial news data from the same period for analysis.

B. Data Preprocessing

The collected data underwent preprocessing to ensure quality and maintain consistency. To address the absence of data, a forward filling method was applied. Subsequently, a set of technical indicators, including moving averages and features based on momentum, were calculated. All numerical features underwent Min-Max scaling. This method allows for improved model convergence. The dataset was partitioned into training and testing sets, with 80% allocated for training and 20% for testing. The temporal ordering of the data was preserved during this process. A sliding window, consisting of 60 time steps, was employed to create the input sequences for the predictive model. Sentiment scores from financial news were incorporated with numerical features to provide additional contextual information.

C. Model Architecture and Training

This model, built upon a Bidirectional Long Short-Term Memory (Bi-LSTM) network, aims to understand the temporal dependencies inherent in financial time-series data. The architecture employs stacked Bi-LSTM layers, which feed into fully connected layers designed for regression output. This structural approach allows for processing sequential data and subsequently generating a continuous prediction. The model underwent training using the Adam optimizer, employing Mean Squared Error (MSE) as the loss function during this process. A baseline LSTM model was implemented under similar conditions for comparative analysis.

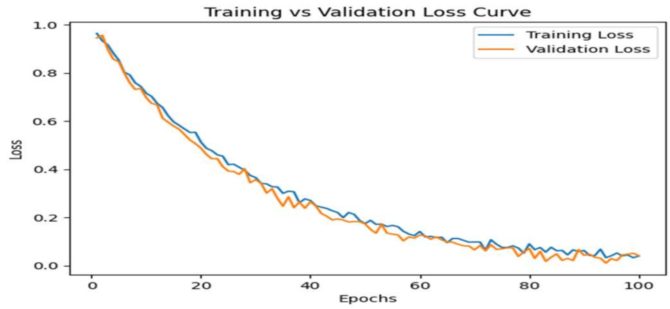

Fig. 3. Training and validation loss curves demonstrating model convergence and generalization performance.

D. Evaluation Metrics

The performance of the model was evaluated with various standard regression metrics. These included the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). This analytical strategy facilitated an in-depth analysis of its predictive accuracy in relation to this case study. Directional accuracy acted as a measure to assess the model's ability to predict the path of price changes.

E. Comparative Performance Analysis

Table I. presents the comparative performance of the baseline LSTM, Bi-LSTM, and the proposed SafeGrowth model.

TABLE I
COMPARATIVE PERFORMANCE OF PREDICTION MODELS

Model	RMSE	MAE	MAPE (%)	R^2
LSTM	312.47	241.83	1.24	0.8612
Bi-LSTM	198.64	153.27	0.79	0.9231
SafeGrowth	144.19	112.56	0.57	0.9587

The Bi-LSTM model demonstrates a noticeable enhancement compared to the baseline LSTM, probably due to its ability to analyze temporal relationships in both directions in financial time-series data. This methodological advance seems to provide a more thorough understanding than previous approaches. The SafeGrowth model provided the best performance across all evaluation metrics, showing a 27% reduction in RMSE versus the baseline model. Integrating sentiment analysis enhances predictive capability by incorporating external market signals derived from financial news. This enables the model to adapt to abrupt shifts in market conditions driven by investor sentiment, thus enhancing predictive accuracy and generalizability.

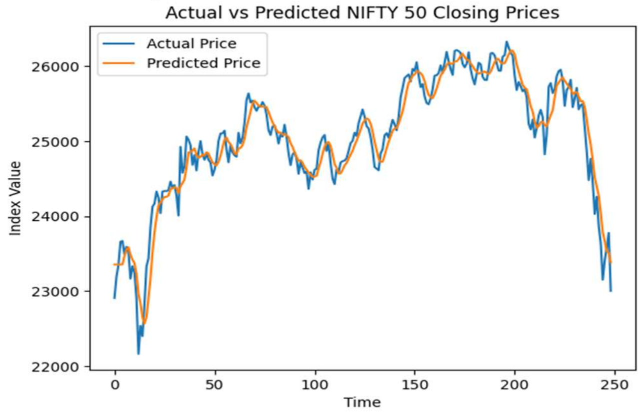

Fig. 4. Comparison of actual and predicted NIFTY 50 closing prices using the proposed Bi-LSTM-based SafeGrowth model.

F. Portfolio Optimization Results

The proposed system's effectiveness, based on CVaR optimization, was accessed by comparing it to a traditional equal-weight portfolio strategy.

TABLE III
PORTFOLIO OPTIMIZATION RESULTS

Metric	Equal Weight	SafeGrowth	Improvement
Annual Return (%)	10.4	16.2	+5.8
Volatility (%)	18.7	13.1	-5.6
Sharpe Ratio	0.556	1.237	+122%
Max Drawdown (%)	-22.3	-12.8	Improved
CVaR (95%)	-2.14	-1.37	Reduced

The results indicate that SafeGrowth improves returns while significantly reducing risk.

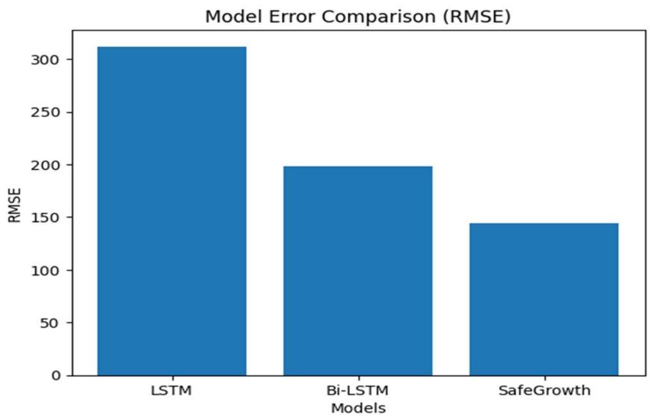

Fig. 5. Portfolio performance comparison between Equal Weight and SafeGrowth strategies showing improved return and reduced risk.

G. Discussion

The experimental analysis illustrates the effectiveness of the SafeGrowth framework in enhancing prediction accuracy and portfolio outcomes. A key factor that allows the Bi-LSTM model to address the limitations of a regular LSTM is its capacity to recognize patterns in data from both earlier and later contexts within a sequence. This ability provides it with a performance edge, especially when the order of inputs is critical. By including sentiment data, we improve predictions, particularly during periods of high market volatility. CVaR integration into portfolio optimization improves risk management by minimizing extreme losses, which is a major advantage. The system achieves greater returns with managed risk, as confirmed by a better Sharpe ratio and smaller drawdown. The proposed system shows substantial potential in intelligent portfolio management. Future work will involve an extended evaluation across longer time horizons beyond the current one-year test period. The directional accuracy of the SafeGrowth model further validates its practical applicability in real-world trading scenarios.

VII. CONCLUSIONS AND FUTURE DIRECTIONS

This paper introduces SafeGrowth, a portfolio management framework. It combines time-series forecasting, sentiment analysis, and risk-aware optimization to aid in sound investment decisions. The proposed system integrates a Bidirectional Long Short-Term Memory (Bi-LSTM) model to assess temporal aspects within financial data. This method works alongside sentiment signals gathered from financial news, aiming for improved prediction accuracy. Including Conditional Value at Risk (CVaR) allows for the reduction of potential losses during times of market uncertainty. This helps in managing risk.

The experimental data indicated that the suggested method yielded a decrease in prediction error and enhanced risk-adjusted returns when compared to established baseline models. The findings provided insights into the method's comparative performance. The research underscores the importance of this method within academic circles. Incorporating sentiment analysis alongside numerical financial data offers a more complete insight into market dynamics. This analytical approach, coupled with a Conditional Value-at-Risk (CVaR)-based framework for portfolio selection, aims to provide an approach to balancing returns with potential risks. This strategy helps in making decisions about risk and return trade-offs. The framework provides a structured approach for considering these factors. These results demonstrate that SafeGrowth is a practical and scalable approach for contemporary portfolio management.

Despite its efficacy, the empirical basis for this current investigation exhibits certain limitations. The assessment is conducted within a restricted timeframe. The system does not explicitly consider transaction costs or market impact. A key limitation of this analysis is that reliance on a single market index may restrict the ability to generalize findings to diverse asset classes and global financial landscapes.

The next phase of this research will involve extending the current framework to include multiple asset portfolios. This development will incorporate additional data sources, such as macroeconomic indicators, to improve the general applicability of the models. A critical aspect of this future work will be to enhance model robustness by exploring advanced architectural designs. Future research endeavours will explore the implementation of real-time deployment strategies. This will involve examining the incorporation of transaction cost modelling to more accurately reflect practical constraints. We will also investigate enhanced explainability techniques. The goal is that these refinements will contribute to greater user trust and a clearer interpretability of decisions made by AI systems.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Prof. Jaymala Chavan for their invaluable guidance, continuous support, and insightful suggestions throughout the course of this research. The authors also extend their appreciation to Dr. Sharvari Govilkar, Head of the Department of Computer Engineering, for their encouragement and support.

Furthermore, the authors are thankful to Dr. Sandeep Joshi, Principal of Pillai College of Engineering, for providing the necessary facilities and a conducive environment for carrying out this work. The support and resources provided by the institution played a vital role in the successful completion of this research.

REFERENCES

[1] E. Cicekyurt and G. Bakal, “Enhancing sentiment analysis in stock market tweets through BERT-based knowledge transfer,” Computational Economics, 2025. doi: 10.1007/s10614-025-10901-8.

[2] T. Jayanth, A. Manimaran, and G. Siva, “Enhancing stock price forecasting with a hybrid SES-DA-BiLSTM-BO model: Superior accuracy in high-frequency financial data analysis,” IEEE Access, vol. 12, pp. 173618–173637, 2024. doi: 10.1109/ACCESS.2024.3502175.

[3] S. Wang, “A stock price prediction method based on BiLSTM and improved Transformer,” IEEE Access, vol. 11, pp. 104211–104223, 2023. doi: 10.1109/ACCESS.2023.3296308.

[4] L. Bacco, L. Petrosino, D. Arganese, L. Vollero, M. Papi, and M. Merone, “Investigating stock prediction using LSTM networks and sentiment analysis of tweets under high uncertainty: A case study of North American and European banks,” IEEE Access, vol. 12, pp. 122239–122248, 2024. doi: 10.1109/ACCESS.2024.3450311.

[5] P. Chen, Z. Boukouvalas, and R. Corizzo, “A deep fusion model for stock market prediction with news headlines and time series data,” Neural Computing and Applications, vol. 36, pp. 21229–21271, 2024. doi: 10.1007/s00521-024-10303-1.

[6] J. R. Chowdhury and H. Liu, “Automated stock price prediction using machine learning: A comprehensive review,” arXiv preprint, arXiv:2408.12408, 2024

[7] A. Peivandizadeh et al., “Stock market prediction with transductive long short-term memory and social media sentiment analysis,” IEEE Access, vol. 12, pp. 87110–87130, 2024. doi: 10.1109/ACCESS.2024.3399548.

[8] N. Singh, W. Jha, P. Gupta, and K. K. Gupta, “Artificial intelligence driven portfolio management,” in Proc. IEEE Recent Advances in Intelligent Computational Systems (RAICS), Kothamangalam, India, 2024, pp. 1–6. doi: 10.1109/RAICS61201.2024.10689727.

[9] W. M. Shaban, E. Ashraf, and A. E. Slama, “SMP-DL: A novel stock market prediction approach based on deep learning for effective trend forecasting,” Neural Computing and Applications, vol. 36, pp. 1849–1873, 2024. doi: 10.1007/s00521-023-09179-4.

[10] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] D. Araci, “FinBERT: Financial sentiment analysis with pre-trained language models,” 2019.

