Grenze International Journal of Engineering and Technology, January Issue

A Comparative Study of Sentiment-Driven Intelligent Systems using FinBERT

Vibhor Singh[1] and Rahul[2]
[1-2]Delhi Technological University/Department of Software Engineering, Delhi, India
Email: vibhor.official0219@gmail.com, rahul@dtu.ac.in

***Abstract*— We highlight the effectiveness of FinBERT in capturing financial sentiment, the superiority of ML models over traditional time-series approaches, and the importance of sentiment-driven feature integration. This paper uses Apple stock data and its associated news from Alpha Vantage and perform a comparative study on stock price prediction using SARIMA and three FinBERT-augmented ML models (SVR, RFR, and GBR). The experimental results demonstrated that FinBERT+GBR achieved the best performance (MAE: 2.5004, MSE: 10.4263, RMSE: 3.2290, R^2: 0.9864), significantly outperforming SARIMA and other ML models. The results underscore the potential of sentiment-augmented ML models for practical applications, though challenges in interpretability and computational efficiency warrant further exploration.**

***Index Terms*— Financial market, News media, Social media, Sentiment analysis, Stock price prediction, Machine learning, Stock price prediction, FinBERT.**

I. INTRODUCTION

The domain of stock market prediction is one of the prominent applications of ML considering its inherent complexity and economic implication. Time series and statistical model was conventional cornerstone of market analysis, but they have limited potential for capturing intricate pattern. Swift evolution of machine learning (ML) techniques, the financial forecasting scenario have undergone significant transformation. All the sentiment induced stock prediction model have prominent social implications, especially in taking educated decision making, providing economic stability. Natural Language processing (NLP) sentiment analysis played a pivotal role for more efficient stock prediction. NLP taps into huge pool of textual data of different sources, comprising of news and social media outlets. The collective mood and opinion of market participants can now be harness using power of NLP's sentiment analysis which is to be fed to predictive model for better forecasting. Fusion of sentiment derived insights with ML algorithms presents a substantial leap which not only surges the predictive power of existing models but also provide nuanced understanding of the psychology of market movements driving factors. Consequently, financial industry witnessing a paradigm shift for the anticipation of stock prices fluctuations, with the support of AI driven sentiment analysis.
Comparative models provide strength to financial stakeholders by utilizing the power of Financial Bidirectional Encoder Representations from Transformers (FinBERT) and ML models to achieve higher prediction results; traditionally these advance analytical processing done through rigorous statistical calculations. Moreover, sentiment infused ML models enhances the financial forecasting of stocks by penalizing speculative trading disseminated misinformation or emotional outrage as model captures intricate complexities.

Grenze ID: 01.GIJET.12.1.36

This paper uses Apple stock (APPL) data and its associated news data with timeline 2023-2025 from Alpha Vantage API and presents a machine learning-based stock prediction model that integrates sentiment analysis using FinBERT, it is a specialized model for financial sentiment analysis that uses BERT. It focuses on enhancing financial stock forecasting by adding investors sentiment data with conventional stock price data. This study takes into consideration traditional time series model like SARIMA for stock price prediction and three FinBERT infused ML models namely Support vector regressor (SVR), Random Forest regressor (RFR), Gradient Boosting regressor (GBR). Eventually all predictive models are compared through regression evaluation metrics like mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), R2(R-squared). Additionally caution for over-reliance on automated predictive models have ethical considerations which should be addressed properly to not amplify market instability. Every necessary regulatory framework should be be complied in a responsible manner for model interpretability and deployment to harness the positive social impact while reducing the risks. Eventually this study helps to build a more aware and equitable financial ecosystem, which in long run benefit society by encouraging sound investment culture and economic resilience.

II. RELATED WORK

D. Shah et al. [1] aim to determine how news sentiments affect stock market movements of pharmacological industry in Nifty Pharma Index namely NSE: ALEKM, NSE: LUPIN, and NSE: SUNPHARMA. Stock data contains intraday (30 min gap) from Investing.com and six-month news article were analyzed from moneycontrol.com. It entails building a 100-word customized pharmaceutical specific dictionary and the predictive model take decision of 'buy', 'sell', and 'hold' based on threshold value of sentiment score derived through customized dictionary. The model achieves directional accuracy of 70.59% for short term prediction. Critics [5] argue that it suffers from limited scalability due to manually created customized dictionary, our work uses a pre-trained financial NLP model FinBERT which dynamically captures sentiments without curating manual dictionary, which offers much more scalability and domain-specific accuracy for APPL stock.

J. Huang et al. [2] utilizes HHPIC (Taiwan stock market) stock data chip indicators like Trading volume, stock price change, margin buy loan etc. Sentiments of social media related to HHPIC was taken from Taiwan's PPT bulletin board system and processed using CKIP parser and NTUSD dictionary to calculate sentiment score. Finally Binary logistic regression induced with backward elimination was used to measure the change extend of stock prices at three levels of p-value (0%, 0.5%, 1%) which subsequently improved accuracy by 4.05 %, 1.35%, 2.70% respectively. It was also concluded that stock price change was negatively correlated with sentiment score suggesting the contrarian market attitude (e.g., selling when others buying and vice-versa). Rivals [3] critique for missing nuanced financial context while using the dictionary, our study differentiates by leveraging FinBERT contextual embeddings, capturing complex sentiment pattern in financial news and compares with multiple models (SVR, RFR, GBR) against SARIMA giving more comprehensive evaluation.

A. Kanavos et al. [3] focuses on the relationship between Twitter tweets sentiments and stock price data of Microsoft and Apple leveraging sentiment analysis of Stanford CoreNLP induced with Naïve Bayes Classifier. Lambda architecture with three layers (Batch, Serving, and Speed) have been used to perform sentiment analysis to correlate tweet sentiments with stock prices. Highest accuracy has been achieved by Bigrams, tweet polarity has association with the price movements of Apple and Microsoft stocks. It solely relies on Twitter data, potentially missing broader sentiments from another platform like stockTwits, Reddit etc as pointed by [6]. Our work stands out by providing richer domain specific sentiment as compared to social media by FinBERT offering broader comparative framework.

I. K. Nti et. al. [4] analyzed stock data from the Ghana Stock Exchange (GCB, MTNGH, TOTAL) with tenure ranging from Jan 2010 to Sep 2019 to predict future stock prices over various timeframes (90, 60, 30, 7) days. Investor's sentiment was obtained from Twitter, forum discussion (sikasem.org), headlines from news portals of Ghana and Google Trends, later they all were analyzed through NLTK for polarity and subjectivity score. Multilayer ANN was employed to predict binary movement (rise/fall) of stock prices for 1-90 days ahead. Combined dataset with sentiments from all sources has highest accuracy of 77.12 %, among all sources Twitter outperforming others in accuracy (up to 60.05%) when combined with stock dataset. The results also emphasize the existence of a strong correlation between social media and stock market activities which indicates that investors can make use of online data to better predict the market and make informed investment choices. Limitations include focus on single market, challenges with cultural specificity and short-term sensitivity as suggested by [10], our study employs FinBERT focusing on APPL stock with rigorous comparison of ML models (SVR, RFR, GBR) against SARIMA, addressing gap in multi-model evaluation.

R. Chiong et. al. [5] called for the construction of a multichannel collective network for stock price forecasting using social media sentiment analysis and representing the results in candlesticks chart. Sentiment features were extracted from Twitter using natural language processing (NLP) tools, while historical stock data was converted into candlestick charts to capture price movement patterns. The model integrates both data types using two CNN branches: 1-D CNN for sentiment classification and 2-D CNN to predict binary movement (increase/decrease) of stock prices over look ahead of 4,6,8,10 days. When tested on five major stocks: Google, Tesla, Apple, IBM, and Amazon, the approach outperformed single-data-source models, achieving a 75.38% accuracy (10 day look ahead) for Apple. Longer prediction periods yielded better results than shorter ones with Apple and Tesla showing stronger performances. Critics like [5] emphasizes that use of general-purpose sentiment tools limits its applicability to financial contexts, reinforcing the need for domain-specific models like FinBERT, which our study integrates with ML models to enhance predictive accuracy.

X. Li et. al. [6] studies twelve stock data of Hang Seng Index(Hong Kong) across four sectors (Commerce, Utilities, Finance, Properties) and also utilize the technical indicators (moving averages, MACD, RSI, and MFI) and the news article from FINNET aligned with stock data analysed through four different type of sentiment dictionary .Then build a DL model(2-layer LSTM) to incorporate stock and sentiment data predicting stock price trend(fall, rise, horizontal).Base line models(SVM, MKL) was outperformed by DL model attaining 83.8% accuracy. This comparison study reveals that Loughran-McDonald financial dictionary perform better than remaining dictionaries. It does not contain any event-specific sentiment analysis which can be implemented for broader validation.

N. Das et. al [7] studied how public sentiment influenced stock market predictions of the Corona virus pandemic using LSTM architectures. The study employs Nifty50 stock data from Yahoo Finance and VADER, LR, Loughran-McDonald, Henry, and other seven sentiment analysis tools on scraped data from stock headline, tweets, financial news, Facebook comments etc. The research finds that Linear SVC-based sentiment scores derived from Facebook comments achieve 98.11% accuracy. When sentiment scores were integrated within the stock data, the prediction accuracy of 98.32% was achieved, marking the extreme effect emotions have on the movements of the stock market.

W. J. Liu et. al. [8] set forth a model that considers seven major stock indices (NYSE, FTSE 100, SSEC, Nikkei, NASDAQ, S&P 100, DJI) from Yahoo Finance alongside financial news from website Investing. Financial news was processed with LSTM-CNN model suing Word2vec embeddings and the Financial PharseBank dataset for sentiment classification achieving 8.3% higher accuracy then standalone LSTM. Combination of sentiment attention mechanism and TrellisNet (SA-TrellisNet) uses the news and stock sentiment index, to predict closing prices which outperformed other models like RNN-LSTM and CNN-GRU, showing lower MSE, MAE, RMSE, MAPE and higher R-squared. The architecture is assessed with seven major stock indices, including NASDAQ, and S&P 500, and it competes with existing prediction methods.

Y. Huang et. al. [9] called for the construction of a multichannel collective network for stock price forecasting using social media sentiment analysis and representing the results in candlesticks chart. Sentiment features were extracted from Twitter using natural language processing (NLP) tools, while historical stock data was converted into candlestick charts to capture price movement patterns. The model integrates both data types using two CNN branches: 1-D CNN for sentiment classification and 2-D CNN to predict binary movement (increase/decrease) of stock prices over look ahead of 4, 6, 8, 10 days. When tested on five major stocks: Google, Tesla, Apple, IBM, and Amazon, the approach outperformed single-data-source models, achieving a 75.38% accuracy (10 day look ahead) for Apple. Longer prediction periods yielded better results than shorter ones with Apple and Tesla showing stronger performances.

M. Kesavan et al. [10] collect historical Infosys stock (NSE) data from Moneycontrol and web scrape sentiment data from Facebook posts, Twitter, News headlines and subsequently sentiment scores are generated through NLTK, GloVe embeddings and a lexicon-based sentiment classification. The sentiment induced stock dataset fed to LSTM network with ADAM optimizer, variable lookback period to predict next day closing price percentage change for predicting price fluctuations (decreases, increases, and stagnation). This predictive model achieved accuracy of 96.95 % confirming inducing sentiment data to predictive model results in close alignment with predicted and actual prices. Limitation may include reliance on limited types of stock, focusing on single stock, challenges in detecting sarcasm. Our study differentiates by utilising FinBERT contextual embeddings, capturing complex sentiment pattern in financial news and comapres with multiple models (SVR, RFR, GBR) against SARIMA giving more comprehensive evaluation.

The reviewed studies emphasize on the value of sentiment prediction but also elaborate critical gaps like reliance on general purpose sentiment analysis tools (e.g., VADER, TextBlob) that leg behind to captures complex sentiment pattern, and limited comparison between sentiment augmented ML and time series models.

III. Motivation

To the advancement of ML and NLP has meaningfully reshaped the background of stock market anticipation, particularly by enabling the incorporation of investor sentiment alongside conventional financial indicators. However, existing research still presents notable limitations, chief among them being the underutilization of sentiment models customized to the financial domain and the scarcity of thorough evaluations comparing sentiment-augmented ML approaches with traditional time-series forecasting techniques. This study seeks to link these gaps by suggesting a stock prediction framework that fit in sentiment analysis using FinBERT, a model specifically optimized for interpreting monetary sentiment using BERT. The extracted sentiment signals from financial news are combined with standard stock market data to train and assess the performance of three ML models: SVR, RFR, and GBR in comparison with the classical SARIMA model. The objective is to measure the influence of domain-specific sentiment features to predictive accuracy and to classify the utmost active method for sentiment-informed stock forecast. Through this, the research aspires to offer valuable insights for both scholarly inquiry and real-world financial decision-making.

IV. Problem Domain

Problem domain of this review focuses on financial market prediction, sentiment analysis, and AI which integrate NLP, ML, DL to solve the problem of traditional stock prediction model that uses historical price data and technical indicators which often fails to capture dynamic behavior of stock market as it is unable to include economic indicators, corporate performance, investor's behavior etc.

V. Problem Definition

While sentiment analysis has been increasingly applied to stock market prediction, many studies rely on general-purpose sentiment models (e.g., VADER, TextBlob) that fail to capture the nuanced sentiment in financial texts. These models often miss domain-specific terminology and context, leading to suboptimal sentiment features for predictive modeling. There is a lack of studies leveraging specialized financial sentiment models, like FinBERT, to improve the correctness of stock price predictions. This study addresses the gap by utilizing FinBERT, a model for financial sentiment analysis which uses BERT, to extract high-quality sentiment features from financial news. By integrating FinBERT-derived sentiment with conventional stock features, the proposed model (combining SARIMA, SVR, RFR, and GBR) captures domain-specific market sentiment, improving prediction correctness as demonstrated by lower MAE, MSE, RMSE, and greater R^2 compared to baseline models.

VI. Problem Formulation and Solution Methodologies

Sentiment analysis is recognized as a game-changer in stock prediction, few studies systematically compare the performance of multiple ML models incorporating sentiment data against traditional time series models. This limits the understanding of which models best leverage sentiment features and under what conditions, hindering practical adoption in financial forecasting. The proposed study fills this gap by steering an inclusive contrast of four predictive models: SARIMA (time series) and three ML fusion models (SVR, RFR, GBR) all incorporating FinBERT-derived sentiment from financial news. The use of standardized evaluation metrics (MAE, MSE, RMSE, R^2) provides a robust benchmark to recognize the furthermost effective model for sentiment augmented stock forecast, offering actionable insights for financial practitioners.

The research was carried out in a Python-driven environment designed to support consistent and scalable outcomes. The hardware system used to meet intensive computational tasks like data preparation, sentiment score generation, model deployment was carried with an Intel core i7 CPU, 16GB RAM, and an NVIDIA GPU. The main programming language employed in this study is Python (version 3.8 or later), in association with specialized libraries like Numpy, Pandas, Scikit Learn etc. Stock data of APPL was retrieved from yfinance library, while the APl stock related news article was fetched with the help of API provided by Alpha Vantage. Hugging Face transformer framework provide FinBERT which performed sentiment analysis for stock news data. Scikit learn library have been deployed for ML predictive models like SVR, RFR, and GBR, in addition with evaluation metrics. Statsmodels provide implementation of SARIMA time series model, in addition Pandas and Numpy provide effective data processing and handling.

Results of visualization for comparing evaluation metric of models was provided by Matplotlib and Seaborn. Dataset tenure for both stock and sentiment have been taken from 03-02-2023 to 03-02-2025 giving a comprehensive timeframe for analysis.

The proposed methodology involves a systematic pipeline for stock price prediction, integrating sentiment analysis with machine learning. Fig. 1 in the below section illustrates the workflow, which includes data collection, preprocessing, sentiment analysis using FinBERT, model training, and performance evaluation. The following detail describes each component of the proposed work.

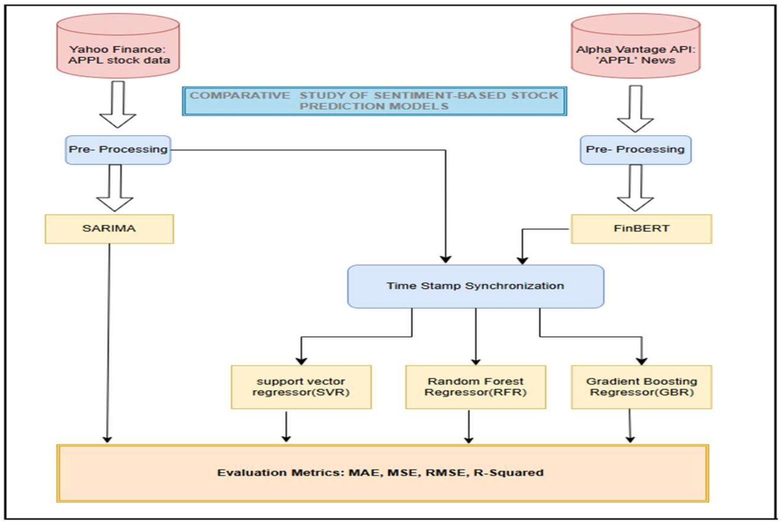

Figure 1. Workflow for comparative study

There are two kinds of dataset taken for this comparative study: Stock dataset contains Apple stock (APPL) which is traded on NASDAQ has been taken for stock prediction using yahoo finance (yfinance) python library. Five values of APPL stock have been taken i.e., 'open', 'high', 'low', 'close', 'volume' for the period of two years (03-02-2023 to 03-02-2025). Sentiment dataset contains Alpha Vantage provide stock news for related stock ticker like 'APPL' for Apple Inc. through its API. We have used its API and collected stock news related to 'APPL' ticker only for same period as stock data.

Data preprocessing involves replacing missing values in stock dataset due to non-trading days are managed by forward filling method to maintain time series data continuity. Normalization of numerical features have been done using Min-max scaling to standardize their ranges, while maintaining their compatibility with ML algorithms. Close price of next day has been selected as target variable, while remaining fields like open, low, high, volume taken as input features. News article have been cleaned by eliminating elements that are not relevant like special characters, HTML tags, stop words, to emphasize on important content. Words of news articles have been divided into tokens by tokenization that aligns with the input requirements of FinBERT. Each news article title and summary text are fed to FinBERT, which in turn classifies each article sentiment score on timestamp order. Later sentiment score aggregation is done for each day to align temporal resolution of numerical stock data. Finally, each day sentiment score and numerical stock data features is integrated which will act as a unified dataset for every FinBERT induced ML models for comparison. The predictive models that are used are as follows:

SARIMA: The SARIMA model is a statistical time-series forecasting method that covers the ARIMA model by incorporating seasonal patterns. It is designed to collect both non-seasonal and seasonal trends, cycles, and dependencies in time-series data, making it suitable for financial forecasting tasks like stock price prediction. In this study, SARIMA serves as the baseline model, trained solely on historical stock data without sentiment features, to provide a point of comparison for the sentiment-augmented machine learning models.

FinBERT+SVR: SVR is a ML method based on Support Vector Machines, adapted for regression tasks. In this study, SVR is augmented with FinBERT-derived sentiment scores to forecast AAPL stock prices, harnessing its ability to model non-linear relationships in high-dimensional data. The combination of FinBERT sentiment and stock features aims to capture both quantitative market dynamics and qualitative investor sentiment.

FinBERT+RFR: RFR is an ensemble machine learning method that pools multiple decision trees to increase predictive correctness and reduce overfitting. In this study, RFR is trained on a combined dataset of AAPL stock features and FinBERT-derived sentiment scores, leveraging its robustness in modeling non-linear relationships and high-dimensional data.

FinBERT+GBR: GBR is an ensemble machine learning technique that builds sequential decision trees to minimize residual errors, offering high predictive accuracy for regression tasks. In this study, GBR is trained on the combined dataset of AAPL stock features and FinBERT-derived sentiment scores, achieving the best performance among all models

Finally, all the regression evaluation metrics (MAE, MSE, RMSE, R²) are compared and visualized, test data that is 20% of total dataset is plotted as next day predicted vs actual closing price.

VII. Results and Comparison

The performance of the four predictive models—SARIMA, FinBERT+SVR, FinBERT+RFR, and FinBERT+GBR was evaluated on the test dataset, comprising 20% of the AAPL stock data from 03-02-2023 to 03-02-2025. The models were assessed using MAE, MSE, RMSE, and R² to measure predictive accuracy and explanatory power. The results are summarized below in Table I highlighting the comparative performance across all metrics.

TABLE I. REGRESSION METRICS OF PREDICTIVE MODELS

Model Name	MAE	MSE	RMSE	R2
SARIMA	6.9316	83.0328	9.1122	0.1969
FinBERT+SVR	3.8762	23.5417	4.8520	0.9487
FinBERT+RFR	3.1527	14.5136	3.8097	0.9859
FinBERT+GBR	2.5004	10.4263	3.2290	0.9864

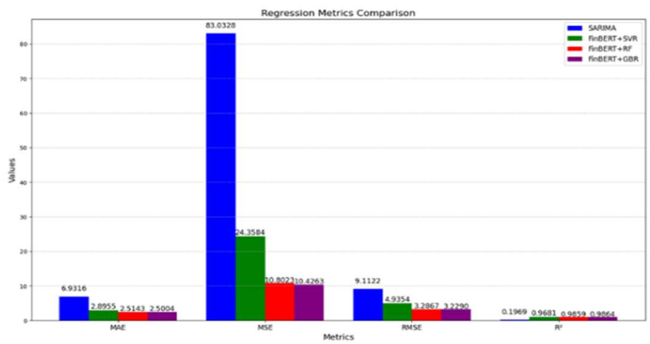

Figure 2. Bar chart of the performances of various models

Fig. 2 illustrate the bar chart of comparative performance, with FinBERT+GBR consistently showing the lowest error metrics and highest R² score, followed closely by FinBERT+RFR, then FinBERT+SVR, and significantly outperforming SARIMA.

The SARIMA model, serving as the baseline, exhibited the poorest performance across all metrics. With an MAE of 6.9316, MSE of 83.0328, and RMSE of 9.1122, it indicates significant prediction errors, reflecting its limited ability to capture the non-linear and volatile patterns in AAPL stock prices. The R² score of 0.1969 suggests that only 19.69% of the variance in the stock prices is explained by the model, underscoring its inadequacy for accurate forecasting in this context as depicted in Fig.3.

Figure 3. Line chart of next day closing price (in US dollors) test results (SARIMA)

Regression metrics of FinBERT+GBR have best performed with lowest value across all metrics with MAE of 3.1527. MSE of 14.5136, and RMSE of 3.8097. GBR have outperformed all other models in handling nonlinear relationships and high dimensional data as it can be seen through R2 of 0.9859 as it explained 98.59% variance of data, explaining highest predictive capability as shown in Fig. 4. The gradient boosting approach, with incorporation of FinBERT induced sentiments have increased the capability of model to effectively capture market dynamics, making it most competitor among other tested models in the study.

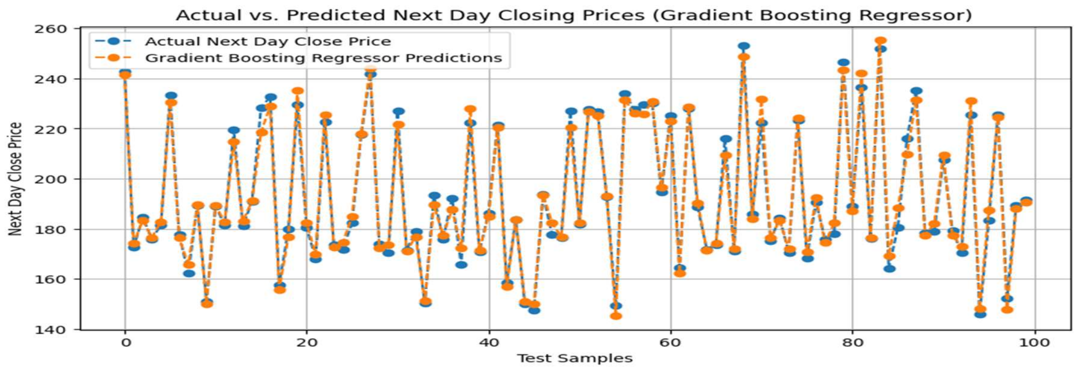

Figure 4. Line chart of next day closing price (in US dollors) test results (GBR)

Regression metrics of FinBERT+RFR have significantly decreased with MAE of 3.1527. MSE of 14.5136, and RMSE of 3.8097. RFR is very robust in handling nonlinear relationships and high dimensional data as it can be seen through R2 of 0.9859 as it explained 98.59% variance of data as shown in Fig 5. Incorporation of FinBERT induced sentiments have increased the capability of model to effectively captures market dynamics, making it a strong competitor among other tested models. The FinBERT+SVR model, which integrates sentiment scores from financial news with stock features, showed substantial improvement over SARIMA. Reduction in error can be seen in the evaluation metrics of MAE of 3.8762, MSE of 23.5417, and RSME of 4.8520. Significant role of adding FinBERT induced sentiment can be seen through improved R2 score of 0.9487 that reflects significant 94.87% variance in prices of stock as shown in Fig 6. But its performance is surpassed by other FinBERT induced ML models.

VIII. CONCLUSION

This study focuses on the effectiveness on sentiment integration with machine learning for stock price prediction by investigating comparative analysis of traditional time series model SARIMA with three FinBERT induced ML models i.e., SVR, RFR, GBR. The superior performance of FinBERT+SVR, FinBERT+RFR, and FinBERT+GBR over SARIMA highlights the effectiveness of FinBERT in capturing nuanced financial sentiment. Unlike general-purpose sentiment models (e.g., VADER, TextBlob), which often miss domain-specific terminology, FinBERT's pre-training on financial texts enables it to extract accurate sentiment scores from news articles.

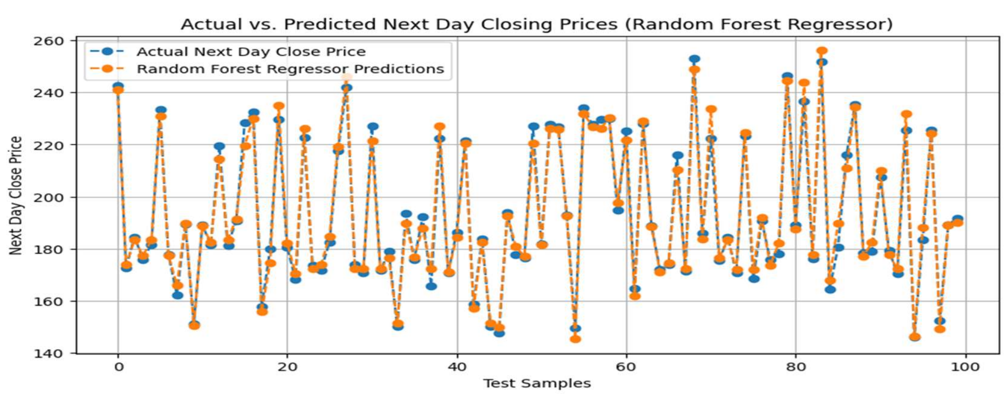

Figure 5. Line chart of next day closing price(in US dollors) test results (RFR)

The low error metrics (e.g., MAE of 2.5004 for FinBERT+GBR) and high R^2 scores (up to 0.9864) demonstrate that FinBERT-derived sentiment features significantly enhance predictive accuracy, addressing the gap in the limited use of domain-specific sentiment models.

This comparative study shows that FinBERT induced models have significantly beat SARIMA, moreover FinBERT+GBR gains the best performance amongst all FinBERT based models. The results of study proof the effectiveness of FinBERT for capturing financial domain specific investor's sentiment and capability of ML models in tackling complex intricate patterns.

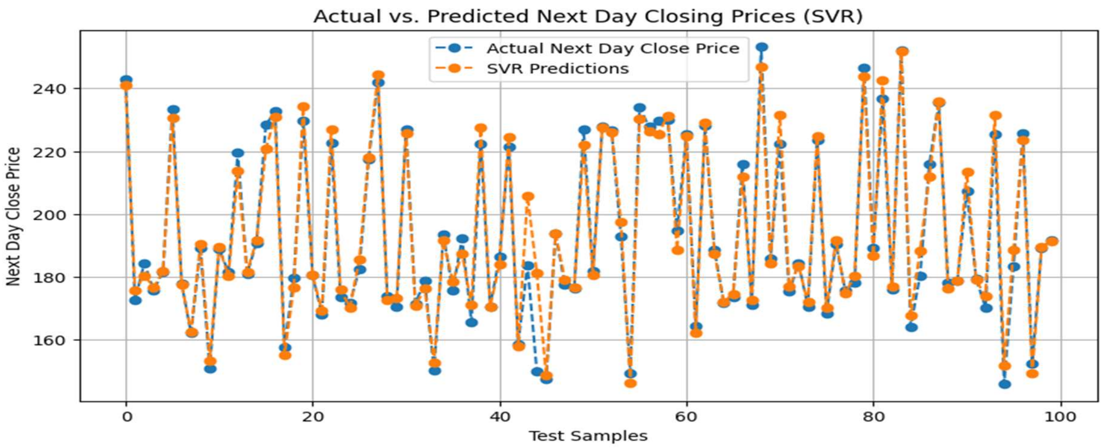

Figure 6. Line chart of next day closing price (in US dollors) test results (SVR).

FUTURE WORK

This study will surely open several paths for future advance research that want to explore sentiment-based stock prediction. Firstly, addition of multiple data sources can enhance sentiment feature set like social media platform (twitter, stocktwits), macroeconomic indicator (inflation, GDP, interest rates), and multimodal data (audio from earning calls. Market analysis video). Secondly, make use of advance deep learning models like LSTM, Transformers induced models, or hybrid models like CNN+LSTM which are capable of obtaining complex pattern and temporal dependencies in financial forecasting of stock data.

Thirdly proposed models can be optimized using hyper parameters which can be done through automated techniques like grid search, Bayesian optimization, and genetic algorithm, which ultimately results in reducing computational overhead.

Fourthly, proposed models can take input from diverse range of stocks across various sectors and different market conditions like emerging markets, recessionary market etc. Fifthly, proposed model should work on interpretability of model by incorporating explainable AI techniques like SHAP, LIME, and PDP etc. Lastly for fast processing of real-time sentiment analysis, low latency pipelines should be made to encompass the practical high frequency trading utility scenario.

REFERENCES

[1] D. Shah, H. Isah, and F. Zulkernine, "Predicting the effects of news sentiments on the stock market," 2021 IEEE International Conference on Big Data (Big Data), pp. 4705–4708, Dec. 2018.

[2] J. Huang and J. Liu, "Using social media mining technology to improve stock price forecast accuracy," Journal of Forecasting, vol. 39, no. 1, pp. 104–116, Jun. 2019.

[3] A. Kanavos, G. Vonitsanos, A. Mohasseb and P. Mylonas, "An Entropy-based Evaluation for Sentiment Analysis of Stock Market Prices using Twitter Data," 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA, Zakynthos, Greece, 2020, pp. 1-7.

[4] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "Predicting stock market price movement using sentiment analysis: Evidence from Ghana," Applied Computer Systems, vol. 25, no. 1, pp. 33–42, May 2020.

[5] R. Chiong, Z. Fan, Z. Hu, and S. Dhakal, "A novel ensemble learning approach for stock market prediction based on sentiment analysis and the sliding window method," IEEE Transactions on Computational Social Systems, vol. 10, no. 5, pp. 2613–2623, Aug. 2022.

[6] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," Information Processing & Management, vol. 57, no. 5, p. 102212, Feb. 2020.
[7] N. Das, B. Sadhukhan, T. Chatterjee, and S. Chakrabarti, "Effect of public sentiment on stock market movement prediction during the COVID-19 outbreak," Social Network Analysis and Mining, vol. 12, no. 1, Jul. 2022
[8] W.-J. Liu, Y.-B. Ge, and Y.-C. Gu, "News-driven stock market index prediction based on trellis network and sentiment attention mechanism," Expert Systems with Applications, vol. 250, p. 123966, Apr. 2024.
[9] Y. Huang, and T.-T. Ho "Stock price movement prediction using sentiment analysis and CandleStick chart representation," Sensors, vol. 21, no. 23, p. 7957, Nov. 2021.
[10] M. Kesavan, J. Karthiraman, R. T. Ebenezer, and S. Adhithyan, "Stock Market Prediction with Historical Time Series Data and Sentimental Analysis of Social Media Data," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 477–482, May 2020.