

Sentiment-Driven Stock Price Prediction: Analysing Green Finance News Using Large Language Models for Tesla

Amir Lorvand¹^a, Halil Yetgin¹^b, Serengul Smith¹^c, Duaa Alkubaisy²^d and Luca Piras³^e

¹Department of Computer Science, Middlesex University, The Burroughs, London, NW4 4BT, U.K.

²Institute of College of CS & IT, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

³Center for Cybersecurity, Fondazione Bruno Kessler (FBK), Trento, Italy

Keywords: Stock Price Prediction, Sentiment Analysis, Green Finance, Large Language Models, Long Short-Term Memory, Recurrent Neural Network, Artificial Intelligence.

Abstract: The rise of green finance has had a significant impact on the way investors evaluate a company's performance, prompting them to consider environmental, social and governance (ESG) criteria alongside traditional financial metrics. Nevertheless, research investigating the potential of ESG-related news to improve stock price prediction models is still limited. This study examines the effectiveness of sentiment analysis derived from green finance and ESG-related news to improve stock price forecasts specifically for Tesla Inc. Historical stock price data was combined with sentiment scores extracted from ESG-related news articles using FinBERT. Two recurrent neural network models (RNN) and two long short-term memory (LSTM) models were developed and evaluated, both with and without the inclusion of sentiment features. The results show that sentiment features have only a modest impact on predictions during stable market conditions, while they provide significant improvements during periods of market volatility. Specifically, during a volatile period when Tesla's share price fluctuated between \$340 and \$480, the inclusion of sentiment features in the LSTM model reduced the root mean square error (RMSE) from 9.87 to 4.06. This significant reduction underscores the improved accuracy achieved by incorporating textual sentiment data. Overall, these results emphasise the importance and potential benefits of combining traditional financial indicators with textual data analysis, especially in times of increased market uncertainty.

1 INTRODUCTION

Accurately predicting stock prices remains a cornerstone of financial decision-making, playing a vital role in guiding investment strategies, portfolio optimisation, and risk management. Traditionally, such predictions rely heavily on technical indicators and fundamental metrics derived from company performance, macroeconomic data, and market trends. However, these models often neglect qualitative signals such as market sentiment, which can capture the psychological and behavioural aspects of investor decision-making. The increasing digitalisation of financial information and the real-time availability of

news and media content have elevated the role of sentiment analysis as a complementary tool in financial forecasting (Das et al., 2024).

Recent developments in natural language processing and machine learning have enabled the extraction of actionable insights from unstructured textual data such as financial news, analyst reports, and social media. These advancements support the design of sophisticated models capable of capturing linguistic subtleties and mapping them to market behaviours, thereby improving the accuracy and timeliness of stock price predictions (Das et al., 2024).

Amid growing concerns over climate change and corporate responsibility, green finance has emerged as a central theme in global investment discourse. It encompasses financial products and frameworks such as green bonds and sustainability-linked loans, aiming to promote environmental and social goals through capital markets. As institutional and retail investors increasingly consider environmental, social, and gov-

^a <https://orcid.org/0009-0003-1011-2727>

^b <https://orcid.org/0000-0003-0994-5839>

^c <https://orcid.org/0000-0003-0777-5637>

^d <https://orcid.org/0000-0001-5920-7856>

^e <https://orcid.org/0000-0002-7530-4119>

ernance (ESG) factors in their decision-making, the narrative surrounding green finance has begun to exert measurable influence on asset valuation (Pagano et al., 2018). Understanding how these sustainability narratives affect market perceptions and pricing mechanisms is therefore of strategic importance.

This study investigates the effect of green finance sentiment on Tesla's stock price by integrating sentiment analysis with deep learning models. Unlike conventional approaches that rely primarily on numerical financial indicators or historical prices, this research incorporates sustainability-oriented news coverage to assess how ESG-related narratives impact investor behaviour and market valuation. The main objective is to quantify the influence of green finance sentiment and evaluate which machine learning models are most effective in translating textual cues into predictive signals for stock price movement.

The research is guided by the following questions:

- How can sentiment derived from green finance news be leveraged in forecasting Tesla's stock performance?
- Is sentiment analysis particularly useful during periods of high price volatility?
- Which machine learning architecture best integrates textual and financial features to enhance predictive accuracy?

These questions are addressed through a data-driven framework, which fuses quantitative time series data with sentiment scores obtained from ESG-themed financial news.

Key challenges in this area include the technical complexity of financial language, the prevalence of misinformation, especially on social media platforms, and source bias in news reporting, all of which may distort sentiment interpretation (Allcott et al., 2019). Despite these limitations, the study offers a novel contribution by incorporating green finance sentiment into a hybrid prediction model, providing valuable insights for traders, financial analysts, and policymakers aiming to align investment strategies with sustainability goals.

The rest of the paper is structured as follows. Section 2 presents a review of the literature related to sentiment analysis, stock price prediction, and green finance. Section 3 describes our method, including data collection, sentiment classification and model development. Section 4 outlines summary of our key findings, limitations and related work, whilst Section 5 concludes the paper with suggestions for future research.

2 BACKGROUND

This section provides an overview of the main concepts and existing research relevant to our study. First, the main sentiment analysis techniques are presented, including both lexicon-based and machine learning approaches, followed by a discussion on the role of ESG sentiment and green finance in influencing investment behaviour. Finally, machine learning architectures, in particular RNNs and LSTMs, and their application in predicting financial time series are discussed.

2.1 Sentiment Analysis Techniques

Sentiment analysis, a central task within natural language processing (NLP), focuses on identifying and extracting subjective information from text data. It is often used to assess public opinion in areas such as marketing, politics and finance. Broadly speaking, sentiment analysis methods can be categorised into lexicon-based and machine learning (ML)-based approaches.

Lexicon-based techniques rely on predefined sentiment dictionaries that link words to sentiment scores. For this purpose, tools such as VADER are often used, which summarise the scores to create an overall sentiment classification. Although computationally efficient and easy to interpret, lexicon-based methods often have problems with contextual interpretation, sarcasm and domain-specific language (Kumar et al., 2020).

ML-based approaches such as Naive Bayes (Berrar, 2019) and Support Vector Machines (SVM) (Bhavsar and Panchal, 2012) tend to be superior to lexicon-based methods, especially when trained on domain-specific labelled datasets. However, these algorithms require robust feature engineering. In this context, representation techniques such as Term Frequency Inverse Document Frequency (TF-IDF) (Qaiser and Ali, 2018), Word2Vec (Mikolov et al., 2013) and GloVe (Pennington et al., 2014) have been developed to capture semantic relations between words more effectively and thus improve classifier performance.

Deep learning techniques have further advanced the field, especially for tasks involving complex sequential or contextual information. Recurrent neural architectures such as the Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) networks (Cho et al., 2014) have proven to be effective in modelling sequential dependencies in texts. More recently, transformer-based models such as BERT and domain-specific variants such as FinBERT have out-

performed traditional techniques by utilising large-scale pre-training and attention mechanisms (Chen, 2024).

2.2 ESG Sentiment and Green Finance

The emergence of green finance and environmental, social and governance (ESG) indicators has redefined the way investors assess company value. ESG sentiment refers to the attitudes and perceptions expressed in news and reports about a company's sustainability practises, which can influence market behaviour and stock valuations (Rudianto et al., 2024).

Green Finance promotes sustainable business practises through financial instruments such as green bonds and ESG screening and encourages investors to consider long-term environmental and social impacts. Companies with high ESG performance are generally perceived as more resilient and are likely to increase investor confidence and financial returns (Rudianto et al., 2024), (Egorova et al., 2022). In this context, sentiment analysis tools were applied to ESG-related news to assess market reactions and investment dynamics (Banerjee et al., 2025).

Tesla Inc. is a notable example of this. With its focus on electric vehicles and renewable energy, Tesla is closely aligned with ESG goals. Therefore, news related to sustainability, environmental regulations or green innovations have a disproportionate impact on Tesla's stock movements. This sensitivity underscores the growing importance of ESG sentiment in financial forecasting and investment decisions.

2.3 Deep Learning in Financial Forecasting

Predicting stock prices is a complex challenge due to the non-linearity, volatility and time dependencies inherent in financial time series. Traditional statistical models often assume linear relationships and stationarity, which limits their effectiveness in real-world scenarios. Deep learning methods, especially those designed to process sequential data, have proven to be powerful alternatives.

RNNs are designed to capture sequential patterns by retaining hidden states that propagate over time and allow the model to take historical context into account (Liu et al., 2024). This makes them well suited for financial forecasting where price movements depend on time-distributed factors. However, it is known that standard RNNs suffer from the vanishing gradient problem, which makes learning over long sequences difficult and limits their ability to model long-term dependencies.

To solve this problem, LSTM networks have a gating mechanism with input, forget and output gates which regulates the flow of information and allows the model to retain relevant knowledge over longer time intervals (Varadharajan et al., 2024). This architecture makes LSTMs particularly effective for modelling financial time series, where long-term dependencies and irregular events are crucial for accurate predictions (Al-Selwi et al., 2024).

Recent research has shown that integrating sentiment characteristics derived from ESG-related news into LSTM-based models can improve predictive accuracy in volatile market (Chen, 2024). This underlines the value of combining quantitative time series data with qualitative textual sentiment in financial modelling. Such integration offers improved explanatory power, especially in the context of green finance, where investor behaviour is closely linked to public narratives and ESG-related developments.

3 METHODS

The subsequent subsections delineate the data collection methodology and expound upon the application of sentiment analysis utilizing FinBERT (Yang et al., 2020). Our selection of FinBERT was predicated on its specialized architecture; as a derivative of BERT (Devlin et al., 2019), it has been explicitly fine-tuned on financial corpora, thereby enhancing its efficacy within this domain. Following this, the study details the data pre-processing protocols and elucidates the procedural framework employed for the development of the predictive models.

3.1 Data Collection

For the purposes of this study, data from two primary sources were used: historical stock price data from Tesla Inc. and news articles on green finance. The stock price data was retrieved through the Yahoo Finance API, while the news articles were collected through the GNews API¹, a subscription-based service that provides access to historical news content. The data collection covered a three-year period, from February 2022 to February 2025, and captured a wide range of financial developments, sentiment swings and sustainability-related discourse relevant to Tesla. This time period was chosen to capture both short-term volatility and long-term patterns in share price behaviour while maintaining a sufficiently large but computationally manageable dataset.

¹<https://gnews.io>

The share price dataset includes the following attributes: Date, Open, High, Low, Close and Volume. These fields are commonly used in time series modelling and serve as key indicators of market activity and investor behaviour.

To collect relevant news articles, a series of targeted search queries were developed to reflect Tesla's commitment to sustainability and green finance. The selected Boolean search term was as follows:

(Tesla AND carbon emissions) OR
(Tesla AND sustainability) OR
(Tesla AND green finance) OR
(Tesla AND renewable energy) OR
(Tesla AND battery technology) OR
(Tesla AND clean energy) OR
(Tesla AND ESG)

These keywords have been carefully selected to ensure the relevance of Tesla's environmental, social and governance (ESG) activities. Due to the length restriction of the GNews API (maximum 200 characters), the most relevant terms with the highest reach were prioritised. The API searches the title, description and full content of articles, maximising the likelihood of obtaining comprehensive results.

During the three-year data collection period, approximately 7,700 news articles were retrieved. Only 35 days within the dataset contained no articles, which is reasonable given the typical publication cycles, including weekends and public holidays. Overall, the resulting corpus provides a solid basis for evaluating the relationship between ESG sentiment and stock price movements in the case of Tesla.

3.2 Sentiment Analysis by FinBERT

Before conducting the sentiment analysis, the title, description and full content of each news article were merged into a single, unified text string. This step was taken to ensure a comprehensive representation of the sentiment expressed in all parts of the article. By integrating all available textual components, the sentiment model is better able to capture nuanced opinions and implicit cues that may appear in different sections of the article.

The resulting text strings were then pre-processed using FinBERT's special tokeniser. This tokeniser automatically converts all characters to lower case, standardises the formatting and splits the text into sub-word units suitable for input into the model. The tokeniser also applies padding and truncation as necessary to ensure that all input matches the fixed input length of the model. FinBERT sets a token limit of 512 so that any input that exceeds this limit is truncated

and only the first 512 tokens are used for inference.

Once tokenisation was complete, the processed input was passed to the FinBERT model to perform sentiment inference. The model generates a series of logits, i.e. raw and non-normalised values, for three sentiment categories: positive, neutral and negative. These logits are then run through a softmax function that normalises them into probabilities that sum to one. This conversion allows the scores to be interpreted as confidence levels for each sentiment class.

To enable numerical analysis, the sentiment results were converted into continuous sentiment scores using a weighted formula:

$$\text{Sentiment Score} = (-1 \times P_{\text{negative}}) + (0 \times P_{\text{neutral}}) + (1 \times P_{\text{positive}}) \quad (1)$$

This formula ensures that the sentiment scores are continuously between -1 and $+1$, with negative values indicating a negative sentiment, positive values indicating a positive sentiment and scores around zero reflecting a neutral tone. This method provides a more granular and analytically tractable measure of sentiment than categorical labelling alone.

Since multiple articles may be published on the same day, the sentiment score was calculated for each article individually and then averaged to obtain a daily sentiment score. In cases where no articles were available for a particular day, a sentiment score of zero was assigned to maintain the continuity of the time series and provide a neutral basis for missing data. This approach ensures the consistency of the dataset and enables robust time series modelling of sentiment-related trends.

3.3 Data Pre-Processing

Thirty-five days within the dataset had no associated sentiment scores as there were no relevant news articles. To ensure the continuity of the time series and avoid missing values, the sentiment value for these days was formed from the arithmetic mean of the sentiment scores of the previous and following day. This simple interpolation method ensures temporal consistency without introducing artificial volatility.

In addition to the raw daily sentiment scores, new features have been developed to better reflect longer-term sentiment trends, which are often more indicative of market movements than short-term fluctuations. Specifically, moving averages of sentiment were calculated over three different time frames: 7 days, 14 days and 30 days. These moving averages provide a smoothed representation of sentiment dynamics and help to identify sustained trends that can

influence investor behaviour and price performance.

To ensure consistency between the characteristics and improve the stability of the model, all numerical input variables were standardised to a common scale. The sentiment-related features already between -1 and 1 , according to their scoring methodology. However, the stock price features, including *Close*, *High*, *Low*, *Open* and *Volume*, were normalised using the Min-Max Scaler technique. This approach scales each value according to the minimum and maximum of its respective column so that all features lie in a range between -1 and 1 .

It is important to note that the *Close* price, which serves as the prediction target, was also normalised during training but inverse transformed after prediction to recover the real monetary values of Tesla's stock price. This step is important to interpret the model results in a meaningful financial context.

Finally, the entire dataset was split into three different subsets to enable robust model evaluation: 66% was used for training, 20% for testing and the remaining 14% for validation. This stratification ensures that the model can be effectively trained and simultaneously evaluated with unseen data to assess its generalisability and predictive performance.

3.4 Model Development and Performance Evaluation

This section outlines the model design, training configurations, and performance evaluation of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) architectures in forecasting Tesla's stock closing prices. The primary objective is to assess the influence of environmental, social, and governance (ESG) sentiment by comparing model performance on two datasets: one including sentiment scores and one excluding them.

Both RNN and LSTM architectures require sequential input data, which was achieved by transforming the dataset into rolling time series windows. A function was implemented to generate these time-ordered sequences, which serve as inputs to the models.

3.4.1 RNN Architecture and Training

The RNN model consists of an input layer, one or more recurrent layers (tuned via grid search), a fully connected dense layer, and an output node. The following configuration yielded the best performance when the sentiment-enhanced dataset was used:

- Number of Neurons in the RNN Layers: 128
- Number of Stacked RNN Layers: 3

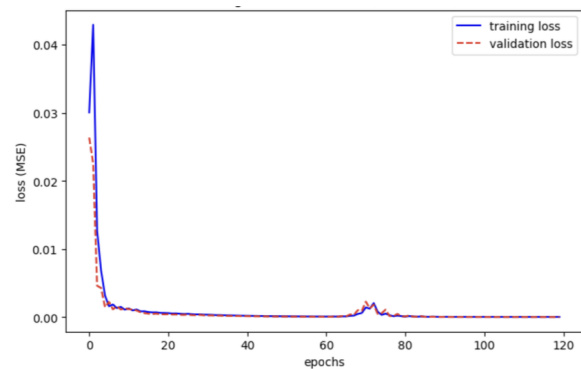


Figure 1: Training vs Validation Loss (RNN with Sentiment Scores).

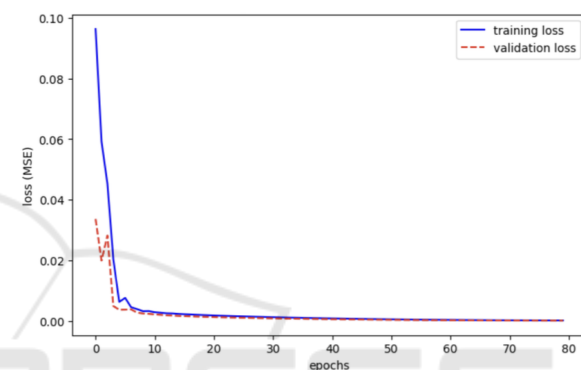


Figure 2: Training vs Validation Loss (RNN without Sentiment Scores).

- Learning Rate: 0.0005
- Batch Size: 64
- Number of Epochs: 120
- Criterion: SmoothL1Loss
- Optimiser: Adam

As illustrated in Figure 1, training and validation losses decrease sharply and converge, suggesting that the model learns effectively without overfitting.

The RNN was also trained on the dataset excluding sentiment scores. The best hyperparameters differed slightly:

- Number of Neurons in the RNN Layers: 64
- Number of Stacked RNN Layers: 3
- Learning Rate: 0.0005
- Batch Size: 64
- Number of Epochs: 80
- Criterion: MSELoss
- Optimiser: Adam

As shown in Figure 2, the loss curves for training and validation remain closely aligned, indicating stable learning.

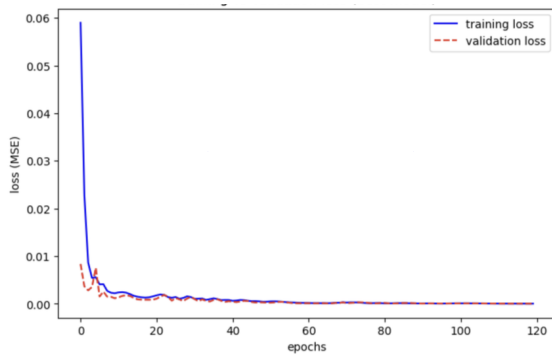


Figure 3: Training vs Validation Loss (LSTM with Sentiment Scores).

3.4.2 LSTM Architecture and Training

The LSTM architecture follows a similar structure, with stacked LSTM layers replacing the RNN cells. When trained on the sentiment-enriched dataset, the optimal configuration was found to be:

- Number of Neurons in the LSTM Layers: 128
- Number of LSTM Layers: 2
- Learning Rate: 0.0007
- Batch Size: 32
- Number of Epochs: 120
- Criterion: Smooth1Loss
- Optimiser: Adam

The training and validation losses remained very low and closely aligned throughout (Figure 3), indicating efficient learning without signs of overfitting.

The LSTM model was then trained on the dataset without sentiment scores, using a slightly deeper configuration:

- Number of Neurons in the LSTM Layers: 128
- Number of LSTM Layers: 3
- Learning Rate: 0.0007
- Batch Size: 32
- Number of Epochs: 120
- Criterion: Smooth1Loss
- Optimiser: Adam

Figure 4 shows a similar pattern of effective learning, with convergence in both training and validation losses within the first 10 to 15 epochs.

3.4.3 Quantitative Evaluation on Full Test Set

The LSTM model incorporating sentiment scores achieved the best overall performance, with an RMSE of 2.92 and an R^2 of 0.9987. The inclusion of sentiment features appears to have provided the model

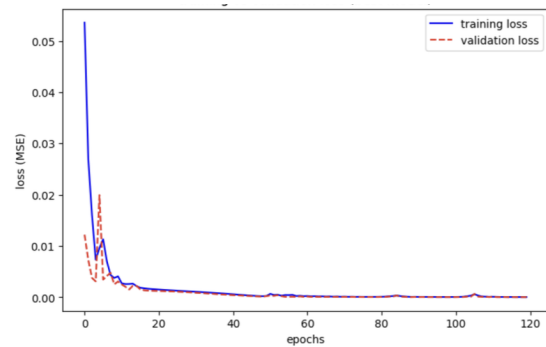


Figure 4: Training vs Validation Loss (LSTM without Sentiment Scores).

with informative context that improves its ability to capture market dynamics, as outlined in Table 1.

3.4.4 Prediction and Real-World Interpretation

Since the models were trained on normalised values, the predicted outputs were inverse-transformed to obtain stock prices in real-world scale. Figures 5 and 6 display the predictions from the LSTM models with and without sentiment scores, respectively. While both track the general trend well, the sentiment-informed model more closely follows abrupt price movements.

3.4.5 Robustness in Volatile Market Conditions

To further assess robustness, all models were evaluated on a volatile subset of the test data (days 80 to 130), which featured sharp price fluctuations. Results presented in Table 2 indicate a marked drop in accuracy for models without sentiment scores.

For instance, the RMSE of the sentiment-informed LSTM model was 4.06, compared to 9.86 for the LSTM without sentiment, suggesting that sentiment features help mitigate error under high volatility. Figure 7 illustrates that the model trained with sentiment features is more responsive to rapid changes in market conditions.

4 SUMMARY AND RELATED WORK

In this section, we summarise the main findings and limitations of the study and also discuss relevant work in the existing literature in order to contextualise our contributions.

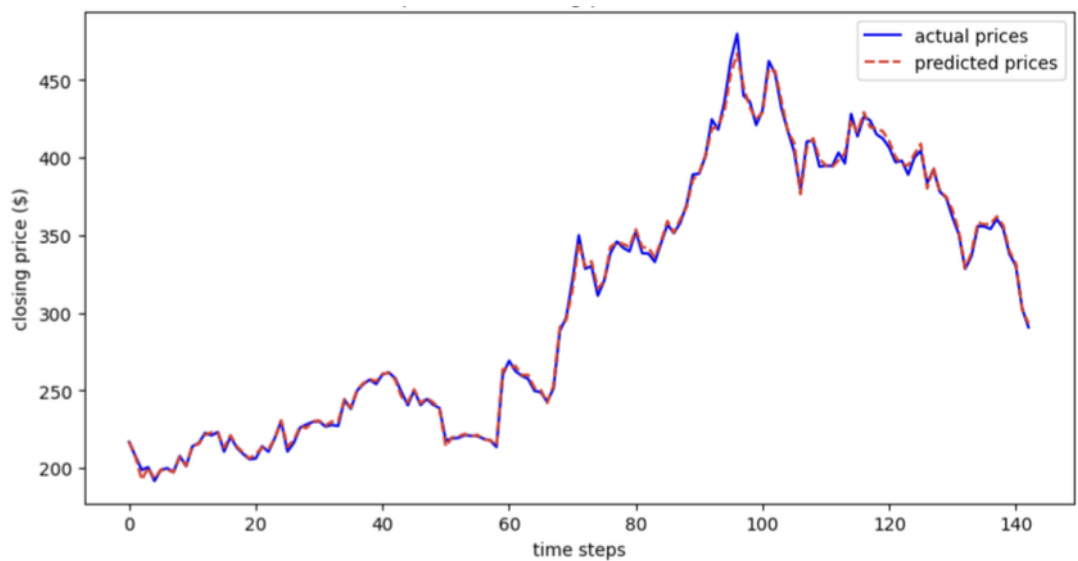


Figure 5: Predicted Closing Price, LSTM including Sentiment Scores.

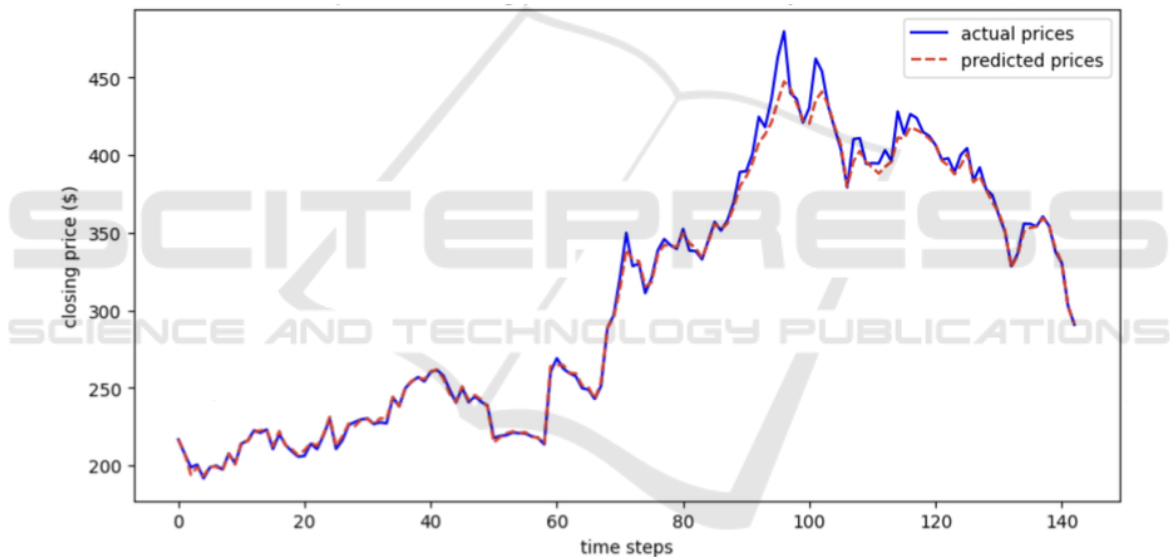


Figure 6: Predicted Closing Price, LSTM excluding Sentiment Scores.

Table 1: Model Performance Comparison.

Model	MSE	RMSE	MAE	R^2 Score
RNN (Sentiment Included)	10.14	3.18	2.17	0.9985
RNN (Sentiment Excluded)	46.56	6.82	4.36	0.9932
LSTM (Sentiment Included)	8.54	2.92	2.16	0.9987
LSTM (Sentiment Excluded)	35.24	5.93	3.12	0.9948

4.1 Summary of Insights and Threats to Validity

The findings suggest that incorporating ESG senti-
ment improves both prediction accuracy and model

responsiveness in volatile scenarios. However, sev-
eral caveats and potential threats to validity should be
acknowledged:

Data Scope: The dataset spans three years, which
may not fully capture long-term economic cycles or

Table 2: Model Performance Comparison on Volatile Subset.

Model	MSE	RMSE	MAE	R^2 Score
RNN (Sentiment Included)	22.84	4.77	3.66	0.9800
RNN (Sentiment Excluded)	122.58	11.07	8.94	0.8926
LSTM (Sentiment Included)	16.43	4.06	3.32	0.9856
LSTM (Sentiment Excluded)	97.38	9.86	6.37	0.9147

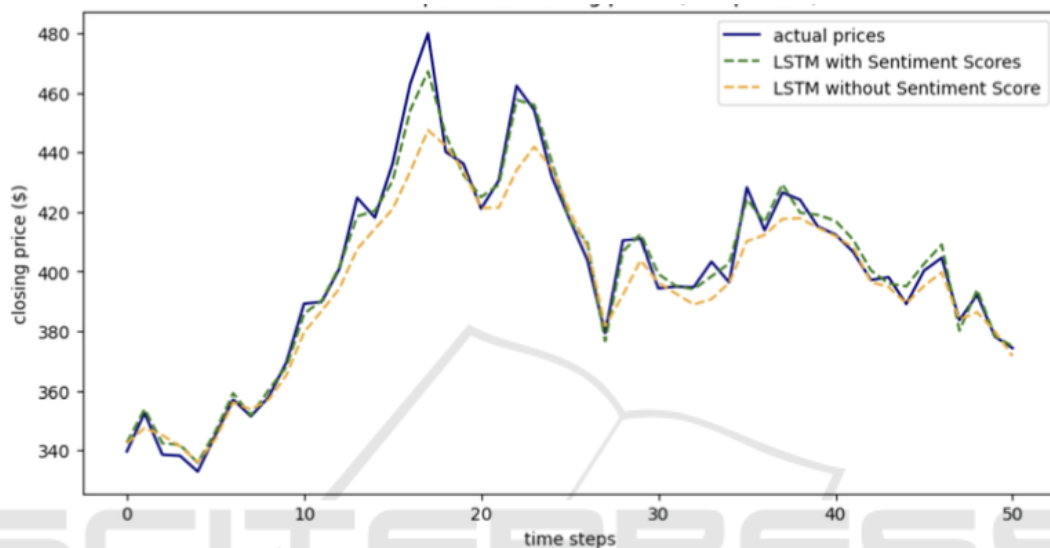


Figure 7: LSTMs Comparison in Volatile Period.

black swan events. Additionally, it is limited to a single stock (Tesla), which could affect the generalisability of the results. It should also be noted that it is limited to GNews API only.

Sentiment Representation: Sentiment scores are aggregated daily and rely on keyword-matched articles. This method might miss contextual nuances or important news items not captured by predefined search terms.

Model Bias and Overfitting: Although validation losses are closely aligned with training losses, overfitting cannot be completely ruled out, particularly due to hyperparameter tuning based on grid search.

Interpretability: While LSTM models provide excellent performance, their black-box nature limits interpretability. Future studies might consider combining LSTM with attention mechanisms or SHAP analysis to uncover underlying decision processes.

Market Noise: Stock prices are influenced by numerous exogenous variables (e.g., macroeconomic policies, investor sentiment, regulation), which are not incorporated in the current model.

Despite these limitations, the results show promis-

ing directions for integrating ESG sentiment into financial prediction tasks. Enhancing sentiment analysis with more fine-grained metrics or incorporating multi-source data (e.g., social media, earnings reports) could lead to more robust predictive frameworks. Overall, sentiment-informed models not only outperform their counterparts on the full test set, but also generalise better under volatile conditions. This supports the hypothesis that green finance sentiment serves as a valuable predictor when stock prices are influenced by ESG-related narratives.

While improvements may appear incremental in aggregate metrics, the ability to respond effectively to market instability is crucial in real-world trading environments. Incorporating such sentiment signals into forecasting systems may enhance risk-adjusted performance for ESG-sensitive stocks like Tesla.

4.2 Related Works

Researchers report that models are more accurate when sentiment data and technical metrics are combined (Das et al., 2024). Banerjee *et al.* (Banerjee et al., 2025) used ESG and investors' sentiment from

X (Twitter) with historical prices to develop a machine learning model to predict Dow Jones Industrial Average (DJIA) and S&P 500 companies, and this integration improved the accuracy. Additionally, Lee *et al.* (Lee et al., 2024) used the same approach by integrating sentiment scores from ESG-related news into technical indicators. They fed them into a Bi-LSTM model to improve the prediction accuracy of the S&P 500 index.

Mohan *et al.* (Mohan et al., 2019) developed a model to predict the S&P 500 companies' stock prices by sentiment analysis, but it did not perform well when the market was volatile. On the other hand, Guo *et al.* (Guo et al., 2020) found that ESG-related news has a significant effect on predicting market volatility. They showed that integrating ESG news into a deep learning model (ESG2Risk) increases the market's volatility prediction accuracy.

5 CONCLUSION AND RESEARCH DIRECTIONS

In this study, a comprehensive deep learning framework for stock price prediction using RNN and LSTM architectures was presented. The models were trained and evaluated on two datasets: one containing Tesla's historical stock data in combination with sentiment scores derived from green financial news using the FinBERT model, and another containing only the historical stock data. In addition, a subset of the test data characterised by high volatility was selected to investigate the robustness of the models under uncertain market conditions.

The results show that although the sentiment signals of ESG-related news do not provide significant improvements in stable market phases, they gain considerably in value when market volatility increases. In particular, the LSTM model trained on data with sentiment data achieved a root mean square error (RMSE) of 4.06, compared to 9.86 for the same model trained without sentiment input. Considering that Tesla's share price fluctuated between \$200 and \$480 during the period analysed, with the volatile subset between \$340 and \$480, this improvement is financially significant and demonstrates the potential of sentiment data to improve prediction performance in dynamic market environments.

This research is at the intersection of financial time series forecasting, natural language processing (NLP) and deep learning. However, it is important to acknowledge some limitations that may have influenced the results. First, the analysis focused exclusively on Tesla, which limits the generalisability of

the results to broader ESG-sensitive sectors or other companies. Second, the sentiment data was sourced through a single API provider (GNews), which is constrained by subscription costs and a limited query length. This may have led to incomplete coverage of the sentiment landscape. In addition, social media platforms such as X (formerly Twitter) and Reddit, which often reflect investor sentiment in real time, were not included in the sentiment analysis.

Another limitation relates to the capabilities of the FinBERT model itself. Due to the maximum token input limit of 512, only the first sections of each article were analysed, potentially excluding relevant sentiment cues from later sections of the text. In terms of financial data, the models relied solely on daily closing prices and disregarded potentially informative macroeconomic indicators, company-level financial data or technical indicators. Finally, the static nature of the sentiment scores (calculated once per article and aggregated daily) may have limited their responsiveness to the dynamics of the news in real time.

Future studies could address these limitations in several ways. First, the scope of the analysis can be expanded to include multiple companies from different industries with varying levels of ESG engagement, allowing for cross-comparison and more robust generalisation. In addition, the inclusion of social media sentiment, particularly on platforms such as X and Reddit, can provide a more immediate picture of public opinion and market psychology. The use of more advanced NLP models, e.g. transformer-based architectures such as GPT, T5 or DeBERTa, could potentially capture deeper semantic representations and longer dependencies in textual data, thus improving the accuracy of sentiment classification.

Furthermore, although our LSTM model showed a marked reduction in RMSE during volatile periods, future iterations can include rigorous statistical testing (e.g., Diebold-Mariano tests) to confirm the significance of these improvements.

In addition to improving the sentiment pipeline, the inclusion of real-time financial indicators such as trading volumes, volatility indices and macroeconomic announcements could improve predictive power. Real-time deployment and backtesting of these models in live trading environments would also provide practical insights into their operational viability and risk-adjusted performance.

In addition, future work could explore more granular sentiment representations, such as fine-grained polarity, topic-specific sentiment or emotion classification, to capture subtle market signals. Exploring the utility of alternative architectures, including gated recurrent units (GRUs), attention mechanisms,

and hybrid models that combine time-series encoders with transformers, could also lead to performance improvements. Finally, expanding sentiment data beyond ESG issues to general economic, technological or geopolitical news could provide a broader perspective on market drivers, especially in times of systemic uncertainty.

REFERENCES

- Al-Selwi, S. M., Hassan, M. F., Abdulkadir, S. J., Muneer, A., Sumiea, E. H., Alqushaibi, A., and Ragab, M. G. (2024). RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. *Journal of King Saud University - Computer and Information Sciences*, 36(5):102068.
- Allcott, H., Gentzkow, M., and Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research and Politics*, 6(2):2053168019848554.
- Banerjee, S., Aggarwal, D., and Sengupta, P. (2025). Do stock markets care about esg and sentiments? impact of esg and investors' sentiment on share price prediction using machine learning. *Annals of Operations Research*, pages 1–40.
- Berrar, D. (2019). Bayes' theorem and naive bayes classifier. In Ranganathan, S., Gribskov, M., Nakai, K., and Schönbach, C., editors, *Encyclopedia of Bioinformatics and Computational Biology*, pages 403–412. Elsevier.
- Bhavsar, H. and Panchal, M. H. (2012). A review on support vector machine for data classification. *International Journal of Advanced Research in Computer Engineering and Technology (IJARCET)*, 1(10):185–189.
- Chen, T. (2024). Sentiment analysis in green finance with llms. In *Proceedings of the 8th International Conference on Economics Management and Green Development*.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Das, N., Sadhukhan, B., Chatterjee, R., and Chakrabarti, S. (2024). Integrating sentiment analysis with graph neural networks for enhanced stock prediction: A comprehensive survey. *Decision Analytics Journal*, 10:100417.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding.
- Egorova, A. A., Grishunin, S. V., and Karminsky, A. M. (2022). The impact of esg factors on the performance of information technology companies. *Procedia Computer Science*, 199:339–345.
- Guo, T., Jamet, N., Betrix, V., Piquet, L. A., and Hauptmann, E. (2020). Esg2risk: A deep learning framework from esg news to stock volatility prediction. *arXiv preprint arXiv:2005.02527*.
- Kumar, S., De, K., and Roy, P. P. (2020). Movie recommendation system using sentiment analysis from microblogging data. *IEEE Transactions on Computational Social Systems*, 7(4):915–923.
- Lee, H., Kim, J. H., and Jung, H. S. (2024). Deep-learning-based stock market prediction incorporating esg sentiment and technical indicators. *Scientific Reports*, 14(1):10262.
- Liu, H., Cheng, J., Cao, J., and Katib, I. (2024). Preassigned-time synchronization for complex-valued memristive neural networks with reaction–diffusion terms and Markov parameters. *Neural Networks*, 169:520–531.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., and Anastasiu, D. C. (2019). Stock price prediction using news sentiment analysis. In *Proceedings of the 2019 IEEE 5th International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 205–208, Newark, CA, USA.
- Pagano, M. S., Sinclair, G., and Yang, T. (2018). Understanding esg ratings and esg indexes. In *Research Handbook of Finance and Sustainability*, pages 339–371. Edward Elgar Publishing.
- Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar.
- Qaiser, S. and Ali, R. (2018). Text mining: Use of tf-idf to examine the relevance of words to documents. *International Journal of Computer Applications*, 181(1):25–29.
- Rudianto, P., Darmono, D., and Tontowi, A. J. (2024). Sustainable investment strategies: Analyzing the interplay between esg ratings, dividend policy, financial performance, and investor trust. *International Journal of Business, Law and Education*, 5(2):2892–2901.
- Varadharajan, V., Smith, N., Kalla, D., Kumar, G. R., Samaah, F., and Polimetla, K. (2024). Stock Closing Price and Trend Prediction with LSTM-RNN. *Journal of Artificial Intelligence and Big Data*, 4(1):1–13.
- Yang, Y., UY, M. C. S., and Huang, A. (2020). Finbert: A pretrained language model for financial communications.