Proceedings of ICFTBA 2025 Symposium: Financial Framework's Role in Economics and Management of Human-Centered Development
DOI: 10.54254/2754-1169/2025.GL28719

Speculation in Tech Stocks: Investigating Whether Incorporating User Sentiment and Lagged Effects Enhances Short-Term Tech Stock Trends

Jiayin Lai

Shanghai Jianping High School, Shanghai, China
lplusy1010@gmail.com

Abstract. Economic factors have become increasingly pivotal globally, driving significant changes in stock market trends. However, inaccurate analysis can lead to missed opportunities and poor decisions. This research explores whether user sentiment and lagged effects, such as past stock price movements and delayed sentiment impact, enhance short-term tech stock trends prediction. The objective is to develop classification models that predict stock price increases and decreases, improving model's performance through sentiment analysis and key stock market features, while identifying the most effective model for this task. This study analyzes four major U.S. tech companies (Apple, Amazon, Microsoft and Tesla) using GPT-3.5 Turbo for sentiment analysis and XGBoost, LSTM, and GRU for prediction. Results demonstrate that integrating sentiment and lagged features enhances predictive performance for some models. LSTM achieves the highest overall improvement, as it is particularly effective at capturing temporal dependencies in sequential data. While XGBoost and GRU show improvements for certain companies, their results remain relatively unstable. These findings highlight the effectiveness of sentiment and lagged features emerges in certain conditions, with LSTM regarding as the most robust model for short-term stock trend forecasting.

Keywords: Machine learning, Sentiment Analysis, Lagged features, Natural Language Processing, Stock Price Prediction

1. Introduction

Owing to its volatility and the influence of multiple determinants, the financial market makes stock trend forecasting a persistent and intricate research problem. Traditional models primarily rely on quantitative data like price and volume, often overlooking qualitative signals such as investor sentiment and temporal dependencies in historical price patterns-known as lagged effects. Gradually, with advances in natural language processing (NLP), new opportunities have emerged to improve prediction accuracy by integrating sentiment from unstructured text. This study explores whether combining user sentiment and lagged features with machine learning and deep learning models can enhance short-term stock trend prediction, particularly for major technology firms.

DOI: 10.54254/2754-1169/2025.GL28719

To explore this question, it is essential to understand how machine learning models have evolved and improved to handle complex stock market data. Conventional time series approaches, such as the AutoRegressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, often face difficulties when handling the nonlinear characteristics of financial markets [1,2]. Recently, combining sentiment analysis with machine learning and deep learning has opened new avenues, leveraging news and social media sentiment to better capture market shifts and enhance prediction accuracy. The earliest application of sentiment analysis was pioneered by Bollen et al., who used sentiment scores as independent variables and closing prices as the dependent variable in regression analysis. Their study achieved a minimum mean squared error of 1.79% and a highest directional accuracy of 86.7%. This groundbreaking work, based on sentiment lexicons and simple classification, revealed a significant correlation between Twitter sentiment fluctuations and market volatility, laying the foundation for financial sentiment analysis [3].

Subsequent studies explored more complex methods to address these limitations. Unlike lexicon-based methods, according to the study done by Deveikyte et al., they used Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT) models capable of capturing complex sentiment patterns, achieving an average accuracy of 65% [4]. Darapaneni et al. applied two models: LSTM using historical prices and a Random Forest based on sentiment analysis combined with macroeconomic factors. They used them to predict Indian stock market. This attempt aimed to forecast overall price trends rather than a single day’s price [5]. Talazadeh and Peraković proposed the Sentiment-Augmented Random Forest (SARF), integrating FinGPT sentiment analysis with Random Forest. SARF improved accuracy by 9.23% and reduced errors compared to traditional Random Forest and LSTM models, demonstrating sentiment’s value in handling market emotions [6].

As research deepens, many researchers began to combine various sentiment analysis tools to improve stock market prediction accuracy, which was known as multimodal sentiment analysis. For example, Li and Ming proposed a Deep Q-Learning Model (DQS) that successfully integrated sentiment analysis with reinforcement learning. This model achieved an increase in returns in portfolio management, and all five machine learning classifiers showed almost perfect accuracy after in-depth tuning of model parameters in terms of binary classification [7].

Besides some conventional and mainstream models and methods, many scholars applied Gated Recurrent Unit (GRU), which is a simpler form of LSTM, into their studies. For example, Mamillapalli et al. proposed a model named GRU_vader, which integrates lexicon-based sentiment analysis, such as Valence Aware Dictionary and sEntiment Reasoner (VADER) with a GRU neural network, aiming to predict stock prices. Findings demonstrated that integrating sentiment analysis considerably boosted the predictive accuracy of stock models, particularly for GRU_vader. This is a relatively prestigious and reliable research which used GRU and sentiment analysis [8]. However, for XGBoost, few researches applied this model in stock market prediction. Shi et al. integrated a hybrid model in their study, including an attention-based CNN-LSTM and XGBoost. Finally, their results showed that the hybrid model is more effective and the prediction accuracy is relatively high, which affirmed the effectiveness of both LSTM and XGBoost [9].

While combining sentiment analysis with machine learning models has proven valuable in capturing market mood, lagged effect appeared as an important factor as well. For instance, Patel et al. incorporated lagged features into machine learning models to better reflect historical stock fluctuations. These features help capture time dependencies in price behavior, offering a more holistic view of market trends [10]. Expanding on this, Moews at al. developed a deep learning

DOI: 10.54254/2754-1169/2025.GL28719

framework that uses lagged correlations from historical data to predict directional trend shifts. Their experiment showed that including lagged features significantly boosted forecasting accuracy, especially in volatile markets [11]. Extending this line of inquiry, Moews and Ibikunle investigated intraday lagged correlations within the S&P 500 index using deep learning architectures. They trained models on short-horizon financial time series and demonstrated that learning from lagged intraday correlations improved prediction robustness in both calm and volatile market states [12]. In a related effort, Curme et al. explored the "lead–lag" effect in stock markets from both a theoretical and empirical perspective. Their study introduced a framework to quantify lead–lag patterns among stocks, revealing that such patterns exhibit temporal consistency and can be systematically exploited [13].

In conclusion, while previous studies have highlighted the individual predictive potential of sentiment analysis and lagged price features, a relatively small number of essays examined their combined effect on stock trend forecasting. Building upon this foundation, the present study investigates whether integrating sentiment signals derived from financial texts and historical lagged features can enhance the accuracy of short-term trend predictions for leading technology firms such as Microsoft, Amazon, Tesla and Apple. Sentiment scores are obtained by using advanced NLP techniques and combined with time-dependent stock data. After that, they are input into XGBoost, LSTM, and GRU models to compare their accuracy by different indicators. The following section outlines the dataset, preprocessing steps, feature engineering pipeline, model architecture, and evaluation metrics employed in this research.

2. Methodology

2.1. Data resource

This dataset used in this study was published on Kaggle by H. Yukhymenko, titled Stock Tweets for Sentiment Analysis and Prediction. Collected between September 30^th, 2021, and September 30^th, 2022, the dataset covers tweets for the 25 most-followed stock tickers on Yahoo Finance, along with associated stock price and trading volume information. The dataset owned an average sample quantity of 253 financial information samples per firm, and the target variable was set to be Price_Up, to determine whether the stock market will rise or fall on the next day. To ensure the representativeness and quality of the research data, this study selected four large-cap and highly discussed technology companies, including Microsoft, Amazon, Tesla and Apple [14].

2.2. Variable introduction

The variables used in this study were categorized into four groups: market variables, sentiment variables, time-related variables, and the target variable. In the original dataset, market variables included basic price indicators such as Open (the stock's opening price), High (the highest price of the day), Low (the lowest price), Close (closing price), Adj Close (adjusted closing price accounting for dividends and splits), and Volume (trading volume). Their ranges are demonstrated in Table 1. Additional statistical features were derived from price data to capture short-term dynamics, including rolling_mean and rolling_std (10-day moving average and standard deviation of closing prices), max_price and min_price (maximum and minimum prices over a rolling window), volume_mean and volume_std (mean and standard deviation of volume), as well as Q1, Q3, IQR (first quartile, third quartile, and interquartile range), and skewness (distribution asymmetry of recent closing prices). Sentiment variables were generated using GPT-3.5 Turbo that assigned a

DOI: 10.54254/2754-1169/2025.GL28719

continuous sentiment_score to each user review. To account for delayed sentiment impact, lagged versions of the sentiment score, including sentiment_score_LAG1, sentiment_score_LAG2, and sentiment_score_LAG3, which represented sentiment scores lagged by one, two, and three trading days, respectively. Time-related variables mainly consisted of lagged price features, including Close_LAG1, Close_LAG2, and Close_LAG3. corresponding to the closing prices lagged by one, two, and three trading days. Finally, the target variable Price_Up is a binary indicator that denotes whether the stock price increased on a given day-with 1 representing an increase, and 0 indicating no change or a decrease. This variable served as the output for binary classification models used in the predictive task.

Table 1. Distribution range of original variables by company

Stock Name	Open (USD)	High (USD)	Low (USD)	Close (USD)	Adj Close (USD)	Volumn (Shares)
AAPL	130.070-182.630	132.390-182.940	129.040-179.120	130.060-182.010	129.664-180.960	41000000-195432700
AMZN	102.750-185.635	104.580-188.107	101.260-183.786	102.310-184.803	102.310-184.803	35754000-272662000
MSFT	236.810-344.620	239.950-349.670	234.410-342.200	236.410-343.110	235.746-339.925	15042000-90428900
TSLA	207.950-411.470	217.973-414.497	206.857-405.667	209.387-409.970	209.387-409.970	35042700-188556300

2.3. Method introduction

To evaluate the impact of user sentiment and lagged effects on short-term stock trend prediction, this study constructed a structured modeling framework. The float chart is demonstrated below (Figure 1). After integrating stock price data with processed sentiment scores and temporal features, a multi-phase procedure—consisting of feature engineering, model training, and evaluation—was performed. Lagged features were created for both market and sentiment variables to capture potential delayed reactions from retail investors and market participants.

This study applied three types of machine learning models—XGBoost, LSTM, and GRU. A core innovation of this methodology was the combination of language model–derived sentiment data (via GPT-3.5 Turbo) with financial time-series information. This integration enabled the models to exploit both behavioral signals and historical market patterns.

During model training, in order to minimize this risk, regularization is essential. In this study, both L_1 (Lasso) and L_2 (Ridge) regularization methods were applied to counter the overfitting caused by these resampling techniques. L_1 regularization encourages feature selection by forcing some coefficients to zero, effectively reducing the number of features used in the model. Meanwhile, L_2 regularization penalizes large coefficients, promoting smaller, more generalized parameter values. By combining these two regularization methods, the model benefits from feature sparsity and reduced complexity, which enhances its ability to generalize and perform well on unseen data.

In optimizing the XGBoost, LSTM, and GRU models, Grid Search method was used for parameter optimization. Initially, manual hyperparameter tuning was performed, adjusting key parameters such as the number of decision trees, learning rate, and tree depth for XGBoost, and the number of layers, units per layer, and dropout rate for LSTM and GRU. Then, to systematically

DOI: 10.54254/2754-1169/2025.GL28719

search for the best parameters, Grid Search was applied, exploring all possible parameter combinations within the pre-set ranges.

Model performance was evaluated using precision, recall, F1-score, and AUC to assess classification effectiveness, and the most effective model was identified. By systematically integrating multi-modal lagged features, the method aimed to reveal whether sentiment dynamics and delayed effects could significantly improve prediction outcomes over traditional models relying solely on historical prices.

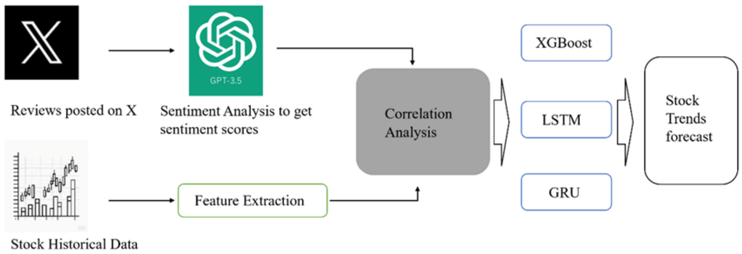

Figure 1. Overview of stock trends prediction processes

3. Results and discussion

3.1. Data pre-processing and sentiment analysis

The dataset was preprocessed by handling missing values using median imputation, removing duplicate records, and treating outliers via the IQR method. Outliers were identified and replaced with the median to preserve data distribution and reduce bias, ensuring model stability and reliable analysis.

Sentiment scores were generated using GPT-3.5 Turbo, with scores categorized into seven ranges: Very Positive ($0.8 \leq$ score ≤ 1.0), Positive ($0.5 \leq$ score < 0.8), Slightly Positive ($0.2 \leq$ score < 0.5), Neutral ($-0.2 \leq$ score < 0.2), Slightly Negative ($-0.5 \leq$ score < -0.2), Negative ($-0.8 \leq$ score < -0.5), and Very Negative ($-1.0 \leq$ score < -0.8), which provided more granular sentiment categories (Fig. 2). A prompt was designed in order to derive a more accurate result. Daily average sentiment scores per company were computed from X platform posts. Lagged features for sentiment scores and stock prices at 1, 2 and 3-day intervals were created to capture delayed market effects. Numerical features, including prices, volumes, and sentiment scores etc., were normalized using Min-Max scaling to ensure consistent feature ranges, critical for stable neural network training. Missing values were imputed with median values.

DOI: 10.54254/2754-1169/2025.GL28719

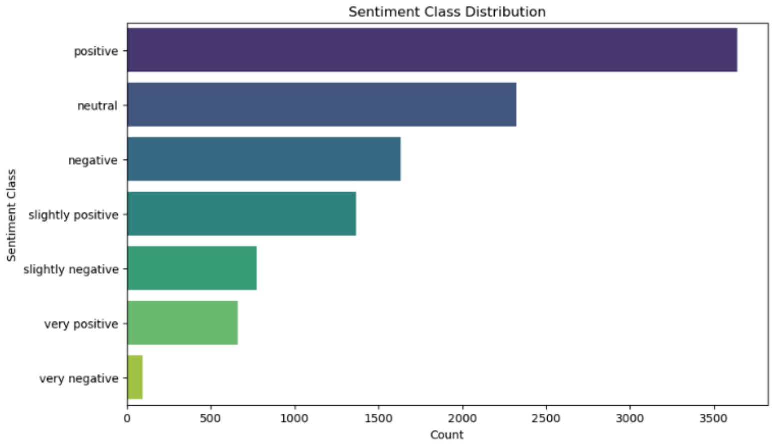

Figure 2. Sentiment class distribution

3.2. Correlation results

Correlation heatmaps were plotted to calculate the correlation between various features for further analysis of the four companies involved. The analysis of these heatmaps showed that most sentiment-related variables, including Average Sentiment Score and its lagged versions, exhibited relatively weak linear correlations with price changes, as reflected by correlation coefficients well below 0.5. This suggests that, overall, sentiment features alone may not strongly influence short-term price movements in a linear fashion. Nonetheless, these visualizations provide useful insight into the structure of the data and help identify potential variables for further modeling. In the following training process, the selection of sentiment and lagged features was guided not solely by correlation strength but also by empirical improvements in model performance during preliminary testing. By iteratively evaluating different combinations of sentiment lags, we retained those that contributed positively to predictive accuracy. These refined features were subsequently integrated into the XGBoost, LSTM, and GRU models to examine their impact on short-term stock price prediction.

DOI: 10.54254/2754-1169/2025.GL28719

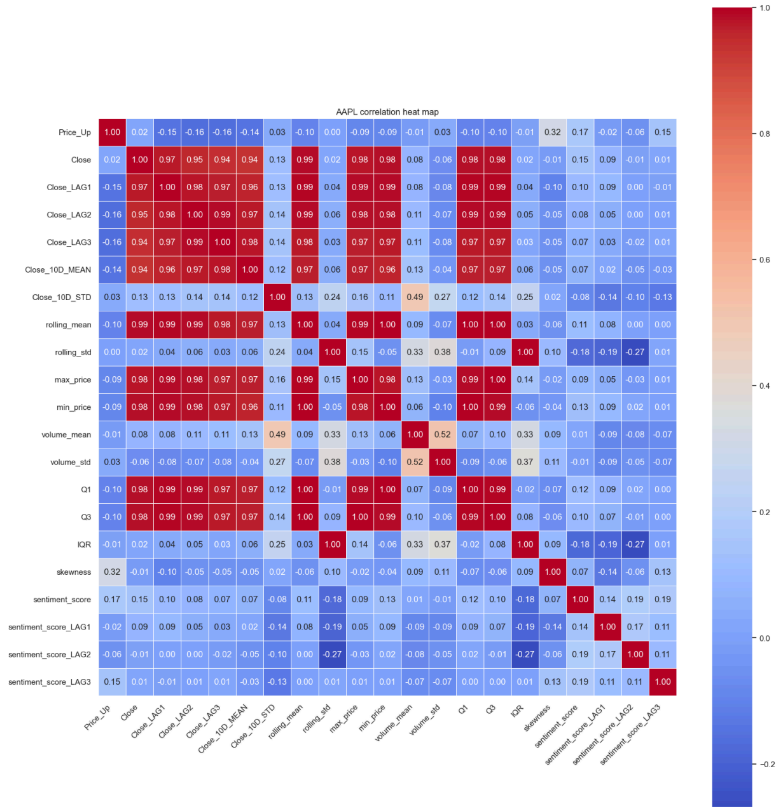

Figure 3. AAPL correlation heatmap

The correlation heatmap of Apple is illustrated in Figure 3. For a more detailed visualization of feature interactions of the other three companies (Amazon, Microsoft and Tesla), see Appendix.

3.3. Model performance

This study selected XGBoost, LSTM, and GRU as the core prediction models to explore the role of sentiment analysis and lagged effects in short-term stock price prediction. These three models each have advantages in handling time series data and are capable of capturing key factors that influence stock market fluctuations from different perspectives. The main objectives of this study were: (1) To assess whether incorporating sentiment scores and lagged variables can improve the predictive performance of the models, thereby validating the value of market sentiment and historical information in short-term stock price predictions. (2) To compare the performance of different models after integrating sentiment variables and lagged features, in order to identify the optimal prediction method. Model performance was evaluated using four key classification metrics: Precision, Recall, F1 score, and AUC (Area Under the ROC Curve). The formulas are given by:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

DOI: 10.54254/2754-1169/2025.GL28719

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

3.3.1. Xgboost

XGBoost was chosen for its ability to handle complex non-linear relationships and its robustness with tabular data. Its high performance in terms of both speed and accuracy, along with its feature selection capabilities, make it a powerful tool for capturing the intricate interactions between sentiment scores, lagged effects, and stock price movements. XGBoost is particularly effective at handling imbalanced datasets and managing missing data, which are common challenges in financial data analysis.

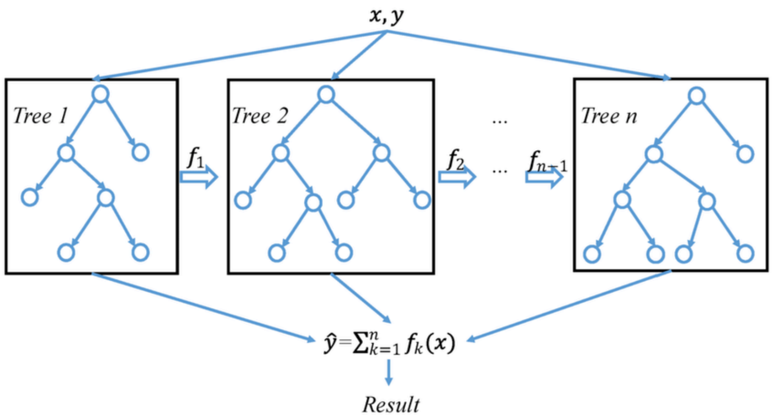

Figure 4. The general architecture of XGBoost [15]

For each iteration, the loss function is defined as:

$$L(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \Omega(f) \tag{4}$$

$l(y_i, \hat{y}_i)$ is the loss function that represents the error between the true value y_i and the predicted value $\hat{y}_i$. Where $\Omega(f)$ is the regularization term, which prevents overfitting. In each iteration, the model models the residuals of the previous round and updates the weak learner's parameters. The final prediction is defined by:

$$\hat{y} = \sum_{t=1}^{T} \eta_t \cdot h_t(x) \tag{5}$$

Where $h_t(x)$ is the prediction of the t-th tree, and η_t is the learning rate of the t-th tree.

DOI: 10.54254/2754-1169/2025.GL28719

Table 2. Results of XGBoost for stock trends prediction

	Data without sentiment and lagged features				Data with sentiment and lagged features			
	Precision	Recall	F1 score	AUC	Precision	Recall	F1 score	AUC
AAPL	0.571	0.667	0.615	0.643	0.744	0.674	0.707	0.749
AMZN	0.600	0.667	0.632	0.706	0.676	0.694	0.685	0.736
MSFT	0.703	0.743	0.722	0.772	0.611	0.629	0.620	0.711
TSLA	0.552	0.821	0.660	0.614	0.653	0.821	0.727	0.778

The XGBoost model demonstrated significant improvement after incorporating sentiment scores and lagged features. Specifically, in the predictions for Apple and Tesla, the model achieved consistent improvements across all evaluation metrics (Precision, Recall, F1 score, AUC), with Apple's precision increasing from 0.571 to 0.744 and Tesla's AUC rising from 0.614 to 0.778, highlighting the importance of sentiment scores and lagged features in short-term stock price prediction. For Amazon, the performance of the XGBoost model also improved with the inclusion of sentiment scores and lagged features, particularly in terms of F1 score and AUC, showing a noticeable increase. For Microsoft, however, the model's performance declined after incorporating these features, with all metrics showing a drop. This result suggests that the inclusion of sentiment scores and lagged features helps the model better capture changes in market sentiment and historical stock price trends for certain companies, though their effectiveness may vary depending on stock-specific characteristics.

3.3.2. Keywords (use style: keywords)

LSTM was selected for its capacity to model sequential data and capture long-term dependencies, a crucial feature when predicting short-term stock trends. Stock prices are influenced by historical patterns, and LSTM's architecture allows it to learn from these sequences over time. This is particularly important when incorporating sentiment analysis, as past sentiment values can impact future stock price movements. LSTM is well-suited for time-series forecasting, which is key for this project's goal of predicting short-term stock trends.

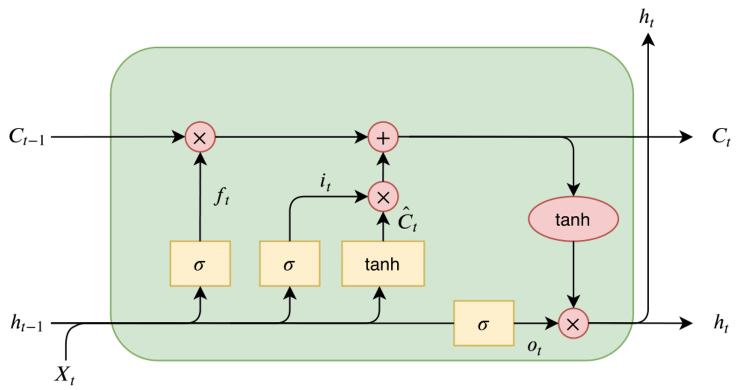

Figure 5. The general architecture of LSTM [16]

Forget Gate (f_t): Determines how much past information to discard.

DOI: 10.54254/2754-1169/2025.GL28719

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f) \tag{6}$$

Input Gate (i_t): Decides how much of the current input to update.

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \tag{7}$$

Candidate Memory Cell ($\widetilde{C}_t$): Determines the memory update.

$$\widetilde{C}_t = tanh\Big(W_C \bullet [h_{t-1}, x_t] + b_C\Big) \tag{8}$$

Memory Cell Update (C_t): Combines past memory and new input.

$$C_t = f_t \bullet C_{t-1} + i_t \bullet \widetilde{C}_t \tag{9}$$

Output Gate (o_t): Controls how much of the memory should be output.

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \tag{10}$$

Final Hidden State (h_t): Output the hidden state.

$$h_t = o_t \bullet \tanh(C_t) \tag{11}$$

Table 3. Results of LSTM for stock trends prediction

	Data without sentiment and lagged features				Data with sentiment and lagged features			
	Precision	Recall	F1 score	AUC	Precision	Recall	F1 score	AUC
AAPL	0.784	0.784	0.784	0.913	0.861	0.838	0.849	0.946
AMZN	0.806	0.784	0.795	0.913	0.750	0.811	0.779	0.857
MSFT	0.861	0.861	0.861	0.910	0.800	0.778	0.789	0.882
TSLA	0.878	0.900	0.889	0.950	0.868	0.825	0.846	0.946

The LSTM model demonstrated generally strong performance both before and after incorporating sentiment scores and lagged features. Specifically, in the predictions for Apple, the model achieved improvements across all evaluation metrics (Precision, Recall, F1 score, AUC), with precision increasing from 0.784 to 0.861 and AUC from 0.913 to 0.946, highlighting the value of sentiment scores and lagged features in short-term stock price prediction. For Tesla, the model also maintained high performance after incorporating these features, with only a slight decrease in recall but stable

DOI: 10.54254/2754-1169/2025.GL28719

AUC (from 0.950 to 0.946), suggesting that the added features preserved the model's predictive capability. For Amazon, however, the model showed a slight drop in precision and AUC after including the sentiment-related features, though recall improved. Similarly, for Microsoft, the performance declined across all metrics, indicating that the inclusion of sentiment scores and lagged features did not benefit the LSTM model for that stock. This result suggests that while sentiment scores and lagged features can enhance LSTM performance in some cases, their effectiveness may vary depending on company-specific dynamics and model sensitivity, and in some cases, may even introduce noise that hinders predictive accuracy.

3.3.3. Figures and tables

GRU was chosen due to its simplicity and computational efficiency in handling sequential data. Unlike LSTM, GRU has fewer parameters and can be trained faster while still maintaining good performance in modeling temporal dependencies. Given the relatively short-term nature of the stock trend predictions required in this study, GRU offers a balance between accuracy and efficiency, making it an ideal choice for this task.

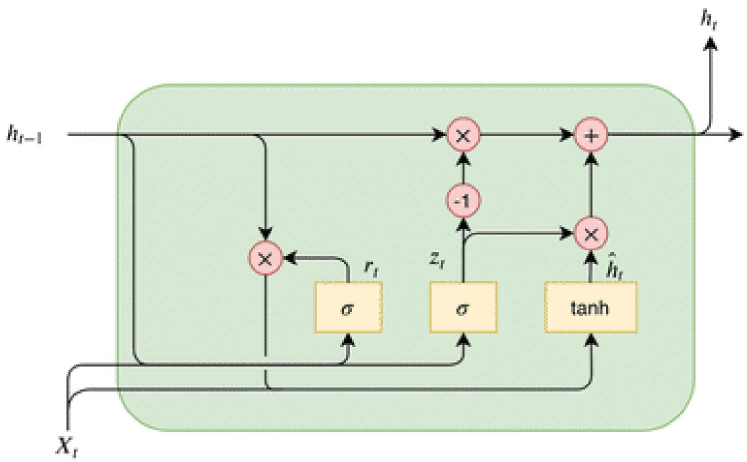

Figure 6. The general architecture of GRU [16]

Reset Gate (r_t): Decides whether to reset the previous hidden state.

$$r_t = \sigma(W_r \bullet [h_{t-1}, x_t] + b_r) \quad (12)$$

Update Gate (z_t): Controls how much of the previous hidden state to retain.

$$z_t = \sigma(W_z \bullet [h_{t-1}, x_t] + b_z) \quad (13)$$

Candidate Hidden State ($\widetilde{h}_t$): Represents a new candidate hidden state.

$$\widetilde{h}_t = tanh\Big(W_h \bullet [r_t \bullet h_{t-1}, x_t] + b_h\Big) \quad (14)$$

DOI: 10.54254/2754-1169/2025.GL28719

Final Hidden State (h_t): Combines the update gate and candidate hidden state.

$$h_t = (1 - z_t) \bullet h_{t-1} + z_t \bullet \widetilde{h}_t \tag{15}$$

Table 4. Results of GRU for stock trends prediction

	Data without sentiment and lagged features				Data with sentiment and lagged features			
	Precision	Recall	F1 score	AUC	Precision	Recall	F1 score	AUC
AAPL	0.775	0.838	0.805	0.797	0.763	0.784	0.773	0.770
AMZN	0.750	0.375	0.500	0.618	0.792	0.475	0.594	0.668
MSFT	0.800	0.606	0.690	0.745	0.760	0.576	0.655	0.718
TSLA	0.684	0.634	0.658	0.641	0.714	0.732	0.723	0.689

The GRU model exhibited mixed performance after incorporating sentiment scores and lagged features. Specifically, in the predictions for Tesla, the model achieved consistent improvements across all evaluation metrics (Precision, Recall, F1 score, AUC), with F1 score increasing from 0.658 to 0.723 and AUC from 0.641 to 0.689, suggesting that sentiment scores and lagged features helped improve predictive capability for this stock. For Amazon, the model also showed improvement, particularly in F1 score (from 0.500 to 0.594) and AUC (from 0.618 to 0.668), indicating a better balance between precision and recall after feature augmentation. However, for Apple and Microsoft, the inclusion of sentiment-related variables led to a slight decrease in performance across most metrics, with Apple's AUC dropping from 0.797 to 0.770 and Microsoft's from 0.745 to 0.718. This result suggests that while sentiment scores and lagged features may benefit GRU performance for certain stocks, their overall effectiveness is limited. Due to GRU's relatively weaker ability to capture long-term dependencies, these added features may also introduce noise and reduce model performance in some cases.

4. Conclusion

The comparison of the three models (XGBoost, LSTM, and GRU) showed that incorporating sentiment and lagged variables generally improved prediction accuracy, although the degree of improvement varied across models and companies. Among them, XGBoost demonstrated the most obvious improvement on most stocks, significantly outperforming the other models and showcasing its strong ability to capture nonlinear relationships and complex interactions. The LSTM model, while showing relatively modest improvements, already had a high baseline performance and maintained good stability and generalization after adding the new features. In contrast, the performance of GRU was relatively unstable; for some stocks, incorporating sentiment and lagged variables even led to performance degradation, indicating its weaker capability to model such features.

This disparity can largely be attributed to differences in model architecture. XGBoost, based on a tree ensemble method, excels at handling unstructured and nonlinear data, making it more capable of uncovering complex patterns behind stock price fluctuations. LSTM possesses powerful temporal memory capabilities, enabling it to better capture dynamic patterns in time series data and respond sensitively to lagged variables. As a simplified version of LSTM, GRU, although more

DOI: 10.54254/2754-1169/2025.GL28719

computationally efficient, has fewer parameters and thus limited expressive capacity. When faced with high-dimensional, nonlinear financial data, GRU may introduce noise instead, which adversely affects prediction performance.

Apart from the disparity among different models, this study offers new insights into financial market sentiment analysis and short-term stock prediction, making the following key contributions: (1) It used GPT-3.5 Turbo for a more delicate and sufficient sentiment analysis and compiled prompt words to improve accuracy. (2) It confirmed the critical role of market sentiment and lagged features in short-term tech stock predictions, addressing the gap in understanding how non-immediate factors influence stock trends and supporting feature selection for prediction models; (3) It systematically compared the performance of XGBoost, LSTM, and GRU in stock trends prediction, highlighting the advantages of XGBoost in nonlinear modeling and the potential of LSTM and GRU in capturing temporal dependencies, providing empirical evidence for model selection in practice; (3) It integrated sentiment analysis with machine learning and deep learning models, proposing a framework for short-term stock market forecasting based on sentiment analysis and lagged effects. These contributions not only extend current research but also offer new perspectives for investment decision-making. In particular, the findings offer new insights into how market sentiment can enhance the accuracy and interpretability of stock prediction models.

Despite the promising results, there are several limitations to this study. First, the sentiment data were sourced from social media, which may introduce noise and bias, potentially affecting the reliability of the predictions. This phenomenon is well-demonstrated in relatively unstable results of part of the companies. Future research could address this by incorporating multiple data sources, such as financial news and analyst reports, to enhance the accuracy and breadth of sentiment analysis. Second, this study focused on a limited number of tech companies, and industry-specific differences may influence the generalizability of the findings. Expanding the research to include more sectors and markets would provide further validation of the models' effectiveness. Moreover, macroeconomic factors (e.g., policy changes, market liquidity) were not incorporated in this study, even though they can significantly affect stock price movements. Future research could integrate these external variables to improve the model's comprehensiveness and accuracy. Overall, these findings not only enhance the methodological toolkit for financial time series forecasting but also provide practical guidance for investors and researchers seeking more reliable and interpretable stock market prediction models.

References

[1] Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time series analysis: Forecasting and control (3rd ed.). Holden-Day.
[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. https: //doi.org/10.1016/0304-4076(86)90063-1
[3] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1–8. https: //doi.org/10.1016/j.jocs.2010.12.007
[4] Deveikyte, J., Slezak, D., & Kalita, A. (2022). A sentiment analysis approach to the prediction of market volatility. Frontiers in Artificial Intelligence, 1, Article 836809. https: //doi.org/10.3389/frai.2022.836809
[5] Darapaneni, N., Bansal, A., & Kumar, S. (2022). Stock price prediction using sentiment analysis and deep learning for Indian markets. arXiv preprint arXiv: 2204.05783.
[6] Talazadeh, S., & Peraković, D. (2024). SARF: Enhancing stock market prediction with sentiment-augmented random forest. arXiv preprint arXiv: 2410.07143.
[7] Li, X., & Ming, Y. (2023). Stock market prediction using reinforcement learning with sentiment analysis. International Journal of Cybernetics and Informatics, 12(1). https: //doi.org/10.5121/ijci.2023.120101

DOI: 10.54254/2754-1169/2025.GL28719

[8] Zhang, Y., & Qiu, Y. (2024). GRUvader: Sentiment-informed stock market prediction. Mathematics, 12(23), Article 3801. https: //doi.org/10.3390/math12233801
[9] Chen, C., Xue, L., & Xing, W. (2023). Research on improved GRU-based stock price prediction method. Applied Sciences, 13(15), Article 8813. https: //doi.org/10.3390/app13158813
[10] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259–268. https: //doi.org/10.1016/j.eswa.2014.10.031.
[11] Moews, B., Herrmann, J. M., & Ibikunle, G. (2018). Lagged correlation-based deep learning for directional trend change prediction in financial time series. arXiv preprint arXiv: 1811.11287.
[12] Moews, B., & Ibikunle, G. (2020). Predictive intraday correlations in stable and volatile market environments: Evidence from deep learning. arXiv preprint arXiv: 2002.10385.
[13] Curme, C., Tumminello, M., Mantegna, R. N., Stanley, H. E., & Kenett, D. Y. (2014). Emergence of statistically validated financial intraday lead-lag relationships. arXiv preprint arXiv: 1401.0462.
[14] Yukhymenko, H. (2021). Stock tweets for sentiment analysis and prediction. Kaggle. https: //www.kaggle.com/datasets/equinxx/stock-tweets-for-sentiment-analysis-and-prediction
[15] Wang, Y., Pan, Z., Zheng, J., Qian, L., & Li, M. (2019). A hybrid ensemble method for pulsar candidate classification. Astrophysics and Space Science, 364(8), Article 139. https: //doi.org/10.1007/s10509-019-3602-4
[16] Ingolfsson, T. M. (2021, November 10). Insights into LSTM architecture: Finding the total number of multiply and accumulate operations in a LSTM layer. Thorir Mar. https: //thorirmar.com/post/insight_into_lstm/

Appendix

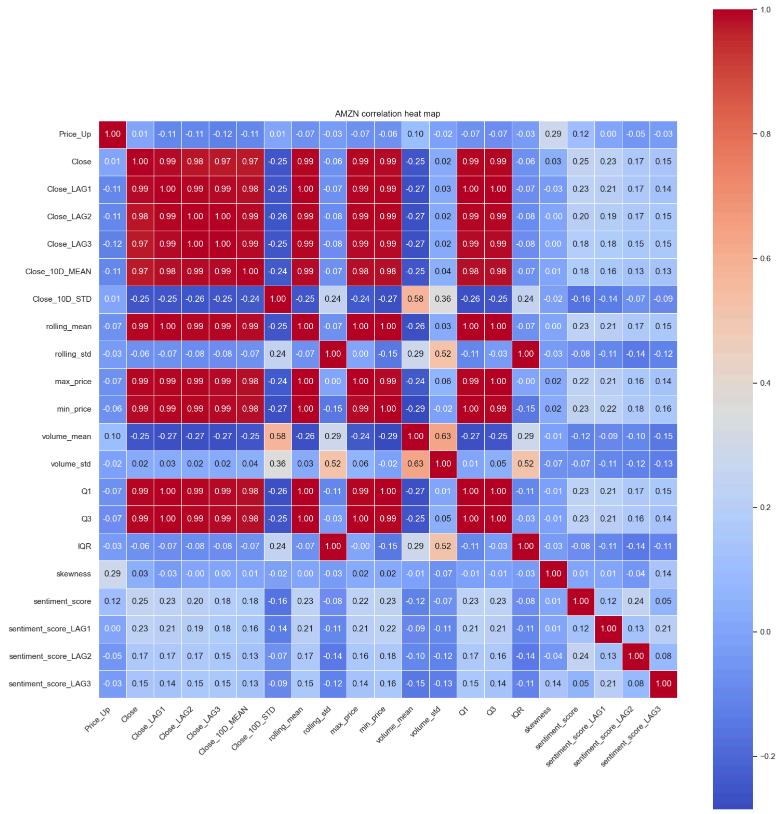

Figure 7. AMZN correlation heatmap

DOI: 10.54254/2754-1169/2025.GL28719

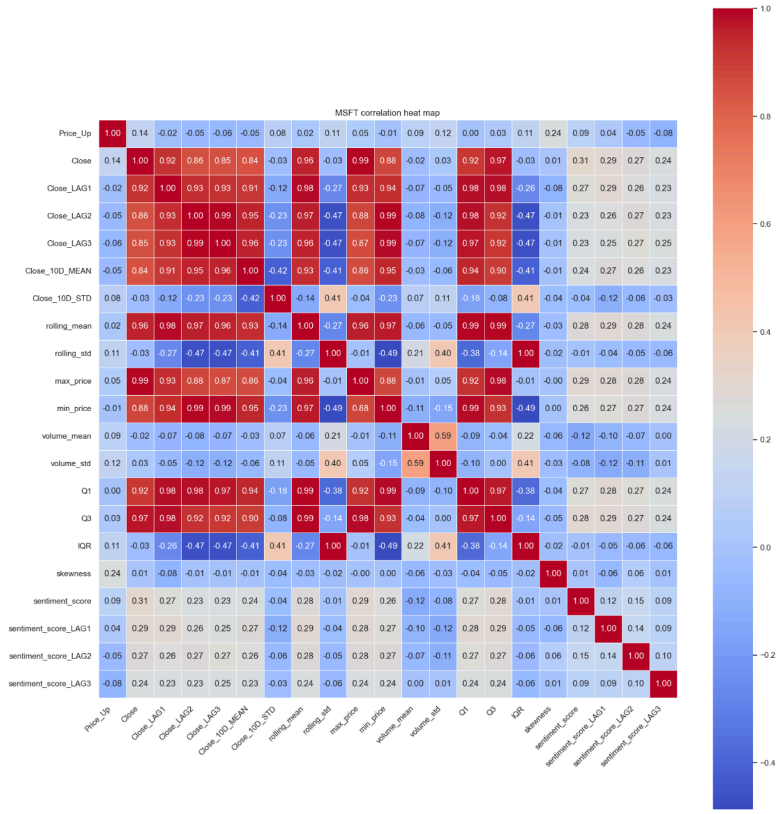

Figure 8. MSFT correlation heatmap

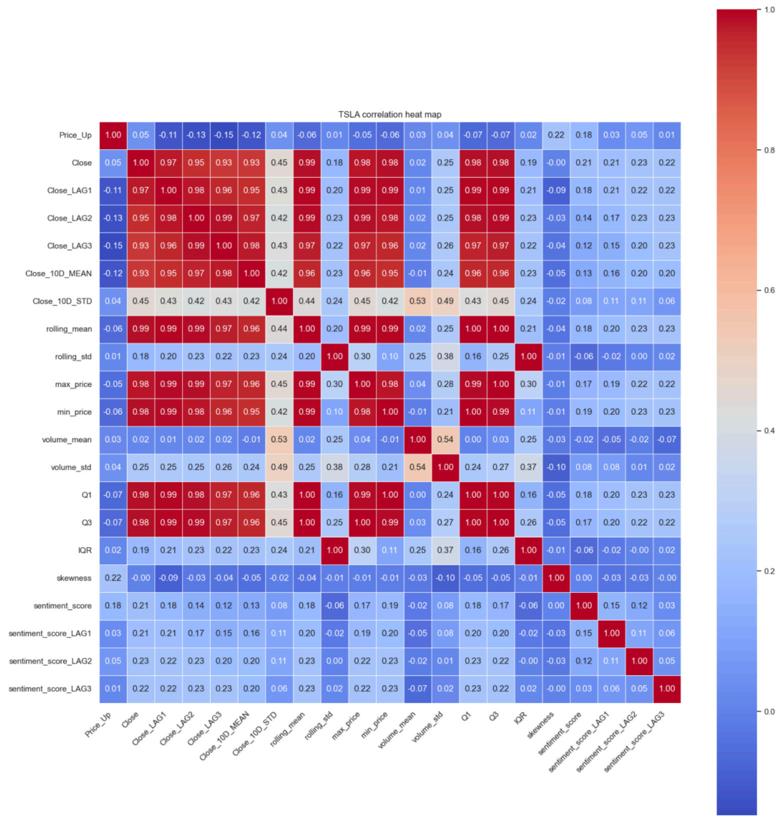

Figure 9. TSLA correlation heatmap