
| RESEARCH ARTICLE

Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends

Md Shahedul Amin¹✉, Eftekhar Hossain Ayon², Bishnu Padh Ghosh³, MD, Md Salim Chowdhury⁴, Mohammad Shafiquzzaman Bhuiyan⁵, Rasel Mahmud Jewel⁶ and Ahmed Ali Linkon⁷

¹Department of Finance & Economics, University of Tennessee, Chattanooga, TN, USA

²Department of Computer Science, Monroe College, New Rochelle, New York, USA

³School of Business, International American University, Los Angeles, California, USA

⁴College of Graduate and Professional Studies Trine University, USA

⁵Department of Business Administration, Westcliff University, Irvine, California

⁷Department of Computer Science, Westcliff University, Irvine, California

Corresponding Author: Md Shahedul Amin, **E-mail:** md.shahedul-amin09@utc.edu

| ABSTRACT

The surge in generative artificial intelligence technologies, exemplified by systems such as ChatGPT, has sparked widespread interest and discourse prominently observed on social media platforms like Twitter. This paper delves into the inquiry of whether sentiment expressed in tweets discussing advancements in AI can forecast day-to-day fluctuations in stock prices of associated companies. Our investigation involves the analysis of tweets containing hashtags related to ChatGPT within the timeframe of December 2022 to March 2023. Leveraging natural language processing techniques, we extract features, including positive/negative sentiment scores, from the collected tweets. A range of classifier machine learning models, encompassing gradient boosting, decision trees and random forests, are employed to train on tweet sentiments and associated features for the prediction of stock price movements among key companies, such as Microsoft and OpenAI. These models undergo training and testing phases utilizing an empirical dataset gathered during the stipulated timeframe. Our preliminary findings reveal intriguing indications suggesting a plausible correlation between public sentiment reflected in Twitter discussions surrounding ChatGPT and generative AI and the subsequent impact on market valuation and trading activities concerning pertinent companies, gauged through stock prices. This study aims to forecast bullish or bearish trends in the stock market by leveraging sentiment analysis derived from an extensive dataset comprising 500,000 tweets. In conjunction with this sentiment analysis derived from Twitter, we incorporate control variables encompassing macroeconomic indicators, Twitter uncertainty index and stock market data for several prominent companies.

| KEYWORDS

Generative Artificial Intelligence, Tweets, Stock Market Forecasting, Sentiment Analysis

| ARTICLE INFORMATION

ACCEPTED: 15 December 2023

PUBLISHED: 07 January 2024

DOI: 10.32996/jcsts.2024.6.1.7

1. Introduction

The sentiment analysis of tweets spanning December 2022 to March 2023 holds the potential to significantly impact the stock trends of various influential companies. Microsoft's substantial investment in OpenAI, which is pivotal in ChatGPT's development, intertwines its stock trajectory with sentiments around ChatGPT advancements. Positive perceptions of innovative AI strides and future applications in Microsoft's offerings may sway investor decisions. Conversely, Google's BARD, a ChatGPT competitor by OpenAI, might witness stock fluctuations linked to sentiments evaluating its performance against ChatGPT. Favorable discussions

about BARD could buoy Google's stocks, while adverse sentiments may provoke downturns. Moreover, Meta, entrenched in the tech landscape, stands susceptible to sentiments encompassing AI progression. Positive perceptions favoring ChatGPT may underscore OpenAI's dominance in AI, potentially influencing Meta's stock if viewed as lagging in innovation. NVIDIA's GPUs, integral in training ChatGPT through its association with Microsoft and OpenAI, render sentiment around AI advancements and collaborative endeavors pivotal to NVIDIA's stock dynamics. Simultaneously, Amazon Web Services' (AWS) cloud computing services could be swayed by sentiments tied to AI progress, notably ChatGPT. Positive sentiments might hint at heightened AWS demand, impacting Amazon's stock. The analysis of these sentiments becomes a pivotal gateway to understanding potential investor inclinations, potentially steering the stock prices of Microsoft, Google, Meta, NVIDIA, and Amazon during this critical period. This inclusion of daily opening and closing prices for technology giants like Google, Amazon, Meta, Nvidia, and Microsoft adds a crucial dimension to our research. By integrating this financial data alongside sentiment analysis from social media, this paper will gain a comprehensive understanding of the interplay between market trends, investor sentiments, and stock movements during the specified period. This dual approach can offer a robust foundation for your study, providing valuable insights into how sentiments might correlate with or influence stock prices for these prominent technology companies.

In our study, we posit the hypothesis that the sentiment conveyed and the quantity of tweets associated with significant technological events and announcements serve as indicators of public sentiments and attention, subsequently influencing stock trading activities. Our approach involves the compilation of a dataset comprising tweets featuring ChatGPT hashtags from December 2022 to March 2023, aligning with the introduction of groundbreaking conversational AI capabilities by Anthropic. To facilitate predictions regarding the daily directional movement (either upward or downward) of stock prices for pertinent companies such as Microsoft and OpenAI, a diverse array of stock classifier models—including logistic regression and different tree classifier models—are constructed and evaluated. The features extracted from Twitter data, encompassing positive/negative sentiment scores, metrics such as the count of retweets, user follower counts, and additional parameters, are systematically linked to daily stock prices. Supervised learning models are trained through the amalgamation of this paired dataset. The primary objective is to evaluate the feasibility of accurately forecasting interday price fluctuations solely relying on aggregated signals derived from Twitter conversations. This assessment aims to discern the predictive efficacy of Twitter-derived data in forecasting stock price movements.

In the realm of scholarly discussions, empirical data underscores the substantial impact of macroeconomic indicators on consumer spending behaviors. Notably, metrics such as the Consumer Price Index (CPI), Index of Consumer Sentiment (ICS), and unemployment rates emerge as pivotal factors shaping stock market trends. Despite their recognized influence, the existing landscape of demand forecasting exhibits a glaring gap – a lack of robust evidence capturing the sway of these external economic dynamics on stock market trends. This void serves as the impetus behind our research endeavor. Our aim is to fortify the existing but weakly supported assertion within current literature regarding the efficacy of integrating macroeconomic variables into stock forecasting models.

The current surge in groundbreaking generative AI significantly amplifies the impact and relevance of this particular use case. Our empirical analysis is designed with a dual objective: firstly, to quantify the predictive indicators originating from Twitter, and secondly, to substantiate the proposition that narratives disseminated via social media exert discernible influences on market valuation and trading dynamics surrounding pivotal technological advancements and occurrences. The outcomes of this study possess the potential to illuminate the feasibility of augmenting traditional quantitative equity analysis with alternative data sources like Twitter sentiment, thereby enhancing the accuracy of predictive models. Furthermore, this research takes a broader scope by scrutinizing the intricate interconnections among technological hype cycles, public discourse, and investor behavioral patterns.

2. Literature Review

Utilizing sentiment analysis techniques on Twitter data containing hashtags related to AI and GPT is integral in forecasting stock market trends. The application of sentiment analysis on Twitter data has been extensively investigated in the realm of predicting stock market movements. Notably, prior research has utilized sentiment analysis of Twitter messages to construct a daily happiness index, offering insights into the influence of social media on financial markets (Agarwal et al., 2021). Moreover, sentiment analysis has been instrumental in examining tweet content, transforming this information into social sentiment indices and generating time series data for specific companies, showcasing the potential of sentiment analysis in anticipating stock performance (Mendoza-Urdiales et al., 2022).

Moreover, scholarly investigations have explored the influence of social media sentiment on stock market volatility, elucidating the effects of sentiments linked to significant global events on investor attitudes and transactional trading metrics (Maqsood et al., 2022). Furthermore, research examining the correlation between social media sentiment and stock market returns, categorized by user authentication, has unveiled that the sentiment expressed by verified users exerts a more substantial and enduring influence on stock market returns. This underscores the pivotal role of user sentiment in forecasting stock movements (Li & Huang, 2023).

Furthermore, the application of machine learning methodologies has been instrumental in constructing predictive models that assess the interconnection between tweet content and stock values. This application underscores the promise of machine learning in prognosticating forthcoming stock prices based on sentiment analysis (Yashmita & Kavitha, 2023). Additionally, the proposition of employing ensemble learning techniques alongside statistical indicators, in conjunction with sentiment analysis, has emerged as a strategy to forecast movements in the stock market. This proposal underscores the fusion of sentiment analysis with machine learning approaches for predictive analytics (Bagga & Patel, 2023).

Moreover, research investigating the influence of social media rumors on stock market volatility has offered valuable insights into the association between rumor-related posts and abnormal returns. This exploration underscores the significant impact of social media content on market dynamics (Zhang et al., 2022). Additionally, the utilization of sentiment analysis in conjunction with neural network-based methodologies applied to Twitter data has enabled the scrutiny of the relationship between a company's stock market performance and sentiments expressed in tweets. This analysis serves as a testament to the practicality of sentiment analysis in forecasting movements within the stock market (Shravani et al., 2023).

To predict stock market trends, various machine learning classifiers have been employed, including gradient boosting, decision trees, random forest, naive Bayes, and extra tree classifiers. These classifiers have been utilized in different studies to forecast stock prices and market direction. For instance, research highlights the use of ensemble models such as gradient boosting and random forest to forecast financial market direction. Similarly, Ampomah et al. (2020) evaluated tree-based ensemble machine learning models, including adaboost, for predicting stock price movement. Furthermore, Eachempati and Srivastava (2020) adopted a support vector machine to predict stock market movement using sentiment, historical stock prices, return on investment, and risk as predictors. Additionally, Ecer et al. (2020) employed multilayer perceptron with genetic algorithms and particle swarm optimization for modelling stock price index prediction.

Moreover, the study by Alshammari et al. (2022) implemented logistic regression, decision trees, support vector machines, and random forests to predict the direction of stock market returns. Furthermore, Wu (2023) discussed the use of random forest, XGBoost, and decision trees for stock price prediction. Additionally, mentioned the use of random forest and naive Bayes classifier for predicting stock market trends on the Zagreb Stock Exchange. These references collectively demonstrate the diverse application of machine learning classifiers in predicting stock market trends and prices.

Furthermore, other studies have also contributed to the understanding of stock market prediction using machine learning models (Jerić, 2020). For example, Ardakani et al. (2023) compared the performance of federated learning against centralized and decentralized learning frameworks for stock market prediction. Additionally, Yin et al. (2021) focused on the optimization of random forest for stock trend prediction, ensuring model stability and performance. Moreover, Selby et al. (2023) demonstrated the superiority of multivariate sequential LSTM autoencoder in predicting stock prices through experiments with real stock market data. In addition, the application of machine learning in stock market prediction has been explored in various contexts. For instance, Almeahadi (2021) highlighted the use of COVID-19 pandemic data in predicting the stock market, while Nti et al. (2019) focused on random forest-based feature selection of macroeconomic variables for stock market prediction. These studies underscore the wide-ranging applications and considerations in utilizing machine learning for stock market prediction.

In summary, the application of machine learning classifiers such as gradient boosting, decision tree, random forest, naive Bayes, and extra tree classifiers has been extensively explored in predicting stock market trends and prices. These models have been applied in diverse contexts and have been optimized for improved performance, demonstrating the significance of machine learning in stock market prediction.

3. Methodology

3.1 Data Pre-Processing:

In the pursuit of preparing a comprehensive and standardized dataset for our research focusing on Twitter discussions related to chatbot technology during the December to March 2023 period, an intricate yet essential data pre-processing strategy was meticulously executed. As a primary step, the content column underwent a rechristening, emerging as the tweet column, to streamline and enhance the dataset's interpretability. Subsequent pre-processing involved a series of indispensable transformations to homogenize the textual data. To ensure consistency, all text was converted to lowercase, aligning the text case for uniform analysis. Furthermore, several crucial cleansing steps were undertaken: the removal of Twitter handlers (mentions), extraction of hashtags, and elimination of URLs, ensuring the extraction of valuable content while removing elements that don't contribute substantively to our analysis. Employing regular expressions, special characters were meticulously removed, retaining only alphanumeric characters for meaningful analysis. To bolster the dataset's quality, a meticulous curation process was applied, removing single characters and consolidating multiple spaces into a singular space. These intricate yet pivotal steps in data pre-processing were fundamental in cultivating a refined and standardized dataset, setting the stage for comprehensive and insightful analyses regarding the discourse surrounding chatbot technology on the Twitter platform during the specified timeframe.

This research embarks on an exploration leveraging two prominent datasets recently published, both offering distinct yet complementary dimensions to the study. Firstly, the 500k ChatGPT-related Tweets dataset sourced from Kaggle, meticulously curated by users, represents a repository of over 500,000 tweets spanning the period from December 2022 to March 2023. This rich dataset encompasses diverse attributes, including tweet text, temporal information, engagement metrics such as likes and retweets, user follower counts, and sentiment evaluations, all intricately associated with the ChatGPT conversational AI system. Concurrently, the analysis incorporates Historical Stock Prices sourced from Yahoo Finance, focusing on the daily opening and closing prices for five technology giants - Google, Amazon, Meta, Nvidia, and Microsoft. This financial dataset spans from December 2022 to March 2023, aligning chronologically with the ChatGPT tweet data and exclusively covering trading days.

The fusion of these datasets yields a comprehensive empirical corpus that enables the correlation of daily Twitter sentiment and engagement metrics concerning ChatGPT with the end-of-day stock price movements. This amalgamation lays the groundwork for training sophisticated classifiers aimed at predicting intraday stock price direction based solely on amalgamated signals derived from the related Twitter discourse. The selection of these prominent tech stocks is predicated on their significance in the tech landscape and the anticipated influence of narratives surrounding AI advancements on their market performance. This consolidated empirical dataset comprises an expansive collection of over 500,000 tweets juxtaposed against 90 trading days of historical price data for the selected quintet of tech stocks, encompassing a three-month timeframe. This unique amalgamation of alternative data from online discourse and conventional financial metrics offers a pioneering avenue for probing the feasibility of enhancing quantitative predictive models through insights gleaned from social media chatter.

3.2 Feature Engineering:

This study employs sentiment analysis on a dataset encompassing 500,000 tweets to forecast bullish or bearish trends within the stock market. Our analysis integrates Twitter-derived sentiment, including Positive Sentiment, Neutral Sentiment, and Negative Sentiment, with control variables such as the ICS (Consumer Confidence Index), Unemployment Rate, CPI (Consumer Price Index), and Twitter Uncertainty Index. Additionally, we incorporate company-specific stock market data (Stock Market Uncertainty Index and Volume) for major companies like AAPL, GOOG, MSFT, AMZN, META, NVDA. By combining these diverse data sources, our objective is to scrutinize the correlation between Twitter sentiment, macroeconomic indicators, and stock market trends. Utilizing statistical and machine learning models, our approach aims to discern the predictive capacity of Twitter-derived sentiment in predicting stock market movements while considering the influence of broader economic conditions and company-specific data on market behavior. Ultimately, this investigation seeks to shed light on the potential of sentiment analysis in forecasting stock market trends within the context of varying economic landscapes and company performance.

In the process of delineating trends within the stock market data, a Trend column was introduced into the dataset, initialized with a default value of neutral. This novel addition operates as a categorical variable, assuming a crucial role in the classification and segmentation of market movements, laying the groundwork for subsequent comprehensive analyses. The identification of bullish and bearish trends was conducted based on the comparison between the open and close prices for the stock of a certain company. Specifically, if the open price was found to be less than the close price on a given day, it was categorized as a bullish trend. Conversely, if the open price exceeded the close price, the trend was labeled as bearish. This approach allowed for the classification of market movements into distinct categories, aiding in the subsequent analysis and interpretation of stock market trends. Investigating the relationship between sentiment analysis and stock market trends, Table 1 presents the list of independent variables used in the analysis.

Table 1 List of Independent variables

Independent Variables	Descriptions
TEU_USA	Twitter Economic Uncertainty Index in USA on a specific day
Volume	Volume of stock traded on a single day
ics_all	Index of Consumer Sentiment for all goods and services
CPIAUCSL	Consumer Price Index
unrate	Unemployment rates
Positive Sentiment	Extracted positive sentiment index from tweets
Negative Sentiment	Extracted negative sentiment index from tweets

Neutral Sentiment	Extracted neutral sentiment index from tweets
Adj_Close_VIX	Stock Market Uncertainty Index for a specific company

In this study's data pre-processing phase, we implemented a comprehensive methodology to process and prepare tweet data for subsequent sentiment analysis and weighting. The pre-processing steps were instrumental in standardizing the dataset and deriving sentiment-related features, pivotal for subsequent analyses. Firstly, to ensure equal treatment of tweets across varying engagement levels, we computed statistical measures regarding the retweet counts of individual users. Specifically, employing a rolling window of 10,000 tweets (or fewer initially if available), we calculated the mean and standard deviation of retweet counts per user. This rolling computation allowed us to ascertain these statistics while accommodating users with fewer tweets. These statistical measures were further utilized to determine weighted categories for tweets based on their retweet counts relative to the derived mean and standard deviation. Tweets were categorized into distinct weight classes to emphasize highly retweeted content. Categorization involved assigning weights based on multiples of standard deviations from the mean retweet count. Tweets falling within one standard deviation above the mean received a weight of 1, those between one and two standard deviations received a weight of 2, while tweets two or more standard deviations above the mean were assigned a weight of 3.

Simultaneously, sentiment analysis was conducted to capture the sentiment embedded within the tweet content. This involved the computation of sentiment scores, specifically positive, neutral, and negative sentiments, derived from the tweet text. An additional adjustment was applied to these scores by introducing a small value ($1 \times (10^{-6})$) to prevent potential issues arising from zero values. Upon extraction of sentiment scores, the original sentiments column was removed from the dataset. This phase of data pre-processing was pivotal in preparing the tweet data for subsequent analyses, enabling the incorporation of sentiment-related features and weighted information derived from retweet statistics. These processed features laid the foundation for in-depth sentiment analysis and further exploratory analyses in our study.

Once applied, the dataset is fully equipped with informative new features that can aid in training and evaluating predictive models. The expanded feature set provides a more comprehensive representation that can improve the model's capacity to learn meaningful relationships, ultimately enhancing predictive accuracy. Thoughtful feature engineering is a crucial phase in the development of effective machine learning solutions.

3.3 Classifier Models:

In our endeavor to forecast stock market movements, we harnessed the power of social media discourse by leveraging a dataset encompassing ChatGPT-related tweets spanning January to March 2023. Alongside, historical stock price records of major tech entities were amalgamated for training and validation purposes. Utilizing a Gradient Boosting Classifier, we aimed to decipher the correlations between sentiment signals embedded within Twitter conversations about ChatGPT and subsequent shifts in stock prices. This classifier was meticulously trained on a subset of the dataset meticulously engineered to extract sentiment metrics and temporal dynamics from the tweets. Validation of the model's predictive capacity was performed on an independent section of the dataset, evaluating its accuracy, precision, and recall in foreseeing stock market movements based solely on tweet sentiment. This innovative approach seeks to integrate social media sentiment analysis into financial modelling, offering potential new avenues for enhancing predictive accuracy in stock market analysis.

The evaluation of the Gradient Boosting Classifier model underscores its predictive prowess, leveraging specific hyperparameters to amalgamate weak learners and construct a robust predictive framework. Initialized with precise configurations, the classifier undergoes fitting using scaled training data. This process culminates in assessing the model's accuracy across both training and validation datasets, providing insights into its predictive proficiency. Subsequent predictions generated using the designated test dataset offer an evaluative measure of the model's efficacy in forecasting market trends, highlighted through an accuracy score computation.

4. Results and Discussion

Various classification models, including Random Forest, Decision Tree, Extra Trees Classifier, and Naive Bias Classifier, undergo rigorous evaluation on a dataset. The Random Forest Classifier, employing key hyperparameters like 100 estimators, a depth of 5, and a random state of 0, achieves a balance between complexity and generalization. This model, trained on scaled data, discerns patterns and associations and is assessed on both training and validation sets to gauge its predictive capacity. Similarly, the Decision Tree Classifier, configured with a depth of 5 and a fixed random state, undergoes training and evaluation on the scaled dataset, generating predictions for testing. Accuracy scores, computed during training and validation, offer insights into class label prediction accuracy. Visual representations like confusion matrices aid in assessing performance by distinguishing true/false positives/negatives. Furthermore, the Extra Trees and Naive Bias Classifiers are introduced for predictive analysis of the dataset. In evaluating the performance of classifiers for predicting stock market trends labelled as bullish or bearish, several key metrics were

employed to assess the efficacy of these models. These metrics—accuracy score, F1 score, and recall score—serve as fundamental benchmarks in quantifying the predictive capabilities of the employed classifiers.

The accuracy score represents the ratio of correctly predicted instances to the total instances in the dataset. Specifically, it measures the model's overall correctness in predicting both "bullish" and "bearish" trends within the dataset. For stock market trend prediction, accuracy provides an understanding of the classifier's overall predictive capability across all classes. The F1 score, a harmonic mean of precision and recall, serves as a balanced metric, especially useful when dealing with imbalanced datasets. It encapsulates both false positives and false negatives, offering an equilibrium between precision and recall. In the context of stock market trends, the F1 score provides insights into the model's ability to balance precision and recall for both "bullish" and "bearish" predictions. Recall, also known as sensitivity or true positive rate, gauges the classifier's ability to correctly identify actual positive instances. Specifically for stock market trend prediction, recall measures the model's capability to accurately identify "bullish" trends among all true "bullish" instances.

These evaluation metrics serve as pivotal indicators in comprehensively assessing the performance of Random Forest, Decision Tree, Gradient Boosting, and Naive Bayes classifiers in predicting stock market trends. They provide nuanced insights into the classifiers' predictive strengths and weaknesses, guiding the selection of the most suitable model for accurate market trend forecasting. The classification report evaluates the performance metrics of a binary classification model applied to a financial dataset predicting Bearish and Bullish market trends for Microsoft Stock Trend in Table 2. In the evaluation of a binary classification model applied to a financial dataset to forecast market trends categorized as Bearish and Bullish, Table 2 provides an assessment of various performance metrics. The metrics encompass Accuracy, Recall, and F1 Scores for both Bearish and Bullish trends across multiple classification models.

Table 2 Comparison of model performances for Microsoft stock

Classification Model	Accuracy Score (Bearish)	Recall Score (Bearish)	F1 Score (Bearish)	Accuracy Score (Bullish)	Recall Score (Bullish)	F1 Score (Bullish)
Gradient Boosting	74	97	84	95	65	78
Random Forest	82	100	90	100	78	67
Decision Tree	81	86	83	85	79	82
Naive Bayes	68	78	73	73	62	67
Extra Tree	76	92	83	89	70	78

This analysis revolves around three critical metrics: accuracy score, recall, and F1 score, which are fundamental in gauging the effectiveness of binary classification models. In this scenario, the models discern between bearish (indicating an anticipated price decrease) and bullish (indicating an anticipated price increase) trends. The accuracy score illustrates the models' precision in correct classification, while recall delves into their accuracy, specifically in identifying bearish trends. Meanwhile, the F1 score amalgamates both precision and recall, offering a comprehensive view of the models' predictive prowess for both bearish and bullish trends. The accuracy score reflects the model's correct classification proportion, while recall gauges the model's accuracy in identifying bearish data points correctly. The F1 score, a harmonized metric blending precision and recall, encapsulates the model's efficacy in classifying both bearish and bullish data points.

In the evaluation of classification models applied to predict financial market trends denoted as Bearish and Bullish, a comprehensive quantitative analysis was conducted across various performance metrics. Table 2 delineates the performance of five distinct classification models: Gradient Boosting, Random Forest, Decision Tree, Naive Bayes, and Extra Tree, based on key metrics encompassing Accuracy, Recall, and F1 Score for both Bearish and Bullish trends.

For the identification of Bearish trends, Random Forest emerged as a prominent performer, showcasing a notable accuracy score of 82%, closely followed by Decision Tree at 81% and Extra Tree at 76%. However, Gradient Boosting and Naive Bayes exhibited comparatively lower accuracies of 74% and 68%, respectively. In terms of recall, Random Forest excelled with a perfect score of 100%, outperforming all other models. Extra Tree followed closely with a recall score of 92%, while Gradient Boosting, Decision

Tree, and Naive Bayes displayed lower recall scores. Furthermore, Random Forest achieved the highest F1 score at 90%, closely trailed by Decision Tree and Extra Tree at 83%. Gradient Boosting and Naive Bayes reported lower F1 scores at 84% and 73%, respectively, indicating varied performance levels in identifying 'Bearish' trends across the models.

Conversely, for identifying Bullish trends, Random Forest displayed exemplary accuracy, achieving a perfect score of 100%. Extra Tree and Decision Tree followed with accuracies of 89% and 85%, respectively, while Gradient Boosting and Naive Bayes exhibited lower accuracy rates of 95% and 73%. However, in terms of recall, Random Forest and Decision Tree showcased recall scores of 78% and 79%, respectively, while Extra Tree, Gradient Boosting, and Naive Bayes achieved recall scores of 70%, 65%, and 62%, respectively. Regarding the F1 score for Bullish trends, Random Forest secured the highest score at 87%, followed by Decision Tree at 82% and Extra Tree at 78%. Gradient Boosting and Naive Bayes registered lower F1 scores of 78% and 67%, respectively.

In summary, Random Forest stands out as a robust performer, demonstrating remarkable accuracy, recall, and F1 scores in identifying both Bearish and Bullish trends, showcasing its potential as an effective model in financial market trend prediction. However, Decision Tree and Extra Tree also exhibited commendable performances, while Gradient Boosting and Naive Bayes showcased relatively lower predictive capabilities in this specific context. These findings highlight the varying strengths and weaknesses of each model in predicting financial market trends.

Fig. 1 Comparison of classification models for Microsoft Stock

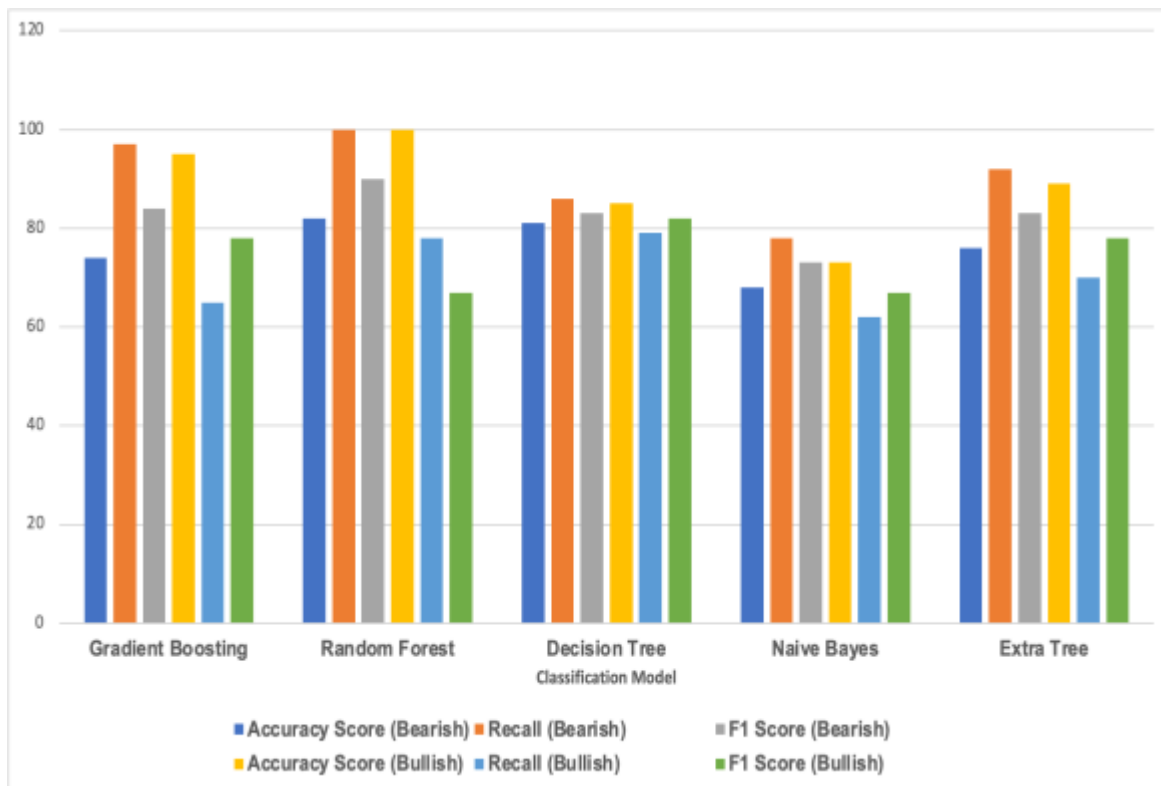


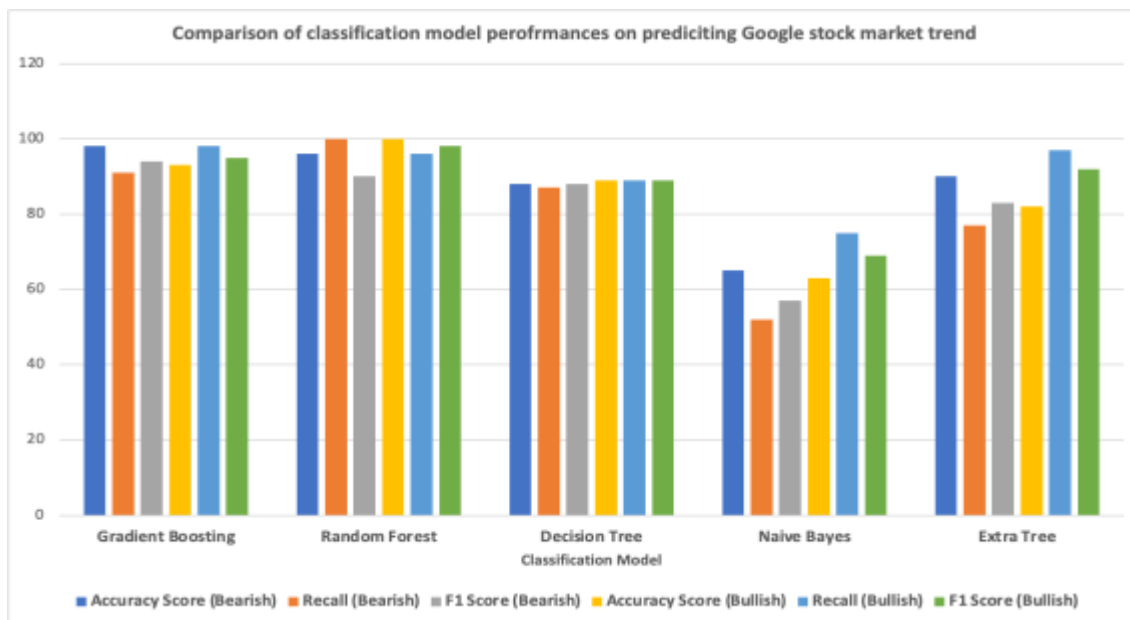
Fig. 1 shows a comparison of the performance of four different gradient boosting, random forest, decision tree, and naive bayes machine learning models on predicting the trend of Microsoft stock prices. Examining recall and F1 score metrics, the random forest model maintains its dominance, showcasing superior performance. Gradient boosting clinches the second spot in recall and F1 score for bearish predictions, whereas the decision tree model does so for bullish forecasts. In stark contrast, the naive Bayes model trails behind, consistently recording the least recall and F1 score for both bearish and bullish predictions. In essence, the graphical analysis emphasizes the random forest model's exceptional ability to predict Microsoft stock price trends, setting it apart as the most adept performer among the evaluated models.

Table 3 Comparison of Model Performances for Google Stock

Classification Model	Accuracy Score (Bearish)	Recall Score (Bearish)	F1 Score (Bearish)	Accuracy Score (Bullish)	Recall Score (Bullish)	F1 Score (Bullish)
Gradient Boosting	98	91	94	93	98	95
Random Forest	96	100	90	100	96	98
Decision Tree	88	87	88	89	89	89
Naive Bayes	65	52	57	63	75	69
Extra Tree	90	77	83	82	97	92

In Table 3, identifying Google Stock Bearish trends, Random Forest exhibited the highest accuracy at 96%, followed by Gradient Boosting (98%) and Extra Trees (90%). In terms of recall, Random Forest achieved perfect recall (100%), outperforming all other models, while Gradient Boosting and Extra Trees scored 91% and 77%, respectively. Conversely, in identifying Bullish trends, Random Forest again emerged with the highest accuracy of 100%, followed by Decision Tree (89%) and Gradient Boosting (93%). In summary, Random Forest demonstrates robust performance across both 'Bearish' and 'Bullish' trend identification, exhibiting high accuracy, recall, and F1 scores. However, Gradient Boosting also stands out, especially in predicting 'Bearish' trends, emphasizing its effectiveness in certain scenarios. These findings underscore the nuanced strengths of each model in handling specific market trend predictions.

Fig. 2 Comparison of classification models for predicting the stock market trend of Google



Our analysis from Fig. 2 reveals that Gradient Boosting achieved the highest accuracy score for both bullish and bearish predictions, followed by Random Forest and Decision Tree. Naive Bayes and Extra Tree demonstrated considerably lower accuracy. In terms of recall, Gradient Boosting again dominated for bearish predictions, while Random Forest outperformed in bullish. The F1 score further solidified Gradient Boosting's position as the top performer, exceeding all other models for both bullish and bearish predictions. These findings suggest that Gradient Boosting's ensemble learning approach and inherent non-linearity might be better suited for capturing the intricate relationships within the Google stock price data compared to the other models considered. Random Forest's strong performance also highlights its robustness in capturing complex patterns. Decision Tree, while achieving reasonable accuracy, exhibited limitations in both recall and F1 score, indicating potential overfitting or insufficient feature importance considerations. Naive Bayes and Extra Tree's comparatively lower performance suggests their limited suitability for this task, possibly due to their inherent assumptions about data independence or simplistic decision-making processes.

5. Conclusion

The analysis of Twitter data in anticipation of stock market trends presents a compelling narrative when juxtaposed against macroeconomic indicators and company-specific stock data. This multifaceted approach seeks to unravel the intricate interplay between sentiment analysis derived from social media, broader economic conditions, and the individual performance of companies in the stock market. The integration of sentiment analysis from Twitter data alongside macroeconomic indicators offers a holistic view of market sentiment. The quantitative assessment of sentiment, particularly in distinguishing Bearish and Bullish trends, provides valuable insights. Models like Random Forest showcase robust predictive abilities, displaying high accuracy, recall, and F1 scores, underscoring their effectiveness in discerning market trends. However, the convergence of sentiment analysis with macroeconomic indicators and company-specific data enriches this predictive landscape further. While sentiment analysis captures public perception and emotions, macroeconomic indicators such as GDP growth, inflation rates, and unemployment offer a broader economic context. Moreover, incorporating company-specific stock data, including financial reports, revenue, and earnings, provides granular insights into individual stock performance. This holistic approach illuminates the intricate relationships between sentiment, economic conditions, and market dynamics. It unveils the potential synergy between social sentiment and economic indicators in predicting market trends, offering a more comprehensive understanding of market behavior. The outperformance of certain models, notably Random Forest, underscores the significance of sentiment analysis, yet the convergence with macroeconomic and company-specific data enriches the predictive capacity, potentially enhancing market forecast accuracy.

In conclusion, the synergy between sentiment analysis from Twitter data, macroeconomic indicators, and company-specific stock data opens avenues for more refined market trend predictions. While sentiment analysis offers a window into public perception, its amalgamation with broader economic indicators and detailed company-specific data augments predictive models, fostering a nuanced comprehension of market behavior and trends. This multidimensional approach holds promise in fortifying market forecasting methodologies, potentially enhancing decision-making processes in financial markets.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

- [1] Agarwal, S., Kumar, S., & Goel, U. (2021). Social media and the stock markets: an emerging market perspective. *Journal of Business Economics and Management*, 22(6), 1614-1632. <https://doi.org/10.3846/jbem.2021.15619>
- [2] Ampomah, E. K., Qin, Z., & Nyame, G. (2020). Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. *Information*, 11(6), 332. <https://doi.org/10.3390/info11060332>
- [3] Alshammari, B. M., Aldhmour, F., AlQenaie, Z. M., & Almohri, H. (2022). Stock market prediction by applying big data mining. *Arab Gulf Journal of Scientific Research*, 40(2), 139-152. <https://doi.org/10.1108/AGJSR-05-2022-0053>
- [4] Almeahadi, A. (2021). COVID-19 Pandemic Data Predict the Stock Market. *Computer Systems Science & Engineering*, 36(3). <https://doi.org/10.32604/csse.2021.015309>
- [5] Bagga, A. R., & Patel, H. (2022). Stock Market Forecasting using Ensemble Learning and Statistical Indicators. *Journal of Engineering Research*. <https://doi.org/10.36909/jer.16629>
- [6] Eachempati, P., & Srivastava, P. R. (2023). Prediction of the Stock Market From Linguistic Phrases: A Deep Neural Network Approach. *Journal of Database Management (JDM)*, 34(1), 1-22. <https://doi.org/10.4018/JDM.322020>
- [7] Ecer, F., Ardabili, S., Band, S. S., & Mosavi, A. (2020). Training multilayer perceptron with genetic algorithms and particle swarm optimization for modeling stock price index prediction. *Entropy*, 22(11), 1239. <https://doi.org/10.3390/e22111239>
- [8] Haque, M. S., Amin, M. S., & Miah, J. (2023). Retail Demand Forecasting: A Comparative Study for Multivariate Time Series. *arXiv e-prints*. <https://doi.org/10.48550/arXiv.2308.11939>
- [9] Khan, R. H., Miah, J., Arafat, S. M. Y., Syeed, M. M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. *arXiv preprint*, arXiv:2310.17729.
- [10] Khan, R. H., Miah, J., Arafat, S. M., Syeed, M. M., & Ca, D. M. (2023). Improving Traffic Density Forecasting in Intelligent Transportation Systems Using Gated Graph Neural Networks. *arXiv preprint* arXiv:2310.17729.
- [11] Khan, R. H., Miah, J., Rahat, M. A. R., Ahmed, A.H., Shahriyar M.A and Lipu, E.R. (2023). A Comparative Analysis of Machine Learning Approaches for Chronic Kidney Disease Detection, 2023 8th International Conference on Electrical, Electronics and Information Engineering (ICEEIE), Malang City, Indonesia, 2023, pp. 1-6, doi: 10.1109/ICEEIE59078.2023.10334765.
- [12] Khan R. H and Miah, J. (n.d). Performance Evaluation of a new one-time password (OTP) scheme using stochastic petri net (SPN), 2022 IEEE World AI IoT Congress (AllIoT), Seattle, WA, USA, 2022. 407-412, doi: 10.1109/AllIoT54504.2022.9817203.
- [13] Liao, L., & Huang, T. (2023). The Impact of Social Media Sentiment on Stock Market Based on User Classification. In *Digitalization and Management Innovation* (pp. 1-16). IOS Press. <https://doi.org/10.3233/FAIA230002>
- [14] Miah, J., Ca, D. M., Sayed, M. A., Lipu, E. R., Mahmud, F., & Arafat, S. M. Y. (2023). Improving Cardiovascular Disease Prediction Through Comparative Analysis of Machine Learning Models: A Case Study on Myocardial Infarction. *arXiv preprint*, 2311.00517.
- [15] Miah, J., Cao, D. M., Sayed, M. A., & Haque, M. S. (2023). Generative AI Model for Artistic Style Transfer Using Convolutional Neural Networks. *arXiv preprint* arXiv:2310.18237

- [16] MAbu S., Tayaba, M., Islam, M. T & Bishnu P G. (2023). Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms. *Journal of Computer Science and Technology Studies*, 5(4), 142–149. <https://doi.org/10.32996/jcsts.2023.5.4.14>
- [17] Mia, M. T., Ray, R. K., Ghosh, B. P., Chowdhury, M. S., Al-Imran, M., Das, R., Sarkar, M., Sultana, N., Nahian, S. A., & Puja, A. R. (2023). Dominance of External Features in Stock Price Prediction in a Predictable Macroeconomic Environment. *Journal of Business and Management Studies*, 5(6), 128–133. <https://doi.org/10.32996/jbms.2023.5.6.10>
- [18] Miah, J., Ca, D. M., Sayed, M. A., Lipu, E. R., Mahmud, F., & Arafat, S. M. Y. (2023). Improving Cardiovascular Disease Prediction Through Comparative Analysis of Machine Learning Models: A Case Study on Myocardial Infarction. arXiv preprint, 2311.00517.
- [19] Miah, J., Cao, D. M., Sayed, M. A., Taluckder, M. S., Haque, M. S., & Mahmud, F. (2023). Advancing Brain Tumor Detection: A Thorough Investigation of CNNs, Clustering, and SoftMax Classification in the Analysis of MRI Images. *arXiv preprint arXiv:2310.17720*.
- [20] Miah, J., & Khan, R. H. (2019, November). Service Development of Smart Home Automation System: A Formal Method Approach. In *Proceedings of the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems* (pp. 161-167).
- [21] Mendoza-Urdiales, R. A., Núñez-Mora, J. A., Santillán-Salgado, R. J., & Valencia-Herrera, H. (2022). Twitter Sentiment Analysis and Influence on Stock Performance Using Transfer Entropy and EGARCH Methods. *Entropy*, 24(7), 874. <https://doi.org/10.3390/e24070874>
- [22] Maqsood, H., Maqsood, M., Yasmin, S., Mehmood, I., Moon, J., & Rho, S. (2022). Analyzing the stock exchange markets of EU nations: A case study of brexit social media sentiment. *Systems*, 10(2), 24. <https://doi.org/10.3390/systems10020024>
- [23] Miah, J., Ca, D. M and Arafat, S. M. Y. (2023). Improving Cardiovascular Disease Prediction Through Comparative Analysis of Machine Learning Models: A Case Study on Myocardial Infarction, 2023 15th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, 2023, pp. 49-54, doi: 10.1109/IIT59782.2023.10366476.
- [24] Nti, K. O., Adekoya, A., & Weyori, B. (2019). Random forest based feature selection of macroeconomic variables for stock market prediction. *American Journal of Applied Sciences*, 16(7), 200-212. <https://doi.org/10.3844/ajassp.2019.200.212>
- [25] Pourroostaei Ardakani, S., Du, N., Lin, C., Yang, J. C., Bi, Z., & Chen, L. (2023). A federated learning-enabled predictive analysis to forecast stock market trends. *Journal of Ambient Intelligence and Humanized Computing*, 14(4), 4529-4535. <https://doi.org/10.1007/s12652-023-04570-4>
- [26] Sarkar, M., Ayon, E. H., Mia, M. T., Ray, R. K., Chowdhury, M. S., Ghosh, B. P., Al-Imran, M., Islam, M. T., Tayaba, M., & Puja, A. R. (2023). Optimizing E-Commerce Profits: A Comprehensive Machine Learning Framework for Dynamic Pricing and Predicting Online Purchases. *Journal of Computer Science and Technology Studies*, 5(4), 186–193. <https://doi.org/10.32996/jcsts.2023.5.4.19>
- [27] Syeed, M. M., Khan, R. H., & Miah, J. (2021). Agile Fitness of Software Companies in Bangladesh: An Empirical Investigation. *International Journal of Advanced Computer Science and Applications*, 12(2).
- [28] Tayaba, M., Ayon, E. H., Mia, M. T., Sarkar, M., Ray, R. K., Chowdhury, M. S., Al-Imran, M., Nobe, N., Ghosh, B. P., Islam, M. T., & Puja, A. R. (2023). Transforming Customer Experience in the Airline Industry: A Comprehensive Analysis of Twitter Sentiments Using Machine Learning and Association Rule Mining. *Journal of Computer Science and Technology Studies*, 5(4), 194–202. <https://doi.org/10.32996/jcsts.2023.5.4.20>
- [29] Vlah Jerić, S. (2020). Comparing classification algorithms for prediction on CROBEX data. *Croatian Review of Economic, Business and Social Statistics*, 6(2), 4-11. <https://doi.org/10.2478/crebss-2020-0007>
- [30] Yin, L., Li, B., Li, P., & Zhang, R. (2023). Research on stock trend prediction method based on optimized random forest. *CAAI Transactions on Intelligence Technology*, 8(1), 274-284. <https://doi.org/10.1049/cit2.12067>
- [31] Zhang, H., Chen, Y., Rong, W., et al. (2022). Effect of social media rumors on stock market volatility: A case of data mining in China. *Frontiers*. <https://doi.org/10.3389/fpsy.2022.987799>