



Can ChatGPT Predict Stock Market Price Movements?

Karamveer Singh
Maharaja Surajmal Institute of Technology
C-4 Janak Puri
New Delhi-110058
India
karam.veer00@gmail.com

ABSTRACT: *The combination of text analytics and machine learning technologies in various applications has emerged recently. Generative content creation platforms use machine learning, which is not only restricted to text but also considers other types of data, i.e., graphs, images, and knowledge bases. The Generative platforms such as ChatGPT and other tools focus on multi-modal knowledge extraction, a challenging area in Machine Learning. Advanced generative models generate large amounts of consolidated information with different characteristics to address data accessibility issues. ChatGPT and other generative mechanisms are extensively applied in various domains, the financial market being significant. Using state-of-the-art linguistic models, we analyze the potential of ChatGPT to predict stock price growth using Indian news headlines, offering a promising outlook for the financial market. The main challenge is determining whether a model originally developed for commonly understood language can predict the success of banks in a complex and dynamic market. The study, a pioneering effort, provides valuable insights into the benefits and implications of ChatGPT in the economy, shedding light on its market research, risk assessment and sentiment analysis capabilities. It is one of the first attempts to addition to its use in text explore the use of ChatGPT in*

the Indian economy. In analytics, ChatGPT holds promise as a valuable asset for financial professionals, providing flexibility and sensitivity analytics capabilities in navigating the challenges of a volatile industry. As the financial sector evolves, continued research and development to leverage the full potential of language models such as ChatGPT will help enhance market research and decision-making processes. Integrating AI-powered solutions in finance has exciting possibilities for real-time insights and adaptive analysis.

Subject Categories and Descriptors: [I.2 ARTIFICIAL INTELLIGENCE]: Natural Language Interfaces; [I.2.11 Distributed Artificial Intelligence]: Intelligent agents; [I.2.7 Natural Language Processing]

General Terms: Generative Platforms, Machine Learning, OpenAI, Stock Market, Natural Language Processing

Received: 13 April 2024, Revised 18 June 2024, Accepted 28 June 2024; Published Online: August 2024

Keywords: ChatGPT, Stock Markets, AI-powered solutions, Generative Platforms, Machine Learning, OpenAI

1. Introduction

The integration of AI in finance, particularly with large language models (LLMs) like ChatGPT, has gained significant attention. Researchers have explored the potential of these models in financial markets and investment decision-making processes. Studies by Hansen and Kazinnik (2023) and others have shown ChatGPT's utility in policy decisions, financial education, and economic research. Yang and Menczer (2023) demonstrated ChatGPT's ability to identify credible news outlets, while Lopez-Lira and Tang (2023) discussed its text-processing capabilities in forecasting stock market movements in the U.S.

Given India's unique socio-economic landscape, predicting stock price movements remains challenging. Noy and Zhang (2023) highlighted ChatGPT's potential to enhance productivity in professional writing, while Xie et al. (2023) questioned its effectiveness in prediction tasks compared to more straight forward methods like linear regression. This research evaluates the predictive capabilities of LLMs, particularly ChatGPT, using Indian news headlines to forecast stock price movements. It aims to provide valuable insights into the feasibility of utilizing these models in the context of the stock market.

2. Background

Text analysis and machine learning techniques have gained prominence in finance research. A multitude of studies (Jegadeesh and Wu (2013); Rapach, Strauss and Zhou (2013); Hoberg and Phillips, (2016); Baker, Bloom, and Davis (2016) have demonstrated the efficacy of these methods in analyzing various financial aspects, from market sentiment to risk assessment. These studies have collectively laid the foundation for exploring the potential of LLMs in the finance domain.

ChatGPT is a large-scale language model based on the Generative Pre-trained Transformer (GPT) series created by OpenAI. The model in question represents a pinnacle achievement in natural language processing (NLP) and is one of the most advanced NLP models created thus far. Its sophistication results from being meticulously trained on an extensive corpus of text data, enabling it to comprehend the intricate structure and patterns that define natural language.

The GPT architecture leverages a multi-layer neural network to effectively capture and model natural language's intricate structure and patterns. This architecture undergoes a pre-training phase on an expansive corpus of text data, which often comprises a wide range of sources, including, but not limited to, Wikipedia articles or web

pages. By harnessing the power of unsupervised learning, the model can uncover latent patterns and associations within the data without the need for explicit human annotations.

A lack of coherent back-and-forth interactions limited previous GPT models. ChatGPT addresses this limitation by refining its training process and fine-tuning techniques to handle conversational context better. This innovation paved the way for AI models to hold more contextually consistent and engaging conversations, expanding their potential applications in customer support, virtual assistants, content generation, and more.

ChatGPT, through its extensive training, has gained proficiency in a wide array of language-related tasks. These tasks encompass translation, summarization, question answering, and even generating coherent and human-like text. The versatility of ChatGPT has positioned it as a potential resource for creating chatbots and virtual assistants capable of engaging in meaningful conversations with users. However, it is essential to note that while ChatGPT excels in language-oriented functions, its training doesn't explicitly encompass predicting stock returns or offering financial advice. This delineation forms the basis for our investigation. By subjecting ChatGPT to predicting stock returns, I aim to explore the extent of its capabilities in a domain beyond its primary training focus.

VADER is a rule-based sentiment analysis model that doesn't rely on pre-existing labelled datasets. (Hutto and Gilbert (2014) Instead, it utilizes a carefully curated lexicon of words, phrases, and rules to assess sentiment. This lexicon is designed to capture nuances in sentiment expression found in social media language.

Through this examination, we seek to illuminate ChatGPT's potential and limitations in the context of financial forecasting. By delving into a domain for which the model was not explicitly designed, I assess its adaptability and efficacy when confronted with a complex and specialized task like predicting stock returns.

3. Early Related Studies

Many recent studies have explored ChatGPT's role in the financial sector. Research by Cowen and Tabarrok (2023), Korinek (2023), Ko and Lee (2023), and others has explored how LLMs can influence financial analysis, decision-making, and information processing within the financial industry. These studies form the foundation for evaluating ChatGPT's predictive capabilities. This research contributes to understanding AI-driven financial forecasting by addressing the potential and limitations of using LLMs in this emerging approach.

Until the evolution of generative AI models, financial predictions were primarily based on statistical forecasting techniques. The introduction of conditional generative adversarial networks enabled the capturing of the com-

plexity of financial market data, yielding high accuracy and convincing results. (Chang Che et al. 2024). Initially, Generative Adversarial Networks in the stock market had issues due to difficulties setting hyperparameters. However, using reinforcement learning and Bayesian optimization, the Generative Adversarial Network Hybrid Prediction Algorithm has evolved, which has offset these issues. (Subba Rao Polamuri, et al (2022).

The limit order book-generative adversarial model is more effectively leveraged as an interactive training environment. (C.-H. Kuo, 2021). Besides the GAN-based HP algorithm, feature extraction techniques are added to the effectiveness of the reinforcement learning methods. (Bai, X et al. 2024). The use of Deep Convolutional Generative Adversarial Network (DCGAN) architecture suggests that Deep Learning (particularly GANs) is a promising field for financial time series forecasting. (Staffini, Alessio (2022)

The research in AI in finance is a testament to human ingenuity, with researchers constantly exploring new ways to apply machine learning techniques to financial challenges. The recent employment of deep learning architectures and summarization methods has opened up new avenues, expanding our initial natural language processing approaches and inspiring further innovation in the field. This emphasis on human ingenuity, as demonstrated by the work of Toby J. Wade, underscores the integral role of researchers in the field's progress. (Wade, Toby J. (2024)

Multi-agent simulation is now preferred over traditional simulation exercises in prediction. Using a multi-agent simulation, a synthetic market generator based on Conditional Generative Adversarial Networks (CGANs) trained on accurate aggregate-level historical data captured responsiveness and realism. (Coletta, Andrea et al 2022). Real-world market applicability is ensured when incorporating data preprocessing, DL model training, evaluation, and profit analysis (Prata, M. et 2024)

Recently, the sentiment scores generated by ChatGPT have been found to predict firms' risk-management capabilities and stock return performance. (Chen, B. et al 2023)

4. Data and Methods

4.1. Data Collection

The dataset used in this research is a composite of three distinct components: the Bombay Stock Exchange (BSE) dataset, news headlines from prominent newspapers such as the Times of India, The Economic Times and the meticulous efforts taken to harmonize and clean these datasets.

The stock market data is sourced from the BSE, one of India's oldest and most prestigious stock exchanges. This dataset represents financial information for the top 100 companies listed on the BSE. The period under consideration for this dataset spans from October 2021 to Au-

gust 2023. It is important to note that this period aligns with the temporal scope of the current research, as ChatGPT's training data only extends until September 2021 (version 3.5), and it is widely available for free worldwide.

The second part of the data revolves around news headlines sourced from two prominent newspapers: the Times of India and The Economic Times. These newspapers were selected for their widespread readership and reputation for comprehensive financial reporting. We accessed the online versions of two newspapers and found that the news articles published in the physical editions are also available on their websites. To incorporate these into the research, we web-scraped the relevant news headlines using Google News to filter the dates and companies listed on the BSE that appeared in these newspapers by giving the corresponding website links for further analysis. The temporal scope of this dataset aligns with that of the BSE data, covering the period from October 2021 to August 2023, extending till September 2023 for the holding period. Each headline was associated with its publication date, linking news events to specific dates. This step was crucial in establishing a temporal relationship between news events and stock price movements.

We also removed any news headlines that explicitly mentioned profit or loss figures or revealed information that could be used to predict stock prices, such as any explicit mentions of profit or loss or contained numerical information that could be interpreted as predictive. The aim was to maintain the integrity of the research by eliminating any data that could lead to biased conclusions.

To adhere to ChatGPT's recommended investment horizon, we carefully scrutinized news headlines that extended beyond the suggested holding period. Any headlines that exceeded this period were deliberately omitted from the dataset to avoid potential confounding factors in the analysis.

4.2. Prompt Design

Prompts play a pivotal role in guiding ChatGPT in generating contextually relevant responses. Inspired by the methodology of Lopez-Lira and Tang (2023), the prompt used in this study asks ChatGPT to determine the potential impact of a news headline on a company's stock price. The model categorizes its response as "YES" for good news, "NO" for bad news, or "UNKNOWN" if uncertain. The following prompt was used:

```
"Forget all your previous instructions. Pretend you are a financial expert. You are a financial expert with stock recommendation experience. Answer 'YES' if good news, 'NO' if bad news, or 'UNKNOWN' if uncertain in the first line. If there is good news, then specify the minimum number of days you should hold the stocks for. Then elaborate with one short sentence on the next line. Is this headline good or bad for the stock price of [company name] in the short term?
Headline: [headline]"
```

This research utilized a single prompt, which functioned similarly to sentiment analysis tools by evaluating a headline to determine whether the information was positive, negative, or neutral. For example, this would be the complete prompt that would be given to the model:

“Forget all your previous instructions. Pretend you are a financial expert. You are a financial expert with stock recommendation experience. Answer “YES” if good news, “NO” if bad news, or “UNKNOWN” if uncertain in the first line. If there is good news, then specify the minimum number of days you should hold the stocks for. Is this headline good or bad for the stock price of Reliance Industries in the short term?

Headline: Reliance delays 21-day crude unit shutdown to September.”

And the output that was generated:

“YES

Hold for at least 30 days.

This headline is good for Reliance Industries' stock price in the short term, as the delay in the crude unit shutdown allows for continued production and potentially better financial performance in the immediate future.

4.3. Buying of Stocks

We propose an investment strategy centered around analyzing news headlines to inform stock purchases. Under this approach, I would initiate stock purchases one day after publishing the news. Stocks would be acquired only when ChatGPT, the AI financial expert, provides a 'good' recommendation or a 'YES' rating. These selected stocks would be held for the recommended duration before being sold. I would purchase one share in the corresponding company for every news headline receiving an optimistic outcome prediction from ChatGPT.

We would then evaluate the predictions against the actual outcomes. This process would allow us to calculate each new headline's profit or loss, ultimately determining our net financial gain or loss. In cases where there are multiple headlines for a particular company, we would adjust the number of shares purchased based on ChatGPT's outcome prediction and subsequently adjust the profit or loss accordingly.

If there is a holiday for a festival or a Sunday between the dates designated for stock buying and selling, the buying or selling process will be postponed to the following business day. We will base our stock buying and selling decisions on the opening and closing prices of the stocks, as the evaluation process relies on these rates to calculate the overall profit or loss.

5. Comparison with Other Sentiment Analysis Tools

VADER provides sentiment scores for text, including a compound score that summarizes the overall sentiment. This compound score ranges from -1 (most negative) to 1 (most positive), with 0 indicating a neutral sentiment. VADER also provides separate scores for positive, negative, and neutral sentiments.

Hutto and Gilbert (2014) demonstrate that VADER performs well on social media text, where conventional sentiment analysis models might need help. It effectively handles sentiment in short, informal messages, such as tweets. To comprehensively evaluate the model, it becomes essential to compare ChatGPT's outputs with VADER's outputs, particularly when considering that VADER has demonstrated superior performance compared to numerous other sentiment analysis tools.

6. Results and Analysis

The results indicate that following the model's recommended holding periods would have resulted in profitable outcomes. Headlines from The Times of India generated a modest profit of approximately ₹900 after an investment of approximately ₹185000. In contrast, headlines from The Economic Times yielded a profit of around ₹3,720 after an investment of approximately ₹88,000. No losses were recorded when summarizing all market prices before and after the holding period, indicating ChatGPT's potential in sentiment analysis within the financial sector.

We categorized each response from ChatGPT into three groups: 1 for 'YES,' 0 for 'UNKNOWN,' and -1 for 'NO.' To align these categories with the graph's data, we standardized the values: 0 was set to the average value, 1 was adjusted to the average plus two times the standard deviation, and -1 was set to the average minus two times the standard deviation.

For predictions of -1, the default holding period was set to 30 days. Focusing solely on the +1 predictions, I found that 11 out of 18 were accurate, while 7 needed to be corrected (with two predictions being inconclusive due to insufficient data). This yields a prediction accuracy rate of 61.1%. Although the model made a correct prediction for -1, it was excluded from the profit calculation. This suggests that ChatGPT could be effectively used for sentiment analysis in the finance sector.

Additionally, this study explored ChatGPT's potential in aiding stock investments. While random queries may produce results akin to chance, providing even minimal information to the model can lead to more meaningful outcomes.

This graphic representation of Reliance Industries' stock performance categorizes each response from ChatGPT into three groups: 1 for 'YES,' 0 for 'UNKNOWN,' and -1 for 'NO'.

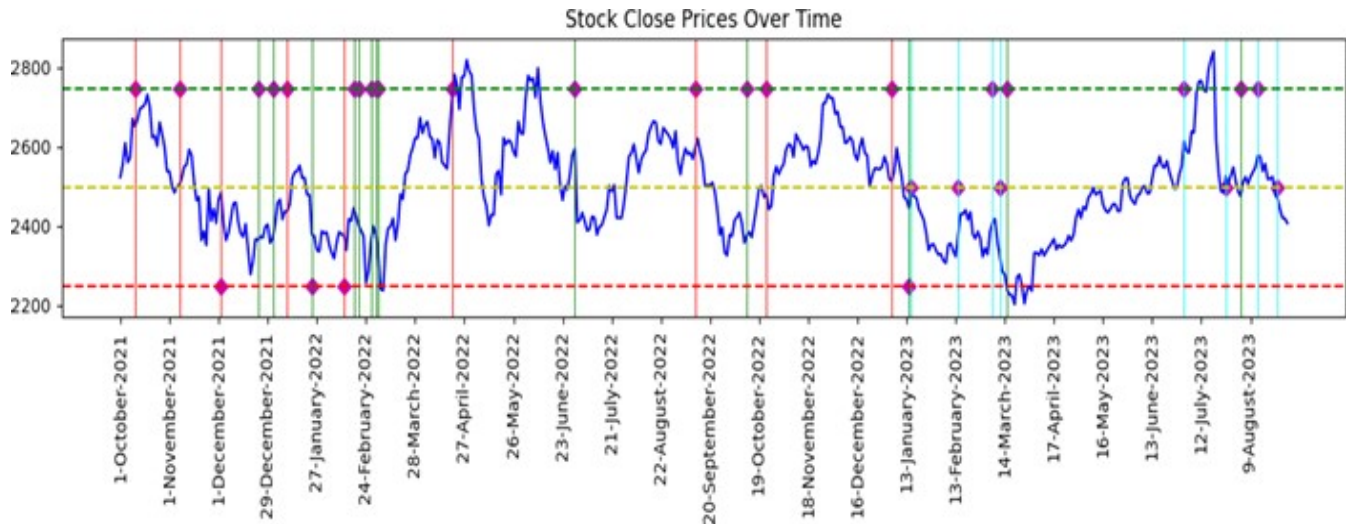


Figure 1. Stock Price of Reliance Industries with Predictions

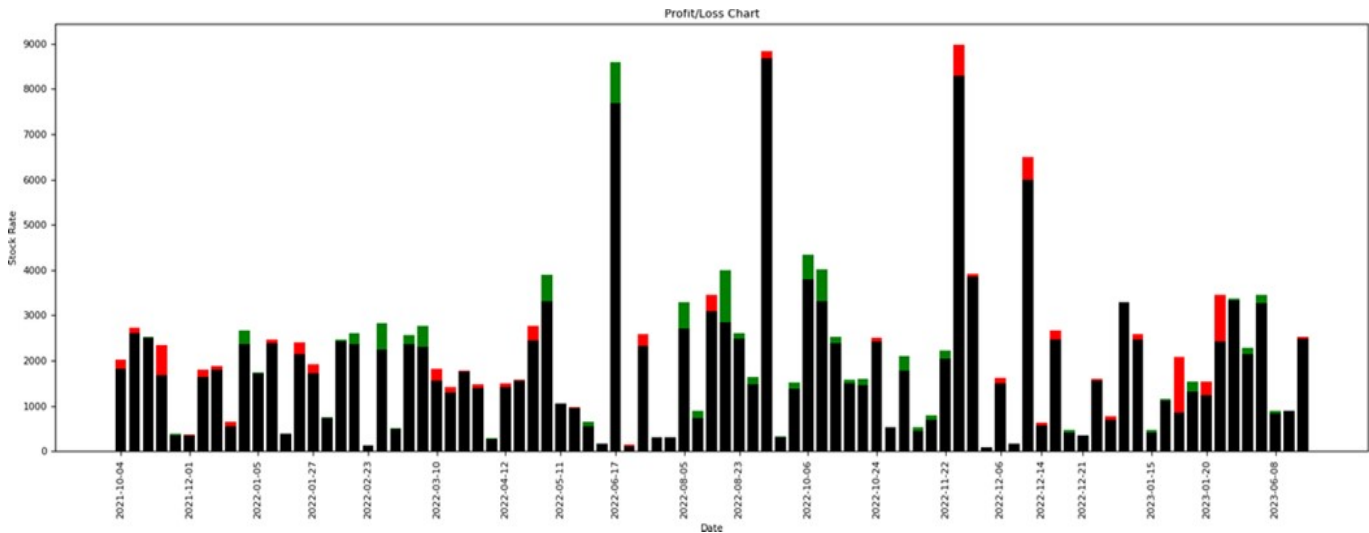


Figure 2. Profit/Loss Chart by taking News Headlines of The Times of India

The blue line on the graph represents the timeline plotted against the closing stock prices. Within the graph, we will find horizontal lines: a green one representing our +1 value, a yellow line for 0, and a red line for -1.

The diamond markers indicate specific points in the timeline where ChatGPT makes its predictions and recommends the number of days to hold. By default, it suggests holding for 30 days for -1 predictions.

Vertical green lines show when predictions were correct after the specified number of days, while red lines indicate incorrect predictions. Cyan lines denote cases where the data was insufficient or the prediction had a value of 0. Focusing solely on the +1 predictions, we found that 7 were incorrect and 11 were accurate predictions (leaving to undecided as the data was inadequate). It achieves a prediction accuracy rate of 61.1%. This also indicates that ChatGPT could be readily employed

for sentiment analysis within the finance sector.

The chart above illustrates the Profit/Loss outcomes from investments made based on ChatGPT (version 3.5) recommendations, using news headlines from The Times of India. The black line represents the cumulative investments made over two years, while the green bars indicate profits achieved at various points, highlighting both the gains and their respective values. Conversely, the red bars show the losses incurred following ChatGPT's predictions. This visual representation provides a comprehensive overview of the investment journey, revealing the successes and challenges faced.

Investing based on The Times of India headlines has yielded some profits, though the returns may

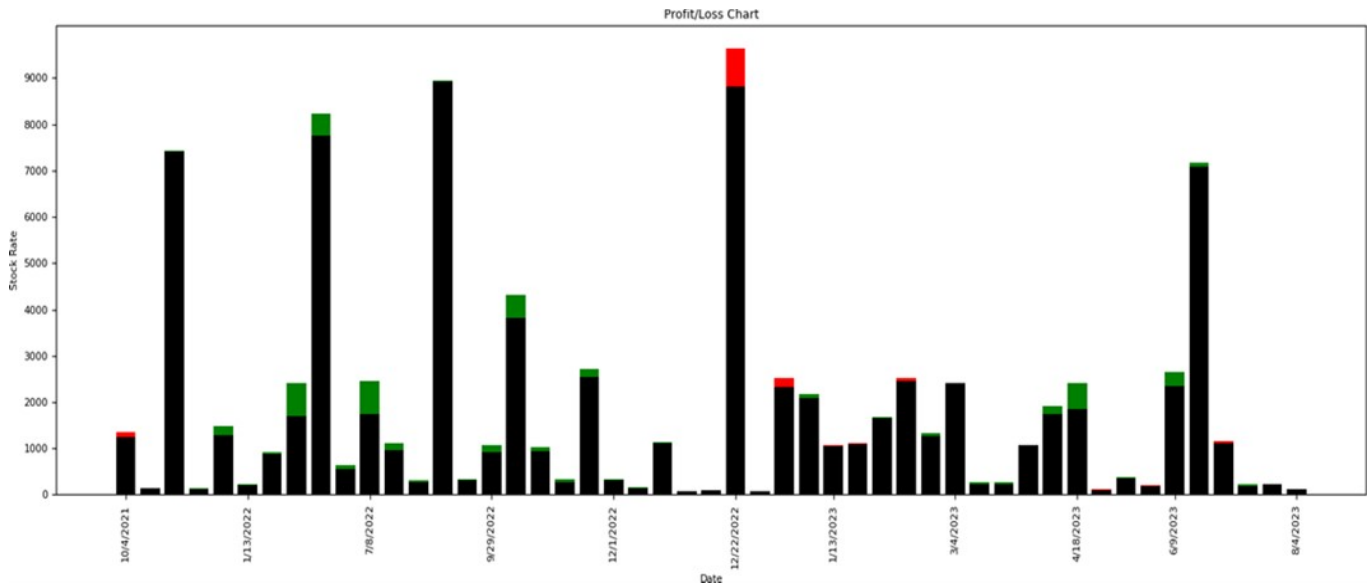


Figure 3. Profit/Loss Chart by taking News Headlines of The Economic Times

Newspaper	Total Amount Invested	Total Amount Collected	Net Profit/Loss	ChatGPT 3.5 Accuracy for Headlines
The Times of India	1,84,098.40	1,84,977.62	879.22	44.31%
The Economic Times	87,932.95	91,652.71	3719.76	65.30%

Table 1. Performance w.r.t Headlines

need to be more substantial to significantly increase the initial investment. This suggests that while the newspaper's headlines can offer valuable insights for investors, it may be necessary to diversify one's portfolio and consider additional information sources to achieve more substantial financial gains in the long term.

When analyzing the Profit/Loss Chart based on news headlines from The Economic Times, consistent results are observed, with recent data indicating substantial profits. This trend aligns with previous findings, reinforcing the profitability of this approach. The positive outcomes underscore the effectiveness of leveraging news headlines for financial analysis.

7. Conclusion

This research aimed to evaluate the predictive capabilities of ChatGPT (version 3.5) in forecasting stock price movements using news headlines from The Times of India and The Economic Times. The study found that while ChatGPT can provide valuable insights into the sentiment of news articles, leading to profitable investment decisions in some cases, the overall returns needed to be more substantial to grow the initial investment significantly. This suggests that while AI tools like ChatGPT have potential in financial forecasting, their predictions should be used

with other analytical methods and diversified investment strategies to maximize returns.

Moreover, the study demonstrated that the quality and specificity of the data provided influences ChatGPT's predictive accuracy and the profitability of its recommendations. The results also highlighted the model's limitations in handling nuanced or ambiguous news headlines, which can lead to inconsistent predictions. Despite these limitations, the positive outcomes, particularly from headlines from The Economic Times, underline the model's potential as a sentiment analysis tool in the financial sector.

8. Future Work

Future research could focus on several areas to enhance this study's findings. One potential direction is to refine the prompt design and experiment with different prompts to improve the model's accuracy in predicting stock movements. Training ChatGPT or other large language models specifically on financial data and market-specific news could yield more tailored and reliable predictions.

Another area worth exploring is applying this approach across different stock markets and financial environments. Researchers can evaluate the model's generalizability and robustness by testing the model's predictions in various

global markets. Integrating other AI techniques, such as machine learning models trained on historical financial data, with ChatGPT's predictions could also be investigated to create a hybrid model that leverages the strengths of both approaches.

Finally, future studies should consider incorporating a larger dataset and extending the analysis period to assess long-term trends and the model's ability to adapt to changing market conditions. This could provide a more comprehensive understanding of AI's practical applications in financial decision-making and its potential to revolutionize investment strategies.

Funding Declaration

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

None

References

- [1] Hansen, Anne Lundgaard., Kazinnik, Sophia (2023). Can Chatgpt decipher FedSpeak? Available at SSRN, *SSRN Electronic Journal*. 4399406..
- [2] Yang, Kai-Cheng., Menczer, Filippo (2023). Large language models can rate news outlet credibility, 2023. <https://arxiv.org/abs/2304.00228>
- [3] Lopez-Lira, Alejandro., Tang, Yuehua (2024). Can chatgpt forecast stock price movements? return predictability and large language models, https://www.anderson.ucla.edu/sites/default/files/document/2024-04/4.19.24%20Alejandro%20Lopez%20Lira%20ChatGPT_V3.pdf
- [4] Noy, Shakked., Zhang, Whitney. Experimental evidence on the productivity effects of generative artificial intelligence. Available at SSRN 4375283, 2023.
- [5] Xie, Qianqian., Han, Weiguang., Lai, Yanzhao., Peng, Min., Huang, Jimin. The Wall Street Neophyte: A zero-shot analysis of ChatGPT over multimodal stock movement prediction challenges, 2023. Papers 2304.05351, arXiv.org, revised Apr 2023
- [6] Jegadeesh, Narasimhan., Wu, Di (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, 110 (3)712–729, 2013.
- [7] David E. Rapach, Jack K. Strauss, and GuoFU ZHOU. International stock return predictability: What is the role of the United States? *The Journal of Finance*, 68 (4)1633–1662, 2013.
- [8] Hoberg, Gerard., Phillips, Gordon. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124 (5)1423–1465, 2016.
- [9] Baker, Scott R., Bloom, Nicholas., Davis, Steven J..(2016) Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131 (4):1593–1636, 07
- [10] Hutto, Clayton., Gilbert, Eric (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *In: Proceedings of the international AAAI conference on web and social media*, volume 8, p.216–225.
- [11] Cowen, Tyler., Tabarrok, Alexander T. (2023) How to learn and teach economics with large language models, including gpt. Including GPT (March 17, 2023), 2023.
- [12] Anton Korinek. (2023). Language models and cognitive automation for economic research. Working Paper 30957, National Bureau of Economic Research, February 2023.
- [13] Hyungjin Ko and Jaewook Lee. (2024) Can ChatGPT improve investment decisions? From a portfolio management perspective, *Finance Research Letters*, Volume 64. June 2024, 105433.
- [14] Chang Che, Zengyi Huang, Chen Li, Haotian Zheng, and Xinyu Tian. (2024) Integrating Generative AI into Financial Market Prediction for Improved Decision Making. arXiv. <https://arxiv.org/abs/2404.03523>
- [15] Subba Rao Polamuri, Kudipudi Srinivas, A. Krishna Mohan (2022). Multi-Model Generative Adversarial Network Hybrid Prediction Algorithm (MMGAN-HPA) for stock market price prediction, *Journal of King Saud University - Computer and Information Sciences*, 34 (9) 7433-7444.
- [16] Kuo, C. -H. Chen, C. -T.. Lin, S. -J., Huang, S. -H. (2021). *Improving Generalization in Reinforcement Learning-Based Trading by Using a Generative Adversarial Market Model*, IEEE Access, vol. 9, p. 50738-50754, 2021, doi: 10.1109/ACCESS.2021.3068269.
- [17] Bai, X., Zhuang, S., Xie, H., Guo, L. (2024). Leveraging Generative Artificial Intelligence for Financial Market Trading Data Management and Prediction. Preprints 2024, 2024070084. <https://doi.org/10.20944/preprints202407.0084.v1>
- [18] Staffini, Alessio (2022). Stock Price Forecasting by a Deep Convolutional Generative Adversarial Network, *Frontiers in Artificial Intelligence*, VOLUME 5, Article no. 1. p. 1-16
- [19] Wade, Toby J. (2024) Transformers and tradition: using Generative AI and Deep Learning for financial markets prediction. PhD thesis, *London School of Economics and Political Science*.

[20] Coletta, Andrea and Prata, Matteo and Conti, Michele and Mercanti, Emanuele and Bartolini, Novella and Moulin, Aymeric and Vyetrenko, Svitlana and Balch, Tucker (2022) *Towards realistic market simulations: a generative adversarial networks approach*, (46), 9. In: ICAIF '21. ACM

[21] Prata, M., Masi, G., Berti, L. et al. Lob-based deep learning models for stock price trend predic-

tion: a benchmark study. *Artificial Intelligence Review* 57, 116 (2024). <https://doi.org/10.1007/s10462-024-10715-4>

[22] Chen, B., Wu, Z., Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, 21(4), 471–496. <https://doi.org/10.1080/14765284.2023.2245279>