

Karamveer Singh

18 September 2023

Can ChatGPT Predict Indian Stock Market Price Movements?

Abstract

The combination of text analytics and machine learning technologies in the context of finance has made a huge difference. This study addresses the need to explore the usability of ChatGPT and other advanced language models in the Indian financial market. Using the power of state-of-the-art linguistic models, we analyze the potential of ChatGPT, a top-notch natural language platform, to predict stock price growth with Indian news headlines. The main challenge is to determine whether a model originally developed for commonly understood language can predict the success of banks in a complex and dynamic market such as India. The study provides valuable insights into the benefits and implications of ChatGPT in the Indian economy, shedding light on its market research, risk assessment and sentiment analysis capabilities. It serves as one of the first attempts to explore the use of ChatGPT in the Indian economy. In addition to its use in text analytics, ChatGPT holds promise as a valuable asset for financial professionals working in India, providing flexibility and sensitivity analytics capabilities in navigating the challenges of a volatile industry. As the financial industry continues to evolve, continued research and development in leveraging the full potential of language models such as ChatGPT will help enhance market research and decision-making processes. The integration of AI-powered solutions in finance has interesting possibilities for real-time insights and adaptive analysis.

RELATED LITERATURE

Text analysis and machine learning techniques have gained prominence in finance research. A multitude of studies (Jegadeesh and Wu, 2013[1]; Rapach, Strauss and Zhou, 2013[2]; Hoberg and Phillips, 2016[3]; Baker, Bloom, and Davis, 2016[4]) 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.

Many papers have delved into ChatGPT to determine the potential of large language models (LLMs) like GPT-3, BERT, T5, XLNet etc, in finance markets and the investment decision-making processs. This could be an emerging area of research that explores how LLMs can influence financial analysis, decision-making, and information processing within the context of the financial industry. Hansen and Kazinnik (2023)[5], Cowen and Tabarrok (2023)[6], Korinek (2023)[7], Ko and Lee (2023)[8] have shown how ChatGPT can perform policy decisions, can be used in financial education, conducting economic research, and can be applicable in investment decisions. Yang and Menczer (2023)[9] demonstrates how ChatGPT can successfully identify credible news outlets, indicating its potential in information verifications and source evaluation. Lopez-Lira and Tang (2023) [10] have discussed the text processing capabilities of recently developed LLMs such as ChatGPT in forecasting stock market movements based in the US. Hutto and Gilbert (2014)[11] introduced VADER, a sentiment analysis tool designed to discern and assign sentiment scores to online comments within the realm of social media, effectively distinguishing between various sentiments.

The unique socio-economic landscape of India makes predicting stock price movements a challenging task. Noy and Zhang (2023)[12] highlighted ChatGPT's potential in enhancing productivity in professional writing jobs, indicating its utility in processing large volumes of news headlines. However, Xie et al. (2023) [13] had other views, indicating ChatGPT might not perform signif-

icantly better than simple methods like linear regression in prediction tasks involving numerical data, raising a question about its effectiveness in a stock market context.

This literature review serves as a foundation for my research, which focuses on evaluating the predictive capabilities of LLMs, particularly ChatGPT, using Indian news headlines to forecast stock price movements. By conducting a comprehensive analysis of existing literature, including various AI-driven financial forecasting approaches, I aim to contribute to the field's understanding of AI-driven financial forecasting, addressing both the potential and limitations associated with this emerging approach. By considering the nuances of Indian news headlines and LLMs' capabilities, my research aims to shed light on the feasibility of utilizing these cutting-edge models in the Indian stock market context. This study holds the potential to provide valuable insights into the practicality and implications of employing LLMs like ChatGPT for predictive analysis in the dynamic realm of Indian finance.

BACKGROUND

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 the field of natural language processing (NLP) and stands as one of the most advanced NLP models created thus far. Its sophistication is a result of being meticulously trained on an extensive corpus of text data, enabling it to comprehend the intricate structure and intricate patterns that define natural language.

The GPT architecture leverages a multi-layer neural network to effectively capture and model the intricate structure and intricate patterns inherent in natural language. Through the utilization of unsupervised learning techniques, this architecture undergoes a pre-training phase on an expansive corpus of text data. This corpus 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.

Previous GPT models were limited by a lack of coherent back-and-forth interactions. ChatGPT addresses this limitation by refining its training process and fine-tuning techniques to better handle conversational context. 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 the generation of coherent and human-like text. The versatility of ChatGPT has positioned it as a potent resource for the creation of 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 the task of 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 labeled datasets. 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, I seek to shed light on the potential and limitations of ChatGPT in the context of financial forecasting. By delving into a domain for which the model was not explicitly designed, I endeavor to assess its adaptability and efficacy when confronted with a complex and specialized task like predicting stock returns.

DATA

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 sourced from the BSE, one of India's oldest and most prestigious stock exchanges. This dataset represents a collection of financial information for the top 100 companies listed on the BSE. The period under consideration for this dataset spans from October 2021 to August 2023. It is important to note that this period aligns with the temporal scope of my research, as ChatGPT's training data only extends until September 2021 (version 3.5) and it is widely available for free across the world.

The second part of our data revolves around news headlines sourced from three prominent newspapers: the Times of India and The Economic Times. These newspapers were selected for their widespread readership and reputation for comprehensive financial reporting. I used Google news to filter all the news on their websites which had the name of the company included and then developed a python script (web scraping) to collect news headlines from their online sources. 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, allowing me to link news events to specific dates. This step was crucial in establishing a temporal relationship between news events and stock price movements.

I have also removed any news headlines that explicitly mentioned profit or loss figures or reveal 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 in nature. The aim was to maintain the integrity of the research by eliminating any data that could lead to biased conclusions.

To adhere to the investment horizon recommended by ChatGPT, I 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.

METHOD OF WORKING

1. Prompt Design

Prompts play a pivotal role. These prompts can vary in length, from a concise line to a comprehensive paragraph, but their primary function remains consistent – to provide specific details that guide ChatGPT in generating contextually relevant responses.

The approach I adopt here draws inspiration from a template utilized in the work of Lopez-Lira and Tang (2023)[10]. The methodology involves a meticulous analysis of both the syntax and semantics of the prompts, generating a range of potential responses, and ultimately selecting the most suitable response based on a variety of criteria, including coherence, relevance, and grammatical correctness.

The prompt that has been designed asks the outcome of a stock for a short period of time as the model would say “UNKNOWN” if it is not specified in the prompt.

I have used the following prompt to give it 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 minimum number of days you should hold the stocks for. Then elaborate with one short and concise sentence on the next line. Is this headline good or bad for the stock price of

_company_name_ in the short term?

Headline: _headline_

The prompt in question is tailored specifically for financial analysis, instructing ChatGPT to evaluate a given news headline and determine its potential short-term impact on a company's stock price. ChatGPT's response should be categorized as either "YES" if the news is deemed favorable for the stock price, "NO" if it is considered detrimental, or "UNKNOWN" if the impact remains uncertain. Furthermore, ChatGPT is tasked with providing a concise one-sentence explanation to support its chosen answer. It is important to note that the prompt specifies that the news headline is the sole source of information provided to ChatGPT. Implicitly, it assumes that the headline contains adequate information for a financial expert to reasonably gauge its influence on stock prices. This tailored prompt serves as a demonstration of ChatGPT's proficiency as a language model in the realm of financial analysis.

For every prompt the model would give either a "YES", "NO" or "UNKNOWN" for example:

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 minimum number of days you should hold the stocks for. Then elaborate with one short and concise sentence on the next line. 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 the stock price of Reliance Industries 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.”

2. Buying of Stocks

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

I would then evaluate the predictions against the actual outcomes. This process would allow me to calculate the profit or loss associated with each new headline, ultimately determining our net financial gain or loss. In cases where there are multiple headlines for a particular company, I 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 occurring between the dates designated for stock buying and selling, the buying or selling process will be postponed to the following business day. I will base my 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.

3. 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 sentiment.

Hutto and Gilbert (2014)[11] demonstrates that VADER performs well on social media text, where conventional sentiment analysis models might struggle. It is shown to be particularly effective at handling sentiment in short, informal messages, such as tweets. To comprehensively evaluate the model, it becomes essential to compare ChatGPT's outputs with those of VADER, particularly when considering that VADER has demonstrated superior performance when compared to numerous other sentiment analysis tools.

RESULTS

I categorized each response from ChatGPT into three groups: 1 for 'YES,' 0 for 'UNKNOWN,' and -1 for 'NO.' To standardize these values relative to the graph's data, I assigned the following numerical values: 0 was set to the average value, 1 was assigned as the average plus two times the standard deviation, and -1 was designated as the average minus two times the standard deviation.

By default, it suggests holding for 30 days for -1 predictions. Focusing solely on the +1 predictions, I found that 7 of them were incorrect and 11 were accurate predictions (leaving two +1 predictions which were undecided as the data was inadequate), It achieves a prediction accuracy rate of 61.1%. Even though the model made a correct prediction for -1, we can disregard it for the purpose of calculating profit. This suggests that ChatGPT could be effectively utilized for sentiment analysis in the finance sector.

This study also examined ChatGPT's potential assistance in stock investments. While random queries to the model may yield outcomes akin to luck, providing even a small amount of information can lead to more meaningful results.

Following the model's recommended holding periods would have resulted in profitable outcomes. Additionally, this research explored whether certain newspapers could generate more informative headlines, enhancing prediction accuracy. By analyzing headlines from The Times of India and The Economic Times, the data revealed a modest profit of approximately 900 after investing approximately 1.85 lakh in the former and a profit of around 3720 after investing approximately 88,000 in the latter.

Notably, there were no losses recorded when summarizing all market prices before and after the holding period.

CONCLUSION

In this paper, I have tried to determine the potential of ChatGPT, in predicting stock prices using Indian news headlines as an input. My findings indicated how ChatGPT could be used in the financial sector in the upcoming times as an sentiment analysis tool as it outperforms other sentiment tools.

This research stands as one of the initial step into exploring ChatGPT's applicability in the Indian financial sector. It has demonstrated that a model not explicitly trained in finance can still yield valuable insights and assist in stock price predictions. While the study has provided promising results, it also underscores the need for further exploration and refinement of large language models in the financial domain.

This research has raised potential future inquiries: What would be the outcome if ChatGPT were trained with this data, and how would it perform in comparison to other Large Language Models like Google AI's Bard and others? In this instance, I utilized ChatGPT version 3.5, but what advantages might be gained from using more advanced versions?

In addition to functioning as a text analysis tool, ChatGPT has the ability to find applications in market research. However, this study has shown how the model can be used in risk analysis, especially in analyzing sentiment in financial markets. If future iterations of the model can better adjust to market conditions and news increasingly, it could be a valuable asset in the Indian economy. This category is characterized by variable sentiment and a rapid flow of input, making ChatGPT's flexibility and sentiment analysis capabilities particularly useful.

FIGURES

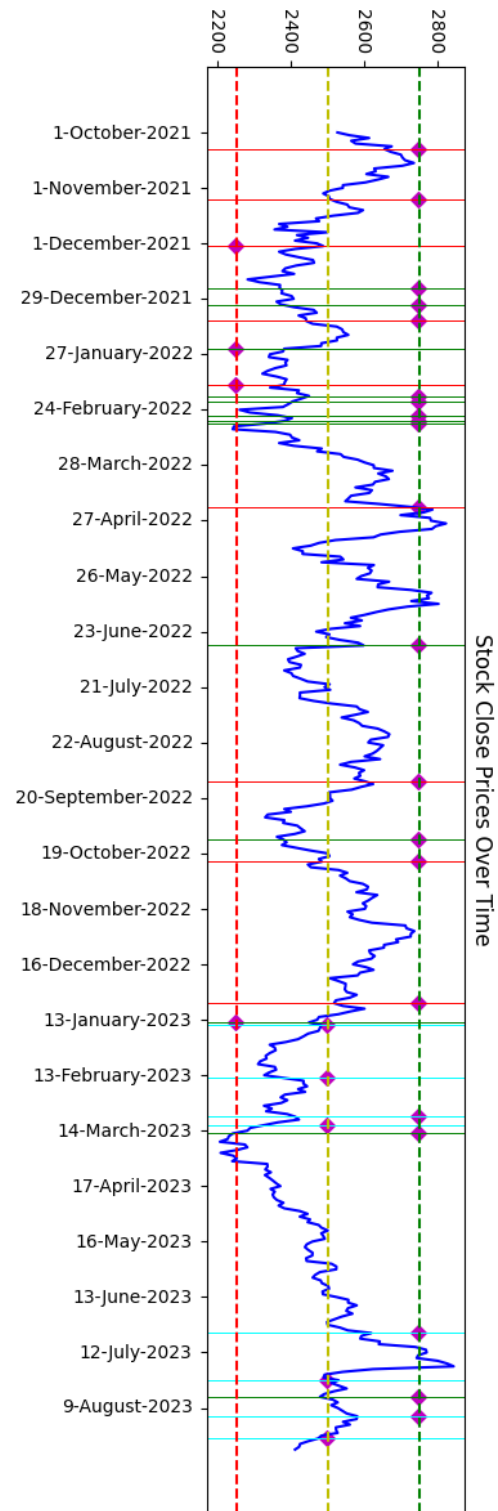


Figure 1: Stock Price of Reliance Industries with Predictions

Depicting the graphical representation of Reliance Industries' stock performance. categorizing each response from ChatGPT into three groups: 1 for 'YES,' 0 for 'UNKNOWN,' and -1 for 'NO.'

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 turned out to be correct after the specified number of days, while red lines indicate incorrect predictions. Cyan lines denote cases where either the data was insufficient or the prediction had a value of 0. Focusing solely on the +1 predictions, we found that 7 of them were incorrect and 11 were accurate predictions (leaving 2 which were 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.

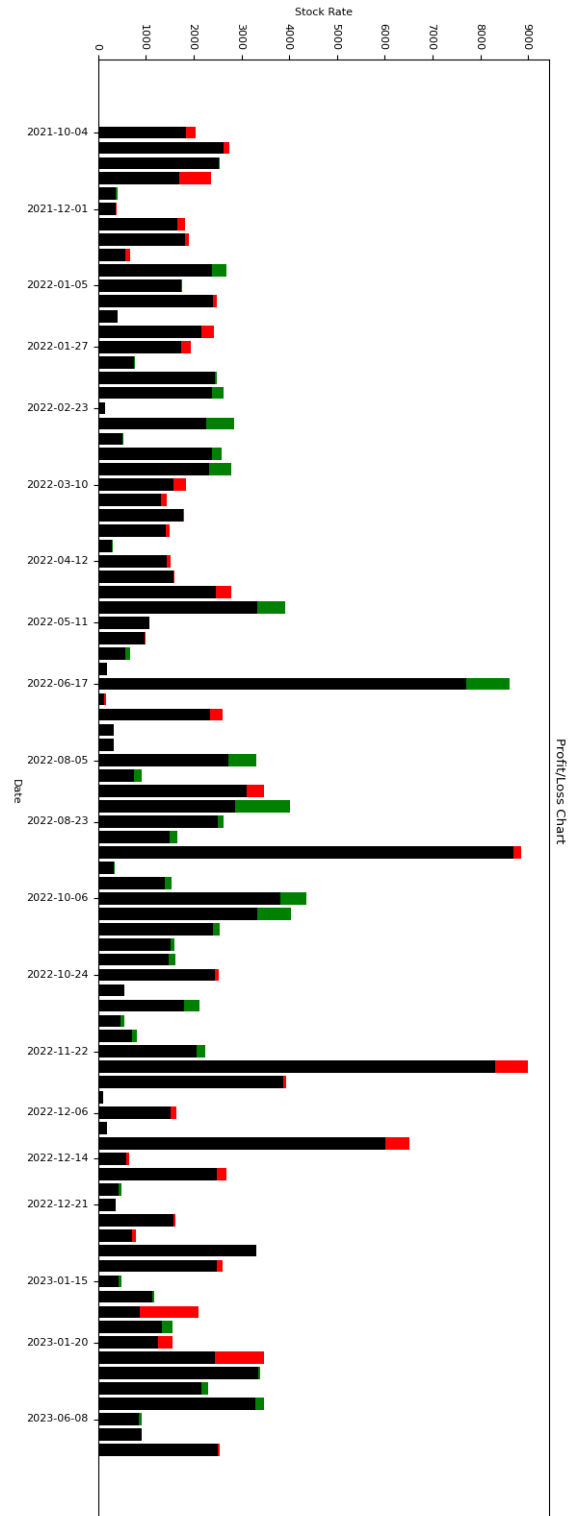


Figure 2: Profit/Loss Chart by taking News Headlines of The Times of India

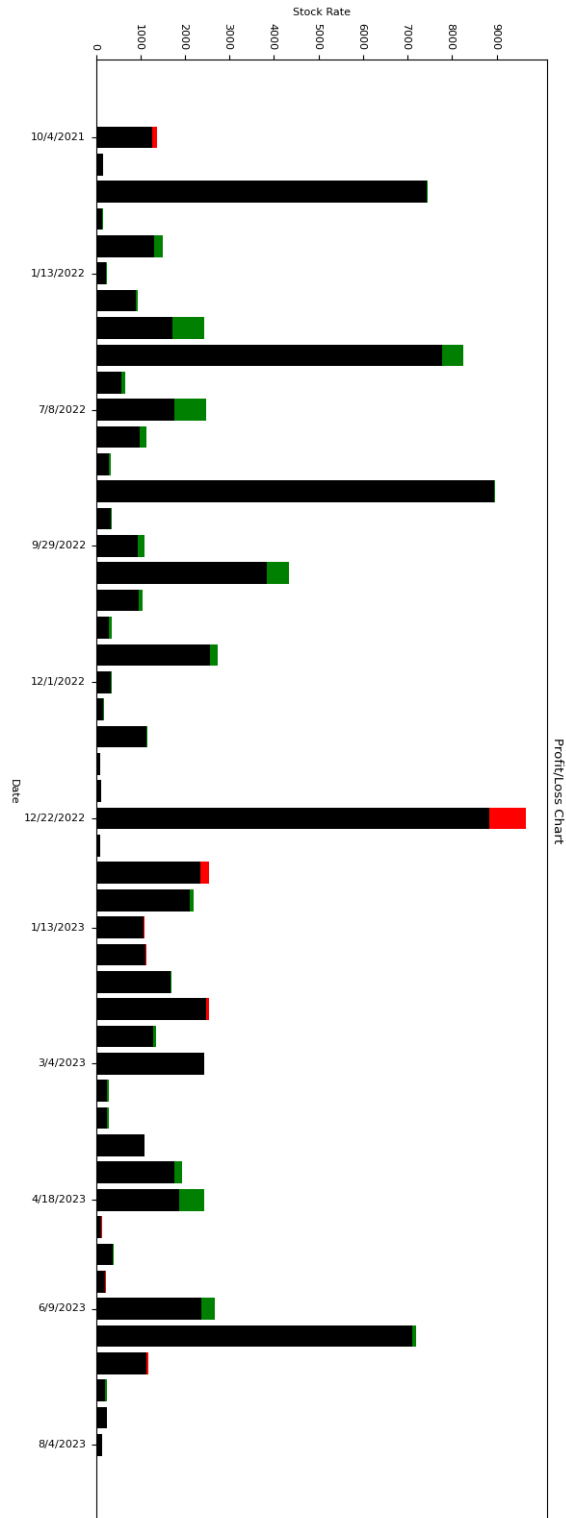


Figure 3: Profit/Loss Chart by taking News Headlines of The Economic Times

The chart displayed above illustrates the Profit/Loss Chart resulting from investments made based on the recommendations provided by ChatGPT (version 3.5) and taking the news headlines from The Times Of India.

The black line represents the cumulative investments made over a span of two years. Meanwhile, the green bars denote the profits realized at specific points in time, showcasing both the gains and their respective values. In contrast, the red bars signify the losses incurred during the course of following ChatGPT's predictions. This visual representation offers a comprehensive overview of the investment journey, highlighting both the successes and challenges encountered along the way.

Investing based on The Times Of India headlines has been shown to yield some profits, although the returns may not be substantial enough to significantly grow the initial investment. This suggests that while the newspaper's headlines can offer valuable insights for investors, diversifying one's portfolio and considering additional sources of information may be necessary to achieve more significant financial gains in the long run.

When analyzing the Profit/Loss Chart based on The Economic Times' news headlines, I observe consistent results, with the latest data indicating substantial profits. This trend mirrors previous findings and underscores the profitability of my approach. This positive outcome underscores the robustness of our strategy in leveraging news headlines for financial analysis.

TABLES

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

Examining the table, it's evident that both The Times of India and The Economic Times have shown notable positivity in stock predictions when assisted by ChatGPT. Notably, The Times of India saw a higher investment in predictions deemed profitable by ChatGPT, but the actual profit was considerably lower. Conversely, The Economic Times had lower investments in comparison but managed to achieve a 4 percent profit. It's worth noting that more headlines were sourced from The Times of India, but they proved to be less relevant, raising the question of whether The Economic Times provides more pertinent financial news. Despite the modest profits, the absence of losses suggests that a more precise approach could potentially maximize profits.

Work Cited

- [1] Narasimhan Jegadeesh and Di Wu. Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3):712–729, 2013.
- [2] 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.
- [3] Gerard Hoberg and Gordon Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.
- [4] Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring Economic Policy Uncertainty*. *The Quarterly Journal of Economics*, 131(4):1593–1636, 07 2016.
- [5] Anne Lundgaard Hansen and Sophia Kazinnik. Can chatgpt decipher fedsspeak? *Available at SSRN*, 2023.
- [6] Tyler Cowen and Alexander T Tabarrok. How to learn and teach economics with large language models, including gpt. *Including GPT (March 17, 2023)*, 2023.
- [7] Anton Korinek. Language models and cognitive automation for economic research. Working Paper 30957, National Bureau of Economic Research, February 2023.
- [8] Hyungjin Ko and Jaewook Lee. Can chatgpt improve investment decision? from a portfolio management perspective. *From a Portfolio Management Perspective*, 2023.
- [9] Kai-Cheng Yang and Filippo Menczer. Large language models can rate news outlet credibility, 2023.
- [10] Alejandro Lopez-Lira and Yuehua Tang. Can chatgpt forecast stock price movements? return predictability and large language models, 2023.

- [11] Clayton Hutto and Eric Gilbert. 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, pages 216–225, 2014.
- [12] Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283*, 2023.
- [13] Qianqian Xie, Weiguang Han, Yanzhao Lai, Min Peng, and Jimin Huang. The wall street neophyte: A zero-shot analysis of chatgpt over multimodal stock movement prediction challenges, 2023.