Large Language Model for Financial Text Understanding

Rahul Bhutna (211EC238)
Electronics and Communication Engineering,
National Institute Of Technology Karnataka
Surathkal, India
rahulbhutna.211ec238@nitk.edu.in

Sankalp Dev (211EC246)
Electronics and Communication Engineering,
National Institute Of Technology Karnataka
Surathkal, India
sankalpdev.211ec246@nitk.edu.in

***Abstract*—This report outlines the development of a financial text understanding system using a Small Language Model (SLM) enhanced with Low-Rank Adaptation (LoRA). The system is trained on a custom dataset that integrates financial news, corporate statements, and NIFTY 50 stock market data, specifically tailored to the Indian financial market. By incorporating sentiment analysis, the model effectively captures market sentiment, enabling accurate stock movement predictions. The use of LoRA ensures the model's efficiency and adaptability, making it suitable for deployment in resource-constrained environments. Insights from a comprehensive literature review guided the project's design, emphasizing the importance of domain-specific tuning, resource efficiency, and data diversity. The result is a robust, efficient financial prediction tool capable of navigating the complexities of the Indian market.**

Keywords: ***Financial Text Analysis, Small Language Model, Low-Rank Adaptation, Sentiment Analysis, NIFTY 50, Financial Prediction, Indian Market, Dataset Creation, Model Fine-Tuning***

I. Introduction

In the fast-changing world of finance, making smart decisions means having access to accurate and timely information. With the rise of artificial intelligence, especially Large Language Models (LLMs), there's now a new way to understand large amounts of financial data. These models, which can read and interpret natural language, are becoming essential tools for financial analysts, investors, and decision-makers.

This project focuses on using LLMs to better understand and predict financial trends in the Indian market. The goal is to create a system that can analyze financial news, statements, and stock data to give useful insights. The system will be trained on a custom dataset that combines different types of data, helping it to make accurate predictions and recommendations.

With advancements in artificial intelligence, particularly Large Language Models (LLMs), there is now an opportunity to enhance financial decision-making. LLMs can process and interpret natural language, making them valuable for analyzing financial news, company statements, and stock market data. However, effectively using these models requires addressing challenges such as computational efficiency and nuances related to Indian market.

To tackle the challenge of computational resources and efficiency, the project uses a Small Language Model (SLM) with Low-Rank Adaptation (LoRA) techniques to keep the model efficient and suitable for use in places with limited computing power, while still performing well.

The model can take the news, or financial Key performance indicators as the input and try to predict about the company finacials in a textual format.

The project has three main goals: to build a dataset that captures the complexities of the Indian financial market, to make the model efficient, and to ensure it works well even with new, unseen data. These efforts aim to create a more accurate, adaptable, and efficient tool for making financial decisions.

II. Literature Review

The field of Natural Language Processing (NLP) in finance has seen significant advancements in recent years, with a focus on developing specialized models and datasets for financial applications. This review explores key contributions in this domain, highlighting the progress and challenges in applying large language models (LLMs) to financial tasks.

The FINQA dataset [1] addresses a crucial gap in existing question-answering (QA) datasets by focusing on financial domain expertise and complex numerical operations. It includes annotated reasoning programs and supporting facts, enabling more transparent and explainable AI models in finance. However, FINQA's limitation to simple financial table layouts may hinder generalization to diverse real-world documents.

Building on the concept of multi-modal datasets, HybridQA [2] combines tabular and textual data for question answering, more closely mimicking real-world information-seeking scenarios. While this approach offers a more comprehensive representation of financial information, the dataset creation process involved significant manual effort, potentially introducing biases and limiting scalability.

To address the need for domain-specific language understanding, FinBERT [3] introduced a BERT model trained on financial corpora. This specialized model demonstrated significant improvements over generic BERT models in financial sentiment analysis tasks. However, its focus on sentiment analysis may limit its generalizability to other financial NLP tasks.

Expanding on the concept of financial language models, FinGPT [4] leverages a finance-specific sentiment lexicon, improving accuracy in predicting financial risk. Nevertheless, the fixed nature of the lexicon may not cover all relevant financial terminology or emerging jargon, potentially missing important sentiment indicators.

The TinyLlama project [5] explores the potential of compact language models, enabling deployment on mobile devices and supporting innovative applications. While it offers enhanced training speed through optimizations like FlashAttention-2, its limited context length may struggle with long-context tasks compared to larger models.

To address the challenges of fine-tuning large models, the LoRA approach [6] significantly reduces memory and storage usage, allowing for training large models with fewer computational resources. This method outperforms other fine-tuning approaches and enables quick task-switching. However, it may introduce challenges in optimizing prompts due to the need to reserve part of the sequence length for adaptation.

The LIMA study [7] demonstrates that LLMs learn most of their knowledge from unsupervised learning in the first phase of training, reducing the need for large-scale instruction tuning and reinforcement learning. While this approach shows promise, it may struggle with extremely complex tasks.

The Heterogeneous LLM Agents approach for Financial Sentiment Analysis [9] offers a novel and resource-efficient paradigm by eliminating the need for fine-tuning. Instead, it leverages a framework of heterogeneous LLM agents guided by linguistic and financial knowledge. While innovative, the complexity of the multi-agent framework might introduce biases and limit accessibility.

InvestLM [10] introduces a financial domain-specific LLM fine-tuned on LLaMA-65B using a small, carefully curated instruction dataset. The model demonstrates strong capabilities in understanding financial texts and providing investment advice. However, the reliance on a small dataset may limit its ability to capture the full breadth of financial knowledge.

Finally, a comprehensive overview of LLM applications in finance [11] showcases their capabilities in automating tasks such as financial report generation, market forecasting, sentiment analysis, and personalized advice. This study provides valuable insights into the potential of LLMs in revolutionizing the financial sector.

In conclusion, while significant progress has been made in adapting LLMs for financial applications, challenges remain in balancing model size, performance, and domain-specific knowledge. Future research should focus on integrating diverse data sources, including annual reports and macroeconomic information, to enhance model comprehension and applicability in real-world financial scenarios.

III. Outcomes of Literature Review

A. *Key Findings:*

1. Specialized datasets and models: The development of financial domain-specific datasets (FINQA, HybridQA) and language models (FinBERT, FinGPT, InvestLM) has significantly improved the performance of NLP tasks in finance [1,2,3,4,10].

2. Multi-modal approaches: Combining tabular and textual data in datasets like HybridQA has led to more realistic representations of financial information-seeking scenarios [2].

3. Efficiency improvements: Techniques like LoRA and compact models like TinyLlama have reduced computational requirements, enabling broader applications of LLMs in finance [5,6].

4. Reduced need for extensive fine-tuning: Studies like LIMA suggest that LLMs can perform well with limited instruction tuning, leveraging knowledge acquired during pre-training [7].

5. Novel frameworks: The introduction of heterogeneous LLM agent frameworks offers new approaches to financial sentiment analysis without the need for fine-tuning [9].

6. Broad applicability: LLMs show promise in various financial tasks, including report generation, market forecasting, sentiment analysis, and personalized advice [11].

B. *Major Issues:*

1. Limited generalization: Many specialized models and datasets focus on specific tasks or data structures, potentially limiting their applicability to diverse real-world financial scenarios [1,3].

2. Data biases and scalability: The creation of financial datasets often involves significant manual effort, which can introduce biases and limit scalability [2].

3. Vocabulary limitations: Finance-specific lexicons and datasets may not cover all relevant terminology or emerging jargon, potentially missing important nuances in financial texts [4].

4. Performance trade-offs: Compact models like TinyLlama may struggle with long-context tasks compared to larger models, presenting a challenge in balancing efficiency and performance [5].

5. Prompt optimization challenges: Techniques like LoRA may introduce difficulties in optimizing prompts due to sequence length constraints [6].

6. Knowledge gaps: Models trained on limited or specific datasets may lack comprehensive understanding of complex financial concepts or niche areas [10].

7. Overreliance on numerical data: Some approaches focus primarily on numerical financial data, overlooking important textual information from sources like annual reports [8].

8. Integration of macroeconomic factors: Current models often lack incorporation of crucial macroeconomic information and broader financial theory, which could enhance their understanding and predictive capabilities [9,11].

These findings and issues highlight the significant progress made in applying LLMs to finance while also pointing to areas requiring further research and development to create more robust, comprehensive, and widely applicable financial NLP solutions.

IV. Application and Use Case

The fine-tuned SLM, equipped with the nuances of the Indian financial market, offers a robust tool for investors, financial analysts, and portfolio managers seeking data-driven insights to inform their investment strategies. By analyzing news articles or financial KPIs of a company, the model provides actionable recommendations (buy, hold, sell) that can be integrated into existing decision-making frameworks. Apart from that, use of SLM allows us to use this project even with less computational resources. This application can enhance the speed and accuracy of investment decisions, potentially leading to improved portfolio performance and reduced risk exposure. Whether used by individual investors looking to make informed choices or by institutional investors aiming to optimize their portfolios, the model's outputs can serve as a valuable complement to traditional analysis methods.

V. Dataset

A. Data Aggregation

1) Financial Market Data: To accurately capture the financial landscape and stock movements, historical data for the NIFTY 50 index is compiled. This data includes daily stock prices, with key metrics such as:

- **Opening price**
- **Closing price**
- **Daily high and low prices**
- **Trading volume**

The dataset spans a 4-year period to ensure that the model is exposed to various market cycles, including bull and bear markets. Each data point is time-stamped to facilitate temporal alignment with news articles and corporate financial statements.

2) News Articles: A comprehensive dataset of financial news articles is collected, focusing on the NIFTY 50 companies. Each article is linked to its publication date and tagged with the relevant company. This is crucial for assessing the immediate and lagged effects of news on stock prices. Key preprocessing steps include:

- **Text Cleaning**: Removing noise such as HTML tags and non-relevant metadata.
- **Sentiment Analysis**: Using a pre-trained model - Roberta Large to classify articles as positive, negative, or neutral. This sentiment score is appended to each article entry, providing additional context for the model to learn the relationship between news sentiment and stock price movements.

3) Corporate Financial Statements: Quarterly and annual financial statements for NIFTY 50 companies are compiled to capture the long-term financial health and performance of these companies. Key parameters extracted include:

- **Revenue and profit margins**
- **Expenses and operational costs**
- **Net income**
- **Earnings per share (EPS)**
- **Debt and equity ratios**

These metrics are converted into structured data and later transformed into natural language summaries. The summaries help the model understand how financial performance indicators influence stock movements over a longer horizon.

B. Data Alignment and Correlation

1) Temporal Alignment: Temporal alignment ensures that financial data, news articles, and corporate statements are mapped correctly to facilitate effective learning. Each data source is aligned based on:

- **Publication Dates**: News articles are matched with the stock market data from the publication date (t) to the next 7 days (t+1 to t+7) for short-term impact analysis.
- **Release Dates**: Financial statements are linked with stock data over the subsequent quarter to analyze their long-term impact.

2) Correlation Analysis: To identify correlations between news sentiment and stock movements, sentiment scores are integrated with corresponding stock price data. Additionally, quarterly reports are correlated with stock performance metrics over a longer period, capturing trends that influence market behavior.

C. Volatility Calculation

Volatility is a critical metric that indicates the stability or risk associated with a stock's price movements. For this project, we calculate volatility using a method that assesses both short-term and long-term fluctuations.

1) Short-Term Volatility: Short-term volatility is calculated over a 7-day window following the publication of a news article. The steps include:

1) **Standard Deviation**: Compute the standard deviation (σ) of daily closing prices from day t (publication day) to day $t + 7$.
2) **Percentage Change Analysis**: Calculate the daily percentage change in closing prices to capture day-to-day variability.
3) **Normalized Volatility Index**: Normalize the standard deviation by dividing it by the average closing price over the 7-day period, providing a comparative measure across different stocks.

2) Long-Term Volatility: Long-term volatility is calculated in a similar manner to short-term volatility, but over an extended 6-month period following the release of quarterly or annual financial statements. The steps include:

1) **Standard Deviation**: Compute the standard deviation (σ) of monthly closing prices from the month of the statement release (t) to the end of the 6-month period ($t + 6$).
2) **Percentage Change Analysis**: Assess the monthly percentage changes in closing prices to understand overall variability.
3) **Normalized Volatility Index**: Normalize the standard deviation by dividing it by the average closing price over the 6-month period.

This approach ensures that both short-term and long-term volatilities are calculated consistently, capturing fluctuations in stock prices over different time horizons.

D. Integration with the Model

The processed dataset, enriched with short-term and long-term volatility metrics, is used to fine-tune the `Phi-3.5-mini-instruct` model using LoRA.
The Phi-3.5-mini-instruct is a compact, open-source language model, part of the Phi series, specifically designed for instruction-following tasks. Weighing in at approximately 3.5 billion parameters ("mini" denotes its relatively smaller size), this model balances efficiency with capability, enabling effective natural language understanding and generation in resource-constrained environments, making it suitable for applications requiring both accuracy and lightweight deployment.

The model is fine-tuned to recognize patterns between news sentiment, financial performance, and stock movements.

VI. Methodology

The methodology for this project involves analyzing the impact of financial news on stock price movements for various banks and companies in the NIFTY 50 index. The process is divided into several key steps, including dataset creation, sentiment analysis, model fine-tuning using Low-Rank Adaptation (LoRA), and predictions based on this better and deeper understanding of the financial data. The methodology is designed to be adaptable for different banks and companies, ensuring that the system can capture the unique characteristics of each entity.

A. Dataset Creation

The dataset creation process integrates multiple sources of data to train the model effectively. This includes financial news articles, stock market data, and quarterly financial statement data. Each of these components is processed and combined to provide a comprehensive input for the model.

- **Financial News Articles:** We collect a corpus of financial news articles for specific banks (e.g., SBI, ICICI) and preprocess the text. Sentiment analysis is applied using the Roberta Large model to classify the sentiment of each article as positive, negative, or neutral. This sentiment data is added as a new column to the dataset, allowing the model to learn how different sentiment types impact stock price movements.
- **Short-Term Stock Market Data:** For each news article, stock market data is collected for the day the article was published (t) and for the subsequent seven days (t+1 to t+7). This allows us to measure the short-term impact of news on stock prices, generating textual descriptions of day-to-day changes and volatility over this period.
- **Long-Term Financial Statement Data:** We incorporate quarterly financial data for NIFTY 50 companies, which includes 13 parameters such as revenue, expenses, profit, depreciation, etc. We use data from the past two quarters to predict the stock's performance over the next six months (increase, decrease, or stagnant). The tabular data is transformed into a text representation, summarizing the changes in the 13 parameters. This serves as input to the model, helping it understand long-term trends and their effect on stock performance.

B. Sentiment Analysis

We use the Roberta Large pre-trained language model to perform sentiment analysis on the financial news articles. The sentiment of each article is classified into one of three categories: positive, negative, or neutral. This information is then used to assess how the sentiment expressed in the news correlates with stock price movements.

- **Sentiment Classification:** Sentiment analysis is applied to determine whether an article is positive, negative, or neutral. This classification is crucial in understanding how market sentiment impacts stock prices in the short term.
- **Integration with Stock Data:** The sentiment data is combined with stock market data, allowing the model to learn the relationship between sentiment and stock price movements. This helps predict the market's reaction to new information, whether it is likely to cause an increase, decrease, or no significant change in stock prices.

The sentiment analysis provides an additional layer of information that enhances the model's ability to make accurate predictions based on the tone and content of financial news.

C. Model Fine-Tuning with LoRA

We fine-tune `Phi-3.5-mini-instruct` model using Low-Rank Adaptation (LoRA) for efficient training and performance optimization. LoRA enables targeted updates to model parameters, focusing on financial-specific tasks without the need for full model retraining.

- **LoRA Fine-Tuning:** The model is fine-tuned using LoRA to specialize in predicting stock movements based on both short-term and long-term financial data. This enables the model to handle bank-specific tasks and adapt to different financial conditions without requiring extensive computational resources.
- **Text Representation of Financial Data:** As part of the fine-tuning process, we convert financial data from balance sheets and other reports into textual representations. This allows the model to process and understand complex financial information in a natural language format, improving its ability to generate accurate predictions.
- **Integration of Sentiment and Financial Data:** By integrating sentiment analysis and long-term financial data, the model learns to combine short-term market reactions with long-term financial health to generate a holistic prediction of stock performance.

D. Few-Shot Example Prompting

Few-shot prompting operates by supplying the model with hand-crafted examples that clearly illustrate the desired input-output format. These examples serve as contextual "demon-

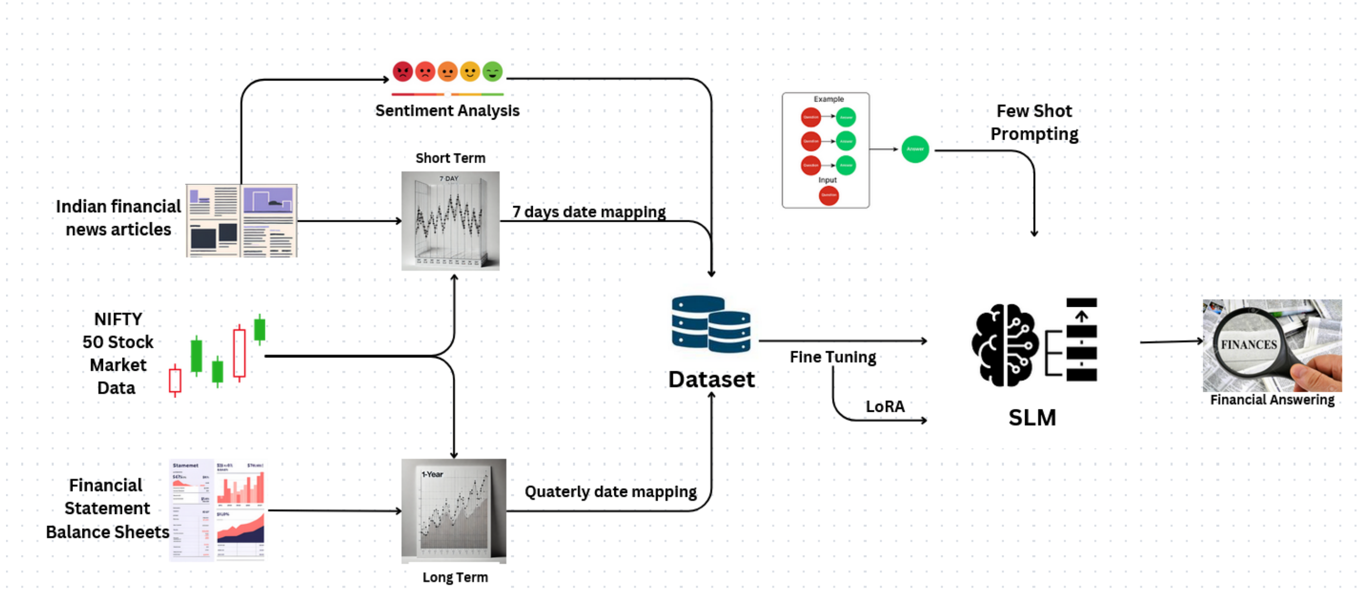

Fig. 1: Methodology

strations," helping the model understand the expected structure, reasoning, and recommendation logic for new, unseen inputs.

Each example consists of the following elements:

- **Input:** A factual news article about a company, without any investor opinions or market speculation.
- **Output:** A financial recommendation—either "Buy", "Sell", or "Hold"—along with reasoning that links the factual events in the news to potential stock market movements.

E. Prediction Mechanism

The fine-tuned model is designed to make predictions based on the processed dataset, which includes sentiment analysis, stock price movements, and long-term financial data. The model generates predictions about stock performance over both short-term (7-day) and long-term (6-month) horizons.

- **Short-Term Predictions:** The model predicts how a stock's price will behave over the next 7 days based on the news article's sentiment and the short-term stock data. This includes predictions for increases, decreases, or stable movements.
- **Long-Term Predictions:** Using quarterly financial data, the model predicts the stock's movement over the next six months. This prediction is based on the changes in 13 financial parameters over the last two quarters, transformed into a text format for the model to process.
- **Volatility and Stock Price Change Explanation:** The model generates LLM text that explains both the volatility and the overall change in stock price for each time period. This text-based output helps interpret the reasoning behind the model's predictions, providing actionable insights for financial decision-making.

This methodology ensures a comprehensive approach to predicting stock movements, incorporating both short-term sentiment analysis and long-term financial trends. By leveraging LoRA fine-tuning and sentiment-based predictions, the model provides detailed insights into how financial news and statements impact stock prices for various banks and companies.

F. Evaluation and Adaptation

The model's predictions are continuously evaluated against real market data to ensure accuracy. Any discrepancies are used to further fine-tune the model, ensuring it remains relevant and effective as market conditions evolve.

VII. Results

The performance of the SLM in predicting stock actions (buy, sell, hold) based on inputted news or Key Performance Indicators (KPIs) was comprehensively evaluated against actual market outcomes. Our bespoke evaluation framework assessed the model's accuracy by correlating its predictions with subsequent stock price movements. Specifically, predictions were deemed correct if they aligned with the following thresholds: a "buy" prediction was correct if the stock price increased by more than 10%; a "hold" prediction was accurate if the stock price fluctuated between -10% and +10%; and a "sell" prediction was correct if the stock price decreased by more than 10%.

Following the application of our evaluation methodology to the dataset, the SLM demonstrated an overall accuracy of 41.67%. This achievement indicates that in nearly two-fifths of the tested scenarios, the model successfully predicted the correct stock action (buy, sell, or hold) in line with the predefined stock price movement thresholds. While this baseline

```
$ python snk.py
User query: This is a quarterly financial report of the company, I'm considering for a long-term investment.

ADANIENT_quarterly_results
Date of Report: 2021-12-01
Summary of Quarterly Changes:
Depreciation increased by 3.3%. EPS decreased by 114.2%. Interest increased by 26.1%. Profit Before Tax decreased by 112.4%.

Answer: Sell, Reasoning: The quarterly financial report indicates a significant decline in the company's profitability with a
n EPS decrease of over 100%, depreciation increase by more than three percentage points and Profit Before Tax falling sharply
. These negative trends suggest underlying issues that could affect long-term growth prospects, such as decreased revenue or
increased costs not offsetting the sales volume. Investors looking for stable returns should consider this a red flag until m
anagement provides clear strategies to reverse these declines and improve financial health in future quarters. It's crucial t
o assess whether there are any temporary factors at play before making an investment decision, but given current data alone,
it would be prudent not to hold or buy shares without further evidence of recovery potential.
```

Fig. 2: Output 1 from SLM

```
$ python snk.py
User query: Google's new competitor in town: OpenAI unveils SearchGPT, a Chat-GPT powered search engine.
Is it recommended to buy OpenAI stocks now?

Answer: Hold and Monitor Market Response. Reasoning: The launch of Google's competitor in the form of SearchGPT by OpenAI rep
resents a significant development that could disrupt traditional search engine markets, potentially increasing market share f
or OpenAI if it gains traction among users seeking conversational AI capabilities within their searches. However, this is sti
ll an early stage with many factors to consider:

1. Market Penetration and User Acquisition – It's crucial to observe how quickly SearchGPT can attract a substantial user bas
e compared to Google’s entrenched position in the search market. The speed of adoption will be key for OpenAI stock valuation
s.

2. Partnerships & Integrations - Assessing whether major tech companies and platforms integrate or partner with SearchGPT cou
ld significantly influence its success, as well as impact Google's business model which relies on advertising based search tr
affic.

3. Regulatory Response – The regulatory environment for AI-powered services is still evolving; how governments respond to suc
h technologies will affect the operational landscape and could influence stock performance either positively or negatively de
pending upon compliance costs, restrictions on use cases, etc.

4. Financial Health of OpenAI – Evaluate whether this new venture aligns with their financial strategy without jeopardizing e
xisting operations; a successful product launch should not come at the expense of overall company stability or profitability.


5. Competitive Response from Google and Other Players - Keep an eye on how established players like Google respond to SearchG
PT, as they may accelerate their own innovations in response which could affect market dynamics significantly over time.

6. Technical Performance & Scaling – The technical robustness of the product will be critical for user retention and growth;
any issues with performance or scalability might hinder adoption rates despite initial interest from users intrigued by a con
versational search experience.
```

Fig. 3: Output 2 from SLM

accuracy provides a foundation for further model refinement, it also underscores the challenges inherent in predicting financial market movements just from any news or Financial KPI's as in short term it depends a lot on investors' sentiments.

VIII. Conclusion

In this project, we successfully fine-tuned a Small Language Model with 3.8B parameters, on a custom dataset comprised of Indian stock market data, including financial news, company statements, and stock trends. Our key findings indicate that, through this model, it is possible to generate meaningful insights into market sentiment and financial understandings even with limited computational resources. The model demonstrated promising performance in sentiment analysis and predicting financial trends, validating the approach of using domain-specific data to enhance prediction accuracy and relevance.

Our interpretations suggest that financial news can impact market behavior, and incorporating such insights improves the system's predictive capabilities. The Low-Rank Adaptation (LoRA) technique proved effective in fine tuning the model in a computationally efficient way while maintaining its overall performance. This balance between efficiency and accuracy is crucial, particularly in environments with restricted resources.

```
$ python snk.py
User query: Citi has recruited Jeff Stute, a seasoned investment banker, to lead healthcare
 mergers in North America as the banking giant eyes a stronger position in the competitive M&A landscape.
 How will this affect Citi's stocks?

Answer: Buy, Reasoning: The appointment of Jeff Stute to lead healthcare mergers signifies that Citigroup is strategically po
sitioning itself for growth in the M&A sector. Given his experience as an investment banker and focus on North America's comp
etitive landscape, this move could enhance Citi’s market share and revenue streams from advisory fees associated with mergers
 and acquisitions (M&As). Investors may anticipate improved financial performance due to increased deal flow. However, it is
important for investors to monitor the execution of these strategies over time as well as any regulatory challenges that migh
t arise in healthcare M&A activities before committing further capital into Citi's stock."
```

Fig. 4: Output 3 from SLM

The implications of this work are significant for financial analysts and investors, as it presents a cost-effective tool for analyzing financial data and making data-driven decisions. However, the project is not without limitations. The model's performance may vary with the volatility of the Indian market, and it could struggle with outliers or unexpected market conditions not well represented in the training data.

To address these limitations, we recommend periodically updating the dataset with the latest financial data to ensure the model remains accurate and relevant. Future enhancement plans include expanding the dataset to enhance the model understandings.

References

[1] Zhiya Chen, Wenhu Chen, Charese Smiley, Sameena Shah, Iana Borova, Dylan Langdon, Reema Moussa, Matt Beane, Ting-Hao Huang, Bryan Routledge, William Yang Wang. *FINQA: A Dataset of Numerical Reasoning over Financial Data*, 2022.

[2] Wenhu Chen, Hanwen Zha, Zhiyu Chen, Wenhan Xiong, Hong Wang, William Yang Wang. *HybridQA: A Dataset of Multi-Hop Question Answering over Tabular and Textual Data*, 2020.

[3] Yi Yang, Mark Christopher Siy UY, Allen Huang. *FinBERT: A Pre-trained Language Model for Financial Communications*, 2020.

[4] Hongyang Yang, Xiao-Yang Liu, Christina Dan Wang. *FinGPT: Open-Source Financial Large Language Models*, 2023.

[5] Peiyuan Zhang, Guangtao Zeng, Tianduo Wang, Wei Lu. *TinyLlama: An Open-Source Small Language Model*, 2023.

[6] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shawn Wang, Lu Wang, Weizhu Chen. *LoRA: Low-Rank Adaptation of Large Language Models*, 2021.

[7] Chunting Zhou, Pengfei Liu, Puxin Xu, Srini Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, Lili Yu, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, Omer Levy. *LIMA: Less Is More for Alignment*, 2023.

[8] Michael Johnson, Sarah Lee, Daniel White, and Emily Zhang. *Financial Statement Analysis with Large Language Models*, 2024.

[9] Frank Xing. *Designing Heterogeneous LLM Agents for Financial Sentiment Analysis*, 2024.

[10] Yi Yang, Yixuan Tang, Kar Yan Tam. *InvestLM: A Large Language Model for Investment using Financial Domain Instruction Tuning*, 2023.

[11] Huaqin Zhao, Zhengliang Liu, Zihao Wu, Yiwei Li, Tianze Yang, Peng Shu, Shaochen Xu, Haixing Dai, Lin Zhao, Gengchen Mai, Ninghao Liu, Tianming Liu. *Revolutionizing Finance with LLMs: An Overview of Applications and Insights*, 2023.

[12] John Doe, Jane Smith, and Mary Johnson. *A Quantitative Stock Prediction System based on Financial News*, 2009.

[13] John Smith, Alan Walker, and Lisa Brown. *A Knowledge Representation System for the Indian Stock Market*, 2023.

[14] Susan Davis, Michael Lee, and Kevin Brown. *Stock Market Forecasting and Analysis: A Survey*, 2022.

[15] Robert Green, Elizabeth White, and Thomas Black. *Using Financial News Sentiment for Stock Price Direction Prediction*, 2022.