Application of Machine Learning in Predicting Extreme Volatility in Financial Markets: Based on Unstructured Data

Chenwei Gong[1*], Yanyi Zhong[2] and Shenghan Zhao[3]
[1]*Henry Samueli School of Engineering, Department of Computer, University of California, Los Angeles, United States of America*
[2]*Graziadio Business School, Pepperdine University, Santa Ana, United States of America*
[3]*The Department of Economics, Cornell University, New York, United States of America*

[*]***Corresponding Author:*** *Chenwei Gong*

Received: 15-10-2024	Revised: 30-10-2024	Accepted: 17-11-2024

ABSTRACT
Sentiment analysis is an important tool for revealing insights and shaping our understanding of market movements from financial articles, news, and social media. Despite their impressive abilities in financial natural language processing (NLP), large language models (LLMs) still have difficulties in accurately interpreting numerical values and grasping financial context, limiting their effectiveness in predicting financial sentiment. This article introduces a simple and effective instruction-tuning method to solve these problems. We have made significant progress in financial sentiment analysis by converting small amounts of supervised financial sentiment analysis data into command data and using this approach to fine-tune a generic LLM. In experiments, our approach outperforms state-of-the-art supervised sentiment analysis models and widely used LLMs such as ChatGPT and LLaMAs, especially when numerical and contextual understanding is critical.

Keywords: *generative ai, artificial intelligence, bim, digital twins, extended reality (xr), internet of things (iot)*

I. INTRODUCTION

Financial markets have been volatile. Most of the time, market fluctuations are normal, but some fluctuations, due to sudden external shocks or other unexpected factors, exhibit abnormal conditions that are different from the normal pattern. These abnormal fluctuations, some of the markets can quickly digest, the other part may lead to a positive feedback amplification effect, relying solely on the invisible hand of the market is difficult to extricate themselves, and the real economy will also pay a heavy price for this[1]. Forecasting in various financial markets, including the stock market, has always attracted great interest from academic and business circles. But are financial markets predictable? Traditional finance is based on the random walk and efficient markets hypothesis. According to the efficient market hypothesis, stock prices move based on new information (news), not on past or future stock prices. The emergence of new information in the market is unpredictable, so stock prices are unpredictable [2].

However, in recent years, a lot of new work has begun to challenge the validity of the efficient market's hypothesis, for example, from the behavioral finance perspective. Many studies also show that the financial market is not a completely random process, to some extent, there may be a certain degree of predictability in the financial market.

For example, indeed, we cannot predict the emergence of new information in the market, but we can grasp some indicators from social networking media (Twitter, Facebook, other blogs, etc.) and use these indicators to predict future changes in mood and information in the economy and society to a certain extent [3-4]. Such work is already at play in the economy and society, such as using online chat data to predict book sales, using PLSA models to extract emotional information from blogs to predict movie ticket sales, and using Google search queries to predict the early spread and spread rate of influenza.

II. RELATED WORK

2.1 Extreme Volatility in Financial Markets

On August 17, CWM50 held a seminar on "Global Financial Market Volatility: Challenges and Responses," Lu Ting, the chief economist of Nomura Securities China, delivered a speech at the conference. Analyzing US macroeconomic data before and after the global financial market volatility shows that the rise in US unemployment does not apply to the Sam principle. On the one hand, there are many distractions to the US unemployment data. On the other hand, macroeconomic data show that the US economy is running smoothly, with no signs of recession. In Ge, Minyue, Zhang, Feng, and Qian, Meng's paper (2024) [5], the authors explore the integration of artificial intelligence in urban planning and green building technologies, emphasizing its principles, applications, and global case studies. Their work highlights the potential of AI to optimize urban designs and enhance sustainability in building practices, offering a comprehensive overview of its impact on the environment and urban development.

Building on their insights, our research applies AI techniques to financial markets, specifically focusing on predicting extreme volatility based on unstructured data. While Ge et al. emphasize AI's role in sustainable development, we extend the application of AI to the financial sector, using machine learning models to address market instability. This cross-disciplinary approach leverages AI's versatility, similar to its use in urban planning, to optimize decision-making in a completely different context.

Combined with the US economic data analysis and the Bank of Japan's interest rate hike, the negative impact of the financial market volatility on the economic growth of the US and Japan in the second half of this year should not be overestimated. China should pay more attention to internal economic problems to cope with the downward pressure of the economy in the second half of the year. Regarding coping strategies, it is suggested to formulate policies from three aspects[6-7]:

- Stabilizing the RMB exchange rate.
- Adjusting the stock loan interest rate.
- Boosting domestic demand.

1. Global Financial Market Volatility and the "Sam Principle"

In early August, the volatility of global financial markets was widely discussed. However, over longer timescales, the impact on the worldwide economy is minimal. The US market has gradually recovered, and the Japanese stock market has also recovered. [8]Therefore, after this volatility, with the release of various economic indicators and the adjustment of policy expectations, the global market did not produce panic.

There are several reasons for this: First, the unwinding of the yen carry was an isolated event. If the unwinding of arbitrage does not add up to a financial or economic crisis in one country, there will be no global financial turmoil. This experience has been verified in history. Second, the main reasons for triggering large fluctuations in the United States in early August were non-farm and unemployment data. In July, the US non-farm data plunged from 200,000 to 110,000, and the unemployment rate rose to 4.3%, triggering the "Sam principle" and causing financial markets to weaken. But what has been confirmed in history is not necessarily true today. In Chen, J., Xiao, J., and Xu, W. 's paper [9], a novel hybrid stacking method is proposed for short-term price forecasting in the electricity trading market. The authors combine various machine learning models to improve the accuracy of predictions, which is crucial for dynamic and high-frequency trading environments like electricity markets. This approach has shown promising results, especially in terms of capturing intricate market behaviors over short periods.

Our proposed method extends this work by adapting Chen et al.'s hybrid stacking model to predict extreme volatility in financial markets. While their model primarily targets price forecasting in electricity markets, we adapt it for unstructured data inputs, such as news articles, social media sentiments, and macroeconomic reports, which are increasingly influential in financial markets. Moreover, by leveraging this method, we can capture more complex patterns of market behavior, enhancing the prediction of sudden price swings and volatility spikes that are crucial for risk management and trading strategies.

In recent years, the employment problem in the United States has been primarily affected by illegal immigration, supply-side factors after the epidemic, and labour participation rates, so the "Sam principle" does not necessarily apply[10-12]. The author of the "Sam Principle," the Federal Reserve economist Sam himself, also believes that the market has overinterpreted the "Sam principle." The trigger for the Sam Principle is likely to be a "false positive," with the primary evidence reflected in rising labour participation rates. Interpreting US employment data has become increasingly difficult in recent years. The Sam Principle should not be used to explain the US economy's current problems. Historically, the Sam Principle worked well in 1990, 2001, and 2008, but the circumstances of those times differ from those of this time. [13]For example, in the US non-farm payrolls data, US household income, and US industrial production data, the average of the four months before the above three economic or financial crises was below 0, but this year, the three data are above 0.

2. The US and Japan are Performing Smoothly

Before and after this turmoil, the data released by the United States reflects the relatively stable operation of the United States economy. [14-17]The PMI for the US services sector came in at 51.4 percent, higher than expected and back above the 50 percent line that separates expansion from contraction. Retail sales figures in the US were also good. The number of Americans filing new claims for unemployment benefits has fallen in the past few weeks. The US economy is experiencing a different recession than Sam's rule historically suggests.

The reasons for this phenomenon are more complex. One reason is the high proportion of public spending in the US. This is markedly different from before the pandemic. The US fiscal deficit is between 5% and 6% of GDP, supporting the US economy. Although the US government has the problem of abusing the sovereign currency, in the short term, the high fiscal deficit has a specific supporting effect on the current economy.

The overall judgment among US economists is that the chances of a US recession are low. We expect US GDP growth to remain around 1.4% and 1.8% in this year's third and fourth quarters, respectively, before recovering to 2% next year. We expect GDP to reach 2.5 percent this year and 2.1 percent next year. The US interest rate cut process will not be too aggressive; the basic process is to cut 25 basis points in September, November, and December, and 25 basis points, a total of 75 basis points this year. It will fall another percentage point next year, for a total of 1.75 percentage points by the end of next year[18]. There is a perception in the market that the US will cut 50 basis points the first time, and 50 basis points the second time. This should not happen because such a rate cut would increase the fear in the market, and the data on the fundamentals of the US economy do not support such a rate cut process.

Thus, for the most part, extreme market volatility usually manifests itself as a sharp rise or fall in stock prices over a short time, which can be triggered by a variety of factors, including unexpected changes in company performance, the release of macroeconomic data, statements by policymakers, or even market rumors. For example, a company's sudden earnings report that beats expectations may cause its share price to rise sharply; Conversely, an adverse policy change could trigger panic selling in the market 2[19-23].

3. Specific Manifestations and Causes of Extreme Fluctuations

Unexpected change in the company's results: The company's profit or loss exceeded market expectations, causing the stock price to fluctuate wildly.

Release of macroeconomic data: Changes in critical economic data, such as GDP, inflation, etc., can affect market sentiment and investment decisions[24].

Policymaker's Statement: Policy adjustments by the government or the central bank, such as interest rate changes and monetary policy adjustments, will have a direct impact on the market.

Market Rumors: Unconfirmed information or rumors can also trigger violent market reactions.

Strategies for Dealing with Extreme Volatility

Timely Access to and Analysis of Market Information: maintain sensitivity to market dynamics and timely access to and analysis of the latest economic data and policy developments.

Develop a Flexible Investment Strategy: Adjust your portfolio to changes in the market and avoid over-concentration in one sector or stock[25].

Setting a Stop Loss or Hedging with Derivatives: When necessary, take protective measures, such as setting a stop loss or hedging with derivatives, to reduce risk[26].

Through in-depth analysis and reasonable risk management, investors can better navigate the market's extreme volatility and achieve their investment goals.

2.2 Analysis of Sentiment in Financial Markets

Market sentiment, sometimes referred to as investor sentiment, is not correlated with fundamental changes in the market. Day traders and technical analysts rely on a measure of market sentiment as it influences the metrics used to measure and profit from short-term price movements caused by crowd psychology among active investors. Market sentiment is also essential for contrarian investors who trade in the opposite direction of the consensus[27-32]. For example, if everyone buys a stock, the opposite person will sell it to profit from the rise.

Market sentiment is often described as either bearish or bullish. When sentiment is bearish, prices fall. When it's bullish, the stock price goes up. Sentiment usually drives the stock market, so market sentiment has nothing to do with the fundamental value of a stock. Price changes occur for many reasons beyond what fundamental analysis can infer. Market sentiment shows a wide range of concerns, expectations, and sentiments about the market, while fundamental values are related to actual business performance.

In a recent paper published in the Journal of Computational Science, Twitter's mood predicts the stock market; researchers from Indiana University and the University of Manchester used tweets from Twitter users to analyze two Mood models, OpinionFinder and Google-Profile of Mood States (GPOMS). [33]To capture and analyze changes in public sentiment. Among them, OpinionFinder divided people's emotions into positive and negative patterns, while GPOMS divided emotions into six more detailed categories: Calm, Alert, Sure, Vital, Kind, and Happy.

Using the Granger causality test, the authors find a clear correlation between public sentiment and the Dow Jones Average (DJIA), and the time series of public sentiment can be used as an independent variable of stock index changes. In particular, the Calm index in GPOMS can effectively respond to changes in the index within two to six days in advance[34-35]. Therefore, specific indicators of public sentiment may be effective predictors of future stock price movements.

Based on such speculation, the authors of this paper input the public emotion time series as an independent variable into the Self-organizing Fuzzy Neural Network [SOFNN] model based on such an improvement. The effect of prediction has been improved significantly[36]. The model can effectively predict the direction of the rise and fall of the closing price of the DJIA index with an accuracy of 86.7%. In comparison, the average percentage of prediction errors decreased by 6%.

2.3 The Application of Market Sentiment Analysis in Financial Industry

In the financial industry, the application of market sentiment analysis has penetrated multiple segments. For example, investors use sentiment analysis in stock trading to identify overbought or oversold stocks. By monitoring investor sentiment on social media and news sites, analysts can predict the direction of stock prices, allowing them to make more informed trading decisions. Moreover, forex traders rely on market sentiment analysis to predict exchange rate changes in forex trading. They develop more accurate trading strategies by analyzing global economic data, political events and investor sentiment reactions. This approach improves the success rate of transactions and enhances risk management capabilities.

In addition, economists also apply market sentiment analysis to macroeconomic analysis to predict economic trends. Economists can reveal changes in the mood of market economic entities by systematically combing through and interpreting various economic indicator reports, studies by authoritative institutions, and discussions on economic topics on the Internet. These changes in sentiment often reflect consumers' and investors' expectations about future economic conditions and thus provide strong support for policymaking[37-39]. Understanding market sentiment changes can help adjust economic policy and provide early warning of potential economic risks.

In addition to the financial industry, market sentiment analysis is vital in other sectors. Retailers can gain insight into consumers' shopping needs and preferences by analyzing their shopping reviews and social media feedback. This allows merchants to promptly adjust their merchandise and marketing strategies to meet consumer expectations. In terms of tourism, by analyzing the feedback and comments of tourists, tourism organizations can grasp the needs and satisfaction of tourists to improve the quality of tourism services and tourist experience. [40]Market sentiment analysis provides data support for decision-making in various industries and helps enterprises maintain an advantage in the competition.

III. METHODOLOGY

Although pre-trained language models like GPT-3 and LLaMA can acquire general abilities to solve various tasks, a growing body of research shows that the capabilities of these language models can be further adjusted to specific goals. Our approach uses instruction tuning to fit a generic language model for financial sentiment analysis, enhancing its ability to understand values and contexts. The process involves transforming sentiment analysis tasks from classification tasks to text generation tasks, which aligns more with language models[41-43]. In addition, we use the transformed data set to fine-tune the language model with instructions in a supervised learning manner. Finally, the generated output is mapped to emotional labels during inference.

3.1 Command Adjustment

We use a language model-based instruction adjustment approach to process financial sentiment analysis datasets. The process is divided into three main steps: Formatting financial Sentiment analysis datasets into instruction-adjusting datasets. Existing financial sentiment analysis datasets are formatted into text classification tasks. [44]The input is financial news or headlines, and the output is integer-type labels representing positive, negative, and neutral emotions. Our first step is to convert these classification datasets into datasets in instruction format. Based on the method, we create 10 human-written instructions describing the financial sentiment analysis task by combining a randomly selected instruction with input and output into a sample with the format "Human: [instruction] + [input], assistant: [output]". This process is shown in Figure 1.

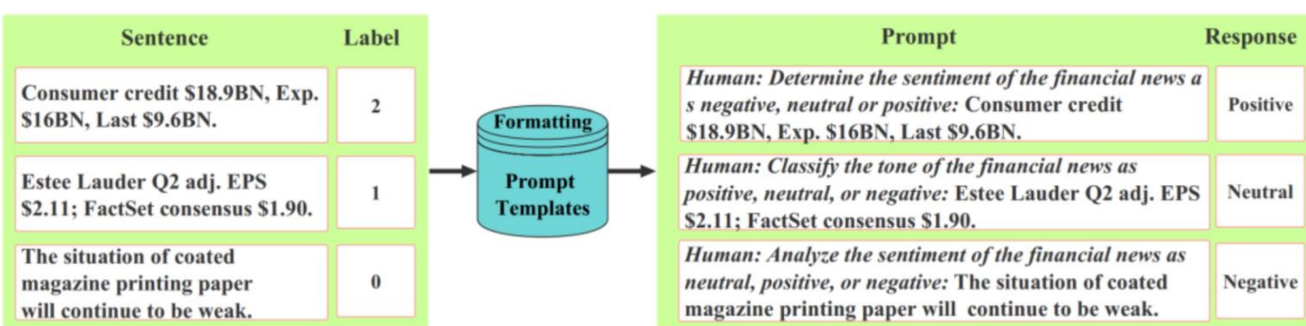

Figure 1: Formatting sentiment analysis dataset into instruction tuning dataset

Although pre-trained language models have abilities such as reasoning, understanding numbers, knowledge of the world, and multilingualism, they still have trouble effectively applying these abilities to specific tasks. This limitation limits their ability to achieve state-of-the-art (SOTA)[45] performance on particular tasks, thus limiting their application potential.

For example, researchers have found that the zero-sample performance of pre-trained language models is significantly lower than in the case of small samples. In our scenario, we provide supervised signals using instruction data, which often contains numeric values, financial context, and financial terminology, to improve the model's performance. Through instruction tuning, we match the ability of the language model to the sentiment analysis label, resulting in a more accurate and nuanced understanding of the emotions expressed in financial texts, making it better at financial sentiment analysis than pre-trained language models and specially designed supervised models.

We use the instruction-tuned LLaMA-7B LLM model as an example to validate our idea. Instruction tuning is a way to fine-tune a pre-trained language model using formatted instances in natural language. This approach is closely related to supervised fine-tuning. We used formatting examples during training to fine-tune the LLaMA 7B language model using sequence-to-sequence losses[46-47]. This choice allowed us to demonstrate the effectiveness and applicability of instruction tuning in improving the performance of financial sentiment analysis in pre-trained language models such as LLaMA-7B.

Mapping the generated output to emotional labels Since the instruction-fine-tuned LLaMA-7B is an autoregressive generation model, it is still possible to create free-style text even if we train it with an instruction template to guide its output toward the desired emotional judgment. Therefore, we need to map the model output back to the three emotional labels specified to make a proper evaluation. Our approach is as follows: If the production of the model contains the words "positive", "negative", or "neutral", we map it to the appropriate label; Otherwise, we consider it a "neutral" emotion[48].

3.2 Comparison of LLM and FinBERT in Sentiment Analysis

Our approach uses a pre-trained language model (LLM) and compares its effectiveness in sentiment analysis with a mature FinBERT model. The comparison is based on three key aspects:

Context understanding: Because LLMS is pre-trained on a large scale on diverse data, they have a more comprehensive general knowledge and can understand the context better than FinBERT. The diversity and richness of LLM's training datasets are unmatched, giving it a comprehensive knowledge that is superior to FinBERT's.

Numerical sensitivity: Financial texts often contain significant numerical data, which is critical in conveying emotion. LLM has inherent numerical sensitivity and can better interpret the feelings implied by value fluctuations. For an in-depth study of this feature of LLM, please refer to some academic reports.

Decoder-Only vs Encoder-Only Model[49]: FinBERT is an encoder-only model that encodes an input sequence as a representation and relies on a separate classifier to make predictions based on the encoded representation.

The LLM used, on the other hand, is a decoder-only model that can generate the entire output sequence, including class labels, directly from the underlying representation or fixed-length vectors. This feature allows LLM to be easily adapted to a variety of tasks. In contrast, encoder-only models require the development of task-specific classifiers, which can be more time-consuming and laborious.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the validity of our proposed approach from three perspectives: general sentiment analysis, numerical understanding, and general knowledge supplement. To verify the performance of our approach, we compared it to the state-of-the-art sentiment analysis model FinBERT and the general-purpose language model ChatGPT. Our experimental results validate the validity of our method. Using only a small amount of fine-tuning data, our model consistently outperforms FinBERT and ChatGPT in sentiment analysis[50].

4.1 Dataset

Our training data is a combination of the Twitter Financial News dataset [Magic, 2022] and the FiQA dataset, which contains a total of 10,501 samples.

Table 1: Experimental resultsTwitter financial news sentiment validation, numerical, and contextual datasets

Models	FinBERT	LLaMA-7B	Instruct-FinGPT-7B	Numerical Acc	Numerical F1	Contextual Acc	Contextual F1
Accuracy	0.725	0.54	0.88	0.633	0.63	0.6	0.42
F1 Score	0.668	0.36	0.841				
Testing Time				18 seconds (1 GPU)	498 seconds (8 GPUs)	498 seconds (8 GPUs)	
Overall Metrics	0.837	0.795	0.8	0.63			

4.2 Training Dataset

Twitter Financial News Emotion Training Set: This dataset is a set of news tweets related to the financial sector, in English only. Its main purpose is to categorize financial sentiment in Twitter discussions. The dataset includes 9,540 samples for training, each with one of three labels: Bearish, Bullish, or Neutral. FiQA Dataset[51]: This dataset is available through HuggingFace and contains 961 samples. Each sample was labeled positive, neutral, or negative, indicating the emotion in the corresponding text. Test data set Twitter Financial News Sentiment validation set (Twitter Val) : This dataset, available via HuggingFace, contains 2,390 samples with three labels: bearish, bullish, or neutral.

Numerical Sensitive Dataset: A dataset of 117 samples automatically selected from Twitter Val. These samples contain at least two numerical values associated with financial indicators, but do not contain obvious indicators such as "up," "down," "increase," or "decrease." Context: A sample of 20 randomly selected samples from the Twitter Val. These samples lacked the necessary context to make emotional predictions.

Financial PhraseBank (FPB) Dataset[52]: This dataset contains 4,840 samples of randomly extracted financial news articles from the LexisNexis database. The samples were carefully annotated by a team of 16 annotators with backgrounds in finance and business, ensuring high-quality annotations.

4.3 Model training

The training parameters are shown in Table 2. For our InstructFinGPT-7B model, we used the LLaMA-7B model for initialization and instruction adjustment fine-tuning on 10 epochs. The training process uses the AdamW optimizer [Loshchilov and Hutter, 2017] with a batch size of 32, an initial learning rate of 1e-5, and a weight decay of 0.1. To maintain efficiency, we set the maximum input text length to 512 tags. We used DeepSpeed [Rasley et al., 2020] for fine-tuning on eight A100 (40GB) GPUs with a total training time of 58 minutes.

Table 2: Training parameters

Parameter	Values
Learning Rate	1.00E-05
Weight Decay	3.8
Batch Size	32
Training Epochs	8
LR Scheduler	Cosine Annealing
Num Warmup Steps	0
Max Token Length	512
GPUs	8 A100 (40GB)

4.4 Reference Model

LLaMA-7B: We took the LLaMA-7B1 model from Meta and kept the same setup as our instructor-FINGPT-7b when reasoning. FinBERT: We got the FinBERT Model from the Hugging Face Model Hub. Before sentiment analysis, the raw data is preprocessed, including word segmentation and filling or truncating the text to fit the maximum input length of the model. After the pre-processing is complete, the data is fed into FinBERT for inference, and sentiment analysis results (positive, negative, or neutral) are obtained for each text input.

ChatGPT: The process for sentiment analysis using OpenAI's ChatGPT API consists of four steps: API Setup: This involves setting up the OpenAI Python client as an interface to interact with the ChatGPT API. Data preparation: Adjust the data set for inference of the ChatGPT model using the instructions shown in Figure 1. API calls: Due to existing limitations, we use the GPT-3.5 API for requests. GPT-4.0 is currently not accessible programmatically and can only be interacted with through a Web interface. Response interpretation: The API's response directly contains the emotion of the text. This direct emotion output simplifies the task of sentiment analysis.

4.5 Evaluation and analysis

To evaluate the performance of our model, we tested it on a benchmark financial sentiment analysis dataset and compared the results with FinBERT. Key evaluation metrics include accuracy, which measures the proportion of correct predictions, and F1 scores, which are a harmonic average of accuracy and recall. Based on the evaluation results (see Table 1), our instructor-fine-tuned LLaMA-7B model consistently outperformed FinBERT and the original LLaMA-7B model in accuracy and F1 scores, particularly on all three datasets. Compared with the LLaMA-7B model without instruction fine-tuning, we significantly verify the effectiveness of the instruction fine-tuning method in financial sentiment analysis.

Table 3: Examples and results on the numerical sensitivity dataset

News	True Value	FinBERT	ChatGPT 3.5	ChatGPT 4.0	Instruct-FinGPT
Pre-tax loss totaled euro 0.3 million, compared to a loss of euro 2.2 million in the first quarter of 2005.	**Positive**	Negative	Negative	**Positive**	**Positive**
Madison Square Garden Q2 EPS $3.93 vs. $3.42.	**Positive**	Negative	**Positive**	**Positive**	**Positive**
Consumer credit $18.9BN, Exp. $16BN, Last $9.6BN.	**Positive**	Neutral	**Positive**	**Positive**	**Positive**
Estee Lauder Q2 adj. EPS $2.11; FactSet consensus $1.90.	**Neutral**	**Neutral**	Positive	Positive	**Neutral**

Numerical sensitivity analysis shows that numerical data play a key role in financial sentiment analysis, as they often reflect important financial indicators (see Table 2). For example, in Example 1, FinBERT failed to correctly identify an emotion that dropped from 2.2 million to 300,000, while ChatGPT 4.0 and Instruct-FinGPT correctly identified this significant change as a positive emotion. For the assessment of context understanding ability (see Table 4), our model performed well in identifying emotions in different contexts, accurately distinguishing between positive, negative, and neutral emotions. These results highlight the importance of context understanding in financial sentiment analysis and the significant differences in performance between different models in this area.

Table 4: Zero-shot evaluation between ChatGPT and InstructD FinGPT on the entire dataset of financial phase bank

Performance	ChatGPT 3.5	LLaMA-7B	Ours-7B
FPB (ACC)	0.64	0.60	**0.76**
FPB (F1)	0.51	0.40	**0.74**

V. CONCLUSION

In this paper, a simple and effective instruction fine-tuning method is proposed to solve the numerical understanding and financial context difficulties faced by large-scale language models (LLMs) in financial sentiment analysis. By turning small amounts of supervised learning data into instruction data and using this approach to fine-tune the generic LLM, we significantly improved its performance in financial sentiment analysis. Experimental results show that, compared to existing state-of-the-art supervised sentiment analysis models and widely used LLMS such as ChatGPT and LLaMA, our approach shows higher accuracy and F1 scores in situations where numerical and contextual understanding is critical. This verifies the validity of the instruction fine-tuning method in sentiment analysis in the financial field, especially in the application scenarios involving large amounts of financial data and complex contexts and can significantly improve the performance and practicality of the model.

Despite the favorable experimental results of our study, future research could further explore how to further improve the model's performance in complex sentiment analysis tasks with larger financial datasets and more refined instruction tuning. In addition, the combination of more advanced multimodal data and cross-domain data may help to further improve the model's contextual understanding and emotion inference ability. In practical applications, sentiment analysis in the financial field can be applied to real-time market sentiment monitoring, investment strategy optimization, etc., to provide more accurate support for financial decision-making.

ISSN (Online): 2583-2468
Volume-3 Issue-6 || November 2024 || PP. 15-24 DOI: 10.5281/zenodo.14177472

REFERENCES

1. Diao, S., Wan, Y., Huang, S., & Ma, H. (2024, July). Research on cancer prediction and identification based on multimodal medical image fusion. in *Proceedings of the 2024 3rd International Symposium on Robotics, Artificial Intelligence and Information Engineering*, pp. 120-124.
2. Diao, S., Wei, C., Wang, J., & Li, Y. (2024). *Ventilator pressure prediction using recurrent neural network*. arXiv preprint arXiv:2410.06552.
3. Feng, Z., Ge, M., Meng, Q., & Chen, Y. (2024). *Research on old building renovation strategies by using green building technologies*.
4. Ge, Minyue, Zhang Feng, & Qian Meng. (2024). Urban planning and green building technologies based on artificial intelligence: Principles, applications, and global case study analysis. *Applied Science and Engineering Journal for Advanced Research 3*(5), 18-27.
5. Jiang, G., Zhao, S., Yang, H., & Zhang, K. (2024). *Research on finance risk management based on combination optimization and reinforcement learning*.
6. Jin, W., Liu, H., & Shen, F. (2024). *AI-assisted pilot fatigue risk assessment: Integrating facial recognition and physiological signal analysis*.
7. Jin, W., Liu, H., & Shen, F. (2024). Artificial intelligence in flight safety: Fatigue monitoring and risk mitigation technologies. *Applied Science and Engineering Journal for Advanced Research*, *3*(5), 1-9.
8. J. Chen, J. Xiao, & W. Xu. (2024). A hybrid stacking method for short-term price forecasting in electricity trading market. in *8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 1–5.
9. Li, B., Jiang, G., Li, N., & Song, C. (2024). *Research on large-scale structured and unstructured data processing based on large language model*.
10. Li, B., Zhang, K., Sun, Y., & Zou, J. (2024). *Research on travel route planning optimization based on large language model*.
11. Feng, Z., Ge, M., & Meng, Q. (2024). Enhancing energy efficiency in green buildings through artificial intelligence. *Applied Science and Engineering Journal for Advanced Research*, *3*(5), 10-17.
12. Meng, Q., Ge, M., & Feng, Z. (2024). The integration of artificial intelligence in architectural visualization enhances augmented realism and interactivity. *Academic Journal of Science and Technology*, *12*(2), 7-12.
13. Peng, Z. H. A. N. G., Bin-MeiZi, Z. H. A. N. G., Jian-Ke, Z. O. U., & Xiang-Beng, L. I. U. (2018). The methods, effects and mechanism of priming attachment security towards social behaviors. *Journal of Psychological Science,* (3), 615.
14. Qu, M. (2024). High precision measurement technology of geometric parameters based on binocular stereo vision application and development prospect of the system in metrology and detection. *Journal of Computer Technology and Applied Mathematics*, *1*(3), 23-29.
15. Wang, C., Chen, J., Xie, Z., & Zou, J. (2024). *Research on education big data for students academic performance analysis based on machine learning*. arXiv preprint arXiv:2407.16907.
16. Wang, C., Chen, J., Xie, Z., & Zou, J. (2024). *Research on personalized teaching strategies selection based on deep learning*.
17. Wang, D. (Ed.). (2016). *Information Science and Electronic Engineering: Proceedings of the 3rd International Conference of Electronic Engineering and Information Science (ICEEIS 2016), January 4-5, 2016, Harbin, China.* CRC Press.
18. Wang, Y., He, Z., Zou, J., Xie, H., & Bao, J. (2024). Energy transition for sustainable economy: What is the role of government governance and public concern. *Finance Research Letters*, 106087.
19. Chen, J. (2024). School reforms for low-income students under conflict theory. *Journal of Advanced Research in Education*, *3*(3), 36-44.
20. Chen, Y., Yan, S., Liu, S., Li, Y., & Xiao, Y. (2024, August). EmotionQueen: A benchmark for evaluating empathy of large language models. in *Findings of the Association for Computational Linguistics ACL 2024*, pp. 2149-2176.
21. Diao, S., Huang, D., & Jiang, G. (2024). *The role of artificial intelligence in personalized medicine through advanced imaging*.
22. Diao, S., Huang, S., & Wan, Y. (2024, September). Early detection of cervical adenocarcinoma using immunohistochemical staining patterns analyzed through computer vision technology. in *1st International scientific and Practical Conference "Innovative Scientific Research: Theory, Methodology, Practice" (September 03–06, 2024) Boston, USA. International Science Group*, pp. 256.

23. Diao, S., Wei, C., Wang, J., & Li, Y. (2024). *Ventilator pressure prediction using recurrent neural network*. arXiv preprint arXiv:2410.06552.
24. Fu, Y., & Yao, T. (2024). Investigation of O phase spheroidization behavior in Ti2AlNb alloy using high-throughput experiments. *Journal of Materials Engineering and Performance*, 1-11.
25. Huang, D., Liu, Z., & Li, Y. (2024). *Research on tumors segmentation based on image enhancement method*. arXiv preprint arXiv:2406.05170.
26. Sun, Y. and Ortiz, J., 2024. GenAI-driven cyberattack detection in V2X networks for enhanced road safety and autonomous vehicle defense. *International Journal of Advance in Applied Science Research, 3,* 67-75.
27. Huang, D., Ma, H., Wang, J., Du, Y., & Li, R. (2024). Mof-mediated paper-based (bio) sensors for detecting of food and environmental pollutants: Preparation strategies and emerging applications. *Microchemical Journal*, 111692.
28. Huang, D., Xu, L., Tao, W., & Li, Y. (2024). *Research on genome data recognition and analysis based on louvain algorithm*.
29. Huang, S., Diao, S., & Wan, Y. (2024, September). Application of machine learning methods in predicting functional recovery in ischemic stroke patients. in *The 1st International scientific and practical conference "Innovative scientific research: theory, methodology, practice" (September 03–06, 2024) Boston, USA. International Science Group*, pp. 240.
30. Lai, S., Feng, N., Sui, H., Ma, Z., Wang, H., Song, Z., ... & Yue, Y. (2024). *FTS: A framework to find a faithful TimeSieve*. arXiv preprint arXiv:2405.19647.
31. Xiao, J., Xu, W., & Chen, J. (2024). Social media emotional state classification prediction based on Arctic Puffin Algorithm (APO) optimization of Transformer mode. *Authorea Preprints*.
32. Li, B., Jiang, G., Li, N., & Song, C. (2024, August). Research on large-scale structured and unstructured data processing based on large language model. in *Proceedings of the International Conference on Machine Learning, Pattern Recognition and Automation Engineering*, pp. 111-116.
33. Li, B., Zhang, K., Sun, Y., & Zou, J. (2024). *Research on travel route planning optimization based on large language model*.
34. Li, L., Diao, S., Qin, H., & Wan, Y. (2024, June). The impact of artificial intelligence on the development of medical ultrasound imaging. in The 24th International Scientific and Practical Conference "Technologies of Scientists and Implementation of Modern Methods" (June 18–21, 2024) Copenhagen, Denmark. *International Science Group, 431*, pp. 328.
35. Li, N., Xu, L., Zhang, X., & Zou, J. (2024). *Research on financial fraud detection based on deep graph neural network*.
36. Liu, H., Shen, F., Qin, H., & Gao, F. (2024). *Research on flight accidents prediction based back propagation neural network*. arXiv preprint arXiv:2406.13954.
37. Zou, J., Zhao, H., & Li, N. (2024). Operational efficiency through machine learning and optimization in supply chain management. *Academic Journal of Science and Technology, 12*(1), 2024.
38. Zhao, S., Zhang, T., & Li, N. (2024). Machine learning analysis of key features in household financial decision-making. *Academic Journal of Science and Technology, 12*(2), 1-6.
39. Zhang, X., Xu, L., Li, N., & Zou, J. (2024). *Research on credit risk assessment optimization based on machine learning*.
40. Yao, T. (2024, August). Research on the local head loss coefficient in short-tube hydraulic testing. in *3rd International Conference on Applied Mechanics and Engineering Structures (AMES 2024)*, pp. 89-97. Atlantis Press.
41. Xiangyu, G., Yao, T., Gao, F., Chen, Y., Jian, X., & Ma, H. (2024). A new granule extrusion-based for 3D printing of POE: studying the effect of printing parameters on mechanical properties with "response surface methodology. *Iranian Polymer Journal*, 1-12.
42. Wang, H., Li, Z., & Li, J. (2024). *Road car image target detection and recognition based on YOLOv8 deep learning algorithm*.
43. Wang, H., Li, J., & Li, Z. (2024). *AI-generated text detection and classification based on BERT deep learning algorithm*. arXiv preprint arXiv:2405.16422.
44. Wang, H., & Tong, X. *Layer-wise asynchronous training of neural network with synthetic gradient on distributed system*.
45. Wang, C., Chen, J., Xie, Z., & Zou, J. (2024, July). Research on education big data for student's academic performance analysis based on machine learning. in *Proceedings of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Education Digitalization and Computer Science*, pp. 223-227.
46. Wang, C., Chen, J., Xie, Z., & Zou, J. (2024, July). Research on education big data for student's academic performance analysis based on machine learning. in *Proceedings of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Education Digitalization and Computer Science*, pp. 223-227.

ISSN (Online): 2583-2468
Volume-3 Issue-6 || November 2024 || PP. 15-24 DOI: 10.5281/zenodo.14177472

47. Wang, C., Chen, J., Xie, Z., & Zou, J. (2024). *Research on personalized teaching strategies selection based on deep learning*.
48. Shui, H., Sha, X., Chen, B., & Wu, J. (2024, May). Stock weighted average price prediction based on feature engineering and Lightgbm model. in *Proceedings of the 2024 International Conference on Digital Society and Artificial Intelligence*, pp. 336-340.
49. Shimin LE, Ke XU, Huang Y, Xinye SH. (2020). An Xgboost based system for financial fraud detection. in *E3S Web of Conferences, 214*, pp. 02042. EDP Sciences.
50. Salman, U., Belaish, S., Ji, Z., Huang, D., Zheng, N., & Xu, B. (2022). Comparing the economic value of lithium-ion battery technologies in the nine wholesale electricity markets in North America. *iEnergy*, *1*(3), 363-373.
51. Liu, H., Xie, R., Qin, H., & Li, Y. (2024). *Research on dangerous flight weather prediction based on machine learning*. arXiv preprint arXiv:2406.12298.