Stock Price Change Prediction Using Prompt-Based LLMs with RL-Enhanced Post-Hoc Adjustments

Qizhao Chen

Graduate School of Information Science, University of Hyogo, Kobe, Japan

af24o008@guh.u-hyogo.ac.jp

Abstract. While most studies apply large language models (LLMs) to financial sentiment analysis, this research explores a prompt-based approach using LLMs, including Deepseek, Gemma, LLaMA, and Qwen, to estimate stock price percentage changes, also known as simple returns. The models are evaluated on a custom dataset comprising historical stock prices and news articles from both target companies and their related entities, such as competitors and business partners. The results indicate that although LLMs can generate reasonably accurate estimates of stock price fluctuations (actual values), they struggle with forecasting the correct price movement direction (i.e., increase or decrease). To address this, this paper proposes a reinforcement learning framework of using Proximal Policy Optimization to improve directional accuracy of LLMs. This study highlights the potential and limitation of LLMs in stock price forecasting and provides insights into possible solutions to improve the LLM predictive performance.

Keywords: LLM, reinforcement learning, stock price prediction, time series forecasting

1 Introduction

Predicting stock prices has always been an important aspect of financial decision making, enabling investors, traders, and analysts to forecast market movements and refine their investment strategies. As financial markets evolve rapidly and global economic systems grow more complicated, accurately forecasting stock prices has become increasingly challenging, yet more essential.

Stock price forecasting is a complex task traditionally conducted through two main methods: technical and fundamental analysis. Technical analysis examines historical stock prices and volume data to discover trends and patterns, while fundamental analysis evaluates a company's financial status, industry position, and macroeconomic conditions to determine its intrinsic value. Recently, due to their ability to process large amounts of structured data and identify hidden relationships that conventional methods may overlook, machine learning models such as the support vector machine (SVM) and Random Forest have been widely used in stock price prediction [1], [2].

However, market movements are increasingly influenced and shaped by unstructured data sources such as news headlines, social media sentiment, and company

G. M. Lee et al. (eds.), *Proceedings of the 2025 4th International Conference on Bigdata Blockchain and Economy Management (ICBBEM 2025)*, Advances in Intelligent Systems Research 195,
https://doi.org/10.2991/978-94-6463-742-7_46

disclosures. These sources provide valuable insights into market sentiment, investor behavior, and emerging trends, but are difficult to analyze using traditional techniques. This is where natural language processing (NLP) plays a crucial role. Many researchers use NLP to extract sentiment from textual data and then use these sentiment data to predict the stock price [3]–[5].

Recently, the rise of large language models (LLMs), such as LLaMA and Deepseek, has significantly transformed NLP applications. Compared to traditional machine learning and deep learning models, LLMs, originally designed to understand and generate human-like texts, incorporate a much wider range of contextual and qualitative data, which allows them to better process the unstructured data such as news articles. In addition, LLMs offer greater flexibility in customizing input prompts to tailor predictions based on various data types, a feature that traditional machine learning models cannot easily replicate. Furthermore, LLMs require less feature engineering due to their ability to understand and extract features directly from raw data.

However, for the application of LLMs in stock price forecasting, the majority of related research only applies LLMs in sentiment analysis using financial news headlines or articles as input and then enters the predicted sentiment and stock price data into some other machine learning or deep learning models such as LSTM and recently Transformer to forecast the stock price. In other words, LLMs are typically used for sentiment analysis rather than time series forecasting. Thus, their effectiveness in time series predictions remains uncertain. If LLMs can also provide accurate time series predictions, it can greatly simplify stock price forecasting. Therefore, the objective of this paper is to examine the effectiveness of prompt-based LLMs in time series forecasting.

The main contributions of this paper are listed as follows.

- This paper applies a prompt-based method to evaluate the performance of generative LLMs in predicting stock price percentage changes, also known as stock simple returns, and demonstrate that LLMs can achieve competitive results in time series forecasting but exhibit limitations in predicting the direction of stock movement.
- Inspired by the concept of post-hoc calibration, I propose a reinforcement learning framework of using Proximal Policy Optimization (PPO) to improve the directional accuracy of the LLM output. The results show that after the PPO adjustment, the directional accuracy increases by about 20%. This framework provides insights into possible solutions to improve LLM prediction.

The remainder of this paper is structured as follows. Firstly, related work is listed (Section 2). The methodology is then described (Section 3). In Section 4, the experimental results are presented. Finally, I conclude with Section 5.

2 Related Work

2.1 Traditional Stock Prediction Methods

Stock price prediction has been a focal point of financial research for decades, with traditional methods falling into three broad categories: technical analysis, fundamental analysis, and machine learning models.

For technical analysis, some popular techniques include simple moving average (SMI), Relative Strength Index (RSI), and Bollinger Band. For example, Brock et al. [6] show the effectiveness of simple technical trading rules in forecasting stock market behavior, demonstrating that moving averages and breakouts of the trading range can generate significant returns. However, some critics argue that technical analysis often fails to consider external factors, such as news events or macroeconomic changes, limiting its predictive power.

For fundamental analysis, research by Fama and French [7] highlights the importance of factors such as company size and book-to-market ratios in predicting stock returns. Although fundamental analysis provides a comprehensive view of a company's financial health, it is often time-consuming and relies heavily on the availability of accurate and timely data.

With the advent of machine learning, researchers have developed more complicated models for stock price prediction. For example, Sang and Li [8] propose AMV-LSTM, an improved LSTM model with an attention mechanism and optimized gating to enhance stock price prediction.

2.2 NLP in Financial Forecasting

The integration of NLP into financial forecasting has opened new avenues for leveraging unstructured textual data, such as news articles, social media posts, and earnings reports. For example, Hajek and Novotny [9] propose a novel stock price prediction approach by integrating news sentiment analysis and investor attention metrics within a Temporal Fusion Transformer framework. Similarly, Mendoza-Urdiales et al. [10] analyze Twitter sentiment's influence on stock prices using Transfer Entropy and EGARCH models and find that negative sentiment has a stronger impact on stock prices than positive sentiment.

The introduction of transformer-based models, such as BERT and GPT, has revolutionized NLP and their applications in finance. These models excel at capturing contextual relationships in text, making them suitable for tasks such as sentiment analysis and document classification. For instance, Chen and Kawashima [11] compare the performances of several LLMs such as GPT and LLaMA with commonly used models such as VADER in financial sentiment analysis and the results show that generative LLMs can outperform the tradition models.

2.3 Prompt-Based Methods

Prompt-based methods represent a significant shift in how LLMs are applied to specific tasks, including financial forecasting. Zero-shot and few-shot learning enable LLMs to perform tasks with little or no task-specific training data. For example, Brown et al. [12] demonstrate the capabilities of GPT-3 in zero-shot and few-shot settings, showing that the model can generalize across different types of tasks with minimal prompting.

In finance, this capability allows LLMs to be applied to tasks such as sentiment analysis and stock price prediction without the need for extensive fine-tuning.

The design of the prompts plays an important role in the performance of LLMs. Prompt engineering involves adjusting inputs that guide the model to generate the desired output. For example, Shin et al. [13] explore the impact of prompt design on model performance, showing that carefully designed prompts can significantly improve accuracy.

3 Methodology

3.1 Data

To evaluate the effectiveness of LLMs in stock price change forecasting, I created a dataset, mainly consisting of two types of data: daily stock prices and news articles. The dataset includes daily stock features (open, high, low, close, and volume) for Apple, HSBC, Pepsi, Tencent, and Toyota, together with corresponding daily news articles for these companies. In addition, news articles related to entities connected to these companies, such as business partners and competitors, are also included, as news about related companies can influence the stock prices of the target companies [14]. The dataset contains approximately 4,000 rows and is available on Hugging Face.

3.2 Prompt Design

A structured prompt is created to guide the LLMs in forecasting the percentage changes of the company's close price in the next day. The dataset, described in Section 3.1 above, will feed the prompt in the position of {stock and sentiment data}. Table 1 is an example of the prompt. The output of the prompt represents the actual value of stock price changes, with the sign indicating the direction of movement. A positive sign indicates an increase in the stock price, while a negative sign indicates a decrease.

3.3 Model Setup

1) Large Language Models: This study employs several advanced LLMs to evaluate their effectiveness in predicting movements of the stock market. The models include Llama 3, Gemma 2, Mistral 7b, Qwen 2.5, and Deepseek R1. These models are chosen for their strong capabilities in understanding financial text, reasoning over numerical data, and generating structured predictions. The model versions are listed in Table 2. Except for Deepseek, all other models will be fine-tuned because the "deepseek-r1-distill-llama-70b" model is from Groq AI [1], which is an American AI company and provides API key to use its models. Fine-tuning of this version of Deepseek model is not available. For all other models, Unsloth[2], an open source fine-tuning library , is used to fine-tune the models.

[1] https://groq.com/

[2] https://unsloth.ai/

Table 1. Prompt for Stock Price Prediction.

Analyze historical stock data and relevant news data to predict whether the closing price of AAPL (Apple) will rise or fall on {next_date}. Consider the following factors: 1. News directly related to Apple (e.g., earnings reports, product launches, executive changes). 2. News about companies with strong business ties or market correlations to Apple, such as major suppliers, competitors, and industry partners. 3. Broader market trends and macroeconomic factors that may impact Apple's stock price. Based on your analysis, respond with only one of the following: If the stock price is predicted to rise, respond with "X/100", where X is the estimated percentage increase. If the stock price is predicted to fall, respond with "X/100", where X is the estimated percentage decrease. For example: If the stock price is predicted to rise by 2.5%, respond with "0. 025". If the stock price is predicted to fall by 1.8%, respond with "-0.018". Do not include any additional commentary, explanations, or disclaimers in your response. Stock and news data for your reference: {**stock and sentiment data**}. Answer:

Unsloth makes fine-tuning two times faster but uses 70% less memory with no accuracy degradation. The fine-tuning Python execution can also be found on the Unsloth website.

Table 2. LLM Versions Used in this Study.

Model	Version
LLaMA	meta-Llama-3.1-8B-Instruct-bnb-4bit
Gemma	gemma-2-9b
Mistral	mistral-7b-instruct-v0.3-bnb-4bit
Qwen	Qwen2.5-7B
Deepseek	deepseek-r1-distill-llama-70b

A brief introduction to each model is given below. **Deepseek** is an open source LLM that leverages a Mixture-of-Experts approach, making it more efficient and cost-effective. **Gemma**, developed by Google, specializes in extracting important insights from text-heavy data sources such as news sentiment and earnings reports, which are significant for market analysis. **LLaMA**, a transformer-based model developed by Meta, is optimized for efficiency and reduced computational demands, which makes it a good candidate for financial forecasting. **Mistral**, recognized for its lightweight yet highly performing structure, provides a balance between computational efficiency and accuracy, making it a reliable choice for processing structured and unstructured financial data. **Qwen** is a series of large language models developed by Alibaba Cloud, designed for tasks such as text generation, code completion, and multilingual processing.

The input format and prompt structure remain consistent across models.

2) ***Reinforcement Learning Model****:* Inspired by the concept of post-hoc calibration, where additional models are used to adjust the predictions made by a primary model, I apply PPO to adjust the LLM predictions, ensuring that the adjusted output better aligns with real-world stock price movements and enhances directional accuracy. However, unlike traditional post-hoc calibration, which modifies confidence scores for classification tasks, I apply PPO to adjust the LLM-predicted output.

PPO is a reinforcement learning algorithm designed to improve both stability and efficiency in policy optimization. It belongs to the family of policy gradient methods that update an agent's decision-making strategy by maximizing the expected rewards. Unlike traditional policy gradient approaches that often suffer from instability, PPO introduces a clipped objective function that prevents overly large updates, ensuring smoother and more reliable learning. This method allows for multiple updates using the same batch of data, striking a balance between sample efficiency and robustness.

The framework is illustrated as follows: First, the LLM-predicted output is split into two sets: one for training the RL agent and the other one for evaluating its performance. The training set is used as part of the state (observation space) in the reinforcement learning environment. The RL agent uses this initial prediction as a reference and learns to make slight adjustments to improve performance. It refines its adjustment strategy over time by receiving rewards for better directional accuracy, with the validation set being used to evaluate the agent's performance after training.

The main logic of this RL framework is to use the LLM-predicted output as part of the initial state or input rather than initializing PPO with a random action or prediction.

More specifically, **State** is defined as the LLM-predicted output with historical close price and technical indicators. **Action** is defined as the adjustment of the LLM-predicted stock price percentage changes. The adjustment is modeled as a continuous space from −0.5 to 0.5. **Reward** is based on improved directional accuracy.

For coding, the environment is created using the Python library Gym, which includes both the state space and the action space. After defining the environment, PPO from the Python library Stable-Baselines3 is used to train the agent.

3.4 Evaluation Metrics

Mean Squared Error (MSE), Mean Absolute Error (MAE), and Directional Accuracy (DA) are used to evaluate the model performances of the LLMs. MSE calculates the average squared difference between predicted and actual values, making it sensitive to large errors due to squaring. MAE measures the average absolute difference, treating all errors with equal weight. DA measures the proportion of times a model correctly predicts the direction of stock price changes (up or down), regardless of the magnitude of the error. The formula is given by:

$$DA = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{(sign(\hat{y}_i)=sign(y_i))}$$

where $\hat{y}_i$ is the predicted percentage change in stock price for the i-th observation, y_i is the actual percentage change in stock price for the i-th observation, $sign(x)$ is a function that returns +1 if $x > 0$, −1 if $x < 0$, and 0 if $x = 0$, $\mathbb{1}_{(\cdot)}$ is the indicator function, which returns 1 if the predicted and actual signs match and 0 otherwise, N is the total number of observations.

Table 3. Prediction Evaluation.

Model Version	**Fine-tuning**	**MSE**	**MAE**	**DA**	**PPO-adjusted DA**
Meta-Llama-3.1-8B-Instruct-bnb-4bit	✓	0.000531	0.0176	0.4601	0.6221
gemma-2-9b	✓	0.000327	0.0129	0.4870	0.6409
mistral-7b-instruct-v0.3-bnb-4bit	✓	0.000336	0.0130	0.4618	0.6233
Qwen2.5-7B	✓	0.000363	0.0136	0.4660	0.6262
deepseek-r1-distill-llama-70b	×	0.000394	0.0152	0.4787	0.6341

4 Results

The Table 3 presents the performance comparison of LLMs in predicting percentage changes in the stock price, evaluated using MSE, MAE and DA. "PPO-adjusted DA" illustrates the LLM prediction results after PPO adjustment, the reinforcement learning technique that aims to improve prediction performance.

The results show that Gemma-2-9B not only has the lowest MSE (0.000327) and MAE (0.0129) but also the highest DA (0.4870), suggesting that it delivers the most accurate predictions among all the models. This is likely due to its model size and fine-tuning. Compared to other fine-tuned models, Gemma-2-9b is a larger model, with 9 billion parameters, striking a balance between model size and computational efficiency. In addition, fine-tuning makes it adapt to particular patterns and nuances of the stock market data, which may be another reason for its better performance.

Although all LLMs achieve relatively low MSE and MAE, their DA is only around 50%, indicating poor performance and highlighting the challenge of applying LLMs in stock price forecasting: Although models may minimize absolute errors, accurately capturing market direction remains difficult.

In addition, compared to other models, DeepSeek-R1-Distill-Llama-70B generates competitive results despite not being fine-tuned. This is likely because its broader knowledge base enables a deeper contextual understanding of stock price movements. Without fine-tuning, the model may rely more on its extensive pretrained knowledge, allowing it to generalize well across different market conditions. This suggests that, in some cases, a well-trained base model can perform competitively even without additional task-specific adjustments.

After adjustment of PPO, the directional accuracy of all LLMs improves by approximately 20%, highlighting the effectiveness of reinforcement learning in improving prediction performance. In particular, after adjustment, Gemma-2-9B still maintains the highest DA (0.6409), closely followed by DeepSeek-R1-Distill-Llama-70B.

5 Conclusion

This study compares the performance of several LLMs in percentage change in stock price movement. The experimental results show that all LLMs can predict the numerical percentage changes with small errors but fail to capture the right direction of the stock movement for about half of the time, which indicates the limitation of applying LLMs in stock price forecasting. However, by applying PPO adjustment, a reinforcement learning technique, the DA of all LLMs improves by about 20%, indicating that although LLMs exhibit limitations in stock price forecasting, the use of hybrid models, a combination of LLM and reinforcement learning in this research, can improve performance. Furthermore, this paper also demonstrates the impact of both model size and finetuning on performance. Models with larger sizes and fine-tuning are more likely to exhibit strong performance.

For future study, I plan to extend the RL framework by incorporating Reinforcement Learning with Human Feedback (RLHF). In this approach, instead of relying solely on rewards from the environment, the model will receive feedback from human evaluators, guiding the RL agent towards more accurate predictions.

References

1. P. Mehta, S. Pandya, and K. Kotecha, "Harvesting social media senti- ment analysis to enhance stock market prediction using deep learning," *PeerJ Computer Science*, vol. 7, p. e476, 2021.
2. C. C. Kao, C. ChiangLin, and K. C. Yang, "Applying random forest algorithm to predicting the stock price trend of ic design companies," in *2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2023, pp. 0098–0102.
3. L. S. Parvatha, D. Naga Veera Tarun, M. Yeswanth, and J. S. Kiran, "Stock market prediction using sentiment analysis and incremental clustering approaches," in *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, 2023, pp.888–893.
4. N. Darapaneni, A. R. Paduri, H. Sharma, M. Manjrekar, N. Hindlekar, P. Bhagat, U. Aiyer, and Y. Agarwal, "Stock price prediction using sentiment analysis and deep learning for indian markets," 2022.
5. A. Churi, D. Chakraborty, R. Khatwani, G. Pinto, P. Shah, and R. Sekhar, "Stock price prediction using deep learning and sentiment analysis," in *2023 2nd International Conference on Futuristic Technologies (INCOFT)*, 2023, pp. 1–6.
6. W. Brock, J. Lakonishok, and B. LeBaron, "Simple technical trading rules and the stochastic properties of stock returns," *The Journal of Finance*, vol. 47, no. 5, pp. 1731–1764, 1992.
7. E. F. Fama and K. R. French, "The cross-section of expected stock returns," *The Journal of Finance*, vol. 47, no. 2, pp. 427–465, 1992.
8. S. Sang and L. Li, "A novel variant of lstm stock prediction method incorporating attention mechanism," *Mathematics*, vol. 12, no. 7, 2024.

9. P. Hajek and J. Novotny, "Beyond sentiment in stock price prediction: Integrating news sentiment and investor attention with temporal fusion transformer," in *Artificial Intelligence Applications and Innovations*, I. Maglogiannis, L. Iliadis, J. Macintyre, M. Avlonitis, and A. Papaleonidas, Eds. Cham: Springer Nature Switzerland, 2024, pp. 30–43.
10. R. A. Mendoza-Urdiales, J. A. Nunez-Mora, R. J. Santillan-Salgado, and H. Valencia-Herrera, "Twitter sentiment analysis and influence on stock performance using transfer entropy and egarch methods," *Entropy*, vol. 24, no. 7, 2022.
11. Q. Chen and H. Kawashima, "Stock price prediction using llm-based sentiment analysis," in *Proceedings of the IEEE BigData 2024*. Washington DC, USA: IEEE, 2024, pp. 4828–4835.
12. T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," 2020.
13. T. Shin, Y. Razeghi, R. L. L. IV, E. Wallace, and S. Singh, "Autoprompt: Eliciting knowledge from language models with automatically generated prompts," 2020.
14. X. Wan, J. Yang, S. Marinov, J.-P. Calliess, S. Zohren, and X. Dong, "Sentiment correlation in financial news networks and associated market movements," *Scientific Reports*, vol. 11, no. 1, Feb. 2021.