FinRLlama - FinRL Contest 2024 Task II

Arnav Grover
grover41@purdue.edu
Purdue University
West Lafayette, Indiana, USA

Figure 1: FinRLlama

Abstract

In response to the task proposed by the FinRL Challenge, this study proposes a novel prompt framework for training large language models (LLMs) adapted with Reinforcement Learning from Market Feedback (RLMF). Our framework incorporates market-specific features and short-term price dynamics to generate more precise trading signals. Traditional LLMs, while competent in sentiment analysis, lack the contextual alignment required for financial market applications. To bridge this gap, we fine-tune the LLaMA-3.2-3B-Instruct model using a custom RLMF prompt design that integrates historical market data and reward-based feedback. Our evaluation shows that this RLMF-tuned framework outperforms baseline methods in signal consistency and achieving tighter trading outcomes.

Conference'24, November 2024, NYC, NY, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

CCS Concepts

• **Computing methodologies → Natural language generation**; **Reinforcement learning**; *Information extraction.*

Keywords

Large Language Models (LLMs), Reinforcement Learning, Financial Sentiment Analysis, Prompt Engineering, Market Feedback, Trading Signals

ACM Reference Format:
Arnav Grover. 2024. FinRLlama - FinRL Contest 2024 Task II. In . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

The application of large language models (LLMs) to financial sentiment analysis represents a significant opportunity for algorithmic trading strategies. While LLMs demonstrate sophisticated language understanding capabilities, their application in financial contexts has been limited by the challenge of incorporating market-specific knowledge and temporal dynamics.

1.1 Background And Related Work

The evolution of financial sentiment analysis has followed several key trajectories in the literature. Early work by Loughran and

McDonald [3] established the importance of domain-specific dictionaries for financial text analysis, highlighting how general-purpose sentiment tools often fail in financial contexts.

Recent developments in prompt engineering have shown promising results across various domains. Wei et al. [7] demonstrated how carefully constructed prompts can elicit domain-specific knowledge from LLMs without fine-tuning, while Vatsal and Dubey [6] provided various methods / frameworks for evaluating prompt effectiveness in various NLP tasks. However, applications in financial sentiment analysis have been limited, with most approaches focusing on model architecture modifications rather than prompt optimization.

1.2 Research Objectives

In response to these limitations, we establish the following objectives for this task:

(1) Develop a novel prompt engineering framework for financial sentiment analysis.
(2) Establish a systematic training methodology for model optimization that adapts dynamically to changing market conditions.
(3) Empirically validate the effectiveness of this framework in improving sentiment-based signal precision and trading performance.

These objectives aim to create a robust framework for generating actionable insights from financial news, advancing LLM utility in financial applications through a combination of prompt engineering and market-aligned learning.

2 Methodology

2.1 Prompt Architecture

The sentiment analysis prompt architecture is designed to generate actionable insights from news headlines to predict stock performance, leveraging a sentiment scale ranging from highly negative to highly positive. This approach is rooted in financial literature that highlights the importance of sentiment analysis in market prediction [3], [2]. The inclusion of market feedback and historical price data aligns with studies that demonstrate the effectiveness of incorporating past performance to refine predictive accuracy [1]. By using adjustable parameters like signal strength and thresholds, the prompt can adapt to different market conditions, a method known to improve model robustness. The instruction to output only a single sentiment score without explanations is a deliberate choice to maintain clarity and speed in real-time trading scenarios. This design ensures that the model provides quick, interpretable insights that are actionable for decision-making, consistent with established practices in financial forecasting.

2.2 Training Process

The model fine-tuning process begins with the base Llama-3.2-3B-Instruct model. The reinforcement learning (RL) component simulates market interactions, where the model outputs sentiment signals, and the model selects trading actions (long, short, or hold). The reward function then evaluates the model's predictions by comparing sentiment scores to actual market performance, assigning

Algorithm 1: Sentiment Signal Scoring Prompt

Input: Signal Bounds: `signal_strength`, Threshold: `threshold`, News Headline: `news`, Price Data: `prices`
Output: Integer between `-signal_strength` and `signal_strength`

[CONTEXT]
Task: Analyze the stock-related news headline and output a sentiment score ranging from `-signal_strength` to `signal_strength` reflecting the sentiment's potential impact on stock performance.

[SENTIMENT SCORING PARAMETERS]
`-signal_strength:` Highly negative sentiment, indicating significant potential for stock price decline.
`-threshold:` Moderately negative sentiment, suggesting minor decline potential.
`0:` Neutral sentiment, indicating no substantial effect on stock price.
`threshold:` Moderately positive sentiment, suggesting potential for slight stock growth.
`signal_strength:` Highly positive sentiment, suggesting significant potential for stock price appreciation.

[MARKET FEEDBACK CONSIDERATIONS]
`Past Market Responses:` Incorporate past market responses to similar news events.
`Market Sentiment Alignment:` Evaluate if the news aligns with or contradicts prevailing market sentiment.
`Historical Price Patterns:` Analyze the historical impact of similar news on stock prices.

[SENTIMENT SCORING EXAMPLES]
"Company X announces layoffs amid economic downturn." Sentiment Score: -8
"Company Y reports record revenue growth in Q1." Sentiment Score: 7
"Market responds positively to Company Z's new product launch." Sentiment Score: 5

[OUTPUT]
Output a single integer sentiment score in the range `-signal_strength` to `signal_strength` based on the analysis.

rewards or penalties based on the accuracy of the sentiment and resultant profits or losses. This process aligns with established RL frameworks in financial applications [5], [4].

The reward function is dynamically adjusted based on the strength of the model's sentiment signal, reinforcing correct predictions and penalizing errors. The function takes into account the model's confidence, incorporating adjustable thresholds to assess market direction. For instance, when the sentiment score exceeds a threshold, the reward varies depending on the actual price movement: long positions are rewarded if a strong positive return is observed, while negative returns despite positive sentiment lead to penalties. Similarly, short positions are rewarded when negative returns align with the sentiment. This system helps the model refine its decision-making over time through feedback loops, gradually improving its accuracy and trading strategies.

The model's fine-tuning process is guided by the Adam optimizer, minimizing the loss function based on the discrepancy between predicted sentiment signals and actual market outcomes. This approach follows deep RL strategies aimed at optimal decision-making, balancing exploration and exploitation to generate robust sentiment-based trading signals.

3 Results

3.1 Experimental Setup

The experimental setup for testing and validating the proposed model's performance spans 2020 to 2023, assessing the accuracy and profitability of sentiment-based trading signals against the baseline. This setup includes news headlines, stock price data, and technical indicators to align sentiment scores with stock movements effectively. Each headline is preprocessed to link with relevant stock price data, and a three-day forward close price is added to facilitate forward-looking impact analysis.

We designate 2020–2022 as the training period and use 2023 exclusively for evaluation. This split enables an assessment of model robustness across diverse market conditions. For both the model and the baseline, each headline generates a buy, hold, or sell signal, with performance measured by cumulative returns, win/loss rate, etc.

3.2 Performance Metrics

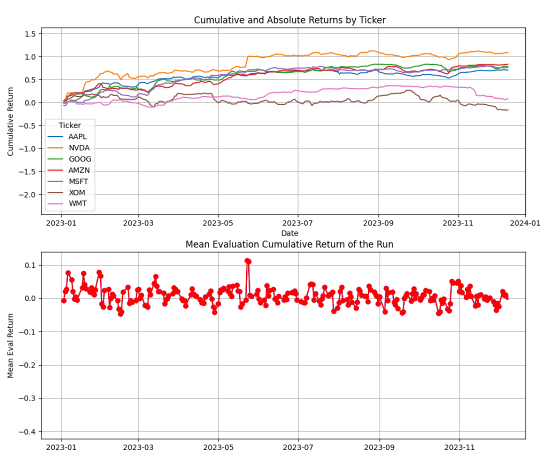

Figure 2: FinRLlama Cumulative Returns Plot

3.3 Comparative Analysis

In Figure 2, cumulative returns across tickers appear to be less volatile. Although NVDA still leads with positive cumulative returns, its gains are not as pronounced, and the spread between the highest and lowest performing stocks is narrower compared to the first plot. XOM still trends downward, but with a less steep decline. This indicates a model response that may be more conservative, possibly due to model finetuning. In Figure 3, the cumulative returns show significant variation among tickers. For instance, NVDA displays consistently higher returns, reaching above 1.5, indicating strong performance. Other stocks like MSFT and GOOG exhibit moderate cumulative returns, staying close to 0.5, while XOM shows a downward trend - temporarily dipping into negative returns. This plot suggests a broader divergence in performance across stocks, with the model interpreting NVDA as markedly outperforming others and XOM underperforming.

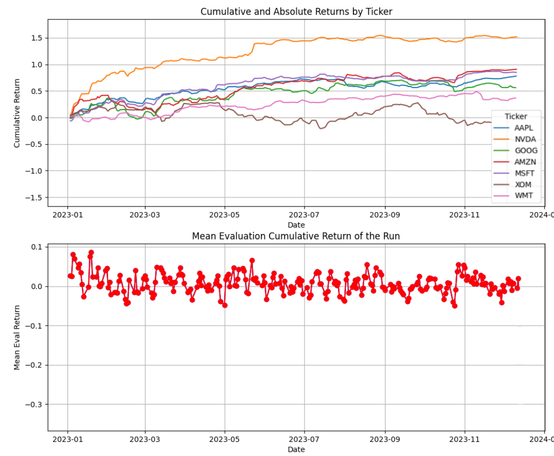

Figure 3: Llama Cumulative Returns Plot

The mean evaluation cumulative return, shown in both second subplots, oscillates around zero in both cases.

In summary, Llama-3.2-3B-Instruct displays a broader range of cumulative returns, indicating higher variability and greater individual gains and losses, while FinRLlama suggests a more conservative approach with reduced volatility in cumulative returns across tickers and smoother mean evaluation.

4 Future Work

Future work to improve the model could focus on refining the reward function to better capture the nuances of financial market dynamics. By incorporating dynamic reward adjustments that account for market volatility and shifts in sentiment, the model could become more responsive to both short-term fluctuations and long-term trends. Enhancing the model's ability to process and integrate historical price data and sentiment trends could improve its prediction accuracy, enabling it to account for delayed market reactions more effectively. Additionally, further fine-tuning with domain-specific financial data would help the model better adapt to the intricacies of market behavior, improving its decision-making accuracy. These improvements could significantly enhance the model's robustness and its ability to generate actionable trading insights.

References

[1] Fahad Alhindi, S. M. Raza, and P. Gupta. 2023. Predicting Economic Trends and Stock Market Prices with Deep Learning and Advanced Machine Learning Techniques. *MDPI Electronics* 13, 17 (2023), 3396. https://doi.org/10.3390/electronics13173396 Explores deep learning methods for predicting stock market trends.

[2] Huina Mao Johan Bollen and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1 (2011), 1–8. https://doi.org/10.1016/j.jocs.2010.12.007

[3] Tim Loughran and Bill McDonald. 2011. When is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66, 1 (Feb. 2011), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x Explores domain-specific dictionaries for financial text analysis.

[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518 (2015), 529–533. https://doi.org/10.1038/nature14236

[5] Shuo Sun, Rundong Wang, and Bo An. 2023. Reinforcement Learning for Quantitative Trading. *ACM Transactions on Intelligent Systems and Technology* 14, 3 (2023), 1–29. https://doi.org/10.1145/3582560

[6] Shubham Vatsal and Harsh Dubey. 2023. A Survey of Prompt Engineering Methods in Large Language Models for Different NLP Tasks. *arXiv preprint* (2023). https://arxiv.org/pdf/2407.12994 Discusses various prompt engineering methods for different NLP tasks.

[7] Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. 2022. Chain of Thought Prompting Elicits Reasoning in Large Language Models. *arXiv preprint arXiv:2201.11903* (2022). https://arxiv.org/abs/2201.11903 Demonstrates how prompt engineering can elicit domain knowledge.