

FinRLlama: A Solution to LLM-Engineered Signals Challenge at FinRL Contest 2024

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



Figure 1: FinRLlama

Abstract

In response to Task II of the FinRL Challenge at ACM ICAIF 2024, this study proposes a novel prompt framework for fine-tuning large language models (LLM) 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 contextual alignment 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; awarded as winner of Task II. You can find the code for this project on [GitHub](#).

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. 2025. FinRLlama: A Solution to LLM-Engineered Signals Challenge at FinRL Contest 2024. 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 [5]. Although LLMs demonstrate sophisticated

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

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

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

language understanding capabilities, their application in financial contexts has been limited by the challenge of incorporating market-specific knowledge and temporal dynamics [8].

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 [6] established the importance of domain-specific dictionaries for financial text analysis, highlighting how general-purpose sentiment tools often fail in financial contexts. In parallel, the development of comprehensive financial reinforcement learning frameworks like FinRL-Meta [4] has provided standardized environments for developing and evaluating trading strategies.

Recent developments in prompt engineering have shown promising results in various domains. Wei et al. [11] demonstrated how carefully constructed prompts can elicit domain-specific knowledge from LLMs without fine-tuning, while Vatsal and Dubey [10] 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, we establish the following objectives for this task:

- (1) Develop a novel prompt framework for sentiment analysis.
- (2) Establish a systematic training methodology for model optimization that adapts dynamically to market conditions.
- (3) Empirically validate the framework's impact on sentiment-based signal precision and trading performance.

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

2 Methodology

2.1 Prompt Architecture

The sentiment analysis prompt architecture generates stock performance predictions from news headlines using a scalable sentiment framework. Building on established sentiment analysis methods in finance [6], [3], the system integrates market feedback and historical data to improve predictive precision [1]. Adjustable parameters enables market adaptation, while single-score output supports rapid trading decisions.

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 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 [9], [7], and follows the task

Algorithm 1: Sentiment Signal Scoring Prompt

Input: Signal Bound, Threshold, News Headline, Price Data
Output: Value in $[-\text{signal_strength}, \text{signal_strength}]$

- 1 [CONTEXT]
 - 2 Task: Analyze the stock-related news headline and output a sentiment score reflecting the sentiment's potential impact on stock performance.
 - 3 [SENTIMENT SCORING PARAMETERS]
 - 4 $-\text{signal_strength}$: Highly negative market sentiment
 - 5 $-\text{threshold}$: Moderately negative market sentiment
 - 6 0: Neutral market sentiment
 - 7 threshold : Moderately positive sentiment
 - 8 signal_strength : Highly positive sentiment
 - 9 [MARKET FEEDBACK CONSIDERATIONS]
 - 10 Past Market Responses: Incorporate past market responses to similar news events.
 - 11 Market Sentiment Alignment: Evaluate if the news aligns with or contradicts prevailing market sentiment.
 - 12 Historical Price Patterns: Analyze the historical impact of similar news on stock prices.
 - 13 [SENTIMENT SCORING EXAMPLES]
 - 14 "Company X announces layoffs amid economic downturn."
Sentiment Score: -8
 - 15 "Company Y reports record revenue growth in Q1."
Sentiment Score: 7
 - 16 "Market responds positively to Company Z's new product launch." Sentiment Score: 5
 - 17 [OUTPUT]
 - 18 Integer sentiment score in range $[-\text{signal_strength}, \text{signal_strength}]$ based on analysis.
-

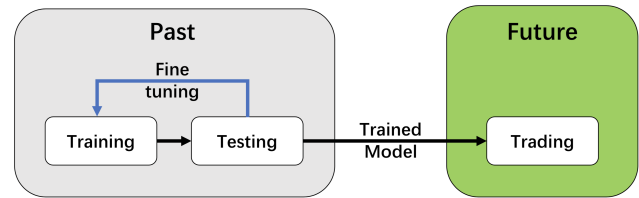


Figure 2: FinRL Train-Test-Trade Pipeline

requirements outlined in the ACM ICAIF FinRL 2024 Competition [2].

The reward function is dynamically adjusted on the basis of the strength of the model's sentiment signal, reinforcing correct predictions, and penalizing errors. The function takes into account the confidence of the model, 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 effectively align sentiment scores with stock movements. Each headline is pre-processed 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 the robustness of the model in 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

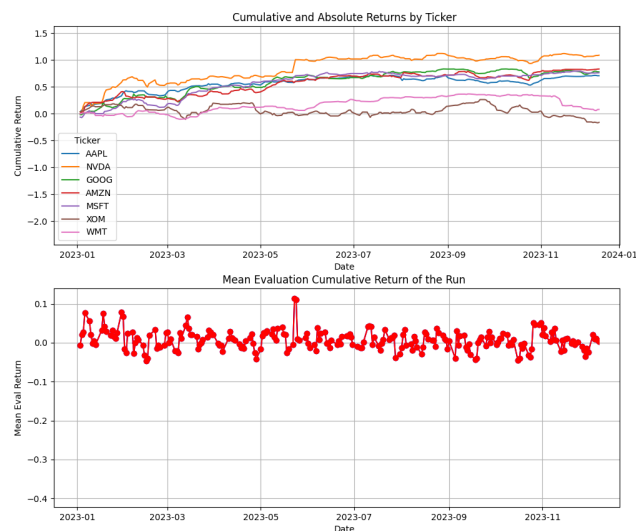


Figure 3: FinRLlama Cumulative Returns

3.3 Comparative Analysis

In Figure 3, the cumulative returns of the tickers appear to be less volatile. Although NVDA still leads with positive cumulative returns, its gains are not as pronounced and the spread between

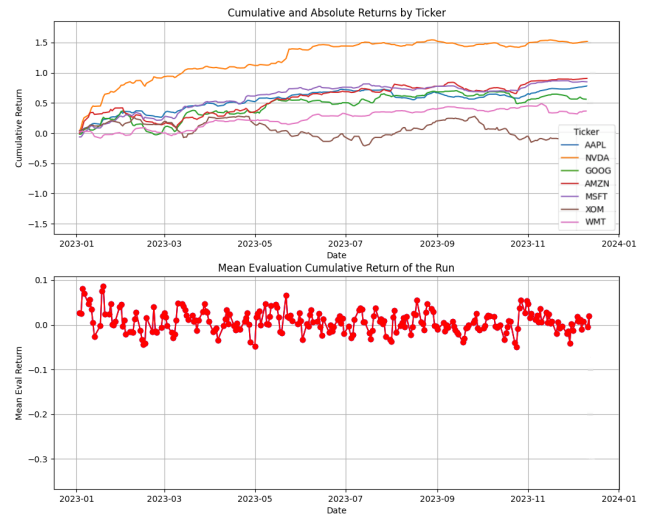


Figure 4: Llama Cumulative Returns

the highest and lowest performing stocks is narrower compared to the first graph. XOM continues to trend downward, but with a less steep decline. This indicates a model response that may be more conservative, possibly due to model fine-tuning. In Figure 3, the cumulative returns show significant variation between tickers. For example, 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 between stocks, with the model interpreting NVDA as markedly outperforming others and XOM underperforming.

The mean cumulative evaluation 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 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, allowing it to account for delayed market reactions more effectively. Furthermore, 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 improve the robustness of the model and its ability to generate actionable trading insights.

5 Acknowledgments

Arnav Grover acknowledges the support of Keyi Wang, Nikolaus Holzer, and Xiao-Yang Liu Yanglet for their invaluable guidance and support throughout the development of this work. We also extend sincere thanks to Yanglet and the organizing team of the FinRL 2024 contest for their exceptional efforts in creating a collaborative and intellectually stimulating platform. Their dedication to fostering innovation and research in the field of open finance has provided an invaluable opportunity to contribute to this evolving domain.

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] Nikolaus Holzer, Keyi Wang, Kairong Xiao, and Xiao-Yang Liu. 2024. Revisiting Ensemble Methods for Stock Trading and Crypto Trading Tasks at ACM ICAIF FinRL Contest 2023-2024. *arXiv preprint* (2024). <https://doi.org/10.48550/arXiv.2501.10709>
- [3] 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>
- [4] Xiao-Yang Liu, Ziyi Xia, Jingyang Rui, Jiechao Gao, Hongyang Yang, Ming Zhu, Christina Dan Wang, Zhaoran Wang, and Jian Guo. 2022. FinRL-Meta: Market Environments and Benchmarks for Data-Driven Financial Reinforcement Learning. *NeurIPS Datasets and Benchmarks* (2022), 1–36. <https://doi.org/10.48550/arXiv.2211.03107>
- [5] Xiao-Yang Liu, Hongyang Yang, Qian Chen, Runjia Zhang, Liuqing Yang, Bowen Xiao, and Christina Dan Wang. 2020. FinRL: A Deep Reinforcement Learning Library for Automated Stock Trading in Quantitative Finance. *NeurIPS Workshop on Deep Reinforcement Learning* (2020). <https://doi.org/10.48550/arXiv.2011.09607>
- [6] 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.
- [7] 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>
- [8] Yuqi Nie, Yaxuan Kong, Xiaowen Dong, John M. Mulvey, H. Vincent Poor, Qingsong Wen, and Stefan Zohren. 2024. A Survey of Large Language Models for Financial Applications: Progress, Prospects and Challenges. *arXiv preprint* (2024). <https://doi.org/10.48550/arXiv.2406.11903>
- [9] 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>
- [10] 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.
- [11] 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.