Enhancing FinRL Trading Agents with Advanced LLM-Processed Financial News: An Improved Approach Using DeepSeek-V3

Satish Chandra
Research Scholar
School of Engineering, Anurag University
Hyderabad, India
gullenasatish@gmail.com

Dr. G. Balakrishna
Associate Professor
School of Engineering, Anurag University
Hyderabad, India
balakrishnacse@anurag.edu.in

***Abstract*—The use of AI techniques to improve stock market related operations has been benefiting traders in the stock market. FinRL is a framework that has been shown to be effective in building Reinforcement Learning (RL) based Trading Agents. Use of LLMs for performance improvement of the Agents is one of the recent initiatives. LLMs are being used to generate signals that could be combined with historical stock prices for better prediction of stock price movement trends leading to profitable trading. This paper aims to develop efficiently engineered prompts to generate DeepSeek-V3-based sentiment signals, integrate them into the FinRL environment, and improve the performance of reinforcement learning trading agents. The code, data, and trading agents are at: https://github.com/SatishChandraPhD/FinRL2025**

***Keywords*—Financial Reinforcement Learning, FinRL, trading agent, prompt, large language model, DeepSeek, stock market, financial news, stock price trend**

I. Introduction

Traditionally, prediction of trends of stock prices had been based on features like Open, High, Low and Close (OHLC) prices data. A recent trend has been to consider alternate features like sentiments. However, with the rising awareness of the importance of adding alternate data for stock price trends prediction several researchers have developed models adding alternate data, like sentiments of market participants. One such approach is to build Reinforcement Learning (RL) Trading Agents trained using various features in addition to OHLC data. Mostapha Benhenda has developed a framework for building such RL based Trading Agents[1]. The framework leverages the Stock Recommendation Prompt of Zihan Dong [2].

This paper performs sentiment analysis by processing financial news using different prompting approaches with DeepSeek-V3 to provide prompt based LLM signals that are used for training FinRL based RL Trading Agents. The main contribution of the paper is to demonstrate the pros and cons of using different prompting approaches for generating sentiment signals using DeepSeek-V3 while building the Agents.

II. Related Works and Research Question

A. *FinRL Trading Agents*

Reinforcement Learning enables learning of effective trading strategies. The AI4Finance Foundation is the world's largest AI finance open-source community. The Foundation has hosted FinRL, which is the first open-source framework for financial reinforcement learning [7].

FinRL has three layers: market environments, agents, and applications. For a trading task (in the top layer), an agent (in middle layer) interacts with a market environment (in the bottom layer), making sequential decisions [7].

B. Large Language Models for Reinforcement Learning of Trading Agents

One of the recent innovations in the application of artificial intelligence in finance, especially for trading, is to include LLM based signals as features, like sentiment and risk [1], [2] for training FinRL based trading agents.

○ *C. Proposed Research Question*

The generation of signals from financial text is one of the key challenges identified in [10]. The proposed research question is "Can advanced prompt engineering and novel integration strategies for DeepSeek-V3 processed financial news signals improve the performance of FinRL-based trading agents compared to the baseline FinRL-DeepSeek approach?"

III. Methodology

A. Incorporation of LLM-Generated Signals into FinRL Environment

To integrate LLM-generated sentiment signals into the FinRL-based trading agents, we made the following modifications to the environment setup:

- Each stock at each time step is represented by its OHLC prices, and the corresponding DeepSeek-V3 generated sentiment score
- The sentiment scores were appended directly to the stock feature vector
- We extended the FinRL data preprocessing pipeline to merge DeepSeek-V3 sentiment signals with trade data on a per-date and per-stock basis
- The RL agent (PPO agent) observes a richer state containing traditional technical indicators and sentiment-based signals at every time step
- No changes were made to the reward function. The change in portfolio value ensures comparability with baseline FinRL agents

B. Pseudocode for Preprocessing

Algorithm: Pseudo code for building RL Trading Agents based on different types of prompting techniques and comparing their performance based on Metrics like Sharpe Ratio

```
START
Step 1: Perform EDA on Nasdaq Stocks Datasets
perform_EDA_on_Nasdaq_stocks_datasets()
Step 2: Use DeepSeek-V3 for Sentiment Score computation
use_cleaned_data_for_sentiment_score_computation
(prompt_type= 4 (Few Shot, Counterfactual, Chain of Thought, Role based as in [1]), num_labels=5)
Step 3: Join Sentiment and Trade Data. Join rows of sentiment data with corresponding rows of trade information based on dates and stock symbol
join_sentiment_score_data_with_trade_data(key= date+tic)
store_joined_data(xyz_joined_datset.csv)
Step 4: Load and Prepare Data for Training and Trading
load_and_prepare_data(xyz_prompt_data, ref_prompt_data)
split_data_into_training_trading_datasets()
Step 5: Create Environment for building Agents
create_env(env_type_based_on_prompting_technique)
Step 6: Train two types of Agents for comparison
train_agent(env,model_name="ppo", total_timesteps=n)
save_model_of_agent1()
save_model_of_agent2()
Step 7: Calculate Performance Metrics
calculate_metrics(df_account_value)
Step 8: Evaluate Agent Performance
evaluate_agent(model, trade_df)
Step 9: Compare the performance of different Trading Agents with Trading Agent built by prompt of type in [1]
combine_display_performance_comparison(metrics_type)
End Process
STOP
```

C. Dataset Selection and Preprocessing

As the project of [1] is the basis of task 1 in FinRL Contest 2025, this paper uses the datasets and framework of [1],[2],[8] and [9]. We have used data like 'sentiment_nasdaq_news.csv' , 'trade_data_2019_2023.csv' in generating sentiment signals from DeepSeek-V3.

D. Model Selection

As DeepSeek-V3 outperformed other LLMs [1] we have considered DeepSeek-V3 as the model for generating sentiment based trading signals.

In our experiment, DeepSeek V3 analyzes financial news from the subset of the FNSPID dataset to generate stock-specific scores (1–5) where a score of 5 means "strong buy," while 1 signals "strong sell."

E. Prompt Engineering for DeepSeek-V3

Role-based Prompting has been used in [1] and [2]. We use Few Shot prompting, Counterfactual prompting and Chain-of-Thought (CoT) prompting approaches to generate signals using the DeepSeek-V3. We reused the prompt used in [1] to generate the benchmark metrics.

F. Market Environment

Keyi Wang, et. al. [10] have discussed the importance of high-quality market environments for developing robust trading agents. The signals generated from DeepSeek-V3 need the right market environment. The stable market based environment, detailed in [10] is our environment of choice.

G. Evaluation Metrics

We have used total return, annual return, annual volatility, Sharpe ratio, maximum drawdown and win rate as the metrics for comparing the performance of models (trading agents) that were built using few shot, counterfactual and CoT prompting approaches with the performance of the agents built using the role based prompting approach as documented in [1].

IV. Experiment and Results

This paper followed the methodology described in the previous section for carrying out the experiment. The results are presented below followed by a brief analysis of the results.

A. Few shot vs Role based prompting approaches

TABLE I. Few Shot TA Vs Role based prompt TA

METRIC	Trading Agent's Prompt Type	
	Few Shot	***Reference (Role based)***
Total Return	0.078088	0.072370
Annual Return	0.120755	0.113106
Annual Volatility	0.120535	0.123613
Sharpe Ratio	1.001830	0.915002
Max Drawdown	0.098987	0.101649
Win Rate	0.395210	0.395210
Rachev Ratio	0.931600	0.904200

The Trading Agent based on the few shot prompting approach consistently outperforms across key financial performance metrics. It achieves a higher total return and a superior annualized return, indicating better long-term profitability. It exhibits lower annual volatility, translating into more stable returns over time. The Sharpe Ratio is higher at 1.00, confirming its superior efficiency in converting risk into return. It has a slightly lower maximum drawdown and a higher Rachev Ratio, indicating better tail-risk behavior and downside protection. The identical win rate for both agents suggests that the performance improvements of the Few Shot model are not due to increased trading frequency but rather more effective decision-making. These results highlight that the Trading Agent based on Few Shot prompting strategy, by incorporating more targeted and context-rich examples, enables the agent to capture sentiment signals more accurately and generate better trading decisions. Its robustness across diverse performance dimensions makes it a preferred approach over the baseline used in the reference paper.

B. Counterfactual vs Role based prompting approaches

TABLE II. Counterfactual TA Vs Reference prompt TA

METRIC	**Trading Agent's Prompt Type**	
	Counterfactual	***Reference***
Total Return	0.078088	0.072370
Annual Return	0.120755	0.113106
Annual Volatility	0.120535	0.123613
Sharpe Ratio	1.001830	0.915002
Max Drawdown	0.098987	0.101649
Win Rate	0.395210	0.395210
Rachev Ratio	0.841400	0.808800

The analysis of the performance of the trading agent based on counterfactual approach is better again and is similar to the few shot based trading agent. The Counterfactual prompting approach yields more stable, risk-adjusted performance and better tail-risk behavior, making it particularly well-suited for risk-averse investors or applications requiring robust downside protection.

C. CoT vs Role based prompting approaches

TABLE III. Chain of Thought TA Vs Reference prompt TA

METRIC	**Trading Agent's Prompt Type**	
	Chain of Thought (CoT)	***Reference***
Total Return	0.068061	0.057534
Annual Return	0.108283	0.094133
Annual Volatility	0.133273	0.139077
Sharpe Ratio	0.812485	0.676839
Max Drawdown	0.112923	0.118230
Win Rate	0.395210	0.395210
Rachev Ratio	0.841400	0.808800

Overall, the trading agent based on CoT prompting approach produces more stable, risk-adjusted returns, making it an attractive choice for growth-oriented investors who still require controlled risk exposure.

As the objective of this paper is to come up with a proof of concept that needs to be developed quickly we have used google's colab environment for conducting the experiments.

V. Conclusion and Future Work

Importantly, we incorporated DeepSeek-V3-generated sentiment signals directly into the FinRL environment as additional state features, enabling the agent to learn from both price data and textual sentiment information. The above results are based on an approach that uses just the PPO algorithm. Experiments can be run using different algorithms as well as certain combinations leading to the ensemble approach for training trading agents. Such approaches might give better results. Further, different types of prompts have led to the generation of sentiment signals that perform differently. Experiments can be run using a combination of different prompting approaches that could generate better signals, resulting in better performance.

For future work we could look at the hybrid approaches for building trading agents and generating sentiment signals using an LLM like DeepSeek, as mentioned earlier in this section. Also, we could consider additional aspects like explainability analysis, bias detection and fine-tuning DeepSeek through domain adaptation

Acknowledgment

We would like to thank the AI4Finance Foundation for the pioneering initiative of development and maintenance of FinRL framework in addition to more such finance related frameworks/tools. We would like to thank Mostapha Benhenda for providing the baseline framework [1], Zihan Dong, et. al. for the prompt signal generation idea and datasets [2] and Keyi Wang, et. al. for the succinct description of market environments in [10]. Further, we would like to thank all the authors listed in the references section and beyond whose contributions have led to the use of AI in finance, especially the ones related to stock trading.

References

1) Mostapha Benhenda, "FinRL-DeepSeek: LLM Infused Risk-Sensitive Reinforcement Learning for Trading Agents," arXiv:2502.07393v1.
2) Zihan Dong, et al., "FNSPID: A Comprehensive Financial News Dataset in Time Series," arXiv:2402.06698v1.
3) Xiao-Yang Liu, "FinRL: A Deep Reinforcement Learning Library for Automated Stock Trading in Quantitative Finance," arXiv:2011.09607v2.
4) John Schulman, et. al., "Proximal Policy Optimization Algorithms," arXiv:1707.06347v2.
5) Xiao-Yang Liu, et. al., "Practical Deep Reinforcement Learning Approach for Stock Trading," arXiv:1811.07522v3.
6) https://github.com/benstaf/FinRL_DeepSeek.
7) https://github.com/AI4Finance-Foundation/FinRL.
8) https://huggingface.co/datasets/Zihan1004/FNSPID/tree/main/Stock_news.
9) https://huggingface.co/benstaf .
10) Keyi Wang, Kairong Xiao and Xiao-Yang Liu Yanglet. Parallel Market Environments for FinRL Contests. arXiv preprint arXiv:2504.02281.
11) X.-Y.Liu ,Z.Xia,J.Rui,J.Gao,H.Yang,M.Zhu,C.D.Wang,Z.Wang, and J. Guo, "FinRL-meta:market environments and benchmarks for data-driven financial reinforcement learning," in International Conference on Neural InformationProcessingSystems,NY,USA,2022.
12) Antonin Raffin, Ashley Hill, Maximilian Ernestus, Adam Gleave, Anssi Kanervisto, and Noah Dormann. Stable baselines3. https://github.com/DLR-RM/stable-baselines3, 2019.
13) Arnav Grover. FinRLlama: A Solution to LLM-Engineered Signals Challenge at FinRL Contest 2024. Conference'24, November 2024, NYC, NY, USA.