FinRL: Adaptive Model Selection for Reinforcement Learning in Stock Trading

1st Vorakorn Kosidphokin
Dept. of Computer Engineering
Kasetsart University, Bangkok, Thailand
vorakorn.ko@ku.th

2nd Phawat Loedtrakunchai
Dept. of Computer Engineering
Kasetsart University, Bangkok, Thailand
phawat.lo@gmail.com

3rd Natthakorn Sinamnuaiphon
Dept. of Computer Engineering
Kasetsart University, Bangkok, Thailand
natthakorn.sin@ku.th

4th Surawit Kuptanon
Dept. of Electrical Engineering
Kasetsart University, Bangkok, Thailand
surawit.k@ku.th

***Abstract*—This paper advances the FinRL-DeepSeek framework by introducing a market-adaptive modeling approach to improve automated stock trading agents. While FinRL-DeepSeek integrates stock prices and financial news using reinforcement learning (RL) and large language models (LLMs), our approach incorporates a dynamic model selection mechanism that adapts to market conditions, such as bull and bear markets determined using Nasdaq index trends. We also replace the single-model structure of FinRL-DeepSeek with multiple specialized models, each tailored to individual stocks. This adaptation enables more precise decision making, improving responsiveness to market fluctuations and enhancing overall trading performance.**

***Index Terms*—FinRL, Reinforcement Learning, FinRL-DeepSeek, Adaptive Model Selection, Stock Trading, Macroeconomic Indicators, Specialized Models**

I. Introduction

Automated stock trading has witnessed significant advancements through the integration of reinforcement learning (RL) and large language models (LLMs). The FinRL challenges participants to develop trading agents capable of effectively responding to market conditions using the FinRL-DeepSeek framework [1]. Prior implementations within this framework have demonstrated that individual models perform optimally under specific market conditions, such as bull or bear markets.

Our work extends the FinRL-DeepSeek framework by incorporating the SMA200 and SMA50 indicators into the decision-making process. We use the trend in the Nasdaq index to determine the market state, guiding trade timing. Rather than using the single-author model of FinRL-DeepSeek, we develop several specialized models, each tailored to a specific stock. This design improves adaptability and responsiveness to individual stock price movements and external market forces.

II. Methodology

A. Stock-Specific Model Selection Based on Market Conditions

Building on the market environment described in Wang *et al.* [2], we reconstruct the FinRL-DeepSeek environment, training RL agents on historical stock prices and financial news. The framework is then augmented through:

- **Integration of Additional Data:** We use the Nasdaq index as a benchmark to assess market health by analyzing bullish and bearish price movements, directly informing trade timing.
- **Adaptive Model Selection:** We develop a mechanism to dynamically allocate the most suitable trading model based on market conditions, using SMA200 and SMA50 to capture market trends.

B. Specialized RL Models for Individual Stocks

Also leveraging the environment from Wang *et al.* [2], we replace the single-model paradigm with multiple specialized RL models, each trading an individual stock. Key advantages include:

- **Stock-Specific Optimization:** Dedicated models are fine-tuned to each stock's historical data, volatility, and sensitivity to macroeconomic events.
- **Improved Trading Decisions:** Independent models enable refined decision making, enhancing trade execution by responding precisely to each stock's unique dynamics.

III. Experiments and Results

A. Experimental Setup

We evaluate our approach on a basket of NASDAQ-100 stocks from January 2019 to December 2024. Each strategy—PPO, CPPO, PPO-DeepSeek, CPPO-DeepSeek, Adaptive Portfolio, Mini PPO, and Mini CPPO DeepSeek. The Adaptive Portfolio dynamically selects PPO during bull markets and CPPO-DeepSeek during bear markets. Performance is measured by cumulative return, CVaR, Rachev ratio, and daily/weekly outperform frequencies, benchmarked against the Nasdaq-100 index.

B. Results

Figure 1 shows the cumulative returns of all evaluated strategies from 2019 to 2024. The plot displays only cumulative return trajectories for each strategy.

979-8-3315-9661-3/25/$31.00 ©2025 IEEE
DOI 10.1109/IDS66066.2025.00023

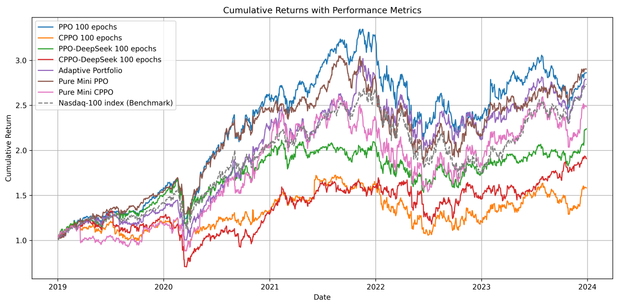

Fig. 1. Cumulative returns of trading strategies from 2019 to 2024.

C. Performance Metrics

Table I summarizes key performance metrics for each strategy:

TABLE I
PERFORMANCE METRICS OF TRADING STRATEGIES

Strategy	IR	CVaR	Rachev Ratio	Daily Outperf.	Weekly Outperf.
PPO (100 epochs)	0.0080	-0.0383	0.9943	0.4912	0.5134
CPPO (100 epochs)	-0.0249	-0.0413	1.0754	0.4610	0.4674
PPO-DeepSeek	-0.0083	-0.0347	0.9070	0.4745	0.4828
CPPO-DeepSeek	-0.0021	-0.0427	0.8797	0.4865	0.4713
Adaptive Portfolio	0.0036	-0.0375	0.9909	0.4796	0.4789
Mini PPO	0.0075	-0.0375	0.9709	0.4857	0.5134
Mini CPPO DeepSeek	-0.0239	-0.0401	1.0100	0.4729	0.4904

Among all strategies, PPO (100 epochs) and Mini PPO deliver the highest weekly outperform frequency (0.5134), suggesting consistent short-term gains. CPPO (100 epochs) achieves the highest Rachev ratio (1.0754), though it underperforms on IR and outperform frequency. PPO-DeepSeek and CPPO-DeepSeek yield mixed results, while the Adaptive Portfolio balances multiple objectives. Mini CPPO DeepSeek displays decent resilience with the highest Rachev after CPPO and reasonable outperform stats.

IV. CONCLUSION AND FUTURE WORK

This paper presents a novel reinforcement learning approach to stock trading, replacing the conventional single-model strategy with multiple specialized models tailored to individual stocks. By integrating macroeconomic indicators into our framework, we improve adaptability and responsiveness to market conditions, leading to more precise trade execution. Experimental results demonstrate that Mini-based portfolios and adaptive switching strategies offer strong long-term returns and favorable risk metrics compared to traditional RL models.

Future work will extend this methodology to futures and options trading and transition from a research-oriented system to a real-time trading framework capable of executing live trades dynamically.

ACKNOWLEDGMENT

This research was supported by the AI Server, Office of Computer Service, Kasetsart University, for providing computational resources. We also thank the FinRL-DeepSeek development team for making their framework publicly available.

REFERENCES

[1] M. Benhenda, "FinRL-DeepSeek: LLM-Infused Risk-Sensitive Reinforcement Learning for Trading Agents," *arXiv preprint arXiv:2502.07393*, 2025.

[2] K. Wang, K. Xiao, and X.-Y. Liu, "Parallel Market Environments for FinRL Contests," *arXiv preprint arXiv:2504.02281*, 2025.