Table 1: Performance comparison of different models on five stocks

Models	TSLA			AAPL			AMZN			NFLX			COIN		
	CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓
Market Baseline															
Random	−25.81	−0.62	62.10	−16.46	−0.23	39.40	−7.80	−0.34	61.20	−3.04	−0.02	32.60	1.27	0.14	57.60
Rule-Based Methods															
MACD (Wang and Kim, 2018)	−46.25	−0.55	74.20	−29.61	−0.42	51.40	−38.14	−0.47	70.30	−27.45	−0.23	47.92	−49.87	−0.54	70.83
RSI (Belafsky et al., 2002)	−45.06	−0.52	72.80	−33.28	−0.48	49.70	−35.10	−0.42	68.20	−26.98	−0.21	46.08	−53.65	−0.63	66.37
Reinforcement Learning Methods															
A3C (Kang et al., 2018)	−86.99	−0.94	96.60	−66.13	−0.75	73.40	−56.63	−0.62	72.80	−35.24	−0.51	64.60	−58.96	−0.69	68.60
DQN (Jeong and Kim, 2019)	−71.25	−0.78	80.62	−47.96	−0.57	56.30	−43.87	−0.48	68.10	0.78	0.11	41.28	−19.25	−0.26	48.41
PPO (Li et al., 2025)	−56.72	−0.48	69.88	−35.50	−0.42	51.40	−28.45	−0.39	64.90	−11.92	−0.26	51.30	9.32	0.34	56.39
LLM-Based Agent															
FinGPT (Yang et al., 2023)	−89.36	−0.98	94.36	−80.37	−1.01	74.60	−60.14	−0.67	70.30	−26.01	−0.40	57.30	−42.63	−0.34	67.09
FinAgent (Zhang et al., 2024)	−65.07	−0.76	85.65	−51.10	−0.69	49.50	−43.29	−0.47	61.20	−18.71	−0.22	41.50	2.76	0.23	56.50
FINMEM (Yu et al., 2024a)	−36.48	−0.45	72.10	−38.94	−0.52	44.82	−27.61	−0.33	66.10	−33.75	−0.37	60.02	−2.13	−0.12	51.28
FINCON (Yu et al., 2024b)	19.67	0.36	59.13	−2.81	−0.06	31.03	7.96	0.12	34.77	6.12	0.17	34.70	21.35	0.63	46.10
FinPos	**62.15**	**0.68**	**42.34**	**36.31**	**0.43**	**27.53**	**30.35**	**0.34**	**18.44**	**28.65**	**1.02**	**20.05**	**54.36**	**0.87**	**34.05**

Table 2: Ablation Results on TSLA, AAPL and AMZN

MTR	QRA	MSP	TSLA			AAPL			AMZN		
			CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓	CR%↑	SR↑	MDD%↓
	✓	✓	18.73	0.35	57.87	12.93	0.20	35.63	9.75	0.21	26.20
✓		✓	53.57	0.49	62.65	29.30	0.31	39.29	27.85	0.43	30.29
✓	✓		58.34	0.63	45.40	34.09	0.39	29.87	28.50	0.32	19.87
✓	✓	✓	**62.15**	**0.68**	**42.34**	**36.31**	**0.43**	**27.53**	**30.35**	**0.34**	**18.44**

assess the contribution of each FinPos component through ablation experiments. Tab. 2 reports results under different module removals, showing that the full model consistently achieves the best performance. Extended ablation results are provided in Appendix D.3.

- **Multi-Timescale Reward (MTR)** is the cornerstone of our position-aware framework. It provides a self-supervised signal that aligns short-term actions with long-horizon outcomes, enabling the agent to internalize the cumulative impact of position decisions.

- **Quantity and Risk Decision Agent (QRA)** is essential for explicit position management. It transforms the agent from a unit-action trader into a continuous-position controller. Removing QRA forces fixed-unit trades, eliminating dynamic exposure control.

- **Market Signal Processing (MSP)** acts as a preprocessing filter that scores raw market data by *relevance* and *importance* before downstream agents analyze them, ensuring that only high-signal information enters the decision pipeline.

5.3.1 Multi-Timescale Reward (MTR)

As shown in Tab. 2, removing MTR causes a severe performance drop across all assets: the CR fall below 20% for all three stocks (vs. 30.4–62.2% in the full model), indicating that MTR is the architectural backbone that enables long-horizon reasoning and strategic consistency in FinPos.

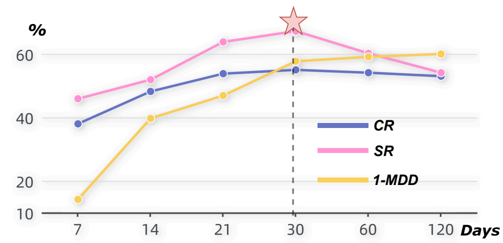

Figure 3: Impact of varying the maximum timescale of the multi-timescale reward on performance metrics.

The effectiveness of MTR depends critically on the choice of timescale for computing the cumulative reward. Our design incorporates three horizons (1, 7, and 30 days) to capture immediate fluctuations, medium-term trends, and long-term stability, respectively. We conduct a sensitivity study by varying the length of the long-term horizon (Fig. 3), finding that: very short windows (7–14 days) yield poor performance, as the agent overreacts to noise without forming stable strategies; performance peaks at 30 days; beyond 30 days, performance declines due to signal dilution, which weakens the reflection mechanism and reduces adaptiveness to regime shifts.

5.3.2 Quantity and Risk Decision Agent (QRA)

Removing QRA leads to a significant degradation in risk control. As shown in Tab. 2, while CR drop moderately (e.g., from 62.2% to 53.6% on TSLA), the MDD worsens dramatically—increasing from 42.3% to 62.7% on TSLA (a 48% relative increase) and from 27.5% to 39.3% on AAPL. This reveals a critical limitation: even with long-horizon awareness provided by MTR, an agent restricted to fixed-unit trades cannot timely adjust its exposure in response to market volatility. Consequently, it lacks the agility needed to hedge positions during sharp, fast-moving drawdowns—precisely the scenarios where position-aware sizing is most valuable.

5.3.3 Market Signal Processing (MSP)

Removing MSP leads to moderate performance degradation (e.g., on TSLA: CR drops from 62.2% to 58.3%, and MDD worsens from 42.3% to 45.4%; see Tab. 2), revealing that raw market data contain substantial noise that distracts downstream agents. Without MSP's relevance filtering, an agent may erroneously treat Elon Musk's tweets about Dogecoin as material signals for TSLA trading, despite their lack of fundamental relevance. Such off-topic social media noise is common in financial news feeds and can trigger spurious trading responses. In contrast, MSP leverages domain-guided heuristics and entity-relevance rules to pre-filter inputs, ensuring that only high-signal information—such as regulatory filings, or company-specific announcements—reaches the LLM analysis pipeline.

5.4 Risk Analysis

Beyond numerical results, Fig. 4 provides focused evidence of FinPos's risk-adjusted performance during a high-volatility period (Mar–Apr 2025), chosen to highlight its ability to manage extreme market conditions—including those triggered by major political events such as the U.S. election cycle. This window offers a stringent test for position-awareness, as agents must adapt rapidly to shifting risk regimes without overreacting. The top-left panel reports the calmar ratio (Appendix B.3), which measures return per unit of maximum drawdown. FinPos achieves a calmar of **1.5**, significantly outperforming all baselines, indicating that its profitability is not driven by excessive risk-taking but by effective control. The bottom-left panel plots cumulative return (CR%) against risk-control ability (1-MDD). FinPos occupies the upper-right frontier—high return with low drawdown—while baselines cluster in regions of either low return or poor risk management. On the right, the time-series curves and exposure-risk overlay reveal the operational advantage of position-aware management: during high-volatility events (marked by vertical dashed lines), non-PA agents exhibit sharp exposure spikes followed by severe drawdowns, while FinPos proactively reduces exposure and stabilizes returns. This demonstrates that position awareness enables not only reactive mitigation but also anticipatory adjustment—critical for navigating real-world market turbulence. Additional analyses for AAPL are provided in Appendix D.2.

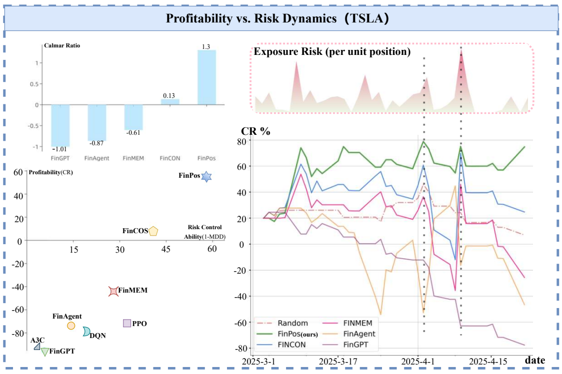

Figure 4: Risk-adjusted performance and exposure dynamics on TSLA during Mar–Apr 2025, highlighting FinPos's advantage under high-volatility events.

6 Conclusion

In this paper, we introduce a position-aware trading task that more closely resembles real market conditions compared to the single-step trading task previously employed by researchers. To address this task, we design FinPos, a novel LLM Trading Agent equipped with position-awareness and risk management capabilities. To meet the demands of position management, an LLM agent must possess heightened market sensitivity, enhanced risk control abilities, and a longer-term perspective. Consequently, FinPos employs a specialized market perception module, a dual-decision agent system, and a multi-timescale reward design to achieve these objectives. Our experiments demonstrate that the position management task imposes higher comprehensive ability requirements on LLM Agents. FinPos explores solutions for trading agents operating in this more realistic market environment.

Limitations

This work is intended for research purposes only. Deploying LLM-based trading systems in real-world financial markets without professional oversight may lead to financial losses, particularly under extreme or unexpected market conditions.

Single-Asset Focus

Although the proposed framework is general, in this paper we adopt a single-asset trading task as the experimental vehicle to validate its effectiveness.Extending the framework to portfolio-based methods and multi-asset allocation settings is an important direction for future work. This will involve addressing cross-asset correlation and rebalancing, their diversified risks, and portfolio-level dynamics, which presents additional complexities that require further model refinement.

Dependency on Prompt Quality:

major limitation of FinPOS lies in its sensitivity to prompt design. In our experiments, we focus on well-established companies with abundant and reliable textual information to ensure stable decision-making. As a result, the system relies on asset-specific prompt configurations that are tailored to the characteristics of individual stocks. This design choice limits the direct applicability of FinPOS to newly listed firms or assets with sparse or highly noisy textual signals, where prompt tuning and data curation become necessary. More broadly, this highlights a structural challenge of LLM-based trading agents: their performance is closely tied to the availability and organization of domain-specific textual inputs, rather than solely to model capacity. Addressing this limitation may ultimately require moving beyond prompt-level engineering. However, model-level adaptation or domain specialization typically demands substantial domain-specific data and computational resources, which we leave to future work.

Reward Design and Multi-Objective Optimization:

In this work, we adopt a lightweight multi-timescale reward design to regulate position risk and trend alignment, prioritizing stability, interpretability, and reproducibility. While reinforcement learning (RL) has shown promise for LLM-based agents, its application in financial markets is challenged by non-stationarity, heavy-tailed risks, and regime shifts, which can lead to overfitting and unstable behavior. Exploring principled RL formulations for learning general financial decision principles remains an important direction for future work.

References

Surjeet Balhara, Nishu Gupta, Ahmed Alkhayyat, Isha Bharti, Rami Q Malik, Sarmad Nozad Mahmood, and Firas Abedi. 2022. A survey on deep reinforcement learning architectures, applications and emerging trends. *IET Communications*.

Peter C Belafsky, Gregory N Postma, and James A Koufman. 2002. Validity and reliability of the reflux symptom index (rsi). *Journal of voice*, 16(2):274–277.

Dominik Benk. 2022. Stock trading using a deep reinforcement learning and text analysis.

Thomas G Fischer. 2018. Reinforcement learning in financial markets-a survey. Technical report, FAU discussion papers in economics.

Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, and 1 others. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Gyeeun Jeong and Ha Young Kim. 2019. Improving financial trading decisions using deep q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Systems with Applications*, 117:125–138.

Qinma Kang, Huizhuo Zhou, and Yunfan Kang. 2018. An asynchronous advantage actor-critic reinforcement learning method for stock selection and portfolio management. In *Proceedings of the 2nd International Conference on Big Data Research*, pages 141–145.

Long Li, Weiwen Xu, Jiayan Guo, Ruochen Zhao, Xingxuan Li, Yuqian Yuan, Boqiang Zhang, Yuming Jiang, Yifei Xin, Ronghao Dang, Deli Zhao, Yu Rong, Tian Feng, and Lidong Bing. 2024. Chain of ideas: Revolutionizing research via novel idea development with llm agents. *Preprint*, arXiv:2410.13185.

Shenjian Li, Mingxuan Yu, and Freddie Dossor. 2025. Option-driven sentiment in finrl: a ppo approach to trading. In *2025 IEEE 11th International Conference on Intelligent Data and Security (IDS)*, pages 62–64. IEEE Computer Society.

Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, and 1 others. 2024. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.