FinRS: A Risk-Sensitive Trading Framework for Real Financial Markets

1st Bijia Liu
Dongbei University of Finance and Economics
Dalian, China
qq1403492787@gmail.com

2nd Ronghao Dang*
Alibaba DAMO Academy
Hangzhou, China
dangronghao.drh@alibaba-inc.com

***Abstract*—Large language models (LLMs) have demonstrated strong reasoning capabilities and are increasingly applied to financial trading. However, existing LLM-based trading agents typically address risk in a reactive manner, relying on post-hoc risk-adjusted metrics, heuristic rules, or prompt-level adjustments, without embedding genuine risk awareness into their decision-making process. As a result, these agents lack an internalized understanding of downside exposure and often behave myopically under market volatility. To address this gap, we propose FinRS, a risk-sensitive trading framework that integrates hierarchical market analysis, a dual-decision agent architecture, and multi-timescale reflection to explicitly align trading actions with both return objectives and downside risk constraints. Extensive experiments demonstrate that FinRS achieves superior returns and improved drawdown stability compared to state-of-the-art baselines, representing a meaningful step toward more robust and risk-aware LLM-driven trading systems.**

***Index Terms*—Trading Agent, Risk-Sensitive, Real Markets**

I. Introduction

In recent years, large language models (LLMs) [1]–[4] have demonstrated significant potential in financial trading [5]. Existing LLM-based trading agents have begun to incorporate risk-related considerations into their design [6]. For instance, FINMEM [7] introduces risk preference instructions within its prompt design; FinAgent [8] evaluates risk only ex post using adjusted performance metrics; and FinCon [9] applies hand-crafted drawdown-control rules. However, these mechanisms remain peripheral add-ons rather than integral parts of the agent's reasoning process. Risk handling is typically decoupled from core decision logic, resulting in strategies that optimize for returns while neglecting systematic downside control. Their limitations stem from three layers of the agent design: risk-control mechanisms, information processing, and decision structure. First, risk control is reactive rather than intrinsic: in most existing agents, risk is not modeled as an explicit decision state but handled through reactive mechanisms (e.g., prompt-level preference switches or post-hoc numerical thresholds), rather than being embedded into the core reasoning process. Because risk is decoupled from core reasoning, agents optimize myopically for returns and exhibit risk-imbalanced behavior under market stress, failing to control exposure or limit drawdowns [10]. Second, information processing lacks depth: most agents follow a "collect-analyze-predict" pipeline, which emphasizes signal aggregation but offers limited structured filtering and causal reasoning. Without domain knowledge and prioritization, agents are easily distracted by noise [11] and miss pivotal risk-relevant events. Third, feedback signals lack multi-horizon risk awareness. Most agents rely solely on next-day price changes for learning, ignoring how today's exposure affects future drawdowns. This short-term feedback loop reinforces myopic behavior and fails to internalize the cumulative nature of risk in sequential trading.

To address the above limitations at the mechanism, information-processing, and decision-structure levels, we propose **FinRS**: a risk-sensitive LLM trading framework for multi-stage decision-making. Its key designs include: (i) **risk-exposure state awareness**, which introduces explicit exposure and volatility state variables into the decision structure, enabling position- and risk-conditioned actions; (ii) **risk-aware reasoning injection**, which embeds risk considerations directly into the LLM's reasoning chain [12], guiding the agent to evaluate uncertainty and downside scenarios alongside immediate profit signals; (iii) **layered information filtering**, which distills risk-relevant cues from financial reports and market news, reducing noise; and (iv) **multi-timescale reflection**, which strengthens the internal risk-control mechanism by combining short-term returns with long-horizon robustness objectives to promote stable behavior across volatile regimes.

II. Related Works

In the financial domain, LLM-based trading agents have evolved from simple prototypes to increasingly complex frameworks. Early work such as FINMEM established a baseline architecture centered around analysis, memory, and decision modules, and introduced a rudimentary notion of risk preference through risk-seeking, risk-averse, and adaptive profiles. Subsequent studies, including FinAgent and FinVision [13], extended this line of research by incorporating multimodal inputs such as numerical, textual, and visual data to better capture market dynamics and historical trading patterns. FINCON and TradingAgents [14] further adopted investment group architectures, embedding multi-agent collaboration and episodic reflection to improve decision stability.

However, existing systems incorporate risk-related components in relatively lightweight forms. For example, FINMEM relies on preference switching, FinAgent reports ex-post risk

*Corresponding author

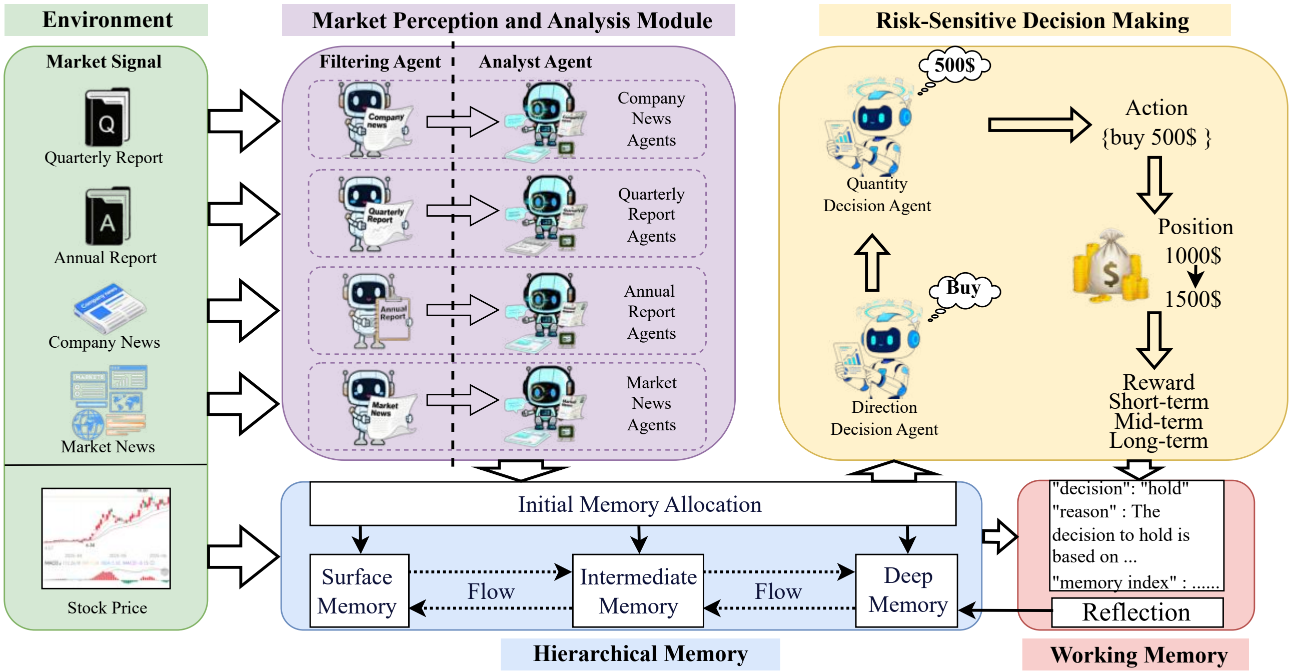

Fig. 1. **Architectural Details of FinRS:** Multi-source market signals are first processed by specialized filtering and analyst agents to extract structured insights. The processed information is then organized through hierarchical memory allocation. A directional agent first determines trade direction, after which a quantity agent computes risk-constrained trade size based on exposure and risk estimates. Multi-timescale reflection then evaluates decisions across different horizons and triggers memory updates, reinforcing consistent risk-aware behavior over time.

metrics, and FINCON introduces a dedicated risk-manager agent that provides heuristic guidance [15]. While these designs introduce elements of risk awareness, risk handling is typically treated as an auxiliary component rather than being tightly integrated into the core reasoning and decision process. This limitation reduces robustness under real-world market volatility. To address this gap, we propose a risk-sensitive framework that embeds explicit risk modeling and control directly into LLM-based trading agents.

III. Architecture of FinRS

We propose FinRS, a risk-sensitive LLM-based trading framework that integrates risk awareness directly into the agent's decision logic rather than treating it as an external constraint. As illustrated in Fig. 1, FinRS adopts a modular architecture with perception, memory, and decision components. Among them, three risk-centric modules play a central role: (1) a market perception and analysis module that extracts risk-relevant signals from prices, news, and reports; (2) a risk-sensitive decision module that injects risk considerations into the LLM's reasoning chain when forming trading actions; and (3) a multi-timescale reflection module that guides the agent to balance short-term returns with longer-term robustness objectives. We next elaborate on these core modules.

A. Market Perception and Analysis Module

Inspired by institutional research workflows, this module (purple in Fig. 1) organizes multiple analyst agents, each specialized in a specific market signal such as news, financial reports, or macro narratives. Incoming raw data are first processed by a screening stage that filters irrelevant or low-impact information and attaches brief rationales for traceability. The retained signals are then forwarded to their corresponding analyst agents, which operate under domain-specific prompt templates to produce structured analytical summaries.

To enhance risk awareness during analysis, we incorporate explicit risk-oriented signals into agent prompts, providing prior guidance on typical downside patterns, volatility signals, and adverse market scenarios. This design steers the LLM to pay attention to risk-relevant evidence rather than relying solely on surface-level sentiment or isolated keywords, thereby mitigating risk-blind interpretations of market information. The resulting analytical outputs are stored in hierarchical memory (blue in Fig. 1), where long-term stable information (e.g., periodic financial reports) is maintained in deeper layers, while transient signals (e.g., breaking news) reside in shallower layers. The memory hierarchy is dynamically updated over time: memories that consistently prove informative for subsequent decisions are consolidated into deeper layers, while less useful memories gradually decay.

B. Risk-Sensitive Decision Module

The risk-sensitive decision module (highlighted in yellow in Fig. 1) is responsible for producing concrete trading actions under dynamic market conditions. Unlike prior LLM-based trading agents that focus primarily on directional prediction,

TABLE I

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	−32.13	−0.91	62.10	−34.21	−0.28	59.40	−52.35	−0.32	61.20	−1.04	−0.02	22.60	−0.22	−0.01	17.60
Rule-Based Methods															
MACD [18]	−49.21	−0.63	74.20	−46.83	−0.57	61.40	−44.92	−0.49	70.30	−7.42	−0.18	27.90	−9.87	−0.21	19.80
RSI [19]	−46.13	−0.59	72.80	−44.28	−0.54	59.70	−42.11	−0.46	68.20	−6.98	−0.16	26.70	−8.65	−0.19	19.30
Reinforcement Learning Methods															
A3C [20]	−92.71	−1.28	96.60	−81.55	−1.14	73.40	−58.24	−0.61	72.80	−5.24	−0.12	23.60	−6.10	−0.09	18.60
DQN [21]	−78.25	−0.77	80.62	−72.48	−0.92	70.30	−52.77	−0.48	68.10	0.78	0.08	22.10	−4.25	−0.06	18.20
PPO [22]	−73.76	−0.56	69.88	−58.33	−0.55	61.40	−43.12	−0.39	64.90	1.72	0.02	21.30	−2.04	−0.04	17.80
LLM-Based Agent															
FinGPT [23]	−95.19	−1.07	94.36	−85.22	−1.52	74.60	−56.87	−0.54	70.30	2.01	0.01	22.30	−3.63	−0.04	17.00
FinAgent [8]	−74.31	−0.89	85.65	−78.00	−1.09	71.50	−49.64	−0.53	61.20	1.12	0.02	21.50	−2.10	−0.03	17.50
FINMEM [7]	−44.03	−0.52	72.10	−45.88	−0.54	47.80	−41.21	−0.36	66.10	0.41	0.05	25.20	0.13	0.01	19.20
FINCON [9]	7.76	0.38	59.13	−16.02	−0.13	29.03	−3.30	−0.02	54.77	7.92	0.63	20.70	7.35	**1.03**	15.10
FinRS	**54.99**	**0.67**	**42.34**	**60.28**	**0.69**	**18.44**	**52.10**	**0.64**	**36.92**	**16.85**	**0.87**	**20.05**	**14.74**	0.92	**14.05**

FinRS incorporates explicit risk-exposure adjustment into the decision process, enabling the agent to reason about both action selection and capital allocation under uncertainty. The module retrieves structured market evidence and historical experience from the hierarchical memory, providing the decision agents with a comprehensive context that reflects current market conditions and accumulated risk-relevant knowledge.

Specifically, FinRS adopts a dual-agent decision structure. The direction agent determines the trading action (buy, sell, or hold) by reasoning over retrieved experience, the recent market dynamics, and the current exposure status. To reduce sensitivity to short-term noise, FinRS integrates 1-day, 7-day, and 30-day price differences as momentum features, the model understands the tension between short-term volatility and medium- to long-term trends, encouraging the agent to balance short-term fluctuations with medium-term trend indications.

Conditioned on the directional decision, the risk-exposure agent adjusts the trade magnitude by considering account exposure, retrieved memory, and the reference signals related to risk. In practice, downside risk is primarily controlled through Conditional Value-at-Risk (CVaR [16]) estimation. We compute CVaR using a rolling 20-day window of realized returns with confidence level α = 5%, and use it as a hard constraint to limit maximum allowable exposure under adverse market conditions. In parallel, a scaled Kelly criterion [17] is used as a soft reference signal, producing a fractional position reference based on the recent empirical return mean, variance, and confidence of the directional agent. The CVaR estimate defines a hard upper bound on the allowable downside risk. Within this risk bound, the agent determines the final trade quantity by jointly considering the retrieved memory, market trends, and the reference signals provided. By combining profit-seeking guidance with explicit downside risk constraints, FinRS enables risk-aware trade magnitude adjustment under volatile market conditions.

C. Multi-Timescale Reflection Module

To promote stable and risk-sensitive trading behavior, FinRS incorporates a multi-timescale reflection mechanism. At each timestep, the agent's action is evaluated using future price movements over multiple horizons, providing feedback that reflects both potential return opportunities and downside risk exposure. This design prevents the agent from overreacting to short-term noise while encouraging awareness of longer-term market.

To embed risk-awareness, we design a trend-aware reward function leveraging multi-scale market momentum signals. Specifically, for each timestep t, we compute three future price trends: $M_t^s = price[t+1] - price[t]$: 1-day trend (short-term), $M_t^m = price[t+7] - price[t]$: 7-day trend (mid-term), $M_t^l = price[t+30] - price[t]$: 30-day trend (long-term). We define the multi-timescale momentum score M_t as:

$$M_t = M_t^s + M_t^m + M_t^l \tag{1}$$

This score represents the aggregated expected trend across multiple timescales (e.g., 1-day, 7-day, and 30-day). The reward at time t, $Reward_t$ is defined as follows:

$$Reward_t = \begin{cases} -(M_t)^2, & position_t = position_{t-1} \\ position_t \times M_t, & otherwise \end{cases} \tag{2}$$

$$position_t = position_{t-1} + d_t \times q_t \tag{3}$$

where d_t and q_t denote the direction and quantity of the trading decision.

This design serves two purposes. First, the agent is encouraged to take positions aligned with consistent trends across multiple timescales, balancing short-term fluctuations against longer-term movements. Second, when the agent keeps its position unchanged, a volatility-sensitive penalty proportional to M_t^2 is applied. This discourages passive exposure during highly volatile periods, where inaction could lead to a certain risk of negative consequences.

IV. EXPERIMENTS

A. Experimental Setup

(i) Datasets. We use real-world financial data from authoritative sources: (1) Stock prices, obtained from Yahoo Finance, include daily open-high-low-close-volume (OHLCV) data; (2)

company and macro news from Finnhub; (3) 10-Q (quarterly reports) and 10-K (annual reports), [24] from the SEC EDGAR API. All data are processed into daily time series.

(ii) Baselines. We compare FinRS with four LLM-based trading agents (FinGPT [23], FinMem [7], FinAgent [8], FinCon [9]), three deep reinforcement learning (DRL) [25] methods (A2C, PPO, DQN), two rule-based strategies (MACD, RSI), and market baselines (Random). All compared models are evaluated using identical data splits and executed under the same market conditions.

(iii) Metrics. We report Cumulative Return [26] (CR%), Sharpe Ratio [27] (SR), and Maximum Drawdown (MDD%), which respectively reflect profitability, risk-adjusted efficiency, and downside risk. CR reflects overall profitability, while SR measures the risk-adjusted return by showing how much excess return it generates per unit of total risk. MDD quantifies the worst-case loss over the trading horizon, which in the position-aware trading task directly reflects exposure and vulnerability under held positions.

(iv) Implementation. All LLM agents use GPT-4o [3], strictly follow the inference settings reported in their original papers. DRL baselines are implemented using the FinRL framework [28]. Training covered Jan 2024-Feb 2025; testing spanned Mar 2025-Apr 2025, covering the U.S. election period and other major macroeconomic events.

B. Main Results

We evaluate FinRS on five representative stocks (TSLA, AAPL, AMZN, NFLX, COIN), covering diverse sectors and volatility levels. As shown in Table I, FinRS consistently achieves competitive cumulative returns while exhibiting markedly improved drawdown control and risk-adjusted performance compared to baseline methods. In particular, FinRS demonstrates strong robustness on stocks that experienced high volatility during the test period, such as TSLA, AAPL, and AMZN. Other LLM agents and DRL methods occasionally suffered large losses, generated negative returns, or failed to outperform simpler benchmark strategies during turbulent market periods. In contrast, FinRS consistently maintained cumulative returns above 50% while effectively controlling drawdowns. For relatively stable stocks like NFLX and COIN, FinRS achieved returns more than twice those of competing approaches. These results demonstrate that FinRS's risk-sensitive design enhances profitability while maintaining stable trading behavior across diverse market conditions, showing resilience to volatility and uncertainty.

C. Ablation Studies

We evaluated the contribution of each FinRS component through ablation experiments on three representative assets (TSLA, AAPL, AMZN). The results in Table II show clear degradation once any module is removed.

- **Risk-Sensitive Reasoning (RS):** A reasoning mechanism that embeds risk awareness into the decision process by modeling uncertainty, volatility, and potential downside exposure.
- **Financial Insight Prompting (FIP):** A targeted prompting strategy designed to enhance LLMs' financial reasoning capability. It emphasizes causal chains, market trends, and probabilistic inference, guiding the agent toward more structured and economically grounded analysis.
- **Market News Integration (MN):** Aggregates and incorporates macroeconomic and geopolitical information into the agent's reasoning process, allowing decisions to reflect broader market context.
- **Multi-Timescale Reflection (MTR):** A self-reflection mechanism that adds a lightweight self-assessment after each decision.

We next analyze each module in detail from three aspects: performance impact, risk-control behavior, and decision-pattern changes, to understand how different components contribute to robustness and profitability.

1) Risk-Sensitive Reasoning (RS): As the core mechanism that injects exposure and volatility awareness into the decision process, RS enables the agents to balance expected returns against downside risk. The ablation results in Tab. II show that removing RS leads to the most severe performance degradation across all components. On TSLA, cumulative return drops sharply from 54.99% to 39.19%, while maximum drawdown deteriorates from 42.34% to 82.89%. Similar declines are observed on AAPL and AMZN, where CR decreases by more than 14-15 points once RS is removed. These results indicate that without explicit risk-sensitive reasoning, the agent tends to take overly aggressive actions under volatile conditions, resulting in amplified drawdowns and unstable returns. Overall, RS plays a central role in aligning trading decisions with risk-aware objectives, making it essential for both profitability and drawdown control.

2) Financial Insight Prompting (FIP): Financial Insight Prompting primarily affects the quality and structure of the agent's reasoning process. As shown in Tab. II, the introduction of FIP consistently improves both the return stability and the drawdown control across assets. In particular, TSLA's CR decreases from 54.99% to 41.57%, and SR falls from 0.67 to 0.49. AAPL records a 11.7-point drop in CR, indicating that risk-aware prompting plays a critical role in stabilizing trading behavior under volatility. To investigate the impact of prompt design on the depth of agent finance insight, we divide the content of prompts into eight key dimensions: Role Setting, Task Breakdown, Depth Analysis, Multi-Dimensional Analysis, Logical Reasoning, Contextual Consideration, actionability, and Clarity & Conciseness. For each dimension, we design prompts with different levels of refinement-Low, Medium, and High-and conduct graded experiments. The results are shown in Fig. 2. The results indicate that Task Breakdown, Depth Analysis, and Multi-dimensional Analysis are the most influential dimensions for overall agent performance. Notably, Depth Analysis plays a particularly crucial role in financial reasoning: since the quality of financial analysis is significantly influenced by deep insights, particularly those concerning causal chains and risk-reward tradeoffs. In contrast, Multi-Dimensional Analysis presents

TABLE II
ABLATION RESULTS ON TSLA, AAPL AND AMZN

RS	FIP	MN	MTR	TSLA CR%↑	TSLA SR↑	TSLA MDD%↓	AAPL CR%↑	AAPL SR↑	AAPL MDD%↓	AMZN CR%↑	AMZN SR↑	AMZN MDD%↓
	✓	✓	✓	39.19	0.45	82.89	45.96	0.49	36.10	36.98	0.43	73.71
✓		✓	✓	41.57	0.49	48.78	48.57	0.54	21.24	39.38	0.50	41.76
✓	✓		✓	45.41	0.53	55.40	52.78	0.60	24.13	43.67	0.55	49.53
✓	✓	✓		48.68	0.59	59.07	57.36	0.66	25.73	47.24	0.60	51.50
✓	✓	✓	✓	**54.99**	**0.67**	**42.34**	**60.28**	**0.69**	**18.44**	**52.10**	**0.64**	**36.92**

a more nuanced pattern. Financial analysis naturally spans across multiple perspectives, such as macro vs. micro and short-term vs. long-term. While moderate prompting improves the model's ability to cover multiple factors, overly elaborate prompts may overwhelm the LLM, causing it to "overthink," lose focus, and eventually degrade performance-an effect we refer to as the information burden. Among the eight dimensions, Contextual Consideration shows the most pronounced performance shift. Without explicit prompting, the model tends to overlook this dimension-for instance, it may treat tariff adjustments as entirely unrelated to stock prices. Once guided, however, the agent develops stronger contextual reasoning: it can link major events (e.g., political shifts, regulatory changes) to firm-level responses (e.g., strategic adjustments to tariffs), and further to market outcomes. This ability allows the model to generate insights that align more closely with the dynamics of the real-world market. Overall, FIP aims to activate risk-relevant prior knowledge and suppress financial hallucinations, thereby improving robustness in noisy market environments.

3) Market News Integration (MN): This module mainly contributes contextual risk signals beyond firm-level information. Removing MN leads to consistent performance degradation across all evaluated assets. On TSLA, cumulative return decreases to 45.41%, while the maximum drawdown increases from 42.34% to 55.40%. Similar declines in CR are observed in AAPL (-7.5) and AMZN (-8.4).

Beyond numerical degradation, we observe systematic changes in the agent's decision behavior when macro-level signals are excluded. Without MN, the agent tends to rely primarily on firm-level fundamentals and recent price trends, often maintaining optimistic positions despite rising systemic risk. By contrast, incorporating macroeconomic and geopolitical signals allows the agent to reassess firm-level optimism under broader market stress, leading to more conservative actions such as exposure reduction or position reversal. These results indicate that MN improves robustness under volatile and event-driven market conditions.

4) Multi-Timescale Reflection (MTR): By incorporating feedback across multiple horizons, Multi-Timescale Reflection improves temporal consistency in decision-making. Removing MTR consistently degrades both return and risk-control performance. In TSLA, CR decreases from 54.99% to 48.68%, while MDD increases from 42.34% to 59.07%. AAPL and AMZN exhibit similar deterioration, with higher drawdowns

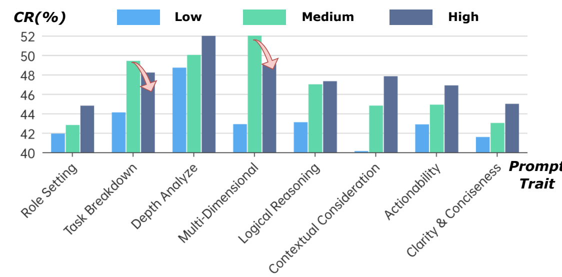

Fig. 2. Prompt ablation across eight characteristics and three emphasis levels.

and reduced cumulative returns. These results indicate that incorporating feedback from multiple temporal horizons helps stabilize decisions and improves drawdown management under long-horizon market fluctuations.

In general, the full FinRS model achieves the best CR, SR, and MDD across all assets. The ablation results confirm that each module contributes complementary benefits, with RS being the most critical driver of robustness and profitability.

V. CONCLUSION

We propose FinRS, a risk-sensitive LLM-based trading framework designed for multi-stage sequential decision-making in realistic financial markets. FinRS explicitly addresses the coordination between trading actions, current positions, and evolving market conditions. To this end, FinRS integrates rolling-window risk estimation and multi-timescale market signals with hierarchical information processing and a dual-decision agent architecture, enabling the agent to dynamically adjust trading actions while accounting for portfolio exposure and potential downside risk. Experimental results demonstrate that incorporating risk-aware reasoning substantially improves both profitability and robustness. FinRS not only enhances predictive performance, but also effectively reduces drawdowns and maintains stable behavior under high market volatility. Our findings highlight the value of embedding risk-aware reasoning in LLM-based trading agents, taking an important step toward improving robustness and stability in trading decisions.

REFERENCES

[1] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al., "Chatgpt for good? on opportunities and challenges of large language models for education," *Learning and individual differences*, vol. 103, pp. 102274, 2023.

[2] Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, et al., "Deepseek-v3 technical report," *arXiv preprint arXiv:2412.19437*, 2024.

[3] Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al., "Gpt-4o system card," *arXiv preprint arXiv:2410.21276*, 2024.

[4] An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al., "Qwen3 technical report," *arXiv preprint arXiv:2505.09388*, 2025.

[5] Lawrence Revsine, Daniel W Collins, and W Bruce Johnson, *Financial reporting & analysis*, McGraw-Hill, 2021.

[6] Bijia Liu and Ronghao Dang, "Finpos: A position-aware trading agent system for real financial markets," *arXiv preprint arXiv:2510.27251*, 2025.

[7] Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W Suchow, and Khaldoun Khashanah, "Finmem: A performance-enhanced llm trading agent with layered memory and character design," in *Proceedings of the AAAI Symposium Series*, 2024, vol. 3, pp. 595–597.

[8] Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, et al., "A multimodal foundation agent for financial trading: Tool-augmented, diversified, and generalist," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024, pp. 4314–4325.

[9] Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yuechen Jiang, Yupeng Cao, Zhi Chen, Jordan Suchow, Zhenyu Cui, Rong Liu, et al., "Fincon: A synthesized llm multi-agent system with conceptual verbal reinforcement for enhanced financial decision making," *Advances in Neural Information Processing Systems*, vol. 37, pp. 137010–137045, 2024.

[10] Malik Magdon-Ismail and Amir F Atiya, "Maximum drawdown," *Risk Magazine*, vol. 17, no. 10, pp. 99–102, 2004.

[11] Fischer Black, "Noise," *The journal of finance*, vol. 41, no. 3, pp. 528–543, 1986.

[12] Zixuan Ke, Fangkai Jiao, Yifei Ming, Xuan-Phi Nguyen, Austin Xu, Do Xuan Long, Minzhi Li, Chengwei Qin, Peifeng Wang, Silvio Savarese, et al., "A survey of frontiers in llm reasoning: Inference scaling, learning to reason, and agentic systems," *arXiv preprint arXiv:2504.09037*, 2025.

[13] Sorouralsadat Fatemi and Yuheng Hu, "Finvision: A multi-agent framework for stock market prediction," in *Proceedings of the 5th ACM International Conference on AI in Finance*, 2024, pp. 582–590.

[14] Yijia Xiao, Edward Sun, Di Luo, and Wei Wang, "Tradingagents: Multi-agents llm financial trading framework," *arXiv preprint arXiv:2412.20138*, 2024.

[15] Eve A Kerr, Brian S Mittman, Ron D Hays, Albert L Siu, Barbara Leake, and Robert H Brook, "Managed care and capitation in california: how do physicians at financial risk control their own utilization?," *Annals of Internal Medicine*, vol. 123, no. 7, pp. 500–504, 1995.

[16] R Tyrrell Rockafellar, Stanislav Uryasev, et al., "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.

[17] Edward O Thorp, "Portfolio choice and the kelly criterion," in *Stochastic optimization models in finance*, pp. 599–619. Elsevier, 1975.

[18] Jian Wang and Junseok Kim, "Predicting stock price trend using macd optimized by historical volatility," *Mathematical Problems in Engineering*, vol. 2018, no. 1, pp. 9280590, 2018.

[19] J Welles Wilder, *New concepts in technical trading systems*, Greensboro, NC, 1978.

[20] Qinma Kang, Huizhuo Zhou, and Yunfan Kang, "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*, 2018, pp. 141–145.

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

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

[23] Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang, "Fingpt: Open-source financial large language models," *CoRR*, vol. abs/2306.06031, 2023.

[24] Adedoyin Tolulope Oyewole, Omotayo Bukola Adeoye, Wilhelmina Afua Addy, Chinwe Chinazo Okoye, Onyeka Chrisanctus Ofodile, and Chinonye Esther Ugochukwu, "Automating financial reporting with natural language processing: A review and case analysis," *World Journal of Advanced Research and Reviews*, vol. 21, no. 3, pp. 575–589, 2024.

[25] Adrian Millea, "Deep reinforcement learning for trading—a critical survey," *Data*, vol. 6, no. 11, pp. 119, 2021.

[26] Robert A Ariel, "A monthly effect in stock returns," *Journal of financial economics*, vol. 18, no. 1, pp. 161–174, 1987.

[27] William F Sharpe, "The sharpe ratio," *Streetwise–the Best of the Journal of Portfolio Management*, vol. 3, no. 3, pp. 169–85, 1998.

[28] Xiao-Yang Liu, Hongyang Yang, Qian Chen, Runjia Zhang, Liuqing Yang, Bowen Xiao, and Christina Dan Wang, "Finrl: A deep reinforcement learning library for automated stock trading in quantitative finance," *arXiv preprint arXiv:2011.09607*, 2020.