

FinPos: A Position-Aware Trading Agent System for Real Financial Markets

Bijia Liu

Dongbei University of Finance and Economics
qq1403492787@gmail.com

Ronghao Dang*

Alibaba DAMO Academy
dangronghao.drh@alibaba-inc.com

Abstract

The exceptional potential of large language models (LLMs) in handling text information has garnered significant attention in the field of financial trading. However, most existing trading agents operate under intraday, independent unit-based trading tasks, where decisions are made as isolated directional actions, and thus lack awareness of continuous position management. Therefore, we propose a position-aware trading task designed to simulate a more realistic market. To address this task, we propose FinPos, a position-aware trading agent system designed to explicitly model and manage continuous positions. FinPos enhances position awareness through three key mechanisms: (1) professional-level interpretation of heterogeneous market information; (2) a dual-agent decision structure that separates directional reasoning from risk-aware position adjustment; and (3) multi-timescale reward signals, allowing the agent to internalize position awareness through experiential feedback rather than static instructions alone. Extensive experiments demonstrate that FinPos surpasses state-of-the-art trading agents in the position-aware trading task, which closely mirrors real market conditions. More importantly, our findings reveal that LLM-centered agent systems exhibit a vast, largely unexplored potential in long-term market decision-making.

1 Introduction

An explosion of market information, facilitated by evolving communication technologies, highlights the growing significance of automated trading systems. In the past decade, deep reinforcement learning (DRL) agents (Fischer, 2018; Shakya et al., 2023) gained considerable attention due to their ability to handle large financial datasets. However, their inherent difficulty in integrating unstructured text information (Benk, 2022) limits their applicability primarily to technical analysis and compromises their interpretability (Balhara et al., 2022).

Recently, large language models (LLMs) (Hurst et al., 2024; Liu et al., 2024) have demonstrated remarkable potential in handling text-modal information and solving complex tasks, such as reasoning and decision-making (Zhai et al., 2025). This suggests that LLM-centered agent systems have the potential to transcend the inherent capability boundaries of traditional DRL trading models. Specialized LLM agents designed for financial trading, like TradingAgents (Xiao et al., 2024), FINMEM (Yu et al., 2024a), and FinAgent (Zhang et al., 2024), have already achieved groundbreaking results. However, as shown in the left panel of Fig. 1, most existing LLM-based trading agents are formulated under *single-step trading tasks*: each day is treated as an isolated decision episode, with discrete actions and fixed unit sizes. Although some studies mention "position" in prompts, the *single-step trading tasks* fundamentally undermines continuous position management. In contrast, real-world trading is inherently multi-step and stateful: a trader's current holdings directly shape future risk, opportunity, and decision logic. To bridge this gap, we propose a **position-aware trading task**, which more closely aligns with real-world trading markets than previous setups.

As shown in the right panel of Fig. 1, existing agent architectures struggle to sustain high yield rates when tasked with position management. This is mainly because they lack three essential capabilities: (1) **Explicit position representation and exposure control**: Without a continuous position state, existing agents cannot perform risk-aware, dynamic exposure control. (2) **Long-term planning capability**: Trading tasks that are re-initialized daily are prone to developing agents with a myopic trading perspective. (3) **In-depth market analysis ability**: Short-term trends are often identifiable through market sentiment, yet a much deeper analysis is essential for understanding long-term growth potential.

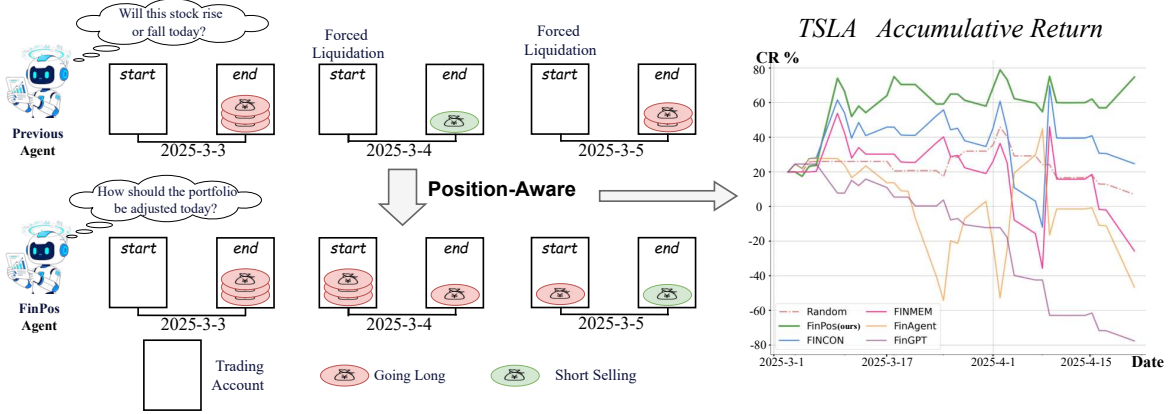


Figure 1: With the introduction of position awareness, the agent must not only predict current market trends but also manage the remaining positions in the account. Agents developed for tasks without position awareness are inadequate for addressing the new challenges posed by position-aware trading tasks.

To address these challenges, we propose **FinPos**, a position-aware trading agent designed to manage portfolio positions in realistic financial markets. FinPos integrates a dual-agent decision architecture with multi-timescale reward mechanism, where a Direction Decision Agent selects discrete actions (buy/sell/hold) and a Quantity and Risk Decision Agent determines transaction volumes.

A defining challenge in **long-term planning capability** is that the risk and return consequences of an exposure decision are inherently delayed and cumulative—they unfold over multiple trading days rather than within a single step. To explicitly encode this temporal dependency into the agent’s reasoning, FinPos adopts a multi-timescale reward market signal during training, which jointly considers daily immediate profit-and-loss (PnL) and medium- to long-term cumulative returns. This guides the Direction Agent to make non-myopic action choices that balance short-term responsiveness with strategic consistency.

To address **explicit position representation and exposure control**, the multi-timescale reward framework also feeds long-horizon performance signals into the reflection process of the Quantity and Risk Decision Agent. This enables the agent to internalize the concept of continuous position and understand how its sizing decisions impact portfolio dynamics. Consequently, the Quantity and Risk Decision Agent dynamically determines the specific transaction volume to buy or sell—subject to a CVaR-based upper bound on single-transaction exposure to ensure capital safety.

By observing the reasoning chain of trading agents, we observe that current trading agents of-

ten analyze market information superficially and lack causal reasoning abilities. For instance, when trading Tesla stock, the agent may overestimate the relevance of news such as "Musk invests in company A" while dismissing significant news like "major changes in U.S. tariff policy" as irrelevant to trading decisions. This issue arises because the information analysis agent lacks fundamental common sense about the specific forms of information. To enhance **in-depth market analysis ability**, we build upon the multi-agent information processing paradigm, go further by injecting domain-specific financial knowledge and causal principles through tailored prompting. This enables FinPos to interpret market signals not just by sentiment, but by fundamental relevance and risk implications.

We conduct extensive experiments using data from multiple real stocks to demonstrate the effectiveness of the FinPos agent in authentic market environments. FinPos outperforms several state-of-the-art (SOTA) financial agents and is the first agent equipped with investment position management capabilities. Moreover, a series of ablation studies underscored the critical importance of risk management, long-term planning, and in-depth market analysis for the agent’s ability to manage investment positions effectively.

2 Related Works

2.1 LLM-Based AI Agents

The evolution of LLMs capabilities has led to a growing interest in LLM-based agents. Recently, the LLM agents have expanded into various domains, including music (Yu et al., 2023), health-

care (Wang et al., 2025), and research (Li et al., 2024). These applications illustrate the vast potential of LLM agents in multiple facets of human society. We aim to bring this potential to the financial markets, making LLM agents key players in investment activities.

2.2 Financial LLM-Based Agents

The evolution of financial agent architectures has exhibited significant progress through successive innovations and expansions in cognitive modeling and system integration. FINMEM (Yu et al., 2024a) was the first to establish a fundamental trading agent architecture centered around analysis, memory, and decision modules. Building upon this foundation, FinAgent (Zhang et al., 2024) integrate multimodal input to enrich their information sources. FINCON (Yu et al., 2024b) and TradingAgents (Xiao et al., 2024) introduced the concept of investment group architectures into agent systems to enhance decision-making reliability. Fundamentally, these agents primarily address a simplified financial game focused on predicting next-day price movement. In real markets, however, agents must not only needs to determine the investment direction, but also decide when to adjust exposure and how much to allocate. Therefore, this paper adopts a position-aware trading task that more closely approximates real market environments.

3 Task Definition

3.1 Single-Step Trading Task

We define a baseline environment as a *Single-Step Trading Task*, where the agent’s position is automatically liquidated at each timestep, and no position is carried over. Within this framework, the agent makes a trading decision $a_t \in \{-1, 0, 1\}$ at each timestep t , corresponding to selling, holding, or buying, respectively. Each decision assumes immediate liquidation at the next timestep, thereby making each action independent, with no continuation of position states. The cumulative return R is the sum of the logarithmic returns at each step:

$$R = \sum_{t=1}^N r_t, \quad r_t = a_t \times \log \frac{\text{price}[t+1]}{\text{price}[t]} \quad (1)$$

This setup is widely adopted in studies such as FINMEM (Yu et al., 2024a) due to its simplification of trading processes, which stabilizes the

training environment and allows more straightforward model evaluation. However, this simplification comes with significant limitations: it completely ignores the continuity of holding positions and the gains and losses arising from holding positions in real trading, thus cannot capture the associated compound returns or the risk exposure carried across periods. Without modeling positions, the agent has no basis for risk control or for detecting market trends, both of which are central to real trading. More critically, by reducing trading to short-term, step-by-step predictions, such a setup underutilizes the natural strengths of LLMs in reasoning over longer horizons. Consequently, advancing financial trading agents towards practical application requires a task more closely mirrors real market conditions.

3.2 Position-Aware Trading Task

Compared to the aforementioned simplified setup, this paper proposes a trading task configuration that aligns more closely with real market behavior, *Position-Aware Trading Task*. In this environment, the agent makes a trading decision on each trading day (each timestep t based on the closing price) and explicitly maintains its position state. The agent’s decisions no longer involve automatic liquidation; rather, it must consider the continuation and adjustment of current positions in future timesteps. Returns R are calculated from the cumulative logarithmic returns based on position:

$$R = \sum_{t=1}^N r_t, \quad r_t = \text{pos}_t \times \log \frac{\text{price}[t+1]}{\text{price}[t]} \quad (2)$$

Where pos_t represents the agent’s current position state, dynamically updated based on past trading decisions. This setup not only more accurately models the continuity and trend dynamics of holding positions in trading but also allows the agent to naturally incorporate key capabilities in strategy learning, such as position control, risk exposure management, and entry-exit timing. The *Position-Aware Trading Task* provides the agent with a more challenging and practically meaningful learning objective, offering stronger potential for strategy generalization and real-world applicability.

Additionally, this task better aligns with the capability characteristics of trading LLM agents. Compared to traditional quantitative models, LLMs excel at extracting long-term trends and underlying

causal structures from complex semantic information rather than performing high-frequency precise numerical optimization. This capability gives them greater potential in mid-to-long-term strategy formulation. Under the *Single-Step Trading Task* setup, the agent primarily focuses on the price direction of the next step, which somewhat forces the LLM struggling to build coherent strategies or cognition. By introducing position continuity and strategy coherence, the *Position-Aware Trading Task* provides LLM agents with a more fitting testing platform, fully unlocking their potential in long-term decision-making and complex information integration. Furthermore, this setup enables agents to construct an internally consistent position management logic centered around fundamentals, market sentiment, and macro semantic information. The combination of long-term and semantic-driven strategies is precisely where LLMs hold their advantages in financial trading tasks.

4 Architecture of FinPos

The architecture of FinPos (Fig. 2) consists of three core modules. The **market signal processing module** adopts a two-level agent hierarchy, consisting of filtering agents and analyst agents, to distill raw market data into structured, decision-relevant signals. The **trading decision module** employs a dual decision framework that explicitly decouples quantity decision-making from directional determination. The **multi-timescale reward module** adopts a multi-timescale reward feedback mechanism to experientially internalize the long-term risk and return implications of position changes. The details of these three modules are presented in the subsequent sections, while all agent prompts are provided in the Appendix A.

4.1 Market Signal Processing and Analysis

FinPos draws inspiration from institutional investment workflows. It assigns specialized agents to distinct information domains forming a division of labor similar to that in private equity firms. For each domain, we further establish a two-tier structure: Signal Processing Agents and Analysis Agents.

Signal Processing Agents: These agents are responsible for preprocessing and filtering high-noise, heterogeneous market data streams. Through domain-specific prompts and heuristic rules, the agents clean, compress, and prioritize information by relevance and importance. For example,

large volumes of low-value or weakly related news are downweighted or discarded, while impactful macroeconomic policies or firm-level events are highlighted and passed on to subsequent modules.

Analysis Agents: These agents analysis the filtered data. We observe that large language models often exhibit financial hallucinations in market scenarios, making it difficult for them to accurately capture the key factors influencing stock prices. They tend to rely solely on explicit keywords in news while overlooking the underlying market logic. For example, an LLM may fail to recognize that news about the S&P 500 is strongly correlated with the performance of leading U.S. stocks. To mitigate this issue, we explicitly inject financial knowledge into the prompts, thereby strengthening their causal awareness. As financial reasoning is gradually integrated into the prompts, the agents evolve from mechanical summarizers into analysts with genuine financial reasoning ability, capable of generating deeper insights and producing more accurate interpretations of market dynamics.

Subsequently, as illustrated in Fig. 2, all analytical results are aggregated into a Hierarchical Memory Module. In this module, important long-term information (e.g., annual reports) is allocated to deep memory, while volatile short-term information (e.g., corporate news) is stored in shallow memory. The hierarchy of memory is not static but dynamically adjusted through post-decision reflection: memories that repeatedly prove their validity are gradually migrated into deeper layers, thereby increasing their weight in future decision-making.

4.2 Dual Trading Decision

In real-world trading, an account’s current holdings directly reflect its risk exposure. Professional traders adjust their buy–sell decisions dynamically based on existing positions to maintain a relative balance between risk and return. FinPos adopts a dual-agent framework that embeds position awareness into both **the Direction Decision Agent** and **the Quantity and Risk Decision Agent**. By integrating multi-source information retrieved from the memory layer, the dual-agent framework enables a decision-making mechanism that more closely mirrors practical trading logic.

4.2.1 The Direction Decision Agent

Incorporating position awareness requires integrating longer-term information. The direction decision agent combines structured memory (news, fi-

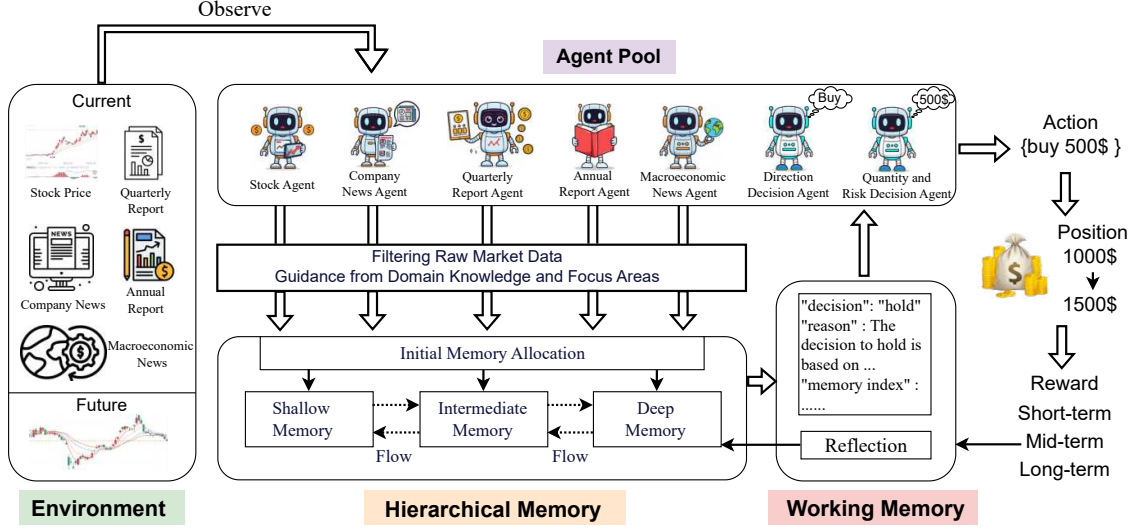


Figure 2: **Architectural Details of FinPos:** Initially, multiple analysis agents leverage domain knowledge to gather diverse information from the environment, subsequently storing it in the memory module. The memory module utilizes a memory allocator to distribute the acquired information across memory layers of varying depths. Subsequently, the most pertinent information for the current decision is placed into working memory, where dual decision agents generate trading actions, while multi-timescale reward signals support reflective updates that consolidate experiential knowledge into deeper memory layers.

financial reports, and past actions) with the current portfolio state and explicitly specifies the strategic intent of each action—whether it represents a long-term position-building move or a short-term tactical adjustment to exploit local trends. This explicit articulation not only enhances interpretability for human supervisors but also provides downstream quantity agents with a clear strategic context. To prevent the agent from over-relying on short-term fluctuations, we design prompts that encourage it to factor in longer horizons, thereby gradually cultivating its long-term planning capability.

4.2.2 Quantity and Risk Decision Agent

FinPos introduces the Quantity and Risk Decision Agent, which determines the specific trade size after a directional decision has been made. This agent determines trade sizes through an explicit risk-aware mechanism that jointly incorporates the current position state, structured memory, and CVaR-based risk references (Conditional Value at Risk (Rockafellar et al., 2000); see Appendix B.2), with per-step transaction sizes capped by the 95% CVaR to control exposure under volatility. This design ensures that every trade size is grounded in the agent’s current position and risk tolerance, making exposure control an explicit, first-class component of FinPos’s decision process.

4.3 Multi-Timescale Reward Design

To guide the agent toward non-myopic trading behavior, we design a trend-aware reward function that leverages *multi-timescale market signals*. Specifically, we calculate three future price trends in each timestep t . $M_t^s = price[t + 1] - price[t]$: 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 score M_t as:

$$M_t = M_t^s + M_t^m + M_t^l \quad (3)$$

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

$$Reward_t = \begin{cases} -(M_t)^2, & pos_t = pos_{t-1} \\ pos_t \times M_t, & otherwise \end{cases} \quad (4)$$

$$pos_t = pos_{t-1} + d_t \times q_t \quad (5)$$

d_t and q_t denote the direction and quantity of the purchasing decision at the current time step, respectively. Crucially, this reward is used only during training-time simulation, no future-dependent signals are accessed at test time.

It serves two complementary roles in guiding the dual-agent architecture: **(1) Guiding the Direction Agent:** The term $pos_t \times M_t$ provides positive

reinforcement when the agent’s directional decision (buy/sell/hold) aligns with the multi-horizon trend M_t . This encourages the Direction Agent to move beyond short-term noise and base decisions on longer-term market signals. **(2) Guiding the Quantity Agent:** During LLM reflection, the Quantity Agent uses this reward to learn when to adjust exposure. Through this process, the agent internalizes the concept of position as a dynamic state. In our experiments, we find that during periods of high volatility, agents tend to maintain an inactive state. To address this, we introduce a quadratic penalty on $|M_t|$, whenever the position remains unchanged. This discourages missed opportunities in volatile markets while also preventing excessive trading in stable periods.

5 Experiments

5.1 Experimental Setup

5.1.1 Datasets

Our research utilizes actual financial data sourced from publicly authoritative providers and encompasses various forms of market information. (1) Stock prices, obtained from Yahoo Finance, include daily open-high-low-close-volume (OHLCV) data. (2) Company news, retrieved using the Finnhub stock API, includes articles’ related company name, publication date, headline, and summary text, processed for sentiment and semantic relevance. (3) Macroeconomic news, also retrieved via the Finnhub stock API, contains publication dates, headlines, and summary texts for each article. (4) 10-Q (quarterly reports) and 10-K (annual reports), accessed through the U.S. Securities and Exchange Commission (SEC) EDGAR API. These documents provide financial information that publicly listed companies are required to disclose, and standardized into a daily time series format.

5.1.2 Comparative Methods

To assess the effectiveness of FinPos, we conduct comparative experiments with four LLM-based trading agents: FinGPT (Yang et al., 2023), FINMEM (Yu et al., 2024a), FinAgent (Zhang et al., 2024), and FINCON (Yu et al., 2024b); three deep reinforcement learning (DRL) methods (A2C, PPO, DQN); two rule-based methods (MACD, RSI); and two market baselines (Random). All compared models are evaluated using identical data splits and executed under the same market conditions. Additional details are provided in Appendix C.

5.1.3 Evaluation Metrics

We evaluate performance using cumulative return (CR%), sharpe ratio (SR), and maximum draw-down (MDD%). 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. CR has been introduced earlier (see Eq. (2)), and the formulas for SR and MDD are provided in Appendix B.1.

5.1.4 Implementation Details

All LLM trading agents are deployed using GPT-4o (Hurst et al., 2024) with a temperature setting of 0.7. All models are evaluated under the position-aware trading task formally defined in Sec. 3.2, ensuring profits and risks are calculated with explicit holding dynamics. Training covered Jan 2024–Feb 2025; testing spanned Mar–Sep 2025, covering the U.S. election period and other major macroeconomic events, making it a representative and challenging evaluation setting.

5.2 Main Results

We evaluated FinPos on five representative stocks: TSLA, AAPL, AMZN, NFLX, and COIN. As shown in Tab. 1, FinPos consistently outperforms all baseline methods, achieving higher CRs, SRs, and lower MDDs across all assets. On high-volatility stocks such as TSLA, AAPL, and COIN, FinPos exhibits strong robustness under turbulent market conditions. While DRL baselines (A3C, DQN, PPO) and LLM-based agents (FinGPT, FinAgent) suffer substantial losses and severe draw-downs (e.g., A3C loses over 80% on TSLA with an MDD of 96.6%), FinPos consistently maintains positive returns, achieving CRs of 62.15% and 36.31% on TSLA and AAPL, respectively, with well-controlled downside risk. For assets with clearer trends, such as AMZN and NFLX, FinPos continues to deliver stable and risk-adjusted performance. FinPos achieves a CR of 30.35% on AMZN with an MDD of 18.44%, and a CR of 28.65% on NFLX with a high SR of 1.02, outperforming all baseline agents in both return and risk control.

5.3 Ablation Studies

We select TSLA, AAPL, and AMZN as test assets due to their rich coverage of news articles, earnings events, and macroeconomic sensitivities. We

Table 1: Performance comparison of different models on five stocks

Models	TSLA			AAPL			AMZN			NFLX			COIN		
	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow
<i>Market Baseline</i>															
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
<i>Rule-Based Methods</i>															
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
<i>Reinforcement Learning Methods</i>															
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
<i>LLM-Based Agent</i>															
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% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow
	✓	✓	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 pre-processing 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

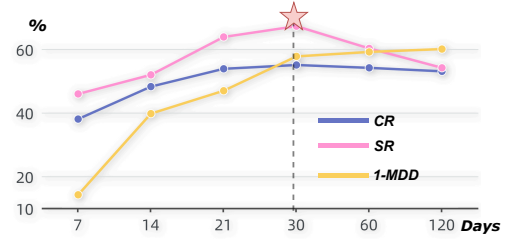


Figure 3: Impact of varying the maximum timescale of the multi-timescale reward on performance metrics.

the full model), indicating that MTR is the architectural backbone that enables long-horizon reasoning and strategic consistency in FinPos.

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

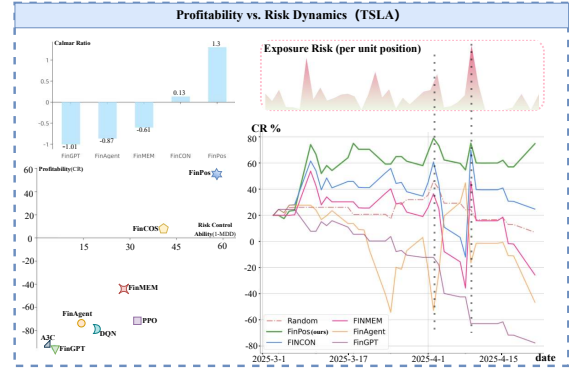


Figure 4: Risk-adjusted performance and exposure dynamics on TSLA during Mar–Apr 2025, highlighting FinPos’s advantage under high-volatility events.

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.

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*.
- Gyeun 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*.

- 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. *arXiv preprint arXiv:2011.09607*.
- R Tyrrell Rockafellar, Stanislav Uryasev, and 1 others. 2000. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42.
- Ashish Kumar Shakya, Gopinatha Pillai, and Sohom Chakrabarty. 2023. Reinforcement learning algorithms: A brief survey. *Expert Systems with Applications*, 231:120495.
- Jian Wang and Junseok Kim. 2018. Predicting stock price trend using macd optimized by historical volatility. *Mathematical Problems in Engineering*, 2018(1):9280590.
- Ziyue Wang, Junde Wu, Chang Han Low, and Yueming Jin. 2025. Medagent-pro: Towards multi-modal evidence-based medical diagnosis via reasoning agentic workflow. *arXiv preprint arXiv:2503.18968*.
- Yijia Xiao, Edward Sun, Di Luo, and Wei Wang. 2024. Tradingagents: Multi-agents llm financial trading framework. *arXiv preprint arXiv:2412.20138*.
- Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang. 2023. [Fingpt: Open-source financial large language models](#). *CoRR*, abs/2306.06031.
- Dingyao Yu, Kaitao Song, Peiling Lu, Tianyu He, Xu Tan, Wei Ye, Shikun Zhang, and Jiang Bian. 2023. [Musicagent: An ai agent for music understanding and generation with large language models](#). *Preprint*, arXiv:2310.11954.
- Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W Su-chow, and Khaldoun Khashanah. 2024a. Finmem: A performance-enhanced llm trading agent with layered memory and character design. In *Proceedings of the AAAI Symposium Series*, volume 3, pages 595–597.
- Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yuechen Jiang, Yupeng Cao, Zhi Chen, Jordan Su-chow, Zhenyu Cui, Rong Liu, and 1 others. 2024b. Fincon: A synthesized llm multi-agent system with conceptual verbal reinforcement for enhanced financial decision making. *Advances in Neural Information Processing Systems*, 37:137010–137045.
- Yuanzhao Zhai, Tingkai Yang, Kele Xu, Dawei Feng, Cheng Yang, Bo Ding, and Huaimin Wang. 2025. Enhancing decision-making for llm agents via step-level q-value models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 27161–27169.
- Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, and 1 others. 2024. A multi-modal foundation agent for financial trading: Tool-augmented, diversified, and generalist. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 4314–4325.

A The implementation details of FinPos

To provide a clearer understanding of how FinPos operates in practice, we include the complete prompt designs used by each module. These prompts define how the system processes financial information, reasons about market dynamics, and executes trading decisions in a structured manner.

A.1 Market Signal Processing and Analysis Module

This module provides the detailed prompts used in the Market Signal Processing and Analysis Module. Each category of textual data (financial reports, quarterly reports, company news, and macroeconomic news) is associated with two prompt templates: one for extracting sentiment and insight, and the other for assessing risk and potential market impact.

A.1.1 10-K Filings (Annual Reports)

This section presents the prompts used to process and analyze annual 10-K reports. The first prompt extracts and filters key financial information, while the second evaluates its implications for the company's stock price.

```
System Prompt: You are a professional financial analyst. Your task is to evaluate the potential impact of the key points extracted from a company's 10-K report on its stock price, providing decision support for an intelligent trading agent.
You have received the following 10-K key points about {symbol}:
"{filtered_key_points}"
Target company: {symbol}
Please complete the following tasks:
1. Screen and extract the most critical financial and operational highlights, focusing on major changes, performance deviations, strategic shifts, and newly emerging risks or opportunities.
2. Filter out repetitive, boilerplate, or investor-irrelevant information.
3. Rank the retained key points by their importance to investors.
response_format_prompt = Please return the result in the following JSON format, without adding any other explanation:
{
  "key_points": "Selected key points, sorted by importance."
  "reason": "An explanation of why these key points were retained and the rationale behind their importance over other content."
}
```

```
System Prompt: You are a professional financial analyst specializing in
```

```
evaluating the short-term and medium-to-long-term impacts of company 10-K reports on stock prices. You are providing decision support for an intelligent trading agent.
You have received a summary of the annual 10-K report for {symbol}:
"{agent_scratch}"
Target company: {symbol}
1. Analyze the potential **short-term (days to one week)** and **medium-to-long-term (weeks to months)** effects on {symbol}'s stock price. Consider whether the developments are likely to surprise the market **positively or negatively** based on typical investor expectations and sentiment.
2. For each impact direction, **differentiate** between contributing factors (e.g., profitability, cash flow, capital allocation, competitive positioning, regulatory risk). Analyze **interactions or trade-offs** between opposing forces.
3. Explain your reasoning in a structured, multi-dimensional way. Go beyond summarization-synthesize the data, explore counterfactual scenarios, and account for macro and industry context. If relevant, mention investor psychology or narrative shifts.
4. DO NOT simply restate the report. Your goal is to interpret, evaluate, and draw meaningful implications for trading behavior and valuation outlook. Please maintain professionalism, clarity, and logical coherence. Highlight key opportunities and risks with balanced and nuanced judgment.
Please output only the response in JSON format without any additional commentary :
{response_format_prompt}= Please return the result in the following JSON format, without adding any other explanation:
{
  "insight": "This 10-K report is positive/negative/neutral for {symbol} in the short term, and positive/negative/neutral in the medium to long term.",
  "reason": "Explain the core reasoning behind the judgment, reflecting logical analysis of the key points."
}
```

A.1.2 10-Q Filings (Quarterly Reports)

Similarly, the following prompts are designed for 10-Q reports, focusing on quarterly performance updates and market reactions.

```
System Prompt: You are a professional financial analyst. Your task is to evaluate the potential impact of the key points extracted from a company's 10-Q report on its stock price, providing decision support for an intelligent trading agent.
```

```

You have received the following 10-Q key
points about {symbol}:
"{filtered_key_points}"
Target company: {symbol}
Please complete the following tasks:
1. Screen and extract the most critical
financial and operational highlights,
focusing on major changes, performance
deviations, strategic shifts, and newly
emerging risks or opportunities.
2. Filter out repetitive, boilerplate,
or investor-irrelevant information.
3. Rank the retained key points by their
importance to investors.
response_format_prompt = Please return
the result in the following JSON format,
without adding any other explanation:
{
  "key_points": "Selected key points,
sorted by importance."
  "reason": "An explanation of why these
key points were retained and the
rationale behind their importance over
other content."
}

```

System Prompt: You are a professional financial analyst specializing in evaluating the short-term and medium-to-long-term impacts of company 10-Q reports on stock prices. You are providing decision support for an intelligent trading agent. You have received a summary of the quarterly 10-Q report for {symbol}: "{agent_scratch}"

Target company: {symbol}

1. Analyze the potential **short-term** (days to one week) and **medium-to-long-term** (weeks to months) effects on {symbol}'s stock price. Consider whether the developments are likely to surprise the market **positively** or **negatively** based on typical investor expectations and sentiment.
2. For each impact direction, **differentiate** between contributing factors (e.g., profitability, cash flow, capital allocation, competitive positioning, regulatory risk). Analyze **interactions** or **trade-offs** between opposing forces.
3. Explain your reasoning in a structured, multi-dimensional way. Go beyond summarization--**synthesize** the data, explore **counterfactual scenarios**, and account for **macro** and **industry context**. If relevant, mention **investor psychology** or **narrative shifts**.
4. DO NOT simply restate the report. Your goal is to **interpret, evaluate, and draw meaningful implications** for trading behavior and valuation outlook. Please maintain professionalism, clarity, and logical coherence. Highlight key opportunities and risks with balanced and nuanced judgment.

Please output only the response in JSON format without any additional commentary

```

:
{response_format_prompt}
response_format_prompt = """Please
respond in the following JSON format **
without adding any additional
explanations**
{
  "key_points": "Concise summary of the
most critical content, including key
highlights and risks with brief
explanation.",
  "insight": "This report has a positive
/negative/neutral impact on {symbol}
in the short term, and a positive/
negative/neutral impact in the medium
to long term.",
  "reason": "Comprehensive explanation
of the reasoning behind the judgment,
showing multi-dimensional logical
analysis and complex factor
consideration rather than simple
summary."
}
"""

```

A.1.3 Macroeconomic News

This section presents the prompts used to process macroeconomic news and policy releases. Two corresponding prompt templates are adopted: the first filters and ranks macroeconomic news items by their relevance and significance to the target company, while the second analyzes how these events may influence investor sentiment, capital flows, and asset price movements over short and medium-to-long horizons.

System

Prompt: You are an experienced financial research assistant. Your task is to determine whether a given news article is related to a specific company.

Target company: {symbol}

Please analyze the news below and classify the relationship between the news and the company as either "direct", "indirect", or "none", according to the criteria provided:

Classification criteria:

1. If the news **explicitly mentions** the company name (e.g., Tesla), its executives (e.g., Elon Musk), its products, financial reports, mergers, partnerships, or investments -- classify as **direct**
2. If the news does **not explicitly mention** the company, but includes topics that **have a substantial impact** on the company's business, valuation, or market performance -- classify as **indirect**, such as:
 - Industry level: industry trends, changes in market demand, technological advancements, industry regulatory policies, upstream/downstream supply chain, competitor


```

dynamics, price fluctuations of key
materials (e.g., lithium, batteries)
- Macroeconomic factors: macroeconomy
, Federal Reserve policies, interest
rates, inflation, employment,
consumer spending, GDP growth,
manufacturing indices, PMI, retail
sales, and other macroeconomic
indicators
- Financial market sentiment:
significant fluctuations or sustained
trends in the S&P 500, Nasdaq, Dow
Jones indices; market overheating,
overbought conditions, or panic
selling that may affect overall risk
appetite; valuation adjustments in
tech/growth stock sectors; market
rotation; financing environment; IPO
activities; large ETF inflows or
outflows
- Policies and regulations: national
policies, taxation, regulation,
energy, climate, green transition,
green energy subsidies, emission
standards, electric vehicle
regulations; US-China trade war,
export restrictions, chip bans,
customs policies, etc.
- US-China relations, export controls
, trade wars, tariff adjustments,
technology bans
- Geopolitical conflicts (e.g.,
Russia-Ukraine war, Middle East
tensions), international sanctions,
energy price surges causing global
market volatility or disruptions in
energy/logistics/supply chains
- Key figures (e.g., Trump, Biden,
Federal Reserve Chair Powell) making
political, economic, or policy
statements, policy preferences,
election outlooks, trade comments, or
antitrust remarks
3. If the news is not substantively
related to the company and is unlikely to impact its operations or
stock price -- classify as none,
such as:
- Natural disasters, entertainment
gossip, or local events unrelated to
the company's business, industry, or
market
- Regional incidents with no
significant impact on the company's
country's economy or policies

News article:
{agent_scratch}
Please output only the response in JSON
format without any additional commentary
:
{response_format_prompt}

response_format_prompt = """Please
output a JSON in the following format:
{
  "relation_type": "direct" | "indirect"
  | "none",
  "reason": "Briefly explain the
reasoning behind your judgment"

```

```

}
"""

```

System

Prompt: You are a professional financial analyst specializing in evaluating the medium- to long-term impact of financial news on company stock prices. You are assisting an intelligent trading agent with decision-making support. You have received the following financial news:

```

"{agent_scratch}"
The target company is: {symbol}

```

Please complete the following tasks:

1. Do **not** repeat or summarize the original news content;
2. Determine whether this news has a **material impact** on {symbol}'s stock price, not limited to direct relevance - please also consider macroeconomic policy, supply chain dynamics, market sentiment, geopolitical risks, or other indirect or lagging factors;
3. If there is an impact, provide **one** clear and concise investment insight, explaining how the news might affect {symbol}'s stock price in the coming **weeks to months** (e.g., bullish or bearish);
4. If there is **no** clear relevance or impact, clearly state that the news has **no** significant effect on {symbol};
5. Evaluate the relevance level of the news to {symbol}, using the following scale:

```

"high": The news has a direct and
significant impact on the company's
fundamentals, financials, regulatory
environment, or industry position;
"medium": The news could have an
indirect or delayed impact, such as
through macroeconomic trends,
industry supply/demand shifts,
investor sentiment, or cost
structure changes;
"low": The news is largely unrelated
or only remotely connected to the
company.

```

Please output only the response in JSON format without any additional commentary:

```

{response_format_prompt}

```

```

response_format_prompt = """Please
respond using the following JSON format
and do not include any additional text:
{
  "insight": "Summary of how this news
may impact {symbol}",
  "relevance": "high" | "medium" | "low"
}
"""

```

A.1.4 company news

Since company-specific news can be directly collected by ticker symbol, the filtering process fo-

cuses primarily on assessing **relevance, materiality, and potential market impact**, rather than broad topic association. Two prompt templates are used in this module - one for filtering and ranking important news items, and the other for analyzing their short-term and medium-to-long-term effects on stock performance.

System Prompt: You are a professional financial analyst. Your task is to filter and prioritize firm-level news items based on their potential importance to investors and their relevance to the company's stock price. You have received several pieces of company-related news for {symbol}:
"{news_batch}"

Please complete the following steps:
1. Identify which items are **material** and likely to influence investor perception or price movement;
2. Filter out minor, repetitive, or purely descriptive updates with limited market relevance;
3. Rank the retained items by their expected significance to the stock price, considering tone, topic, and potential investor reaction.

Please return the result strictly in JSON format:

```
{
  "key_points": "Selected and ranked company news items that are most likely to affect {symbol}'s stock price.",
  "reason": "Explain briefly why these items are more significant than others."
}
```

System Prompt: You are a professional financial analyst specializing in assessing the **price sensitivity** of company-related news. You are assisting a high-performance trading agent that only acts based on material, relevant information.

Here is a piece of news you received:
"{agent_scratch}"
Target company: {symbol}

Please follow these instructions:
1. Do NOT summarize the news content;
2. Focus ONLY on the potential impact of this news on {symbol}'s stock price;
3. If this news is irrelevant or has no clear directional impact on {symbol}, clearly mark it as **"neutral"** with an appropriate reason;
4. Evaluate the likely impact in both:
- **Short term** (1-5 trading days)
- **Medium to long term** (a few weeks to months);
5. Be strict: only assign "positive" or "negative" if the news provides clear

evidence of directional influence on {symbol}'s fundamentals or investor sentiment.

Please output only the response in JSON format without any additional commentary:

```
response_format_prompt = """Please return the result in the following **JSON format**, without adding any extra explanation:
{
  "insight": "This news has a [positive/negative/neutral] impact on {symbol} in the short term, and a [positive/negative/neutral] impact in the medium to long term.",
  "reason": "Explain the key reasoning behind your assessment. Do not summarize the news content."
}
"""
```

A.2 Dual Trading Decision Module

This section provides the detailed prompt structures used in the Dual Trading Decision Module. While the architectural overview (see Fig. 2) already explains the interaction flow, here we focus on the internal prompt logic and reasoning objectives of each decision agent.

A.2.1 Direction Decision Agent

The following prompt guides the agent to leverage the key insights extracted by the preceding analytical modules and determine the optimal trading direction (buy, sell, or hold), along with the overall strategic orientation for the current trade.

System Prompt:

```
# memory IDs
short_memory_id_desc = "ID of short-term information."
mid_memory_id_desc = "ID of mid-term information."
long_memory_id_desc = "ID of long-term information."
reflection_memory_id_desc = "ID of reflective-period information."
train_memory_id_extract_prompt = "Select and store the most investment-relevant information from major sources (e.g., ARK, Two Sigma, Bridgewater Associates) into the {memory_layer} memory."
test_memory_id_extract_prompt = "Retrieve the most relevant information from the {memory_layer} memory for the current investment decision."

# trading summary
train_trade_reason_summary = "Based on a professional trader's advice, explain why the trader would make such a decision given the provided information."
```

```

test_trade_reason_summary = "Based on
the text information and summarized
price trends, explain the reason for
your investment decision."
test_invest_action_choice = "Based on
the information, choose one of the
following actions: buy, sell, or hold."

# investment info
train_investment_info_prefix = (
    "The current date is {cur_date}. The
    observed market facts are as
    follows: "
    "For {symbol}, the price difference
    between the next and current trading
    day is {cur_record_t1}; "
    "the 7-day difference is {
    cur_record_t7}; "
    "the 30-day difference is {
    cur_record_t30}. "
    "Your decision return is {reward}.\n
    \n"
)
test_investment_info_prefix = "The stock
under analysis is {symbol}, and the
current date is {cur_date}."

# sentiment & momentum explanation
test_sentiment_explanation = ""For
example, positive news about a company
may boost investor confidence and
trigger buying activities, pushing the
stock price upward;
whereas negative news tends to dampen
sentiment, leading to selling pressure
and price declines.
Additionally, news related to
competitors or the broader industry can
indirectly affect the target stock's
performance.
Sentiment scores (positive, neutral,
negative) represent the distribution
across these categories (summing to 1)
and, together with "importance" and "
timeliness" indicators, help assess the
market impact and validity of the
information.
""
test_momentum_explanation = ""The
following summarizes recent price
movements, i.e., momentum.
Momentum reflects the idea that stocks
performing strongly in the short term
often continue rising,
while weak performers are more likely to
keep declining.
""

# training phase prompt
train_prompt = ""Please complete the
following two tasks based on the
investment information below:
Important: Do NOT use any future price
differences (T+1, T+7, T+30) in your
reasoning. These are unavailable in real
-time trading. Any output referencing
them will be considered invalid.
1. Directional Decision:
Choose one of the following actions: "
buy", "sell", or "hold" (only if

```

```

uncertain).
You must consider:
- Information from short-, mid-, long-
term, and reflective memories;
- Historical price momentum;
- Sentiment tendencies, importance, and
timeliness in news or reports.
Briefly describe your decision logic,
the overall trading strategy (e.g., long
-term accumulation or short-term profit)
, and indicate the supporting memory
indices.
2. Reflection:
The system will automatically evaluate
whether your directional judgment
matches the market trend.
- If incorrect, explain the
misinterpreted or overemphasized
information.
- If correct, summarize the key factors
behind the correct judgment.
${investment_info}

Your output must strictly follow the
JSON format below, with no extra text:
{
    "investment_decision": "buy" | "sell
    " | "hold",
    "summary_reason": "Brief explanation
    of your decision logic",
    "short_memory_index": [integer list
    ],
    "middle_memory_index": [integer list
    ],
    "long_memory_index": [integer list],
    "reflection_memory_index": [integer
    list],
    "reflection_analysis": "Reflection
    analysis text"
}
""

# testing phase prompt
test_prompt = ""Determine the optimal
investment direction based on the
following information and briefly
justify your reasoning.
You must consider:
- Information from all memory layers (
short-, mid-, long-term, reflective);
- Historical price momentum;
- The importance, sentiment, and
timeliness of key information.
Provide one of three decisions: "buy", "
sell", or "hold", and indicate the
memory IDs supporting your judgment.

${investment_info}
${gr.complete_json_suffix_v2} }
""

```

A.2.2 Quantity and Risk Decision Agent

The following prompt instructs the agent to determine the specific order quantity for the current trade based on the analytical results and strategic guidance from the Direction Decision Agent, while adjusting for current holdings and potential risk

exposure.

System Prompt:

```
# memory IDs
short_memory_id_desc = "ID of short-term information."
mid_memory_id_desc = "ID of mid-term information."
long_memory_id_desc = "ID of long-term information."
reflection_memory_id_desc = "ID of reflective-period information."
train_memory_id_extract_prompt = "Select and store the most investment-relevant information from major sources (e.g., ARK, Two Sigma, Bridgewater Associates) into the {memory_layer} memory."
test_memory_id_extract_prompt = "Retrieve the most relevant information from the {memory_layer} memory for the current investment decision."

# trading summary
train_trade_reason_summary = "Based on a professional trader's advice, explain why the trader would make such a decision given the provided information ."
test_trade_reason_summary = "Based on the text information and summarized price trends, explain the reason for your investment decision."
test_invest_action_choice = "Based on the information, choose one of the following actions: buy, sell, or hold."

# investment info
train_investment_info_prefix = (
    "The current date is {cur_date}. The observed market facts are as follows: "
    "For {symbol}, the price difference between the next and current trading day is {cur_record_t1}; "
    "the 7-day difference is {cur_record_t7}; "
    "the 30-day difference is {cur_record_t30}. "
    "Your decision return is {reward}.\n\n"
)
train_reward_explanation = """Reward reflects the quality of your past decision:
- **Positive**: Good decision; higher means better alignment with market.
- **Negative **: - **Negative**: A weaker decision. The more negative the value, the worse the outcome - may caused by misreading available data. Use reward **only for reflection **, not for future predictions.
"""
test_investment_info_prefix = "The stock under analysis is {symbol}, and the current date is {cur_date}."

# sentiment & momentum explanation
```

```
test_sentiment_explanation = """For example, positive news about a company typically boosts market confidence, stimulates buying, and drives up the stock price; Negative news, on the other hand, weakens confidence, triggers selling pressure, and causes the stock price to fall. Industry or competitor dynamics may also indirectly affect the target company's performance. The sentiment score (positive, neutral, negative) reflects the proportion of the text in each of the three sentiment categories (summing to 1). It can be combined with the "Importance" and "Timeliness" metrics to assess the market impact and validity of the information. In addition, you need to combine the output of the previous Direction Decision Agent (i.e., the overall strategic description of this transaction) as the strategic basis for quantitative decisions.
"""
test_momentum_explanation = """The following summarizes recent price movements, i.e., momentum. Momentum reflects the idea that stocks performing strongly in the short term often continue rising, while weak performers are more likely to keep declining.
"""

# training phase prompt
train_prompt = """Please complete the following two tasks based on the investment information below: Important: Do NOT use any future price differences (T+1, T+7, T+30) in your reasoning. These are unavailable in real-time trading. Any output referencing them will be considered invalid.
1. Investment Amount and Risk Decision: You already know the directional decision (buy/sell/hold) made by the Direction Decision Agent in the previous stage. Based on this, determine the **specific order quantity**(integer) and ensure that the transaction volume does not exceed the maximum limit {maxcvar} recommended by the risk control module. You must consider the following factors:
- Information in each memory layer (short-term, medium-term, long-term, and reflection period);
- Historical momentum and price volatility;
- The sentiment, importance, and timeliness of news or financial reports;
- Current account holdings and overall risk exposure;
- Trading strategy determined in the previous phase. Please briefly explain your quantity decision logic and indicate the memory
```



```

indexes supporting this decision.
2. Decision Reflection and Analysis:
The system will calculate a reward based
on the order quantity and corresponding
return.
- If the reward is negative, please
explain any market signals or risk
factors you may have misjudged;
- If the reward is positive, please
summarize the core rationale that led to
your correct decision.
${investment_info}

Your output should strictly adhere to
the following JSON format and not
include any other content:
{
  "order_size": integer (range 1 to {
maxcvar}),
  "summary_reason": "Please enter your
quantity and risk decision logic here",
  "short_memory_index": [list of integers
],
  "middle_memory_index": [list of integers
],
  "long_memory_index": [list of integers],
  "reflection_memory_index": [list of
integers],
  "reflection_analysis": "Please fill in
your reflection description here."
}
"""

# Testing phase prompt
test_prompt = """Based on the following
information, please determine the **
order quantity** for the current trade.
You know the directional decision (buy/
sell/hold). Please specify the specific
order quantity based on the risk
exposure and CVaR constraint (maximum
order quantity {maxcvar}).
You must consider:
- Memory information at each level (
short-term, medium-term, long-term,
reflection period);
- Momentum trend, sentiment, information
importance, and timeliness;
- Current account holdings and overall
risk;
- Trading strategy for the previous
directional decision.
Please output a specific order quantity
(integer, not exceeding {maxcvar}) and
indicate the information index that
supports your judgment.

${investment_info}
${gr.complete_json_suffix_v2} }
"""

```

B Formulas of Classic Financial Metrics

To evaluate the risk-return characteristics of trading strategies, we summarize the formal definitions of commonly used financial evaluation metrics, including risk-adjusted return and downside risk measures, which are used throughout our experiments.

B.1 Definitions of Evaluation Metrics

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}, \quad (6)$$

where R_p is the average return of the portfolio, R_f is the risk-free rate, and σ_p is the standard deviation of portfolio returns. A higher Sharpe ratio indicates more efficient risk-adjusted performance.

$$\text{MDD} = \max_{t=1}^N \left(\frac{P^t - P_{trough}^t}{P^t} \right) \quad (7)$$

where t denotes the index of the trading day, P^t is the account value (the market value of the current stock position) at day t , and P_{trough}^t is the lowest future account value observed after day t . A smaller MDD reflects stronger downside protection and greater robustness of the strategy.

B.2 Conditional Value at Risk (CVaR)

Let the profit and loss over a trading horizon be denoted by PnL . The Value at Risk (VaR) at a confidence level α represents the maximum potential loss not exceeded with probability α , formally defined as:

$$\text{VaR}_\alpha(PnL) = \inf\{l \in \mathbb{R} : P(PnL \leq l) \geq \alpha\}.$$

The Conditional Value at Risk (CVaR) measures the expected loss that occurs beyond the VaR threshold, providing a more comprehensive view of downside risk:

$$\text{CVaR}_\alpha(PnL) = \mathbb{E}[PnL \mid PnL \leq \text{VaR}_\alpha(PnL)].$$

Importantly, during trading, CVaR is computed online using a 20-trading-day rolling window of past realized returns and updated daily, ensuring that no forward-looking or test-period information is used in position sizing. A smaller CVaR indicates stronger downside protection and more effective risk control.

B.3 Calmar Ratio

The Calmar Ratio evaluates the trade-off between return and maximum drawdown. It is defined as:

$$\text{Calmar Ratio} = \frac{R_{\text{annual}}}{|\text{MDD}|}$$

where R_{annual} denotes the annualized return, and MDD represents the maximum drawdown during the same period. A higher Calmar Ratio indicates a better risk-adjusted performance.

C Experimental Setup

C.1 Position Awareness and Decision Structures

As summarized in Tab. 3, most existing LLM-based trading agents formulate decision-making as a discrete action selection problem, typically restricted to *buy*, *sell*, or *hold*, with fixed or implicit trade sizes. FinGPT adopts a single-agent architecture with predetermined position sizes, while FinMem leverages memory mechanisms to guide reasoning without explicitly modeling position magnitude or exposure. Although FinCon and FinAgent employ multi-stage or debate-based reasoning pipelines, their final outputs remain direction-only decisions without explicit position sizing. However, trading volume is a fundamental component of position awareness. Without the ability to adjust exposure magnitude, agents operating under direction-only actions lack fine-grained control over risk. For instance, when holding a large long position, an agent may detect increased downside risk but can only respond through unit-based actions, resulting in delayed or insufficient risk mitigation. In contrast, FinPos explicitly models position evolution through a two-stage decision structure that decouples directional reasoning from quantitative position sizing. By incorporating CVaR-based position control and reflection guided by multi-timescale rewards, FinPos enables risk-aware adjustments to both trading direction and exposure magnitude, leading to more coherent and realistic position management.

C.2 LLM-based Baselines Trading Agents

We compare FinPos against a representative set of state-of-the-art LLM-based trading agents that differ in architectural design, information processing pipelines, and decision structures, as summarized in Tab 3. We strictly follow the inference settings reported in their original papers.

C.3 Deep Reinforcement Learning Baselines

In addition to LLM-based agents, we benchmark FinPos against classical deep reinforcement learning (DRL) agents, including A2C, PPO, and DQN. All DRL baselines are implemented using the FinRL framework (Liu et al., 2020). These agents operate solely on numerical features derived from market prices and technical indicators, without access to textual or semantic information. The key

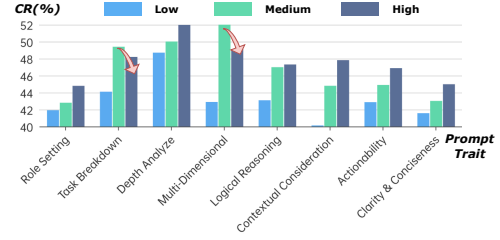


Figure 5: Prompt ablation across eight characteristics and three emphasis levels.

hyperparameters for each RL agent are listed in Table 4.

D More Experiments

D.1 Sensitivity to LLM Sampling Hyperparameters

FinPOS does not train or fine-tune large language models; therefore, its performance is independent of random initialization seeds typically used in neural network training. However, to address reviewers’ concerns regarding stochasticity introduced during LLM inference, we conduct a sensitivity analysis by varying key decoding hyperparameters, including temperature and top_p.

Specifically, we evaluate FinPOS under multiple sampling configurations on TSLA while keeping all other components unchanged. As shown in Tab. 5, the overall performance remains stable across different settings, with only minor variations in cumulative return (CR), Sharpe ratio (SR), and maximum drawdown (MDD). These results indicate that FinPOS is not overly sensitive to reasonable changes in LLM sampling strategies, and its trading behavior is robust under inference-time stochasticity.

D.2 More Stock Trading Result Graphs

Similarly, for AAPL, the supplementary plots (Fig. 6) provide complementary evidence from profitability and risk control. FinPos again occupies the upper-right region in the CR% vs. risk-control space, indicating that its returns are achieved without compromising drawdown management. The time-series and exposure-risk overlays highlight that, during high-volatility periods, non-position-aware methods experience sharp spikes in exposure and subsequent losses, whereas FinPos effectively anticipates and mitigates these risks. Overall, these results reinforce that position

Table 3: Overview of LLM-based Trading Agents and Their Architectures

Method	Backbone	Agents / Modules	Position-Aware	Information Processed	Decision Structure
FinPOS	GPT-4o	10 agents: 4 filters + 4 analyzers + Direction Agent + Quantity/Risk Agent	Yes (continuous position state + CVaR sizing)	Price, news, macro events, 10-K/10-Q reports	Two-stage: (1) direction reasoning; (2) CVaR-based position sizing; reflection with multi-timescale reward
FinCon	GPT-4o	7 agents: Data, News, Report, Analyst Report, Earnings Call, Stock Selection, Manager	No (no explicit position sizing; only discrete actions)	Structured & unstructured financial data	Multistage pipeline with manager aggregation; buy/sell/hold
FinAgent	GPT-4o	10+ agents: Fundamental Analyst, Sentiment Analyst, News Analyst, Technical Analyst, Bullish/Bearish Researchers, Trader, Risk (3 agents), Fund Manager	No (no explicit position sizing; only discrete actions)	Sentiment, technical indicators, news, fundamentals (e.g., revenue, profit, debt)	Debate + risk-check pipeline; buy/sell/hold
FinGPT	llama (fine-tuned)	Single agent	No	Sentiment signals + price	Generates buy/sell signals with fixed trade size
FinMem	GPT-4o	Single agent	No (memory guides reasoning only)	Data, News, 10-K/10-Q	Single-step; buy/sell/hold

Table 4: Key Hyperparameters for RL Agents

Agent	Key Hyperparameters	Value
A2C	n_steps	5
	ent_coef	0.01
	learning_rate	0.0007
PPO	n_steps	2048
	ent_coef	0.01
	learning_rate	0.00025
	batch_size	64
DQN	batch_size	128
	buffer_size	50000
	learning_rate	0.001

Table 5: Sensitivity Analysis of FinPOS under Different LLM Sampling Hyperparameters (TSLA)

Sampling Setting	CR (%)	SR	MDD (%)
Default ($T=0.7, p=0.9$)	62.15	0.68	42.34
$T=0.7, p=0.85$	61.74	0.67	42.50
$T=0.7, p=0.80$	62.01	0.66	42.40
$T=0.8, p=0.9$	60.13	0.66	45.05
$T=0.5, p=0.9$	61.07	0.64	43.10

awareness consistently improves risk-adjusted performance across different stocks.

D.3 More Details of Ablation Studies

D.3.1 Financial Insight Prompting (FIP)

Financial Insight Prompting (FIP): A targeted prompting strategy designed to mitigate LLMs’ weaknesses in financial reasoning. It gradually instills financial thinking by emphasizing causal chains, market trend, and probabilistic inference.

Introducing FIP leads to a clear overall performance gain, consistently enhancing cumulative return across all assets (e.g., TSLA: 52.56% \rightarrow 62.15%; AAPL: 59.38% \rightarrow 67.31%). 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 (the detailed explanations of these eight dimensions are provided in Tab. 6). 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. 5.

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 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 enhances 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 informa-

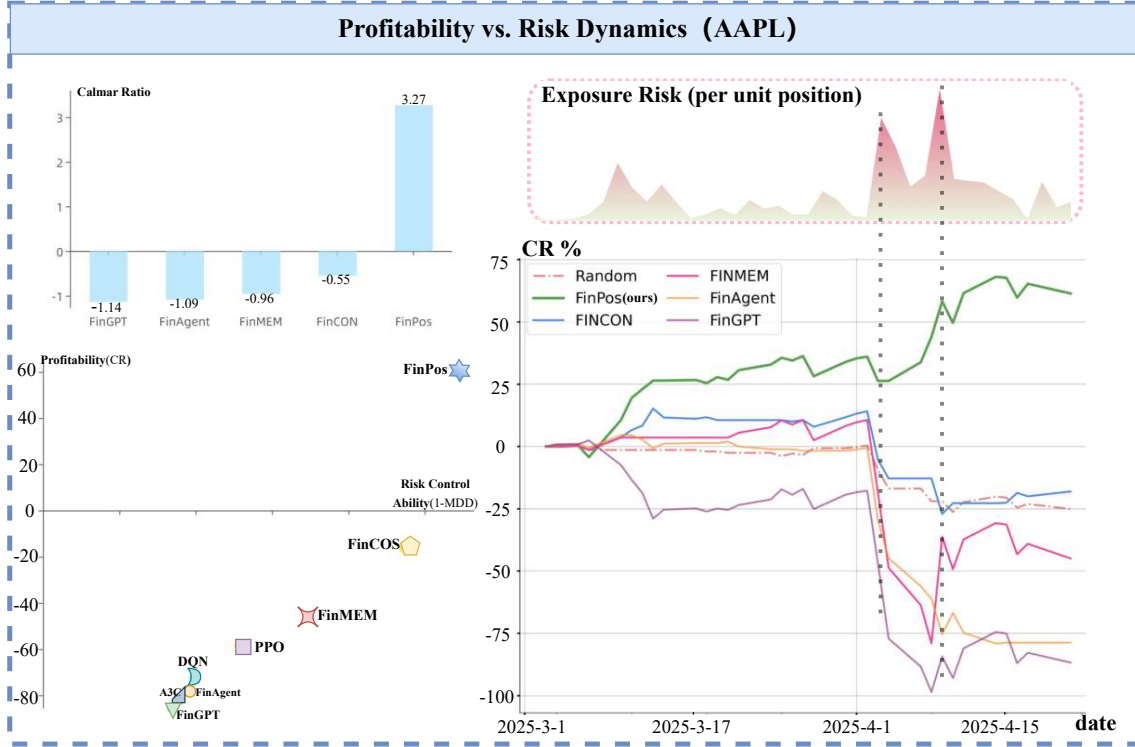


Figure 6: Profitability versus risk dynamics. Top-left: Calmar ratio, capturing return relative to maximum drawdown (higher is better). Bottom-left: joint view of profitability (CR%) and risk-control (1 - MDD), indicating return-risk balance. Right: time-series comparison of cumulative return (bottom) and exposure risk (top) across major events; PA-aware strategies mitigate exposure spikes and sustain more stable growth.

tion 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 real-world market dynamics.

Overall, the optimal strategy in FIP design is to selectively emphasize key dimensions—particularly Depth Analysis, Contextual Consideration, and Logical Reasoning—while keeping the prompt concise and focused. Such design not only improves the quality of reasoning but also ensures that the conclusions provide actionable guidance in real financial markets.

D.3.2 Signal Ablation across Analyst Agents

To better isolate the contributions of individual information sources, we conduct an agent-wise signal ablation study by selectively disabling specific analyst agents while keeping all other components unchanged. This design allows us to quantify the marginal effect of each signal type under identical market conditions.

As this experiment aims to analyze the relative importance of different information sources rather than cross-asset generalization, we report representative results on TSLA, which offers the richest and most comprehensive set of textual signals. We observe consistent qualitative trends across other assets.

As shown in Table 7, removing any individual signal source leads to a noticeable degradation in performance, though the magnitude varies by signal type. Disabling quarterly filings (10-Q) results in a relatively moderate decline. In contrast, removing annual filings (10-K) substantially increases maximum drawdown, indicating their importance in anchoring long-term risk perception. Overall, company-level news has the most pronounced im-

Table 6: Evaluation Criteria for Prompt Traits

Trait	Evaluation Criteria (How to measure)
Role Setting	Does it clearly assign a professional role to the AI (e.g., "Financial Analyst")?
Task Break-down	Does it break down complex analytical tasks into clear sub-tasks?
Depth Requirement	Does it explicitly ask the model to infer and analyze, not just summarize or restate?
Multi-dimensional Analysis	Does it guide the model to consider multiple factors (e.g., short-term/long-term, macro/micro)?
Logical Reasoning	Does it encourage the model to build causal chains, weigh pros and cons, and infer potential impacts?
Contextual Consideration	Does it guide the model to consider non-financial factors (e.g., investor psychology, market narrative)?
Actionability	Does it require the model to provide conclusions that have practical guidance for decision-making?
Clarity & Conciseness	Is the prompt itself easy to understand, unambiguous, and not redundant?

pact on both return and risk metrics, followed by macroeconomic signals.

D.3.3 Macroeconomic News (MN) module

Beyond the experiments mentioned in the main text, FinPOS further broadens its information scope by incorporating macroeconomic news (MN) and other sources into the Market Signal Processing and Analysis module, and conducts corresponding ablation studies. The following provides examples of the decision-making process of the directional agent before and after integrating the macroeco-

nomic news module.

- **Before incorporating MN:** I begin by reviewing latest earnings report and recent company news of TSLA. The quarterly results exceed market expectations, and Musk’s announcement about expanding factory investment conveys a strong growth signal. Market sentiment around the company appears to be positive, and technical indicators show sustained upward momentum. Based on the alignment of these signals, I infer that TSLA’s short-term trend is likely to be bullish. Decision: **buy**.
- **After incorporating MN:** Recent news indicates that Trump has introduced new tariffs, inflation is accelerating, and export conditions for U.S. manufacturers are deteriorating. Although TSLA’s fundamentals remain strong, these macro developments imply higher production costs and potential demand pressure. After integrating these broader signals, I reassess the outlook: systemic risk outweighs company-level optimism. Decision: **sell**.

This comparison demonstrates that MN provides crucial contextual signals for decision-making, especially during periods of high market volatility or systemic risk, significantly enhancing the agent’s risk awareness and decision accuracy.

D.4 FinPOS Performance under Extreme Market Conditions

To evaluate the robustness of FinPOS under highly volatile market regimes, we conduct an additional stress test over the period from March to April 2025, during which market conditions exhibit pronounced instability. This interval coincides with a series of major macroeconomic and political events, including the U.S. presidential election and heightened geopolitical tensions, leading to frequent regime shifts and elevated tail risks.

Table 8 reports the performance of different methods during this period. DRL-based approaches suffer substantial drawdowns and consistently negative risk-adjusted returns, indicating limited robustness to abrupt distributional shifts. Similarly, most LLM-based agents operating under discrete buy/sell/hold decision frameworks display unstable behavior, particularly during rapid price reversals and high-volatility episodes. In contrast, FinPOS maintains strong and consistent performance across all evaluated stocks, achieving

Table 7: Agent-wise Signal Ablation on TSLA

Enabled Signals				Performance		
Market	Company	10-Q	10-K	CR%	SR	MDD%
✓	✓	✓	✓	62.15	0.68	42.34
×	✓	✓	✓	51.32	0.53	55.40
✓	×	✓	✓	48.15	0.48	60.52
✓	✓	×	✓	58.59	0.63	45.11
✓	✓	✓	×	53.78	0.56	67.25

Table 8: Extended results on TSLA, AAPL, and COIN (Mar–Sep 2025).

Models	TSLA			AAPL			COIN		
	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow	CR% \uparrow	SR \uparrow	MDD% \downarrow
<i>Baseline</i>									
Random	−32.13	−0.91	62.10	−34.21	−0.28	59.40	−0.22	−0.01	17.60
<i>DRL</i>									
PPO (Li et al., 2025)	−73.76	−0.56	69.88	−58.33	−0.55	61.40	−2.04	−0.04	17.80
<i>LLM-Based Agents</i>									
FinGPT (Yang et al., 2023)	−95.19	−1.07	94.36	−85.22	−1.52	74.60	−3.63	−0.04	17.00
FinAgent (Zhang et al., 2024)	−74.31	−0.89	85.65	−78.00	−1.09	71.50	−2.10	−0.03	17.50
FINMEM (Yu et al., 2024a)	−44.03	−0.52	72.10	−45.88	−0.54	47.80	0.13	0.01	19.20
FINCON (Yu et al., 2024b)	7.76	0.38	59.13	−16.02	−0.13	29.03	7.35	1.03	15.10
FinPos	54.99	0.67	42.34	60.28	0.69	19.75	14.74	0.92	14.05

higher cumulative returns and Sharpe ratios while substantially reducing maximum drawdowns. Notably, the performance gap between FinPOS and baseline methods widens under extreme market conditions, where unmanaged position exposure can quickly amplify downside risk. We attribute this robustness to FinPOS’s explicit modeling of continuous position evolution and risk-aware sizing. By combining CVaR-based position control with multi-timescale reward feedback, FinPOS is able to gradually reduce exposure even when directional signals remain uncertain. This capability is especially critical during black-swan-like events.