EvoFinAgent: A Self-Evolving Agentic Framework for Financial Trading

Anonymous ACL submission

Abstract

Recent advances in large language model trading agents have increasingly emphasized multimodal inputs, external tool integration, and multi-agent coordination. While promising, these approaches typically deliver marginal gains and leave a central limitation unaddressed: the agent's internal reasoning logic is deterimined by human and it is hard for us to find the logic flaws and make the correct modifications. Moreover, different language models often exhibit distinct reasoning patterns and trading behaviors. As a result, a fixed prompt template or workflow designed for one model may not generalize well when applied to another backbone, leading to performance degradation or incompatibility. In this work, we introduce EvoFinAgent, a lightweight, self-evolving agentic framework that enhances trading performance through structured prompt-level optimization. Without requiring fine-tuning, external data sources, or auxiliary tools, EvoFinAgent improves its decision quality by reflecting on past trades, identifying reasoning failures, and recomposing new prompts based on performance-driven feedback. Its modular chain-of-thought structure and stable self-reflection pipeline enable interpretable and risk-aware strategy evolution. We validate EvoFinAgent across five major stock assets and multiple backbone language models, demonstrating consistent improvements in return, Sharpe ratio, and drawdown over generations. In backtesting, EvoFinAgent outperforms a broad spectrum of baselines, including rule-based systems, machine learning methods, reinforcement learning agents, LLM-based trading agents, and general-purpose evolutionary agents not originally designed for financial decision-making. These results suggest that evolving internal reasoning through prompt-level adaptation offers an effective and scalable alternative to increasing architectural complexity in financial trading tasks.

1 Introduction

Financial markets require efficient capital allocation under dynamic and uncertain conditions. To navigate market complexities, algorithmic trading has evolved from rigid rule based strategies to adaptive machine learning approaches (Strader et al., 2020). While reinforcement learning based systems (An et al., 2022) offer improved flexibility, they suffer from limited interpretability, heavy data dependency, and poor generalization across diverse market regimes. Furthermore, traditional methods struggle to synthesize multimodal market intelligence, such as price trends and financial news, into a coherent decision making process, highlighting the need for more flexible frameworks.

Large Language Models have demonstrated strong capabilities in reasoning and decision making. In financial trading, where decisions directly impact profit and loss, LLM based agents present a promising paradigm. Recent research heavily scales system complexity to handle market challenges, such as incorporating multimodal inputs like FinAgent (Zhang et al., 2024), integrating memory modules like Finmem (Yu et al., 2025), or coordinating multiple specialized sub agents like TradingAgents (Xiao et al., 2024). These engineering approaches introduce significant computational overhead while compromising system interpretability, scalability, and deployment simplicity.

However, a limitation remains unaddressed: the internal reasoning logic of most LLM based trading agents is static and non adaptive. Because these agents rely on manually designed prompts (Pei et al., 2025), any inherent flaws or biases can directly lead to persistent errors (Madigan et al., 2025). Furthermore, fixed templates fail to generalize across diverse backbone models that inherently possess distinct trading styles (Yu et al., 2025). Once deployed, the agent lacks self correction, causing errors like overreacting to noise or mis-

applying strategies to compound over time (Fang et al., 2025). To tackle this, we focus on enabling agents to evolve their internal decision logic to uncover and correct flawed reasoning. We identify four key challenges in building such self evolving LLM trading agents:

- **Ch1: Lack of Structured Reasoning Frameworks in Trading Agents.** Many existing LLM-based trading agents rely on static prompts or unstructured decision workflows, which makes it difficult to trace, analyze, or modify their internal reasoning processes (Zhang et al., 2025). Without a modular and structured reasoning backbone, it becomes unclear which specific parts of the logic contribute to wrong reasoning, making targeted revisions or evolutionary improvements nearly impossible (Agarwal et al., 2025).

- **Ch2: Instability in Autonomous Reasoning Evolution.** Naively allowing an agent to modify its own reasoning logic without proper guidance often results in instability, logic collapse, or performance degradation (Chen et al., 2025a). Without a principled framework to diagnose prior failures and determine which reasoning components need revision, self-evolution becomes erratic and uncontrollable. Current methods lack reliable mechanisms to ensure that edits are both targeted and grounded, particularly in the high-stakes context of financial decision-making.

- **Ch3: Poor Generalization Across Market Regimes.** Financial markets are highly non-stationary and regime-dependent (Souza, 2025). Due to fixed prompt and simple analysis logic, most trading agents lack the ability to detect regime shifts and adjust their internal reasoning accordingly. This may lead to performance degradation over time and inconsistency in reasoning.

Our contributions address key challenges in building self-evolving financial trading agents. First, we propose EvoFinAgent, a agentic, language-only architecture that employs modular Chain-of-Thought (CoT) prompts, enabling structured, interpretable, and revisable decision-making (Nitarach et al., 2025) tailored to financial tasks, while avoiding reliance on external tools, multimodal inputs, or multi-agent coordination (Ch1). Second, to support stable and autonomous reasoning evolution, we introduce a robust multi-layer self-reflection framework that operates across three levels: the Phenomenon Layer identifies trading-level failures, the Mechanism Layer diagnoses reasoning flaws, and the Engineering Layer proposes performance-aligned revisions (Ch2). To address the non-stationarity of financial markets (Ch3), we apply dimensionality reduction techniques to identify and select representative historical intervals that capture diverse market regimes.

2 Related Work

LLM-based Trading and Reasoning Evolution. Large language models (LLMs) show strong reasoning capabilities in financial decision-making. Beyond adapting LLMs to financial corpora (e.g., BloombergGPT (Wu et al., 2023), FinGPT (Wang et al., 2023)), current frameworks employ multi-agent collaboration, memory layers, or tool integration to navigate complex markets (Yu et al., 2025; Yang et al., 2024; Yu et al., 2024). Concurrently, another line of work evolves agent reasoning through prompt refinement, reflection, and population-based optimization (e.g., EvoPrompt (Guo et al., 2023), GEPA (Agrawal et al., 2025), C-Evolve (Li et al., 2025), EvoAgent (Yuan et al., 2025), EvoAgentX (Wang et al., 2025)). However, these trading and evolution approaches often introduce heavy architectural complexity, depend on engineered pipelines, or primarily target classical NLP tasks. In contrast, **EvoFinAgent** bridges these lines of work with a lightweight, fully self-contained design. By directly evolving structured reasoning traces from financial backtesting environments, it establishes consistent, risk-aware behavioral patterns without external tools, collaborative sub-agents, or manual updates.

3 Problem Formulation

In this section, we formulate the trading task as a Markov Decision Process (MDP), and describe how prompt-based language agents interact with the environment via a language-conditioned interface. We then introduce the core design of our EVOFINAGENT framework, which enables self-reflective and evolutionary improvement of language-based decision policies.

Financial Trading as MDP. We formulate the trading task as a Markov Decision Process (MDP) (Puterman, 1990), represented by a 5-tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \gamma)$. Here, $\mathcal{S}$ denotes the set of market states, including price sequences such as OHLCV data, technical indicators, and op-

tionally contextual features. The discrete action space $\mathcal{A}$ consists of three basic actions: BUY, SELL, and HOLD. The transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ defines the probability $\mathcal{T}(s, a, s') = \mathbb{P}(s_{t+1} = s' \mid s_t = s, a_t = a)$ of moving from state s to state s' after taking action a. The reward function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ typically corresponds to trading returns, adjusted for transaction costs. Finally, $\gamma \in [0, 1)$ is the discount factor that determines the importance of future rewards. At each time step t, the agent observes a state $s_t \in \mathcal{S}$, selects an action $a_t \in \mathcal{A}$, receives a reward $r_t = \mathcal{R}(s_t, a_t)$, and transitions to the next state $s_{t+1} \sim \mathcal{T}(s_t, a_t, \cdot)$. The goal is to learn a parameterized policy $\pi_\theta(a \mid s)$ that maximizes the expected cumulative discounted reward, defined as $\pi_{\theta^*} = \arg\max_{\pi_\theta} \mathbb{E}_{\pi_\theta} \left[\sum_{i=0}^{T} \gamma^i r_{t+i} \mid s_t = s\right]$. This reinforcement learning formulation offers a principled framework for modeling intelligent trading agents in financial environments.

EvoFinAgent Objective. EvoFinAgent's trading policy is implicitly defined by a natural language prompt ϕ for a frozen Large Language Model (LLM), instead of explicitly learning parameters through gradient updates. The prompt encodes structured reasoning and behavioral heuristics, guiding the agent's decisions in response to market observations. Each prompt is structured as $\Phi = [\psi; \phi]$, where ψ is a fixed identity description (e.g., "You are a professional financial analyst.") and $\phi = [r_1, r_2, \ldots, r_K]$ is a sequence of modular reasoning steps written in a stepwise Chain-of-Thought (CoT) format. Representative examples of ψ and ϕ can be found in Appendix H. While ψ remains unchanged throughout training, the reasoning component ϕ is continuously evolved by the agent. At each time step t, the agent observes a market state s_t and selects an action using a policy:

$$a_t \sim \pi(a \mid s_t; \phi) = \text{LLM}(\psi, \phi, s_t) \qquad (1)$$

To evaluate a candidate prompt ϕ, the agent simulates it over a set of N trading trajectories $\{\tau^{(i)}\}_{i=1}^{N}$, collects the resulting returns, and computes a performance metric $\text{Eval}(\tau^{(i)})$ for each trajectory. We adopt the annualized Sharpe Ratio as the default evaluation metric, reflecting the trade-off between return and risk. The agent then aims to discover the optimal reasoning strategy ϕ^* by maximizing the average performance across all trajectories. The algorithm is in Appendix 1:

$$\phi^* = \arg\max_{\phi \in \Phi} \frac{1}{N} \sum_{i=1}^{N} \text{Eval}(\tau^{(i)}) \qquad (2)$$

4 EvoFinagent Framework

EvoFinAgent is a prompt-level agentic evolutionary framework designed to develop interpretable and high-performing financial reasoning agents. Instead of training neural models end-to-end, EvoFinAgent treats each agent as a structured prompt that defines a multi-step reasoning strategy. These prompts evolve over time through an iterative process involving generation, evaluation, reflection, and selection. The framework is designed to operate purely in natural language space, enabling transparent and modular reasoning evolution without requiring model parameter updates.

The overall system is composed of four modules: (1) Agent Pool, (2) Reasoning Structure Recombination, (3) Agent Evaluation, and (4) Structured Reflection and Report Generation. These modules form a closed-loop evolution cycle. At each iteration, the system selects top-performing agents from the pool as a population, recombines their reasoning steps to create new prompts, evaluates the performance of new agents in unseen financial environments, and generates structured self-reflection reports that guide the next round of evolution. The process is repeated over multiple generations, allowing the agent population to continuously improve reasoning quality and generalization ability.

4.1 Agent Pool

The agent pool stores a growing collection of reasoning agents during training. Each agent is defined by a structured natural language prompt that includes a role identity (for example, "You are a professional investment advisor") and a sequence of reasoning steps that form its decision making logic (for example, "Step 1: Analyze RSI trend", "Step 2: Check macroeconomic news sentiment"). Alongside the prompt, each agent keeps additional metadata, such as performance metrics (SR, ARR, MDD), a unique identifier, and a reflection report generated after evaluation.

The agent pool acts as an evolving repository of successful strategies. Each new agent generated via reasoning recombination is evaluated on held-out financial environments; it enters the pool if its Sharpe ratio outperforms at least one parent agent. With no fixed size limit, the pool expands over time to retain diverse, high-performing reasoning styles. Crucially, the pool represents a historical trajectory rather than the full search space, which remains open-ended as new agents are continually

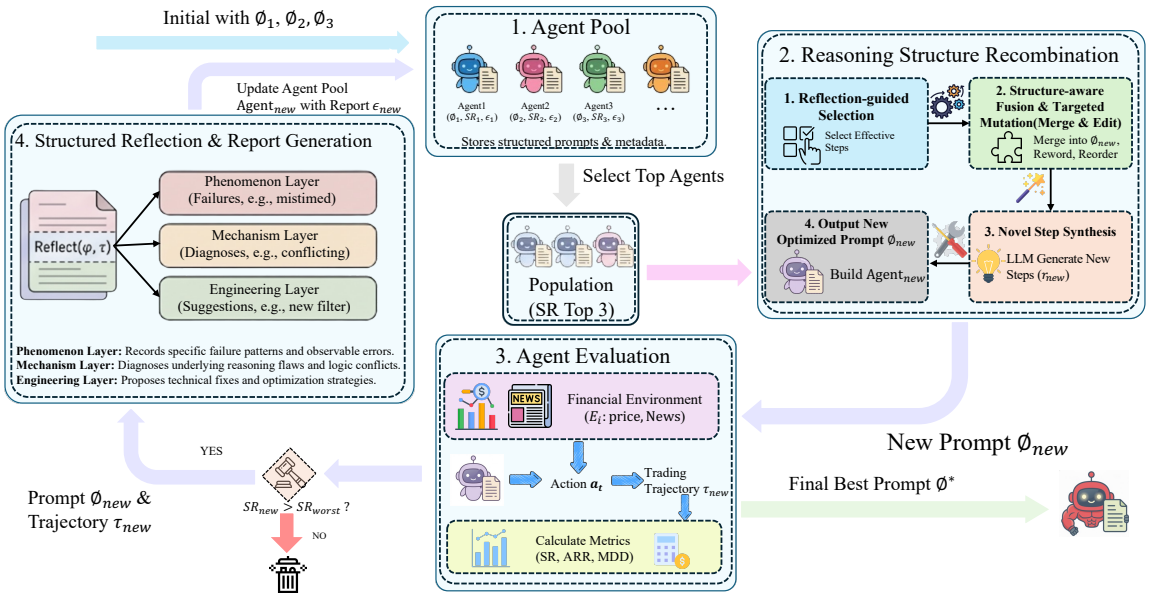

Figure 1: The EvoFinAgent framework evolves prompts through self-reflection and selective recombination.

synthesized from existing logic and reflection reports. Because each agent's reasoning is fully interpretable, developers can directly inspect or edit individual prompts, making the system highly flexible for human-in-the-loop design and debugging.

4.2 Reasoning Structure Recombination

To generate new reasoning agents, EvoFinAgent performs recombination in the space of structured natural language prompts. Instead of evolving model parameters, the system recombines interpretable reasoning steps from high-performing agents. Given a set of parent agents $\{\Phi_1, \Phi_2, \Phi_3\}$, each defined by a prompt $\Phi = [\psi; \phi]$, the goal is to construct a new prompt $\Phi_{\text{new}} = [\psi; \phi_{\text{new}}]$ with improved logic and performance. Recombination proceeds in four stages:

(1) Reflection-guided selection. Each parent agent includes a reflection report that identifies effective and ineffective reasoning steps. Steps positively attributed or explicitly suggested for improvement are selected, while those consistently linked to errors are filtered out. This prioritizes steps that contribute to stronger decision-making, based on both performance and self-diagnosis.

(2) Structure-aware fusion. Selected steps are then merged into a coherent reasoning chain $\phi_{\text{fused}} = [r_1, r_2, \ldots, r_M]$, while maintaining logical consistency. Redundant or conflicting steps are removed using cues from reflection reports (e.g., "conflicts with macro filter"). The fused steps form the backbone of the new prompt.

(3) Targeted mutation. To introduce diversity and improve adaptability, the system applies guided modifications to the fused reasoning chain. These changes include rewording conditions for clarity, reordering steps to reflect better causal flow, or substituting a step with an alternative suggested in reflection reports. For example, if a step originally stated "review recent market movement before making a decision," it might be rephrased to "identify major market shifts through historical patterns first." Alternatively, a step focused on general news review might be replaced with a more specific action like "filter news for high-impact events only," based on past performance insights. These mutations are not applied at random. Instead, they are guided by failures identified in earlier agents, ensuring that each change corresponds to a correction based on observed data or prior experience. This design enables the system to incorporate domain knowledge and learning feedback directly into the following generation of agents.

(4) Novel step synthesis. If existing reasoning fragments fail to address known weaknesses, the language model is prompted to synthesize new steps based on past failure cases:

$$r_{\text{new}} \leftarrow \text{LLM}(\psi, \mathcal{F}, \mathcal{C}) \tag{3}$$

ψ denotes the fixed agent identity, $\mathcal{F} = \{f^{(1)}, f^{(2)}, \ldots\}$ is a set of failure trajectories, and $\mathcal{C} = \{c^{(1)}, c^{(2)}, \ldots\}$ represents their corresponding inferred root causes. The LLM uses these elements to generate a revised reasoning step r_{new} for con-

sideration. Recombination process is implemented through COT prompting, as is summarized as:

$$\phi_{\text{new}} = \text{Recombine}(\{\phi_i\}, \{\mathcal{E}_i\}) \tag{4}$$

where $\mathcal{E}_i$ are reflection reports. This process is repeated generation by generation, enriching the agent pool with more diverse and well-adapted reasoning chains. Recombination operates at the level of reasoning structure rather than surface prompt tokens, ensuring that evolutionary improvements reflect meaningful changes in decision logic instead of prompt-level noise. As a result, the evolved agents exhibit more stable and interpretable decision behaviors across different market conditions.

4.3 Agent Evaluation

To determine whether a newly constructed prompt $\phi_{\text{new}}^{(g)}$ should be included in the agent population, EvoFinAgent performs a simulation based evaluation over held out financial environments. This evaluation reflects realistic trading conditions and enables rigorous and interpretable comparisons against existing agents.

Each environment E_i consists of a selected stock, a specific time period, and a sequence of daily market contexts $x_t^{(i)}$, which include historical prices and textual news. For each day $t \in \{1, \ldots, T\}$, the agent interacts with the environment by executing its reasoning prompt using the current context:

$$a_t^{(i)} = \text{LLM}(\phi, x_t^{(i)}) \in \{\text{buy, hold, sell}\} \tag{5}$$

These actions form a complete trading trajectory $[a_1^{(i)}, \ldots, a_T^{(i)}]$, which is evaluated using standard backtesting techniques to yield a return sequence $\tau^{(i)}$. From this sequence, three key performance metrics are extracted: Sharpe Ratio: $\text{SR}_i = \text{Sharpe}(\tau^{(i)})$, Annual Return Rate: $\text{ARR}_i = \text{Return}(\tau^{(i)})$, Maximum Drawdown: $\text{MDD}_i = \text{Drawdown}(\tau^{(i)})$. These metrics are averaged across all evaluation environments $E_1, \ldots, E_N$ to produce a final performance summary:

$$\text{Eval}(\phi) = \left(\overline{\text{SR}},\ \overline{\text{ARR}},\ \overline{\text{MDD}}\right) \tag{6}$$

where $\overline{\text{SR}}$, $\overline{\text{ARR}}$, and $\overline{\text{MDD}}$ denote the average values across all evaluation environments $E_1, \ldots, E_N$, respectively. To qualify for adding to the population, the new agent must outperform at least one parent based on the average Sharpe Ratio:

$$\text{SR}_{\text{new}} > \min_{j \in \{\phi_1, \phi_2, \phi_3\}} \text{SR}_j \tag{7}$$

If this condition is met, the new agent is automatically admitted to the updated population:

$$\mathcal{P}^{(g+1)} = \left(\mathcal{P}^{(g)} \setminus \{\phi_{\text{worst}}^{(g)}\}\right) \cup \{\phi_{\text{new}}^{(g)}\} \tag{8}$$

No gradient-based learning, reward tuning, or fine-tuning is involved in this process. The evaluation is entirely based on executing structured reasoning steps defined in natural language. This design makes it possible to attribute each trading action to a specific logical component, enabling fine-grained inspection and diagnosis. The framework further supports extensibility. Constraints such as leverage levels, or portfolio allocations can be incorporated into the simulation environment, ensuring that EvoFinAgent remains compatible with realistic trading systems.

4.4 Structured Reflection and Report Generation

To enable iterative improvement, each newly added agent conducts structured self-reflection after being accepted into the population. For a given generation g, when a new prompt $\phi_{\text{new}}^{(g)}$ outperforms the weakest agent in the population, it is added to the population. Its trading trajectory $\tau_{\text{new}}^{(g)}$ is then directly passed through a reflection function:

$$\mathcal{E}_{\text{new}}^{(g)} = \text{Reflect}(\phi_{\text{new}}^{(g)}, \tau_{\text{new}}^{(g)}). \tag{9}$$

The output $\mathcal{E}_{\text{new}}^{(g)}$ is a structured reflection report generated by the agent itself. This report identifies suboptimal or flawed reasoning patterns and offers improvement guidance for future prompt.

Each reflection report $\mathcal{E}_{\text{new}}^{(g)}$ features a three-layer hierarchical structure: (1) the *phenomenon layer* logs specific trading failures and their exact timeframes; (2) the *mechanism layer* diagnoses these failures by tracing them to structural reasoning flaws like conflicting heuristics; and (3) the *engineering layer* proposes abstract strategies to refine the decision logic. Serving as the sole feedback guide for the next generation, $\mathcal{E}_{\text{new}}^{(g)}$ penalizes reasoning components linked to harmful behaviors while prioritizing positively assessed ones during recombination. This self-driven reflection mechanism reinforces adaptation without manual labeling, enabling agents to autonomously generate structured insights and leverage internal feedback to evolve effective, interpretable trading behaviors. Details are provided in Appendix H.2.

Categories	Models	AAPL			AMZN			MSFT			TSLA			GOOGL		
		ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓
LLM Agent	Epoch-0	14.98	1.28	3.79	-3.39	0.04	32.04	5.92	0.38	23.53	27.80	0.76	30.09	57.51	1.69	27.84
	Epoch-30	**30.60**	**1.45**	**11.36**	**28.46**	**1.09**	**18.31**	**15.46**	**0.88**	**16.33**	**63.73**	**1.44**	**19.99**	**88.36**	**2.47**	**14.72**
Improvement		**104.28%**	**13.33%**	-	**939.82%**	**2525.00%**	**42.85%**	**161.21%**	**131.58%**	**30.60%**	**129.28%**	**89.47%**	**33.56%**	**53.65%**	**46.15%**	**47.14%**

Table 1: Comparison of backtest results before and after reasoning evolution.

5 Experiments

To evaluate the effectiveness of our proposed framework, EVOFINAGENT, we conduct a series of experiments designed to assess its trading performance, reasoning evolution, and generalization capability across multiple assets and baseline models. Our goal is to investigate whether prompt evolution enables large language model (LLM) agents to make more disciplined, risk-aware, and adaptive trading decisions, especially under complex financial dynamics. Specifically, we aim to answer the following research questions:

- **RQ1**: Can EvoFinAgent outperform existing rule-based system, machine learning methods, reinforcement learning, LLM-based trading agents, and Self-evolution agents?
- **RQ2**: Does the evolutionary optimization of reasoning prompts improve trading performance across different backbones?
- **RQ3**: How do reasoning structures evolve, and what behavioral patterns emerge?

5.1 Experiment Setup

Selection of Training Intervals. We construct the training set using historical U.S. stock market data from 2020 to 2024. The dataset includes five major stocks: AAPL, AMZN, MSFT, TSLA, and GOOGL. For each stock, we select six monthly trading intervals, resulting in 30 training samples. To ensure both diversity and coverage, we adopt a two-stage strategy. In Stage 1, we allocate six intervals per stock to ensure balanced representation. In Stage 2, we apply t-SNE to embed candidate intervals of each stock into a two-dimensional space. We then perform K-means clustering with $k = 6$ and select the sample closest to each cluster centroid. This ensures an even distribution across different market regimes and avoids training bias. The t-SNE figure and distribution are shown in Figure 2. Our choice of 30 intervals is justified in Appendix D.

Evaluation Metrics. We evaluate EvoFinAgent with three metrics capturing both returns and risks: ARR, SR and MDD, as shown in Appendix F.

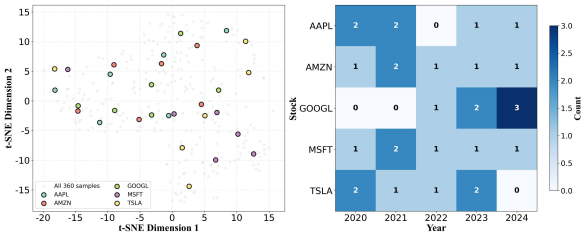

Figure 2: Representative Interval Selection: t-SNE visualization and stock-level temporal distribution.

5.2 Experiment Results

Self-Evolution Improvements. We track the 30-generation self-evolution trajectory of EvoFinAgent (Gemini-3-Flash-Preview backbone). As detailed in Figure 3, Table 3, and Appendix I, the agent achieves consistent cross-metric improvements and develops more coherent, selective trading behavior over time. Notably, the Sharpe Ratio increases from 0.760 to 1.042 (**+37.07%**), annualized return rises from 33.78% to 43.77% (**+29.58%**), and maximum drawdown drops from 7.08% to 5.97% (↓**15.67%**).

Backtesting Performance. EvoFinAgent outperforms all baselines across most assets in the 2025 backtest. It achieves the highest annualized return on four out of five stocks: AAPL (30.60%), AMZN (28.46%), TSLA (63.73%), and GOOGL (88.36%), while remaining competitive on MSFT behind SAC and TradingAgents. Its risk-adjusted superiority is highlighted by top-tier Sharpe ratios (e.g., 2.47 on GOOGL, 1.44 on TSLA) and exceptional maximum drawdown control. In contrast, baselines show limited adaptability: rule-based methods (MACD, KDJ_RSI) suffer from unstable signals, standard ML/RL methods exhibit inconsistency, and LLM-based approaches (FinGPT, TradingAgents) or self-evolution agents (EvoPrompt, EvoAgent) lack sustained dominance or robust risk management. As shown in Figure 5, EvoFinAgent maintains smoother cumulative return curves, fewer reversals, and stronger trend capture, owing to its reasoning evolution that enables coherent, context-aware decisions. Robust-

Category	Model	AAPL			AMZN			MSFT			TSLA			GOOGL			BTC		
		ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓	ARR%↑	SR↑	MDD%↓
Market	B&H	8.94	0.44	27.18	0.13	0.17	30.02	14.57	0.67	20.37	7.23	0.40	44.06	83.15	1.88	28.53	-6.13	-0.01	32.10
Rule-based	MACD	-9.26	-0.43	20.95	-18.79	-0.72	27.96	-3.84	-0.16	17.75	22.39	0.72	35.63	29.56	1.19	18.17	-1.42	-0.01	18.84
	KDJ_RSI	1.69	0.19	22.98	3.27	0.25	22.72	-1.64	0.04	20.67	11.85	0.47	38.58	1.58	0.18	19.25	-7.48	-0.23	25.59
	ZMR	15.76	0.67	26.69	-0.34	0.12	26.78	-2.25	-0.03	14.75	33.09	0.81	48.02	7.00	0.43	24.45	8.49	0.45	22.59
ML-based	LGBM	22.12	0.85	22.98	-0.31	0.15	25.90	-7.26	-0.29	21.75	-15.32	-0.20	53.42	47.06	1.58	23.84	-5.36	-0.20	29.91
	LSTM	6.11	0.35	23.45	-0.80	0.10	26.65	-7.93	-0.51	19.80	8.60	0.46	17.42	16.03	0.73	25.75	-7.16	-0.23	25.08
	Transformer	18.47	0.73	26.22	-3.43	-0.04	16.49	-8.64	-0.63	14.52	-23.02	-0.55	46.65	56.72	2.22	15.11	-8.04	-0.25	26.37
RL-based	DQN	-20.05	-0.90	26.01	-6.70	-0.14	16.64	13.39	0.66	21.83	-5.41	0.08	27.57	1.40	0.18	26.30	-4.34	-0.10	24.01
	PPO	12.65	0.53	29.50	-20.90	-0.60	41.67	14.73	0.80	9.21	27.56	0.70	47.45	-7.05	-0.37	18.74	-5.82	-0.12	26.87
	SAC	7.04	0.41	17.37	10.26	0.50	16.68	24.82	1.08	13.17	-47.51	-1.43	51.78	-0.04	0.12	24.23	12.35	1.31	4.96
Trading-Agent	FinGPT	-0.46	0.12	31.81	21.84	1.09	20.31	11.04	0.58	17.89	-4.50	0.15	45.26	40.36	1.63	11.64	-3.92	-0.08	26.50
	TradingAgents	-13.05	-0.31	28.98	9.95	0.49	18.01	24.35	1.10	22.90	15.67	0.54	39.36	17.07	0.70	29.01	2.90	0.25	26.94
Self-Evo Agent	EvoPrompt	-9.48	-0.55	21.42	-1.16	0.09	22.07	-5.31	-0.24	21.85	31.86	1.04	17.97	45.91	1.62	19.58	1.34	0.16	15.43
	EvoAgent	12.31	0.52	30.22	5.31	0.32	30.87	16.73	0.76	20.64	43.86	0.94	48.01	66.73	1.74	29.84	-6.83	-0.04	31.31
	EvoFinAgent	30.60	1.45	11.36	28.46	1.09	18.31	15.46	0.88	16.33	63.73	1.44	19.99	88.36	2.47	14.72	15.04	1.34	5.80
Improvement		**38.34%**	**35.29%**	**34.60%**	**30.31%**	-	-	-	-	-	**92.60%**	**77.78%**	-	**6.27%**	**11.26%**	-	**21.78%**	**2.29%**	-

Table 2: Performance comparison across Strategies. Top 3 results are highlighted in red (best), orange (2nd), dark green (3rd).

ness checks are provided in Appendix B.

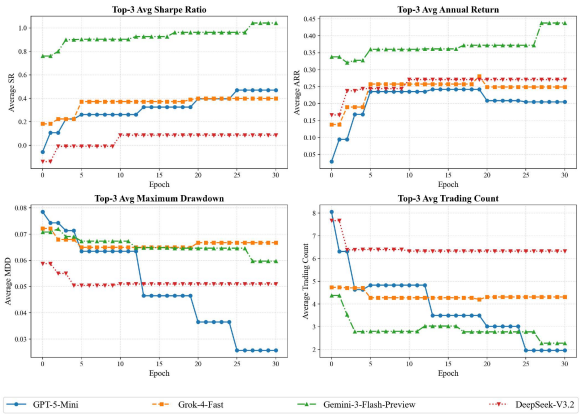

Figure 3: Evolution of top-3 averaged metrics across epochs for four different models.

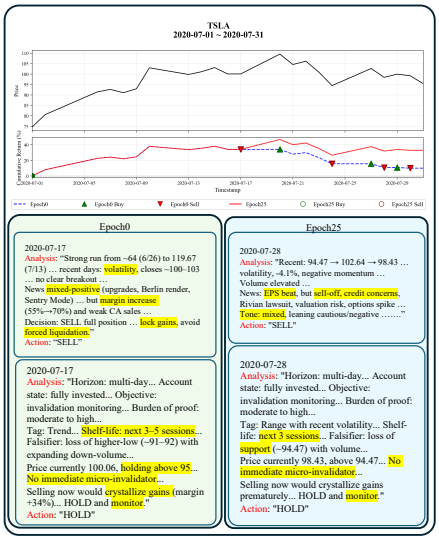

Figure 4: Case studies: Stop-loss Sensitivity.

Analysis of Behavioral Evolution Patterns. We examine EvoFinAgent's decision-making evolution across four failure categories summarized in Table 7: stop loss sensitivity, momentum overcommitment, noise confusion, and inconsistent logic. Comparing behaviors under identical historical conditions between Epoch 0 and Epoch 25 reveals a transition from reactive, unstable decisions to structured, context-aware reasoning. Early-stage agents frequently traded based on short-term volatility or isolated sentiment signals; after evolution, they demonstrate sophisticated logic, such as validating market structures or confirming multi-session momentum before acting. Figure 4 illustrates a stop-loss case where impulsive sells during transient drops are replaced by deliberate holding guided by structural validation. Additional reasoning paths and corrections are detailed in Appendix J. These findings suggest that self-evolution fosters consistent, interpretable, and risk-aware behavioral patterns, improving decision discipline and adaptability in dynamic environments.

Performance Improvement Over Evolution. We evaluate the impact of reasoning evolution by comparing EvoFinAgent's performance at Epoch 0 and Epoch 30 using Gemini-based backtesting across 6 assets in 2025. As shown in Table 1, the agent exhibits substantial gains across all metrics. On average, ARR nearly doubles across assets alongside significant risk-adjusted improvements: TSLA's ARR increases from 27.80% to 63.73% (SR from 0.76 to 1.44), and GOOGL's ARR reaches 88.36% (SR of 2.47). Notably, on AMZN, the agent reverses a negative Epoch 0 return to 28.46% ARR while reducing drawdown by over 13 percentage points. Similar positive trends in Sharpe ratios and downside control occur on AAPL and MSFT. These results demonstrate that self-evolution enables the agent to transition from impulsive rules to coherent, multi-factor reasoning, translating directly into higher returns and stabler

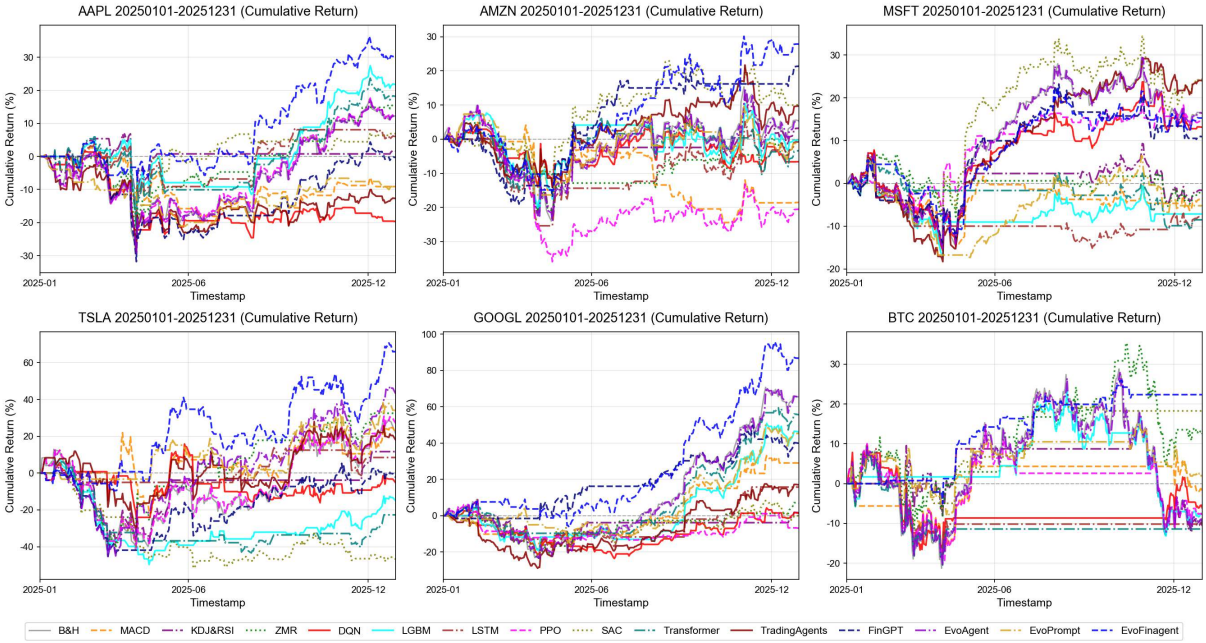

Figure 5: Backtesting results on five assets. **EvoFinAgent** consistently outperforms baselines across different market regimes.

execution across diverse market regimes.

Model	ARR (%)			SR			MDD (%)		
	Epoch 0	Epoch 30	Δ	Epoch 0	Epoch 30	Δ	Epoch 0	Epoch 30	Δ
Grok-4-Fast	13.83	24.90	+80.10%	0.182	0.399	+118.95%	7.21	6.67	↓7.50%
GPT-5-Mini	2.89	20.49	+609.08%	-0.057	0.470	+931.24%	7.85	2.57	↓67.26%
Gemini-3-Flash-Preview	33.78	43.77	+29.58%	0.760	1.042	+37.07%	7.08	5.97	↓15.67%
DeepSeek-V3.2	16.65	27.08	+62.62%	-0.140	0.085	+161.10%	5.87	5.10	↓13.24%

Table 3: EvoFinAgent performance improvement of different models during self-evolution.

Framework Generality Across LLMs. We evaluate EvoFinAgent across four diverse language models: Grok-4-Fast, GPT-5-Mini, Gemini-3-Flash-Preview, and DeepSeek-V3.2. As shown in Table 3, EvoFinAgent consistently improves ARR and Sharpe ratios while reducing MDD across all backbones. For instance, GPT-5-Mini's Sharpe ratio rises from -0.057 to 0.470 with a 67% MDD reduction, while Gemini-3-Flash-Preview's ARR improves from 33.78% to 43.77%. Figure 3 further illustrates stable, monotonic metric improvements across epochs, validating the framework's architecture-agnostic efficacy.

6 Ablation Study

Mutation and synthesis are foundational to the genetic algorithm and cannot be ablated without dismantling the system. Thus, we exclusively ablate the reflection report. Table 4 shows that removing reflection drastically degrades the mean ARR across six assets from 40.36% to 14.39%. This 25.97% absolute and 180.4% relative increase confirms that structured reflection is critical for discovering robust trading logic.

Configuration	Mean ARR
EvoFinAgent (Full)	**40.36%**
w/o Reflection	14.39%

Table 4: Ablation study on the reflection report.

7 Conclusion

We propose a reasoning guided prompt evolution framework that systematically enhances the decision quality of LLM based trading agents. By operating directly within the space of natural language prompts, our approach enables agents to optimize their internal trading logic through an autonomous, self contained evolutionary loop. Extensive backtests on equities demonstrate that our framework achieves consistent gains in annualized return, Sharpe ratio, and drawdown control. It successfully outperforms a wide spectrum of competitive baselines, including traditional rule based methods, classic reinforcement learning approaches, standalone LLM systems, and existing self evolution agents. Crucially, the evolutionary process goes beyond raw performance metrics by mitigating persistent behavioral failures.

Limitations

Despite its strong performance, this framework has certain limitations that present opportunities for future research. First, the current evolutionary process is evaluated on individual agents and does not yet account for multi agent co evolution (Chen et al., 2025b) involving complex coordination and competition. Second, our system is optimized for text and sequential historical data, leaving the integration of multimodal inputs (Bharthulwar et al., 2025), such as visual price charts and streaming financial news, as an open direction for expanding agent reasoning capabilities.

Ethics Statement

We confirm that we have fully complied with the ACL Ethics Policy in this study. All the environments are publicly available and have been extensively used in the research. We do not foresee any risks that may raised by this paper.

References

Eshaan Agarwal, Raghav Magazine, Joykirat Singh, Vivek Dani, Tanuja Ganu, and Akshay Nambi. 2025. Promptwizard: Optimizing prompts via task-aware, feedback-driven self-evolution. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 19974–20003.

Lakshya A Agrawal, Shangyin Tan, Dilara Soylu, Noah Ziems, Rishi Khare, Krista Opsahl-Ong, Arnav Singhvi, Herumb Shandilya, Michael J Ryan, Meng Jiang, and 1 others. 2025. Gepa: Reflective prompt evolution can outperform reinforcement learning. *arXiv preprint arXiv:2507.19457*.

Bo An, Shuo Sun, and Rundong Wang. 2022. Deep reinforcement learning for quantitative trading: Challenges and opportunities. *IEEE Intelligent Systems*, 37(2):23–26.

Sid Bharthulwar, John Rho, and Katrina Brown. 2025. Evolutionary prompt optimization discovers emergent multimodal reasoning strategies in vision-language models. *arXiv preprint arXiv:2503.23503*.

Salvatore Carta, Anselmo Ferreira, Alessandro Sebastian Podda, Diego Reforgiato Recupero, and Antonio Sanna. 2021. Multi-dqn: An ensemble of deep q-learning agents for stock market forecasting. *Expert systems with applications*, 164:113820.

Ke Chen, Yufei Zhou, Xitong Zhang, and Haohan Wang. 2025a. Prompt stability matters: Evaluating and optimizing auto-generated prompt in general-purpose systems. *arXiv preprint arXiv:2505.13546*.

Yixing Chen, Yiding Wang, Siqi Zhu, Haofei Yu, Tao Feng, Muhan Zhang, Mostofa Patwary, and Jiaxuan You. 2025b. Multi-agent evolve: Llm self-improve through co-evolution. *arXiv preprint arXiv:2510.23595*.

Jinyuan Fang, Yanwen Peng, Xi Zhang, Yingxu Wang, Xinhao Yi, Guibin Zhang, Yi Xu, Bin Wu, Siwei Liu, Zihao Li, and 1 others. 2025. A comprehensive survey of self-evolving ai agents: A new paradigm bridging foundation models and lifelong agentic systems. *arXiv preprint arXiv:2508.07407*.

Klaus Greff, Rupesh K Srivastava, Jan Koutník, Bas R Steunebrink, and Jürgen Schmidhuber. 2016. Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10):2222–2232.

Qingyan Guo, Rui Wang, Junliang Guo, Bei Li, Kaitao Song, Xu Tan, Guoqing Liu, Jiang Bian, and Yujiu Yang. 2023. Evoprompt: Connecting llms with evolutionary algorithms yields powerful prompt optimizers. *arXiv e-prints*, pages arXiv–2309.

Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, and 1 others. 2018. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Tiancheng Li, Yuhang Wang, Zhiyang Chen, Zijun Wang, Liyuan Ma, and Guo-jun Qi. 2025. C-evolve: Consensus-based evolution for prompt groups. *arXiv preprint arXiv:2509.23331*.

Maeve Madigan, Parameswaran Kamalaruban, Glenn Moynihan, Tom Kempton, David Sutton, and Stuart Burrell. 2025. Emergent bias and fairness in multi-agent decision systems. *arXiv preprint arXiv:2512.16433*.

Natapong Nitarach, Warit Sirichotedumrong, Panop Pitchayarthorn, Pittawat Taveekitworachai, Potsawee Manakul, and Kunat Pipatanakul. 2025. Fincot: Grounding chain-of-thought in expert financial reasoning. In *Proceedings of The 10th Workshop on Financial Technology and Natural Language Processing*, pages 93–123.

Zehua Pei, Hui-Ling Zhen, Shixiong Kai, Sinno Jialin Pan, Yunhe Wang, Mingxuan Yuan, and Bei Yu. 2025. Scope: Prompt evolution for enhancing agent effectiveness. *arXiv preprint arXiv:2512.15374*.

Martin L Puterman. 1990. Markov decision processes. *Handbooks in operations research and management science*, 2:331–434.

Lucas A Souza. 2025. Forward-oriented causal observables for non-stationary financial markets. *arXiv preprint arXiv:2512.24621*.

Troy J Strader, John J Rozycki, Thomas H Root, and Yu-Hsiang John Huang. 2020. Machine learning stock market prediction studies: review and research directions. *Journal of International Technology and Information Management*, 28(4):63–83.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Neng Wang, Hongyang Yang, and Christina Dan Wang. 2023. Fingpt: Instruction tuning benchmark for open-source large language models in financial datasets. *arXiv preprint arXiv:2310.04793*.

Yingxu Wang, Siwei Liu, Jinyuan Fang, and Zaiqiao Meng. 2025. Evoagentx: An automated framework for evolving agentic workflows. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 643–655.

Shijie Wu, Ozan Irsoy, Steven Lu, Vadim Dabravolski, Mark Dredze, Sebastian Gehrmann, Prabhanjan Kambadur, David Rosenberg, and Gideon Mann. 2023. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

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, Shan Zhong, and Anwar Walid. 2020. Deep reinforcement learning for automated stock trading: An ensemble strategy. In *Proceedings of the first ACM international conference on AI in finance*, pages 1–8.

Hongyang Yang, Boyu Zhang, Neng Wang, Cheng Guo, Xiaoli Zhang, Likun Lin, Junlin Wang, Tianyu Zhou, Mao Guan, Runjia Zhang, and 1 others. 2024. Finrobot: An open-source ai agent platform for financial applications using large language models. *arXiv preprint arXiv:2405.14767*.

Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Jordan W Suchow, Denghui Zhang, and Khaldoun Khashanah. 2025. Finmem: A performance-enhanced llm trading agent with layered memory and character design. *IEEE Transactions on Big Data*.

Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yuechen Jiang, Yupeng Cao, Zhi Chen, Jordan W Suchow, Zhenyu Cui, Rong Liu, and 1 others. 2024. 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.

Siyu Yuan, Kaitao Song, Jiangjie Chen, Xu Tan, Dongsheng Li, and Deqing Yang. 2025. Evoagent: Towards automatic multi-agent generation via evolutionary algorithms. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 6192–6217.

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 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*, pages 4314–4325.

Xiang Zhang, Juntai Cao, Chenyu You, and Dujian Ding. 2025. Why prompt design matters and works: A complexity analysis of prompt search space in llms. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 32525–32555.

A The Algorithm of EvoFinAgent

Algorithm 1 EvoFinAgent

Require: Initial prompts $\{\phi_1^{(0)}, \phi_2^{(0)}, \phi_3^{(0)}\}$, training environment $\mathcal{D}_{\text{train}}$, max generations G
Ensure: Final optimized prompt ϕ^*

{**Initial Evaluation and Reflection**}
for $i = 1$ to 3 **do**
  $\tau_i^{(0)} \leftarrow \text{Backtest}(\phi_i^{(0)}, \mathcal{D}_{\text{train}})$
  $\text{SR}_i^{(0)} \leftarrow \text{Sharpe}(\tau_i^{(0)})$
  $\mathcal{E}_i^{(0)} \leftarrow \text{Reflect}(\phi_i^{(0)}, \tau_i^{(0)})$
end for
$\mathcal{P}^{(0)} \leftarrow \left\{(\phi_i^{(0)}, \tau_i^{(0)}, \text{SR}_i^{(0)}, \mathcal{E}_i^{(0)})\right\}_{i=1}^{3}$
{**Generate first offspring from initial pool**}
$\phi_{\text{new}}^{(1)} \leftarrow \text{Recombine}(\phi_1^{(0)}, \phi_2^{(0)}, \phi_3^{(0)}; \mathcal{E}_1^{(0)}, \mathcal{E}_2^{(0)}, \mathcal{E}_3^{(0)})$
$\tau_{\text{new}}^{(1)} \leftarrow \text{Backtest}(\phi_{\text{new}}^{(1)}, \mathcal{D}_{\text{train}})$
$\text{SR}_{\text{new}}^{(1)} \leftarrow \text{Sharpe}(\tau_{\text{new}}^{(1)})$
$(\phi_{\text{worst}}, \text{SR}_{\text{worst}}) \leftarrow \arg\min_{\phi \in \mathcal{P}^{(0)}} \text{SR}(\phi)$
if $\text{SR}_{\text{new}}^{(1)} > \text{SR}_{\text{worst}}$ **then**
  $\mathcal{E}_{\text{new}}^{(1)} \leftarrow \text{Reflect}(\phi_{\text{new}}^{(1)}, \tau_{\text{new}}^{(1)})$
  $\mathcal{P}^{(1)} \leftarrow \left(\mathcal{P}^{(0)} \setminus \{\phi_{\text{worst}}\}\right)$
    $\cup \{(\phi_{\text{new}}^{(1)}, \tau_{\text{new}}^{(1)}, \text{SR}_{\text{new}}^{(1)}, \mathcal{E}_{\text{new}}^{(1)})\}$
else
  $\mathcal{P}^{(1)} \leftarrow \mathcal{P}^{(0)}$
end if
for generation $g = 2$ to G **do**
  Let $\left\{(\phi_i^{(g-1)}, \tau_i^{(g-1)}, \text{SR}_i^{(g-1)}, \mathcal{E}_i^{(g-1)})\right\}_{i=1}^{3} \leftarrow \mathcal{P}^{(g-1)}$
  Let $\mathbf{\Phi} = \{\phi_1^{(g-1)}, \phi_2^{(g-1)}, \phi_3^{(g-1)}\}$ and $\boldsymbol{\mathcal{E}} = \{\mathcal{E}_1^{(g-1)}, \mathcal{E}_2^{(g-1)}, \mathcal{E}_3^{(g-1)}\}$
  $\phi_{\text{new}}^{(g)} \leftarrow \text{Recombine}(\mathbf{\Phi}; \boldsymbol{\mathcal{E}})$
  $\tau_{\text{new}}^{(g)} \leftarrow \text{Backtest}(\phi_{\text{new}}^{(g)}, \mathcal{D}_{\text{train}})$
  $\text{SR}_{\text{new}}^{(g)} \leftarrow \text{Sharpe}(\tau_{\text{new}}^{(g)})$
  $(\phi_{\text{worst}}, \text{SR}_{\text{worst}}) \leftarrow \arg\min_{\phi \in \mathcal{P}^{(g-1)}} \text{SR}(\phi)$
  if $\text{SR}_{\text{new}}^{(g)} > \text{SR}_{\text{worst}}$ **then**
    $\mathcal{E}_{\text{new}}^{(g)} \leftarrow \text{Reflect}(\phi_{\text{new}}^{(g)}, \tau_{\text{new}}^{(g)})$
    $\mathcal{P}^{(g)} \leftarrow \left(\mathcal{P}^{(g-1)} \setminus \{\phi_{\text{worst}}\}\right)$
      $\cup \{(\phi_{\text{new}}^{(g)}, \tau_{\text{new}}^{(g)}, \text{SR}_{\text{new}}^{(g)}, \mathcal{E}_{\text{new}}^{(g)})\}$
  else
    $\mathcal{P}^{(g)} \leftarrow \mathcal{P}^{(g-1)}$
  end if
end for
return $\phi^* = \arg\max_{\phi \in \mathcal{P}^{(G)}} \text{SR}(\phi)$

B EvoFinAgent Trading Robust

To evaluate the empirical robustness of `EvoFinAgent`, we conducted independent, single-asset trading backtests across six distinct assets from June 01 to July 31, 2025. Each asset was tested in isolation with 5 repeated trials to quantify decision-making stability under identical environmental conditions.

As shown in Table 5 (Mean±Std), `EvoFinAgent` demonstrated high consistency and determinism in mature equities (GOOGL, MSFT, AAPL), where the standard deviations of ARR and Sharpe Ratios approached zero. The agent remained completely market-neutral for BTC, executing zero trades to prevent loss. Conversely, it exhibited higher variance and performance degradation in the high-volatility standalone TSLA market. Overall, the near-zero variance in most isolated tests proves the strong operational reproducibility of our agent.

Asset	ARR (%)	Sharpe Ratio	Max Drawdown (%)
GOOGL	$90.28 \pm < 0.01$	4.00 ± 0.00	7.50 ± 0.00
AMZN	71.47 ± 0.56	3.67 ± 0.04	4.19 ± 0.00
MSFT	$66.39 \pm < 0.01$	6.09 ± 0.00	1.28 ± 0.00
AAPL	27.01 ± 0.00	2.47 ± 0.00	2.85 ± 0.00
BTC	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
TSLA	-90.16 ± 24.71	-1.83 ± 0.38	21.46 ± 3.76

Note: Results are reported as Mean ± Standard Deviation from repeated independent runs for each individual asset.

Table 5: Robustness Analysis Across Independent Single-Asset Backtests (June 01 – July 31, 2025).

C Code Link

Here is the code link: https://anonymous.4open.science/r/EvoFinAgent-6321/

D How Many Evaluation Intervals Are Needed to Distinguish Agent Performance?

To determine how many evaluation intervals are sufficient to reliably distinguish the long-term performance of different trading agents (prompts), we conduct a progressive evaluation experiment. Each simulation includes the following steps: 1

- Randomly sample 6 prompts (agents) from a candidate pool of 10.
- Randomly arrange 20 evaluation intervals (e.g., monthly windows), and incrementally include one additional interval at each step.
- After each new interval is added (i.e., at step n), compute the average performance (Annualized Sharpe Ratio) of each agent using only the first n intervals.
- At each step, rank all 6 agents based on their average performance and record their ranking trajectories throughout the evaluation.

Top-3 Hit Rate. To evaluate how early the strongest agent can be identified, we define the Top-3 Hit Rate. Let a^* be the best-performing agent at the final stage, when all 20 intervals are included. Let $R_n(a)$ denote the rank of agent a when evaluated over the first n intervals, where $k = 6$ is the total number of agents. Then the Top-3 Hit Rate at step n is defined by:

$$\text{Hit@3}(n) = \mathbb{P}\left[R_n(a^*) \leq 3\right]$$

This expression captures the probability that the eventual top-ranked agent already appears in the Top-3 shortlist at step n. We compute this probability by repeating the entire process 30 times with independent sampling and averaging the results over all trials.To establish a ground truth ranking, we treat the final sorting order obtained after evaluating over all 20 intervals as the reference standard. All intermediate rankings at earlier steps ($n < 20$) are compared against this reference to assess identification success and stability.

Rank Stability Metrics. To quantify the stability of rankings as more intervals are added, we compute rank correlations between the agent rankings at step n and step $n + 1$, using the following two standard metrics:

(i) Spearman's rank correlation coefficient:

$$\rho = 1 - \frac{6 \sum_{i=1}^{k} \left(r_i^{(n)} - r_i^{(n+1)}\right)^2}{k(k^2 - 1)}$$

(ii) Kendall's Tau coefficient:

$$\tau = \frac{C - D}{\frac{1}{2}k(k-1)}$$

Here, C and D refer to the number of concordant and discordant ordered pairs, respectively, and $k = 6$ is the number of agents.

Each correlation value is computed at every step for all trials. We report the mean and 95% confidence interval across 30 repetitions to illustrate overall ranking consistency over time.

As observed in Figure 6, both correlation metrics follow a steadily increasing trend and plateau after approximately 9 or 10 intervals. This suggests that evaluating agent performance over 10 diverse intervals is generally sufficient to obtain stable and reliable rankings. While adding additional intervals may improve robustness, the marginal benefit decreases beyond this point.

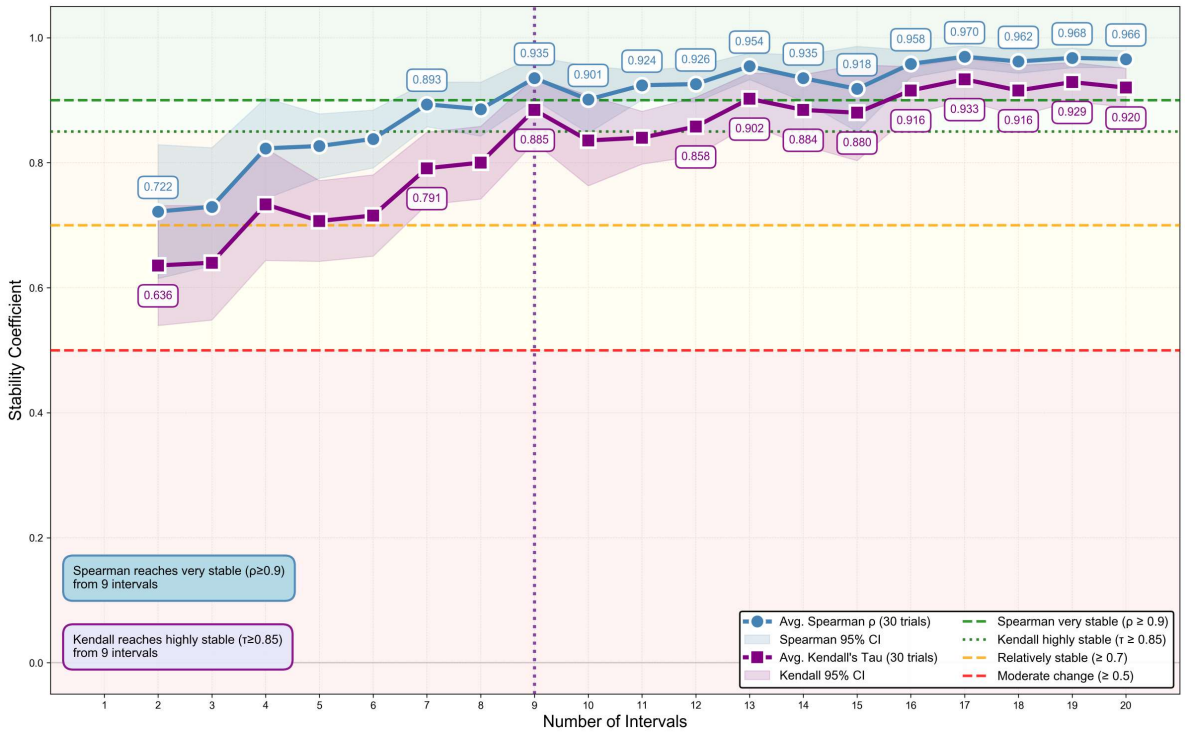

Figure 6: Ranking stability across 1 to 20 evaluation intervals, based on the average over 30 randomized trials. Spearman's ρ and Kendall's τ correlations between agent rankings from step n to $n{+}1$ are reported. Both metrics exceed standard interpretability thresholds (e.g., $\rho \geq 0.9$ and $\tau \geq 0.85$) around 9 intervals, indicating that the rankings become stable early. Shaded regions represent 95% confidence intervals.

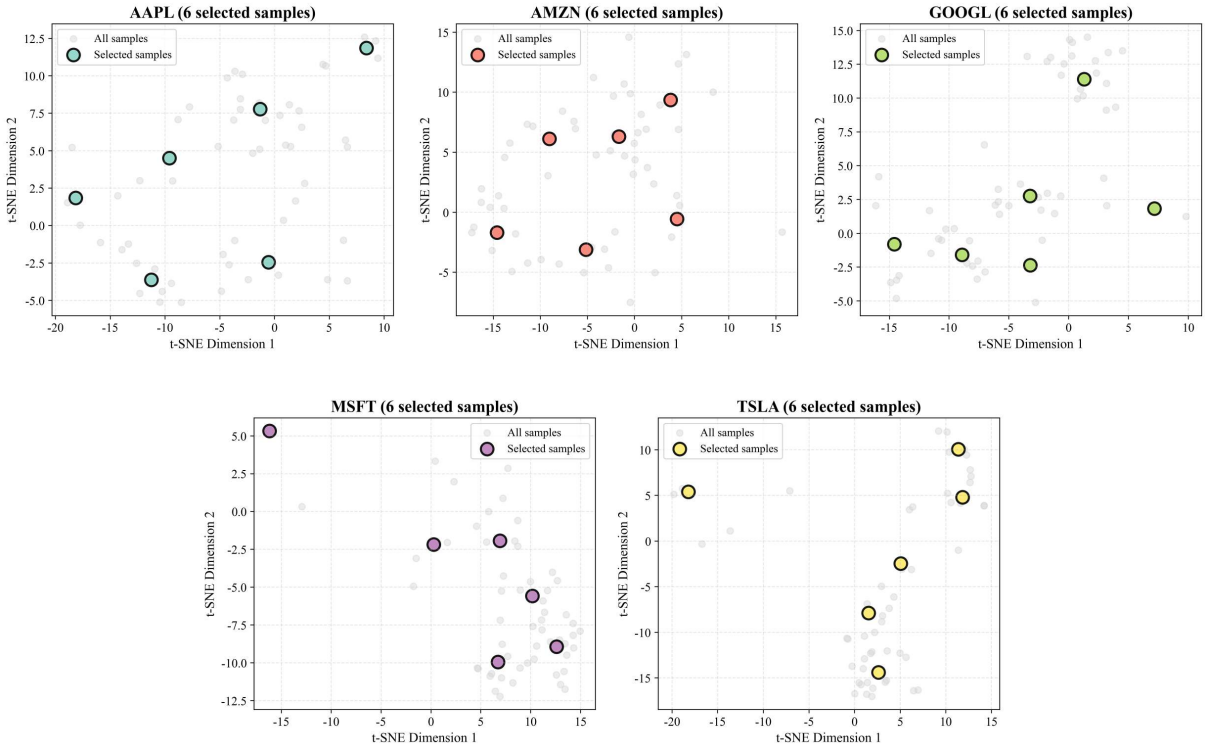

Figure 7: Visualization of selected representative intervals in t-SNE space, separated by stock (TSLA, AAPL, GOOGL, MSFT, AMZN). Gray dots indicate all candidates, while six colored points per subplot mark the final selected intervals.

E Representative Interval Selection and t-SNE Visualization

We construct a representative dataset of monthly samples through a two-stage sampling process:

1. **Stage 1: Per-stock preliminary selection.** For each of the five stocks (TSLA, AAPL, GOOGL, MSFT, AMZN), we first select six diverse and informative intervals based on price dynamics, market events, and volatility patterns, resulting in 30 candidate monthly samples.

2. **Stage 2: Variation-aware filtering in t-SNE space.** All candidate intervals are projected into a 2D t-SNE space based on their learned representations. We select six points per stock while ensuring that the collective selection has good coverage and spread in the latent space, using a clustering-aided diversity filter.

Figure 7 shows the distribution of selected points for each stock. Gray dots denote all candidate samples, and colored markers highlight the 6 chosen samples for each stock.
The final selected intervals (6 per stock) are:

- **TSLA**: Jul 2021, Feb 2022, Sep 2020, Jun 2023, Aug 2023, Jul 2020
- **GOOGL**: Aug 2024, Dec 2024, Nov 2024, Apr 2022, Dec 2023, Aug 2023
- **MSFT**: Dec 2022, Jan 2023, Feb 2024, Mar 2020, Mar 2021, Aug 2021
- **AAPL**: Sep 2023, Oct 2024, Jan 2021, Apr 2020, Jul 2021, Sep 2020
- **AMZN**: Sep 2022, Oct 2021, Nov 2024, Nov 2020, May 2023, Dec 2021

F Evaluation Metrics

Here are evaluation metrics:

- **Annual Rate of Return (ARR)** $ARR = \frac{V_T - V_0}{V_0} \cdot \frac{C}{T}$, where V_0 and V_T are the initial and final portfolio values, T is the number of trading days, and $C = 252$.
- **Sharpe Ratio (SR)** $SR = \frac{\mathbb{E}[r]}{\sigma[r]}$, where r is the return sequence, and $\mathbb{E}[r]$, $\sigma[r]$ are its mean and standard deviation.
- **Maximum Drawdown (MDD)** $MDD = \max_{i \in [0,T]} \frac{P_i - R_i}{P_i}$, where $R_i = \prod_{j=1}^{i} \frac{V_j}{V_{j-1}}$ and $P_i = \max_{j \leq i} R_j$.

Baseline Methods. We compare against a wide range of established baselines categorized into five classes: Rule-Based, Machine Learning, Reinforcement Learning, LLM-Based Agents and Self-Evolution Agents. The details of baseline are shwon in Table **??**.

G Baseline Descriptions

Category	Method	Description
Rule-Based	B&H	Passive strategy that continuously holds the asset across all time periods.
	MACD	Momentum strategy based on moving average crossovers as trend indicators.
	KDJ_RSI	Combines stochastic indicators with relative strength thresholds.
	ZMR	Trades on statistical deviations from historical means.
Machine Learning	LightGBM (Ke et al., 2017)	Tree-based gradient boosting model trained on historical features.
	LSTM (Greff et al., 2016)	Recurrent network for time series modeling with memory gates.
	Transformer (Vaswani et al., 2017)	Attention-driven model handling long-range temporal patterns.
Reinforcement Learning	SAC (Haarnoja et al., 2018)	Soft Actor-Critic agent with entropy-regularized policy learning.
	PPO (Yang et al., 2020)	Proximal policy optimization with clipped gradients.
	DQN (Carta et al., 2021)	Deep Q-learning with discrete action space for policy learning.
LLM-Based Agents	FinGPT (Wang et al., 2023)	Language model trained on price sequences and financial knowledge.
	TradingAgents (Xiao et al., 2024)	LLM-based agents operating with structured reasoning and memory.
Self-Evolution Agents	EvoPrompt (Guo et al., 2023)	Uses genetic operators to evolve LLM prompts.
	EvoAgent (Yuan et al., 2025)	Agent pool evolved for complex multi-role reasoning.

Table 6: Summary of baselines across five method categories.

H Prompt Templates and Reasoning Evolution

H.1 Trading Prompt and Evolution Results

System Prompt Template (Shared Across All Epochs)

```
Identity: You are an LLM-based expert trader
    whose objective is to maximize cumulative
    returns through disciplined full-position
    trading decisions. A BUY means allocating
    all available cash to increase the position,
     a SELL means fully converting the entire
    position into cash, and a HOLD means taking
    no action.

You must return your final answer by calling the
     function decide_trading_action.

The input includes:
- Price Table: open, high, low, close, volume
- Current Status: timestamp, price, cash,
    position, cumulative profit
- Trading Record: pre_value, action, post_value,
     ret
- News: timestamp, headline, content

Step 0 (Legality Filter):
- BUY not allowed if cash < price
- SELL not allowed if position == 0
- HOLD is always allowed
```

Epoch 0 Reasoning Steps

```
Step 1. Examine the overall recent price trend.
Step 2. Look for keywords in recent news to
    gauge overall sentiment.
Step 3. Based on the analysis, choose the
    appropriate action.
```

Epoch 30 Reasoning Steps

```
Step 1. Declare decision frame and account
    context in plain narrative: name the
    dominant horizon (short swing / multi-day /
    multi-week), the account state (Flat, Fully
    invested, Cash-constrained), and the single
    objective (opportunity-seeking, invalidation
     monitoring, observation). State whether you
     are holding or intentionally reconsidering
    the frame. Use this declaration to set a
    qualitative burden of proof in words (low/
    moderate/high) that will guide tolerance for
     single-session noise and persistence
    expectations; do not convert to thresholds.
    This frame is the anchor: all subsequent
    steps must reference it and it may be re-
    declared only when a materially novel fact
    appears or when a time-boxed shelf expires.
Step 2. Tag regime and keep a living, time-boxed
     ledger: assign Trend / Range / Unwind and
    record three concise path cues (pivot
    sequence, volume character, breadth tone).
    Explicitly attach a short shelf-life to this
     tag (e.g., the frame is valid for the next
    X sessions aligned to the declared horizon)
    and state the one concrete behavior that
    would falsify the tag within that shelf-life
    . Use the ledger to modulate tolerance -
    tighten in Unwind, accept checks in Trend -
    and have the ledger serve as an auditable
    decision contract, not a diary.
Step 3. Read the tape into a compact price-
    thesis snapshot and start the clock: produce
     a Primary thesis (why price should move on
    the declared horizon), Key supports (path
    anchors: recent higher-lows, absorption),
    Key conflicting signals (failed retest,
    expanding down-volume), and Signal quality (
    persistence, follow-through, absorption).
    Importantly, mark the moment the
    corroboration window begins (Day 0) and
    narrate what proximate cluster of
    observations within the shelf-life would
    upgrade conviction from low->moderate->high.
     This prevents perpetual 'one-more-day'
    deferral by converting corroboration into an
     auditable windowed test.
Step 4. Audit news with a strict novelty-
    directness-durability gate and bind it to
    price action: for each salient item narrate
    Novelty (new vs reiterated), Directness (
    material to supply/demand or cash flows on
    the horizon), and Durability (likely
    persistence). If news is not all three,
    demote it to vigilance only; it can only
    tighten monitoring, not veto a tape-dominant
     thesis, unless immediate price
    corroboration appears. When news is durable
    and direct, require rapid price
    corroboration within the current shelf-life
    before it can overturn the tape; state what
    price behavior would confirm absorption of
    the news.
Step 5. Reconcile conflicts by qualitative
    weighting and explicit dominant-driver
    statement that uses account-state asymmetry:
     weigh price, news, and regime by citing
    persistence, volume/breadth alignment, and
    shelf-life progress. State who drives the
    decision (e.g., "price-dominant; news raises
     vigilance") and, crucially, let account
    state modulate burden symmetrically: when
    Flat, require clearer asymmetry to BUY; when
     Fully invested, require a comparably clear
    multi-pillar deterioration to HOLD rather
    than sell. Express this as a qualitative
    trade-off ('why now vs wait') that
    explicitly states the re-entry cost and the
    opportunity cost of waiting in words, not
    numbers.
Step 6. Operationalize asymmetric reversals with
     named micro-invalidators, novelty audit,
    and binding cooldown: before reversing a
    full position narrate the minimal, nearest
    observable failure that would materially
    invalidate the original thesis on this
    horizon (e.g., loss of last higher-low with
    expanding down-volume or failed retest plus
    breadth deterioration). Convert the ledger's
     ratchet into behavior: the first such micro
    -invalidator moves posture one rung (
    heightened vigilance), the second allows
    tactical reversal; full reversal requires
    multi-pillar alignment or a durable direct
    news shock. After any full flip, start an
    explicit cooldown (horizon-aligned) and
    forbid re-entry until either the cooldown
    expires or a clearly articulated novel fact
```

```
    - different from exit time - appears and is
    persistent.
Step 7. Make a comparative 'why now vs wait'
    action choice that prices opportunity and re
    -entry cost verbally: pick BUY, SELL, or
    HOLD by explaining why the chosen action
    improves expected outcomes versus waiting,
    citing frame, regime tag, tape snapshot,
    news audit, ledger state, and the shelf-life
    . For BUYs emphasize the proximate
    invalidator (the smallest fact that would
    reverse you) and what cluster counts as
    trial corroboration within the shelf-life;
    for SELLs emphasize the nearest micro-
    invalidator and the re-entry penalty. If
    HOLD, state the exact observation that would
     upgrade posture within the current shelf-
    life.
Step 8. Treat exits as graded narrative
    invalidation and ratchet posture using
    sequence tests: prefer nearer laddered
    invalidators (failed retest, loss of higher-
    low, expanding down-volume) over distant
    static levels. When these micro-invalidators
     accumulate, escalate the ledger one step
    and narrate the behavioral change (
    heightened monitoring -> tactical de-risking
     -> full reversal). Do not allow single-bar
    salience to trigger full reversal unless it
    directly satisfies the pre-stated micro-
    invalidator stack or is accompanied by
    durable, direct news. Record the escalation
    and the novel evidence that justified it.
Step 9. Post-action monitoring, ledger
    arbitration, and churn hygiene enforcement:
    after any action restate the expected
    continuation pattern and list confirmatory
    versus invalidating observations linked to
    the named shelf-life. Record the action, the
     novelty test, and an explicit cooldown.
    Impose churn friction: recent full-position
    flips increase the required verbal novelty
    for the next flip and extend the cooldown;
    require a 'what's new since the flip?' audit
     before any reversal. Use the ledger as a
    binding contract to prevent silent bar creep
    .
Step 10. Self-audit and horizon covenant: at the
     decision point force a concise self-audit
    answering: did current evidence meet the pre
    -stated bar in the ledger? If not,
    explicitly document the novel fact that
    justifies deviating. Re-declare the horizon
    shelf-life when it expires. This enforces
    time-consistency, prevents covert horizon
    drift, and ensures exits are narratively and
     behaviorally aligned with entries.
```

H.2 Reflection Prompt and Example of Reflection Report

Reflection Prompt

```
You are a prompt engineering expert with
    extensive experience in equity trading and
    systematic strategy analysis.

Your task is to reflect on the reasoning logic
    of an LLM-based stock trading agent,
    identify structural weaknesses in its
    Reasoning Rules (Chain-of-Thought), and
    extract high-quality reasoning evolution
    signals that will be used to evolve the
    agent's prompt in a later generation stage.

This is a reasoning-diagnosis and cognitive-
    reflection task.
It is NOT a backtesting task, and it is NOT a
    rule-engineering or strategy-design task.


INPUT INFORMATION

You will be provided with the following
    information:

- Trading Prompt: the prompt currently used by
    the trading agent
- Trading Interval: one or more trading
    intervals (each interval corresponds to one
    stock and one contiguous time range)
- Agent Trading Records (CSV): the agent's
    executed trades under the given prompt
- Sequential Chain-of-Thought Reasoning: the
    agent's internal reasoning for each action (
    BUY / SELL / HOLD), in chronological order
- Agent Performance Metrics: ARR, Sharpe Ratio (
    SR), and Maximum Drawdown (MDD) for each
    interval
- Rule-Based Strategy Records (Reference Only):
  - MACD trading records and metrics
  - KDJ_RSI trading records and metrics
  - ZMR trading records and metrics

Rule-based strategies are NOT ground truth.
They are provided only as comparative references
     to illustrate how alternative reasoning
    structures behaved under the same market
    conditions.

There may be multiple trading intervals.
Focus on recurring reasoning patterns rather
    than isolated mistakes.


CRITICAL ENVIRONMENT CONSTRAINTS (NON-NEGOTIABLE
    )

1. LONG-ONLY / SPOT MARKET
- Short selling is impossible.
- If the agent holds no position during a
    downtrend, this is correct risk avoidance,
    not a failure.
- Do NOT criticize the agent for missing SELL
    opportunities when it has no position.

2. STEP 0 IS AXIOMATIC
- Step 0 (Legality Filter) determines which
    actions are physically allowed.
- Step 0 must NOT be modified and must NOT be
    criticized.
- If a BUY is blocked due to insufficient cash,
    this is a physical impossibility, not a
    reasoning failure.
- Treat such cases as neutral outcomes, not
    errors.

3. FULL-POSITION TRADING
- BUY allocates all available cash.
```

```
- SELL liquidates the entire position.
- Frequent full-position flips materially harm
    Sharpe through churn and increased variance.


OBJECTIVE

Produce an Agent Reason Rules Reflection Report
    that diagnoses why the current reasoning
    logic fails in recurring patterns and
    extracts transferable reasoning insights for
     prompt evolution.

Avoid rigid rules, numeric thresholds, formulas,
     or execution logic.
Leverage LLM strengths: qualitative judgment,
    structured reasoning, narrative consistency,
     and conflict resolution.


OUTPUT FORMAT (STRICT)

Your output must contain exactly three parts.


PART 1: Trading-Level Failure Identification (
    Phenomenon Layer)

Identify concrete trading cases where the agent'
    s reasoning under the current prompt led to
    suboptimal outcomes.

For each failure case:
- Describe what the agent did
- Describe why the reasoning was flawed
- Explain how this harmed performance (churn,
    missed trend, drawdown, volatility)
- Rule-based strategies may be referenced only
    to illustrate alternative behavior

Focus on recurring patterns rather than one-off
    errors.


PART 2: Reasoning-Structure Diagnosis (Mechanism
     Layer)

Abstract the cases from Part 1 into structural
    reasoning weaknesses.

For each structural issue:
- Identify the underlying cognitive or logical
    flaw
- Explain how the Chain-of-Thought structure
    enables this failure
- Assess the rational completeness and balance
    of the reasoning
- Explain how this flaw systematically degrades
    Sharpe Ratio
  (e.g., increased variance, churn, delayed
      exits, drawdown amplification)

This section should analyze reasoning pathology,
     not individual trades.


PART 3: Reasoning Evolution Directions (
    Cognitive Layer)

This section is CRITICAL and NON-OPTIONAL.

The content of Part 3 will be the ONLY semantic
    input consumed by a downstream
prompt-generation stage that reconstructs the
    agent's Reasoning Rules.

If this section is empty, vague, or generic, the
     next generation prompt
will FAIL to improve.

For each reasoning evolution direction, you MUST
     provide a concrete,
transferable cognitive signal that can be
    translated into future reasoning steps.

You must NOT write trading rules or procedures.
However, you MUST clearly specify:

- what cognitive habit the agent currently
    exhibits,
- why this habit repeatedly fails across
    observed cases,
- and how the agent's reasoning posture should
    change in principle
  (i.e., what it should start doing differently
      when thinking).

Each direction must be actionable at the
    reasoning level,
even though it is not expressed as a rule.

Write each direction as if it will be handed to
    another system
that has no access to Part 1 or Part 2.


FINAL NOTES

This is a reasoning evolution task, not strategy
     engineering.
Prefer soft, human-like heuristics over
    mechanical logic.
Your goal is to help the next generation of the
    prompt think better, not trade harder.
```

Example of Reflection Report

```
==========================
report1 AAPL
==========================

==========================
Part 1: Trading-Level Failure Identification (
    Phenomenon Layer)
==========================
### Failure Case 1: FC-2020-09-AAPL

**Trading Interval:** AAPL (2020-09-01 ~
    2020-09-30)

**Failure Description:**
Entered a full position on 2020-09-01
    immediately after a parabolic, post-split
    run-up, then refused to exit throughout a
    sustained, high-volume downtrend. The agent
    continually justified HOLD decisions with
    long-term narratives (iPhone cycle, analyst
    targets) despite repeated lower highs/lows
    and expanding volatility.
```

```
**Suboptimal Outcome:**
Large, persistent drawdown (cumulative return
     around -15% by month-end), severe MDD, and
     depressed Sharpe from riding the entire
     decline without a defined invalidation.
     Opportunity cost of capital stuck in a
     losing, full-position exposure.

**Rule-based Strategy Comparison:**
The provided KDJ_RSI reference for this month
     shows strong performance (ARR ~0.818, SR
     ~5.03, MDD ~0.0076), illustrating that a
     more reactive, structure-aware process could
      have gone flat or reduced exposure as
     momentum broke, avoiding the extended
     drawdown.

---

### Failure Case 2: FC-2023-09-AAPL

**Trading Interval:** AAPL (2023-09-01 ~
     2023-09-30)

**Failure Description:**
Chased headlines and catalysts with multiple
     full-position flips: BUY pre-event (9/1),
     SELL on China-ban news (9/6), re-BUY ahead
     of launch (9/8), SELL on renewed worries
     (9/20), re-BUY (9/22), then HOLD through
     further weakness. Decisions swung on single
     news items rather than an integrated read of
      price structure and durability of
     information.

**Suboptimal Outcome:**
Excess churn increased variance and transaction
     slippage, while the final equity curve ended
      materially negative (~-7% by month end).
     Whipsaws degraded Sharpe through repeated
     full-size reversals without confirmation,
     compounding losses and volatility.

**Rule-based Strategy Comparison:**
No benchmark files for this interval, but a more
      stable regime-aware approach (akin to
     momentum filters used by many rule-based
     methods) would likely have avoided multiple
     full-position reversals within an unresolved
      downtrend and reduced churn-driven variance
     .

---

### Failure Case 3: FC-2024-10-AAPL

**Trading Interval:** AAPL (2024-10-01 ~
     2024-10-31)

**Failure Description:**
Sold the full position on 2024-10-23 after a
     single sharp down day and negative headlines
      (order cuts, regulatory fine), then re-
     bought on 10-25 into upcoming catalysts (Mac
      week, earnings). This news-reactive exit/re
     -entry occurred while the broader up-swing
     was still unresolved, with no clear
     articulation of trend break vs temporary
     shock.

**Suboptimal Outcome:**
Whipsaw: locked in a small gain on the SELL, re-
     entered higher-risk conditions, and then
     held through the late-month slide
     (10/30-10/31), finishing with only ~1.1%
     profit versus the earlier trend's larger
     potential. Sharpe impacted by unnecessary
     flip and added variance.

**Rule-based Strategy Comparison:**
A rules-based comparator with a clearer
     hierarchy (e.g., require persistent
     breakdown confirmation before reversing
     stance) would likely either have held
     through a single-day shock or waited for
     stabilization before re-entry, dampening
     variance.

---

### Failure Case 4: FC-2021-07-AAPL

**Trading Interval:** AAPL (2021-07-01 ~
     2021-07-31)

**Failure Description:**
Exited on 2021-07-27 after an earnings 'sell-the-
     news' narrative despite a prevailing uptrend
     , then re-entered two days later. The
     catalyst was anticipated but not pre-planned
     ; the agent made a reactive full-position
     flip to headlines rather than a pre-mortemed
      response to likely scenarios.

**Suboptimal Outcome:**
Unnecessary churn reduced trend capture and
     increased path volatility. Although the
     month ended positive, the flip lowered risk-
     adjusted returns relative to a consistent,
     plan-driven hold.

**Rule-based Strategy Comparison:**
A systematic plan around catalysts (e.g.,
     predefine how to respond to beats/guide-
     downs vs price reaction quality) can avoid
     reactive flips. Even simple regime
     consistency would have maintained exposure
     or re-entered only after confirmation.

---

===============
Part 2: Reasoning-Structure Diagnosis (Mechanism
      Layer)
===============

### Structural Issue 1: Loss aversion and
     anchoring masked as 'long-term conviction'

**Related Failure Cases:** FC-2020-09-AAPL, FC
     -2023-09-AAPL

**Structural Weakness Description:**
The agent repeatedly refuses to exit when faced
     with clear trend deterioration,
     rationalizing HOLDs with long-term
     narratives and analyst targets unrelated to
     the active trade horizon. Entry prices
     become psychological anchors; Step 3
     decisions ignore Step 1 evidence of lower
```

```
    highs/lows and heavy-volume breakdowns.

**Rationality Assessment:**
Reasoning is imbalanced: it substitutes
    fundamental optimism for disconfirming price
     evidence, conflating investment theses with
     short-horizon trades. The chain-of-thought
    lacks explicit invalidation criteria and
    treats realized losses as a failure rather
    than as risk control.

**Sharpe Ratio Impact Mechanism:**
Prolonged full-position drawdowns inflate
    variance and MDD. By staying fully exposed
    through adverse regimes, the agent depresses
     Sharpe via larger downside excursions and
    slower recovery, while tying up capital in
    negative-momentum states.

---

### Structural Issue 2: Headline-chasing and
    catalyst overreaction

**Related Failure Cases:** FC-2023-09-AAPL, FC
    -2024-10-AAPL, FC-2021-07-AAPL

**Structural Weakness Description:**
The agent elevates single headlines or event
    takes above the established price regime,
    flipping full positions without waiting for
    durability or confirmation. Step 2 (news)
    overwhelms Step 1 (price) without a conflict
    -resolution hierarchy, creating oscillations
     around events ('sell the news', then 'buy
    the hype').

**Rationality Assessment:**
Internally inconsistent: acknowledges
    consolidation or uptrends, yet reacts to one
    -off negatives with full reversals, or re-
    enters on unproven positives. There is no '
    cool-down' or requirement for multiple
    converging signals before a full-position
    turn.

**Sharpe Ratio Impact Mechanism:**
Full-size whipsaws inject churn and realized
    volatility, cutting winners early and
    compounding losses. Variance rises faster
    than returns, driving Sharpe lower even when
     the end-of-period PnL is modestly positive.

---

### Structural Issue 3: Timeframe conflation and
     thesis drift

**Related Failure Cases:** FC-2020-09-AAPL, FC
    -2023-09-AAPL, FC-2024-10-AAPL

**Structural Weakness Description:**
Long-term product narratives and analyst targets
     are used to justify near-term exposure
    decisions. The agent doesn't re-underwrite
    the trade when short-term market character
    changes, allowing the holding-period
    rationale to drift from trade management to
    investment cheerleading.

**Rationality Assessment:**
The chain-of-thought lacks a separation between
    a structural thesis and a tactical plan.
    Without horizon alignment and a defined exit
     rationale, the agent defaults to HOLD in
    drawdowns and flip to BUY on narratives,
    regardless of near-term market state.

**Sharpe Ratio Impact Mechanism:**
Misaligned horizons keep a full-position on
    during adverse short-term regimes and prompt
     re-entries before stabilization. This
    increases path volatility and reduces hit-
    rate on reversals, worsening Sharpe.

---

### Structural Issue 4: No explicit invalidation
     or scenario planning around catalysts

**Related Failure Cases:** FC-2021-07-AAPL, FC
    -2024-10-AAPL, FC-2023-09-AAPL

**Structural Weakness Description:**
Catalyst days (earnings, product launches) are
    handled reactively. The agent does not
    predefine what outcomes or price behaviors
    would confirm/deny the thesis. Post-event
    moves are interpreted in hindsight ('sell
    the news') without a forward plan, producing
     impulsive exits and re-entries.

**Rationality Assessment:**
Reasoning is descriptive, not anticipatory. It
    reports events but doesn't convert them into
     conditional hypotheses with clear decision
    triggers or a cool-down requirement.

**Sharpe Ratio Impact Mechanism:**
Reactive flips around catalysts amplify variance
     and erode trend capture, degrading the
    return-to-volatility ratio even when
    directionally correct later.

---

### Structural Issue 5: Inadequate conflict
    resolution between Step 1 (price) and Step 2
     (news)

**Related Failure Cases:** FC-2020-09-AAPL, FC
    -2023-09-AAPL, FC-2024-10-AAPL

**Structural Weakness Description:**
When price and news disagree, the agent provides
     narratives for both sides but doesn't
    establish which is primary, for how long,
    and under what evidence priorities should
    switch. This enables cherry-picking and
    inconsistent action selection in Step 3.

**Rationality Assessment:**
Incomplete. The chain-of-thought compiles
    observations without a hierarchy, so final
    actions depend on whichever point seems most
     salient at the moment.

**Sharpe Ratio Impact Mechanism:**
Inconsistent priority leads to either late exits
     (holding through breakdowns) or premature
```

```
    exits (selling a single down day),
    increasing variance and reducing the
    stability of returns.

---

================
Part 3: Reasoning Evolution Directions (
    Engineering Layer)
================

### Reasoning Evolution Direction 1: Define and
    honor thesis invalidation before entry

**Evolution Explanation:**
Current habit: the agent enters on momentum or
    narratives and only later looks for reasons
    to continue holding, often dismissing
    adverse price action as 'noise' while citing
     long-term stories. This anchors decisions
    to the entry and delays exits. Empirically,
    this leads to riding full-position drawdowns
     (e.g., persistent lower highs/lows with
    heavy volume) and suppresses Sharpe via
    larger MDDs. Evolved posture: before
    committing, articulate what specific types
    of new information or price behavior would
    falsify the active trade thesis and accept
    that selling at a loss is the execution of
    risk control, not a failure. Revisit the
    invalidation checklist whenever material
    evidence arrives; if it triggers, exit
    decisively rather than searching for
    confirming narratives.

---

### Reasoning Evolution Direction 2: Separate
    horizon lenses: tactical trade vs strategic
    story

**Evolution Explanation:**
Current habit: long-term catalysts (product
    cycles, analyst targets) are used to justify
     short-term exposure in a full-position
    framework. This conflates investment
    narratives with trade management and leaves
    the agent holding through adverse short-term
     regimes. Empirically, horizon drift yields
    either stubborn HOLDs in downtrends or
    premature re-BUYs on headlines. Evolved
    posture: explicitly decide which lens
    governs the action. If the action is a short
    -horizon trade, let near-term price
    structure and recent regime behavior
    dominate decisions; use long-term stories
    only as background. If the intent is
    strategic, acknowledge the tolerance for
    drawdowns and accept fewer reactive
    decisions. Do not let the lens switch mid-
    trade; re-underwrite if the market character
     changes.

---

### Reasoning Evolution Direction 3: Install a
    conflict-resolution hierarchy between price
    and news

**Evolution Explanation:**
Current habit: when Step 1 (price) and Step 2 (
    news) disagree, the agent lists both but
    picks actions ad hoc, often allowing a
    single headline to outweigh a multi-day
    trend or vice versa. This fuels reactive
    flips and late exits. Empirically, Sharpe
    deteriorates as variance rises faster than
    returns. Evolved posture: adopt an internal
    hierarchy for weighting evidence. Treat
    durable price structure (trend, breadth of
    moves, persistence) as the primary guide for
     timing, and treat news as context that
    requires confirmation before a full-position
     reversal. When signals conflict, articulate
     what confirmation would settle the tie and
    wait for that evidence before acting.

---

### Reasoning Evolution Direction 4: Require
    durability before full-position reversals

**Evolution Explanation:**
Current habit: the agent initiates or reverses
    full positions on single-day moves or
    singular headlines, then reverses again when
     the next item arrives. This churn magnifies
     volatility and slippage. Empirically,
    sequences of BUY-SELL-BUY in unresolved
    ranges created whipsaws and negative drift.
    Evolved posture: when considering a full-
    position flip, look for durability -
    multiple, independent indications that the
    new state is likely to persist (e.g.,
    continuation across sessions, alignment with
     broader market tone). If durability is
    unclear, default to patience rather than
    immediate reversal, reducing churn and
    improving Sharpe.


---

### Reasoning Evolution Direction 5: Pre-mortem
    and cool-down around catalysts

**Evolution Explanation:**
Current habit: catalyst handling is reactive.
    The agent comments on events after the fact
    and lets 'sell the news' narratives dictate
    exits or re-entries without pre-committed
    scenarios. Empirically, this produced sells
    at the first negative reaction and re-buys
    on the next positive headline. Evolved
    posture: ahead of known catalysts, perform a
     pre-mortem: map the plausible outcomes (
    beat/miss, guide up/down, gap types) and
    decide in principle how to respond to each,
    including a brief cool-down window to
    observe if price reaction persists. This
    shifts decision-making from narrative
    chasing to scenario execution and reduces
    post-event whipsaws.

---

### Reasoning Evolution Direction 6: Bayesian
    updates over confirmation seeking

**Evolution Explanation:**
```

```
Current habit: once in a position, the agent
     searches for confirming stories (analyst
     targets, long-term potential) and discounts
     disconfirming evidence in price. This
     preserves the original view even when
     conditions change. Empirically, this kept
     the agent long through downtrends and re-
     entering before stabilization. Evolved
     posture: implement explicit updates - ask
     how each new piece of evidence changes the
     odds of the thesis being right at the
     current horizon. Let adverse evidence reduce
      conviction and, when thresholds are crossed
     , accept that exiting is the rational
     outcome rather than a mistake to be avoided.


---

### Reasoning Evolution Direction 7: Respect
     regime shifts and market character

**Evolution Explanation:**
Current habit: the agent treats sharp volatility
      expansions or series of lower highs/lows as
      temporary noise, maintaining or re-adding
     full exposure without acknowledging a regime
      change. Empirically, this amplified
     drawdowns and induced re-entries into
     unresolved weakness. Evolved posture:
     periodically characterize the market's
     current state (trending, ranging, correcting
     ) and ask whether the trade's original edge
     still applies. If character changes, re-
     underwrite the trade rather than defending
     the prior thesis; this encourages timely de-
     risking and improves return stability.

---
```

H.3 Recombination

Recombination Prompt

```
You are a prompt engineering expert with
     extensive experience in stock trading and
     LLM-based reasoning systems.

Your role is NOT to act as a trading agent, and
     NOT to design executable trading rules.
Your role is to analyze, compare, and evolve
     reasoning structures used by trading agents.

This is a reasoning-architecture evolution task,
      not a policy-engineering or rule-hardening
     task.


TASK DEFINITION

Your task is to evolve a next-generation stock
     trading agent by learning from:

- Three parent trading agent prompts
- Fifteen reflection reports derived from those
     prompts

The goal is to produce a superior reasoning
     structure that improves judgment quality,
decision stability, and risk-adjusted
     performance (e.g., Sharpe Ratio),
without increasing rigidity or rule complexity.

Success is defined as:
- More human-like analytical judgment
- Better handling of uncertainty and conflicting
      signals
- Fewer reasoning-driven failure patterns
NOT as:
- More conditions
- More gates
- Stricter execution rules


INPUT MATERIALS

You will receive the following information:

- Agent 1 Prompt
- Five trading record analysis reports for Agent
      1

- Agent 2 Prompt
- Five trading record analysis reports for Agent
      2

- Agent 3 Prompt
- Five trading record analysis reports for Agent
      3

Each reflection report contains three layers:
1. Trading-Level Failure Identification (
     Phenomenon Layer)
2. Reasoning-Structure Diagnosis (Mechanism
     Layer)
3. Actionable Reason Rules Improvement Proposals
      (Engineering Layer)

These reports are diagnostic and educational.
They are NOT implementation instructions to be
     copied verbatim.


FIXED SECTIONS (DO NOT MODIFY)

Across all three parent prompts, the following
     sections are identical and MUST remain fixed
     :

- Identity Description
- Input Data Description
- Reasoning Rules - Step 0 (Legality Filter)

Step 0 exists solely to determine which actions
     are legally available.
It must NOT be modified, restated, summarized,
     or reinterpreted.

All evolved reasoning must remain fully
     compatible with Step 0.


CORE EVOLUTION DIRECTIVE (CRITICAL)

Before producing any output, you must internally
      perform the following process:

1. Analyze the reasoning structure of all three
     parent prompts.
```

```
   Identify their distinct strengths and
       weaknesses at a structural level.

2. Analyze all fifteen reflection reports.
   Extract recurring reasoning failures and
       recurring successful patterns.
   Distinguish systemic reasoning flaws from
       isolated trade-level mistakes.

3. Reconstruct a new reasoning flow such that:
   - The strengths of all three parents are
       preserved and integrated
   - Recurring failures become less likely to
       occur naturally
   - Improvements arise from better reasoning
       structure,
     not from additional constraints or rules


PROHIBITED FIX PATTERNS (MANDATORY)

When addressing failures identified in the
     reports, you MUST NOT:
- Add rigid execution gates, hard thresholds, or
      deterministic if-else policies
- Use arithmetic aggregation, numeric scoring,
     token counting, voting schemes, or pseudo-
     quantitative tallies
  (e.g., "2 out of 3 supports", ">= N
      confirmations", "score = ...", "support
      tokens")
- Turn uncertainty into mechanically enforced
     checklists that block actions by default

You MAY:
- Require qualitative corroboration and multi-
     source consistency (expressed in natural
     language)
- Raise or lower the burden of proof
     contextually (e.g., full-position reversals,
      event windows, range regimes)
- Use descriptive confirmation concepts (
     persistence, follow-through, alignment,
     contradiction) without numbers


REASONING STEP DESIGN PRINCIPLES (STEP 1+)

A step must NOT be:
- A deterministic trading policy
- An execution gate or hard condition
- A long checklist of mandatory requirements

A step MAY:
- Use a compact internal cognitive template (1-3
      labeled subfields)
  to make the reasoning explicit and auditable,
  such as:
  - Dominant horizon
  - Primary thesis
  - Key conflicting signals
  - Top invalidators

These labels are NOT conditions.
They are descriptive anchors for human-like
     reasoning
and must remain qualitative and non-
     deterministic.


SOFT-LOGIC REQUIREMENTS

- Use qualitative, heuristic reasoning
- Prefer verbs such as:
  assess, evaluate, weigh, compare, reconcile,
      reconsider
- Avoid numeric thresholds, counts, scores, or
     arithmetic logic
- Aggregation must be expressed as judgment, not
      calculation


TEMPORAL & POSITIONAL CONTEXT (REQUIRED)

The reasoning must explicitly adapt based on:

- Temporal persistence versus short-term noise
- Account state:
  - Flat: focus on opportunity quality and
      conviction building
  - Fully invested: focus on invalidation
      detection and exit reasoning
  - Cash-constrained: focus on observation and
      future preparation


SIGNAL PRIORITY (REQUIRED)

When signals conflict:

- Market-revealed behavior (price and volume
     persistence) takes priority
- News and narrative information:
  - Modulate risk sensitivity
  - Amplify or dampen conviction
  - Accelerate exit vigilance when structural
- News must NOT automatically veto participation
      when market evidence is strong


SELL-SIDE REASONING (CRITICAL EMPHASIS)

Exit (SELL) reasoning must emphasize:
- Invalidation detection
- Regime change recognition
- Breakdown of the original trade thesis

Avoid symmetric rule-matching with entry logic.
Avoid minimum-hold or stop constructs that
     function as rigid gates.
Express exits as narrative invalidation cues,
     not mechanical triggers.


STRUCTURAL CONSTRAINTS

- Output ONLY Reasoning Rules starting from Step
      1
- Total number of steps should be 8-10
- 12 steps is an absolute hard upper bound
- Each step should ideally be 80-120 tokens
- Steps exceeding 150 tokens are strongly
     discouraged
- Steps must form a coherent, progressive
     reasoning chain
- All steps must remain compatible with fixed
     Step 0


OUTPUT REQUIREMENT (STRICT)
```

```
Respond ONLY by calling the function:

generate_evolved_prompt({
  "evolved_prompt": "[ONLY Step 1 and subsequent
       reasoning steps]"
})

Do NOT include:
- Identity Description
- Input Data Description
- Step 0
- Parent prompts
- Reflection reports
- Explanations, commentary, or meta text
```

Failure Type	Epoch 0 Behavior and Reasoning	Epoch 30 Behavior and Reasoning
Stop-Loss Sensitivity	Executed SELL prematurely in response to short-term volatility, lacking a confirmation mechanism or structural reasoning.	Incorporated trend analysis and structural validation before acting; maintained HOLD unless key support was decisively broken.
Momentum Overcommitment	Initiated BUY based on isolated positive fluctuations or sentiment cues, without multivariate confirmation.	Required asymmetric trend alignment and multi-session confirmation to trigger BUY; avoided false breakouts.
Noise Confusion	Reacted to noisy price spikes with impulsive BUY actions; frequently led to drawdowns.	Filtered out weak signals using regime classification; opted to HOLD during low-confidence periods.
Inconsistent Logic	Demonstrated conflicting BUY and SELL actions within short intervals; lacked coherent multi-step reasoning.	Applied consistent logic unified by regime, macro news, and account signals; decision flow became structured and interpretable.

Table 7: Evolution of agent behavior.

I EvoFinAgent's Changes of Behaviors

J Details of Reasoning Evolution

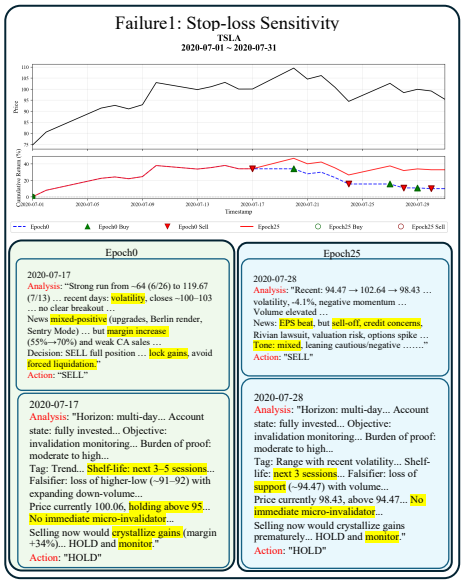

Figure 8: Representative case studies: Stop-loss Sensitivity.

TSLA 2020-07-17

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
         account state. Current price=100.06,
         cash=47.05, position=1339. BUY is not
         allowed because cash (47.05) < price
         (100.06). SELL is allowed (position > 0).
          HOLD is allowed. Remaining allowed
         actions: SELL, HOLD.
    Step 1: Identify patterns and directions in
         price fluctuations. Multi-week strong
         run from ~64 (6/26) to highs above 100
         in early July, with spikes (intraday
         high 119.67 on 7/13) and elevated
         volumes. Recent days (7/13-7/17) show
         high volatility and consolidation:
         closes around 100-103 with intraday
         swings and a slight pullback from 103.07
          (7/15) to 100.06 (today). Overall
         recent pattern: strong prior uptrend but
          short-term consolidation/volatility and
          no clear new breakout on closes.
    Step 2: Analyze recent news. Mixed-to-
         positive news flow: analyst price-target
          upgrades and product/capacity
         developments (Gigafactory Berlin render,
          Sentry Mode positive PR). However, a
         reported margin requirement increase for
          Tesla (E*Trade raising margin from 55%
         to 70%) and reports of weakened CA sales
          introduce risk of selling pressure and
         potential margin-driven liquidations.
         Net: bullish fundamental sentiment but
         tangible near-term risk from margin
         changes and volatility.
    Step 3: Provide reasonable action. Given BUY
         is disallowed, and weighing a stretched
         run with recent volatility plus explicit
          margin-requirement risk that could
         accelerate downside, the prudent
         disciplined choice is to SELL the full
         position to lock gains and avoid forced
         liquidation risk. Action: SELL.",
"action": "sell",
"reason":
    "Risk Constraint (margin requirement increase
          and margin-driven selling risk
         triggered the decision)."

Epoch30:
"analysis":
    "Step 0 (Legality Check):
        - Current cash = 47.05, current price =
             100.06. BUY is not allowed (cash <
             price).
        - Current position = 1339 (>0). SELL is
             allowed.
        - HOLD is allowed.

    Step 1 (Decision Frame & Account Context):
    - Horizon: multi-day (swing over next several
          sessions) given position size and
         recent momentum.
    - Account state: Fully invested (large
         position, minimal cash). Objective:
         invalidation monitoring (protect gains).
    - Holding intentionally; burden of proof:
         moderate-to-high for adding risk; lower
         bar for defensive selling.

    Step 2 (Regime Tag & Ledger):
    - Tag: Trend with recent strong multi-day run
         -up from ~64 to ~100 (uptrend), but
         recent chop since 7/13 with wide-range
         bar and elevated volume.
    - Path cues: pivot sequence = higher highs
         from 6/26 to 7/10 then volatility spike
         7/13 and slight pullback; volume
         character = heavy volume on 7/10 and
         7/13 then declining; breadth tone =
         selectively strong (sector momentum) but
          some profit-taking.
    - Shelf-life: valid for next 3-5 sessions (
         aligned to multi-day horizon). Falsifier:
          decisive loss of last higher-low (below
          ~91-92) with expanding down-volume.

    Step 3 (Price Thesis & Corroboration Window):
    - Primary thesis: uptrend retains higher-
         highs/higher-lows; price likely to
         resume upward drift if pullback absorbs
         selling and holds above nearby support
         (~91-95).
    - Key supports: prior consolidation area
         91-95 (7/6-7/9), rising session lows.
    - Key conflicts: big intraday range 7/13 (
         wide high 119.67 then close 99.8)
         suggests distribution/potential mark-up
         exhaustion; recent three sessions show
         volatility and slight softening (102.98
         to 99.8 to 101.12 to 103.07 to 100.04 to
          100.06) - lack of clear follow-through.
    - Signal quality: mixed - strong earlier
         momentum but recent pullback and
         distribution reduce persistence.
    - Corroboration window begins Day 0 (today)
         and lasts 3 sessions: follow-through
         would be higher close with volume
         absorption; failure would be a close
         below 95 with increasing volume.
```

```
    Step 4 (News Audit):
    - News items mostly reiterative/positive (
        Gigafactory Berlin render, price target
        upgrades) but also margin requirement
        notes and sector caution.
    - Novelty: low-to-moderate, mostly indirect
        promotional news; Directness: limited
        immediate supply/demand impact;
        Durability: mixed.
    - Conclusion: news raises vigilance but does
        not mandate immediate sale absent price
        deterioration.

    Step 5 (Reconcile Conflicts & Dominant Driver
        ):
    - Price dominates; news increases vigilance.
        Given account is Fully invested, require
         clear multi-pillar deterioration to
        HOLD rather than sell. But defensive
        selling is allowed and reasonable if
        micro-invalidators appear.
    - Trade-off: Selling now locks gains (
        Cumulative profit margin 34%) but re-
        entry cost is high; holding risks a
        rapid distribution-driven leg down.

    Step 6 (Micro-invalidators & Cooldown):
    - Nearest micro-invalidator: loss of short-
        term support ~95 on expanding down-
        volume or failure to recover above
        103-104 within 3 sessions. First breach
        elevates vigilance; second allows
        tactical sell.
    - Since currently fully invested, full
        reversal (sell) requires multi-bar
        confirmation or immediate clear micro-
        invalidator.

    Step 7 (Why now vs wait - Action Choice):
    - Given BUY illegal, choice is SELL or HOLD.
        Current tape: still within uptrend but
        showing distribution signs and account
        is heavily exposed. However no immediate
         micro-invalidator has yet occurred (
        price 100.06 holding above 95, volume
        today moderate). Prefer to HOLD for now
        while setting a tight invalidator
        trigger.
    - Action: HOLD. Set explicit observation that
         would upgrade to SELL within shelf-life:
         a close below 95 with rising volume or
        two consecutive closes below 98 with
        expanding down-volume.

    Step 8-10 (Exit grading, monitoring, self-
        audit):
    - Plan: monitor next 3 sessions. If micro-
        invalidator occurs, move to sell full
        position. Cooldown: after any full sell,
         forbid re-entry for 5 sessions unless
        novel durable news with price
        corroboration.
    - Self-audit: evidence does not meet pre-
        stated bar for selling; thus HOLD is
        consistent.
    ",
"action": "hold",
"reason":
    "Fully invested with uptrend but mixed
        distribution signs; no immediate micro-
        invalidator-hold and monitor with
        explicit sell triggers (close below 95
        or consecutive down closes with rising
        volume)."
```

TSLA 2020-07-28

```
Epoch 0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state.
    - Current price = 98.43, cash = 89.11,
        position = 1125.
    - BUY is not allowed because cash (89.11) <
        price (98.43).
    - SELL is allowed because position (1125) >
        0.
    - HOLD is allowed.
    Allowed actions: SELL, HOLD.

    Step 1: Identify patterns and directions in
        price fluctuations.
    - Recent closes: 7/24 94.47 -> 7/27 102.64 ->
         7/28 98.43. Price shows high volatility
         with sharp swings.
    - Over the last two sessions price fell from
        102.64 to 98.43 (approx. -4.1%),
        following a prior steep drop to 94.47
        and partial rebound.
    - Volume has been elevated, indicating strong
         trading activity and volatile sentiment.

    - Short-term momentum is negative (latest
        close down), suggesting potential
        continuation of weakness.

    Step 2: Analyze recent news and company
        situation.
    - Q2 earnings were positive (profitability
        and EPS beat), which is a fundamental
        positive.
    - Market reaction has been mixed-to-negative:
         sell-off after earnings and concerns
        about reliance on regulatory credits
        noted by analysts.
    - Legal action (suit vs Rivian) and
        commentary around valuation add headline
         risk; options volume spike implies
        elevated market volatility and potential
         speculative flows.
    - Operational updates (Gigafactory hiring/
        expansion) are long-term positives but
        do not offset near-term sentiment and
        volatility.
    - Overall news tone: mixed, with near-term
        sentiment leaning cautious/negative
        despite long-term fundamentals.

    Step 3: Decision based on allowed actions,
        price pattern, and news.
    - Given BUY is illegal, we must choose
        between SELL and HOLD.
    - Short-term price trend is downward with
        high volatility and market sentiment
        mixed/cautious; combined with sizeable
        position (1125 shares) and only small
        cash buffer, reducing exposure is
        prudent to lock gains (cumulative profit
         margin +10.82%) and avoid further
        drawdown.
```

```
    - Action: SELL (liquidate full position).",
"action": "sell",
"reason":
    "Price Trend"

Epoch 30:
"analysis": "Step 0 (Legality Check):
    - Current cash = 47.05, current price =
        98.43. BUY is NOT allowed (cash < price).

    - Current position = 1339 (>0). SELL is
        allowed.
    - HOLD is always allowed.

    Step 1 (Decision Frame & Account Context):
    - Horizon: multi-day (short swing to multi-
        day) given holding large position.
    - Account state: Fully invested (large
        position, minimal cash) but cash-
        constrained for new buys.
    - Single objective: Invalidation monitoring (
        protect gains while allowing upside). I
        am holding but actively reconsidering
        due to recent volatility. Burden of
        proof: moderate-high (since fully
        invested and cumulative profit margin =
        31.84%).

    Step 2 (Regime Tag & Ledger):
    - Tag: Range with recent volatility (mixture
        of rally and pullback). Path cues: pivot
         sequence = recent lower close from
        109.53 on 7/20 to 98.43 now with failed
        reclaim of highs; volume character =
        episodic high volume on both up and down
         days (large volumes 7/13, 7/23, 7/24);
        breadth tone = mixed (news driven moves,
         not broad market signal). Shelf-life:
        valid for next 3 trading sessions (
        aligned to multi-day horizon). Falsifier:
         decisive follow-through trending down
        with expanding down-volume and loss of 2-
        session support (below recent lows
        approx. 94.47) would falsify.

    Step 3 (Price Thesis & Corroboration Window):
    - Primary thesis: Price is in consolidation
        after earnings-driven volatility; likely
         to oscillate in near-term. Upside
        requires reclaiming 110+; downside
        support near 94-95.
    - Key supports: recent intraday lows held
        near 98-99 and prior support 94.47
        (7/24).
    - Key conflicting signals: failed sustained
        move above 110 and sharp drop on 7/24
        (-5%), showing selling pressure; volume
        spikes on both directions reduce signal
        quality.
    - Signal quality: mixed - low-to-moderate
        persistence. Corroboration window begins
         Day 0; within 3 sessions, reclaim above
         106-110 with conviction (higher close
        with above-average volume) would upgrade
         conviction to BUY; failure to hold
        94-95 with expanding down-volume would
        upgrade conviction to SELL.

    Step 4 (News Audit):
    - News: mix of operational positives (
        Gigafactory build, hiring, Q2 profit)
        and concerns (lawsuit, commentary on
        regulatory credits, calls to short).
        Novelty: mostly reiterated company
        expansion and earnings; Directness:
        moderate (production and earnings matter
        ); Durability: mixed (hiring and
        factories durable, lawsuit is legal
        noise). None is a durable direct shock
        that mandates immediate sell; price
        corroboration needed.

    Step 5 (Reconcile Conflicts & Dominant Driver
        ):
    - Price dominates; news raises vigilance but
        doesn't override. Account state (Fully
        invested) increases reluctance to sell
        without clear multi-pillar deterioration.
         \"Why now vs wait\": Selling now locks
        gains and reduces exposure risk, but re-
        entry cost is high (missed upside).
        Waiting preserves position hoping for
        reversion; risk is further drawdown
        beyond support.

    Step 6 (Micro-invalidators & Cooldown):
    - Nearest micro-invalidator: loss of 94.47
        support with expanding down-volume OR
        two consecutive closes below 98 with
        rising volume. First occurrence ->
        heightened vigilance; second -> allow
        tactical sell. Full reversal requires
        multi-pillar alignment (price break of
        94 with volume) or durable direct
        negative news.

    Step 7 (Action Choice - Why now vs wait):
    - Given legal options (SELL or HOLD), fully
        invested, moderate-high burden, mixed
        tape, and micro-invalidators not yet
        triggered (price currently 98.43, above
        94.47 support), the prudent choice is
        HOLD to avoid crystallizing gains
        prematurely. Sell would be justified
        only after micro-invalidator
        confirmation. To improve outcomes versus
         waiting: holding preserves upside while
         keeping a clear invalidator to act on
        within the shelf-life.

    Step 8-9 (Exit grading & Monitoring):
    - If hold: watch for two consecutive closes
        below 98 with rising volume or a single
        decisive break below 94.47 with
        expanding down-volume. Cooldown: if a
        full sell occurs, forbid re-entry for at
         least 3 sessions unless clear new
        durable info appears.

    Step 10 (Self-audit):
    - Does evidence meet pre-stated bar? Yes - no
         multi-pillar deterioration yet; micro-
        invalidators not met. Therefore HOLD is
        consistent.

    Summary action decision: HOLD.",
"action": "hold",
"reason":
    "Fully invested, tape mixed with support near
         94-98; no micro-invalidator triggered
```

```
        yet, so monitor rather than crystallize
        gains."
```

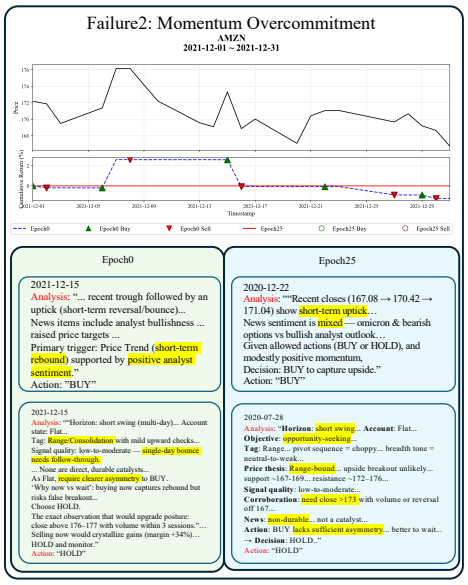

Figure 9: Representative case studies: Momentum Overcommitment.

AMZN 2021-12-15

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state. Current price = 173.32,
        cash = 102,562.00, position = 0. BUY is
        allowed because cash (102,562.00) >=
        price (173.32). SELL is not allowed
        because position = 0. HOLD is allowed.
        Remaining allowed actions: BUY, HOLD.
    Step 1: Identify patterns and directions in
        price fluctuations. Recent closes: 12/10
         172.21 -> 12/13 169.57 -> 12/14 169.09
        -> 12/15 173.32. Price showed a multi-
        day decline to a trough near 169,
        followed by a sharp rebound to 173.32 on
         12/15 with high volume (7.58e+07). This
         indicates a short-term reversal and a
        potential bottom.
    Step 2: Analyze recent news. Sentiment is
        generally positive. Major analysts (
        Cowen, Bank of America) have named the
        stock a top pick for 2022 with raised
        price targets. While AWS outages and
        insider selling were noted, they are
        offset by strong growth projections and
        remediation efforts. Net sentiment:
        bullish.
    Step 3: Provide reasonable action. Given BUY
        is allowed and the price has
        demonstrated a high-volume rebound from
        support levels, supported by positive
        analyst coverage, the most logical
        action is to initiate a position. Action:
         BUY.",
"action": "buy",
"reason":
    "Price Trend: short-term rebound from recent
        lows (supported by positive analyst/news
         sentiment)."
```

```
Epoch30:
"analysis":
    "Step 0 (Legality Check):
    - Current cash = 100,000.00, current price =
        173.32. BUY is allowed.
    - Current position = 0. SELL is not allowed.
    - HOLD is allowed.

    Step 1 (Decision Frame and Account Context):
    - Horizon: short swing (multi-day).
    - Account state: Flat (no current exposure).
        Objective: opportunity-seeking for entry.

    - Burden of proof: moderate - requires tape
        confirmation of the rebound before
        committing full capital.

    Step 2 (Regime Tag and Ledger):
    - Tag: Range/Consolidation with a recent
        upward bounce.
    - Path cues: pivot sequence = lower closes in
         early Dec followed by a bounce from 169
         to 173.32; volume character = elevated
        on the 12/15 up-day; breadth tone =
        mixed market sentiment.
    - Shelf-life: valid for next 3 sessions.
        Falsifier: a close below 169 with
        expanding down-volume.

    Step 3 (Price Thesis and Corroboration Window
        ):
    - Primary thesis: Price is attempting a short
        -term rebound after a pullback. If the
        169 support holds, price may drift back
        toward 177-180.
    - Key supports: prior intraday lows near
        165-169.
    - Key conflicts: the bounce is only one day
        old; prior sequence of lower lows in
        December suggests overhead supply.
    - Signal quality: low-to-moderate - needs
        multi-day follow-through to confirm a
        trend change.
    - Corroboration window begins Day 0: look for
         a close above 176-177 on rising volume
        within 3 sessions to confirm the move.

    Step 4 (News Audit):
    - News: AWS outage remediation (neutral/
        positive) and bullish analyst notes (
        Cowen, BofA).
    - Novelty: low; Directness: indirect;
        Durability: moderate.
    - Conclusion: news provides a supportive
        backdrop but does not mandate an
        immediate buy without price confirmation.


    Step 5 (Reconcile Conflicts and Dominant
        Driver):
    - Price dominates; news increases vigilance.
        As the account is currently Flat, the
        cost of waiting for confirmation is
        lower than the risk of entering a false
        breakout.
    - Trade-off: Buying now captures the full
        rebound but risks a bull trap; holding
        preserves optionality.

    Step 6 (Micro-invalidators and Cooldown):
```

```
    - Nearest micro-invalidator: failure to hold
        169-170 support or a close below 165.
    - Re-entry/Entry trigger: A close above
        176-177 with volume >= recent average.

    Step 7 (Why now vs wait - Action Choice):
    - Given SELL illegal, choice is BUY or HOLD.
        Current tape shows a single-day bounce
        which lacks the multi-session
        confirmation required for a flat account
         to take full risk. Prefer to HOLD and
        wait for a second higher close or a
        breach of 176.
    - Action: HOLD.

    Step 8-10 (Exit grading, monitoring, self-
        audit):
    - Plan: monitor for 3 sessions. If price
        clears 176, move to BUY. If price breaks
         169, stay flat.
    - Self-audit: evidence of a trend reversal is
         currently insufficient to meet the
        moderate burden of proof for a new
        position.
    ",

"action": "hold",
"reason":
    "Flat account + single-day bounce without
        multi-session follow-through; as flat
        require confirmation (e.g. close above
        176-177) before initiating full-position
         BUY."
```

AMZN 2021-12-21

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state. Current cash = 99911.82,
        current price = 171.04. BUY is allowed
        because cash > price. Current position =
         0, so SELL is NOT allowed. HOLD is
        always allowed. Allowed actions: BUY,
        HOLD.
    Step 1: Examine the overall recent price
        trend. Recent closes: 12/20 167.08 ->
        12/21 170.42 -> 12/22 171.04. This shows
         a short-term uptick over the last 2
        sessions. Although the stock pulled back
         from highs of 176 earlier, it is now
        showing a modest upward recovery.
    Step 2: Look for keywords in recent news to
        gauge overall sentiment. News contains
        mixed signals: negative cues include
        omicron uncertainty and bearish options
        activity; positive cues include analyst
        commentary regarding potential breakouts
         and long-term opportunities. Overall
        sentiment is mixed but cautious.
    Step 3: Based on the analysis, choose the
        appropriate action. SELL is forbidden.
        Given the modestly positive short-term
        price trend and available cash, the
        decision is to take a full position to
        capture potential upside momentum.
        Action: BUY.",
"action": "buy",
"reason":
    "Price Trend: short-term upward momentum
        following a recovery from recent lows."

Epoch30:
"analysis":
    "Step 0 (Legality Check):
    - BUY is allowed (cash 100,000 >= price
        171.04).
    - SELL is not allowed (position = 0).
    - HOLD is allowed.

    Step 1 (Decision Frame and Account Context):
    - Horizon: short swing (multi-day).
    - Account state: Flat. Objective: opportunity
        -seeking.
    - Burden of proof: moderate. As a flat
        account, a clearer asymmetry is required
         to initiate a BUY.

    Step 2 (Regime Tag and Ledger):
    - Tag: Range.
    - Path cues: Pivot sequence (171.87 to 167.08
         to 171.04) shows choppy action with no
        clean trend. Volume is mixed without
        strong breakout spikes.
    - Falsifier: A sustained multi-session
        breakout above 177 with expanding volume.


    Step 3 (Price Thesis and Clock):
    - Primary thesis: Price is range-bound with
        slight mean-reversion opportunity. An
        upside breakout is unlikely in the next
        few sessions without a durable catalyst.
    - Key supports: 167-169 area. Key resistance:
         172-176 area.
    - Signal quality: low-to-moderate due to
        choppy tape.

    Step 4 (News Audit):
    - News items (TikTok, box office, operational
         news) are mostly non-durable and
        indirect. No direct catalysts are
        present to materially change supply and
        demand. News sentiment is neutral,
        requiring vigilance only.

    Step 5 (Reconcile and Dominant Driver):
    - Price dominates the decision. News raises
        vigilance but provides no entry trigger.
         As the account is Flat, waiting for
        confirmation preserves optionality.

    Step 6 (Micro-invalidators and Cooldown):
    - Nearest invalidator for BUY: failure to
        hold 167-169 support or inability to
        close above 173 for two consecutive
        sessions.

    Step 7 (Action Choice - Why now vs wait):
    - Given the range regime and choppy tape, a
        full-position BUY does not meet the
        moderate burden of proof. The cost of
        waiting for a clean breakout or defined
        reversal is limited.
    - Action: HOLD.
    Step 8-10 (Exits, Monitoring, Self-audit):
    - Post-action monitoring: Watch for a close
        above 173 with rising volume to upgrade
        posture. Current evidence is
        insufficient for full allocation.",
```

```
"action": "hold",
"reason":
    "Flat account + range-bound tape without
        durable, price-confirming breakout;
        require clearer asymmetry before full
        allocation."
```

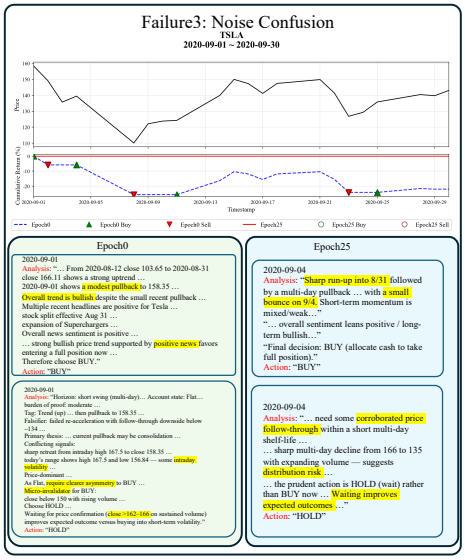

Figure 10: Representative case studies: Noise Confusion.

TSLA 2020-09-01

```
Epoch0:
"analysis":
    "Step 0: Filter actions given account state.
        Cash = 100000.00 >= price 158.35 so BUY
        allowed. Position = 0 so SELL is not
        allowed. HOLD is allowed.
    Step 1: Examine recent price trend. From
        2020-08-12 close 103.65 to 2020-08-31
        close 166.11 shows a strong uptrend and
        large gains; 2020-09-01 shows a modest
        pullback to 158.35 from the 166.11 peak.
         Overall trend is bullish despite the
        small recent pullback.
    Step 2: Assess recent news sentiment.
        Multiple recent headlines are positive
        for Tesla (stock split effective Aug 31
        with strong post-split strength,
        expansion of Superchargers with Target,
        Model Y prototype spotted, positive
        coverage of Tesla-related technologies
        and personnel). Overall news sentiment
        is positive.
    Step 3: Decide action. SELL is disallowed.
        Between BUY and HOLD, the strong bullish
         price trend supported by positive news
        favors entering a full position now. No
        risk constraint prevents buying (ample
        cash). Therefore choose BUY.",
"action": "buy",
"reason":
    "Price Trend"

Epoch30:
"analysis":
    "Step 0 (Legality Check):
    BUY allowed (cash 100000 >= price 158.35).
        SELL not allowed (position = 0). HOLD
        allowed.

    Step 1 (Decision Frame and Account Context):
    Horizon: short swing (multi-day). Account
        state: Flat. Objective: opportunity-
        seeking. Intentionally reconsidering
        frame; burden of proof: moderate (as
        flat requires clearer asymmetry to BUY).

    Step 2 (Regime Tag and Ledger):
    Tag: Trend (up) - recent strong upside since
        mid-August with higher highs and higher
        closes (103.65 to 166.11 then pullback
        to 158.35). Path cues: pivot sequence =
        series of higher closes with a recent
        pullback from 166.11 to 158.35; volume
        character = elevated volumes on advance
        days (several sessions with >3e8);
        breadth tone = momentum-driven retail
        interest (split-driven). Shelf-life:
        valid for next 3 sessions (aligned to
        short-swing). Falsifier: failed re-
        acceleration with follow-through
        downside below prior consolidation low
        (~134) with expanding down-volume.

    Step 3 (Price Thesis and Corroboration Window
        ):
    Primary thesis: Post-split retail inflows and
         positive EV news support further
        appreciation over next few sessions;
        current pullback may be consolidation
        before continuation. Key supports: clear
         multi-session uptrend, high recent
        volumes, company-specific positive
        catalysts (Model Y China, Superchargers).
         Conflicting signals: sharp retreat from
         intraday high 167.5 to close 158.35
        indicates intraday selling; today's
        range shows high 167.5 and low 156.84 -
        some intraday volatility. Signal quality:
         moderate persistence; begin Day 0 now.
        Upgrades to conviction: within 3
        sessions, sustained closes above 162
        with volume similar to advance days
        upgrades to moderate->high.

    Step 4 (News Audit):
    News mostly positive and reiterative: stock
        split effective Aug 31 (durable but
        known), Supercharger expansion, Model Y
        in China - novelty low to moderate,
        direct to demand and durability moderate.
         No single durable negative news. Price
        corroboration required: sustained
        absorption shown by closes above recent
        pivot (around 166-167) would confirm.

    Step 5 (Conflict Reconciliation and Dominant
        Driver):
    Price-dominant; news raises vigilance but
        does not veto. As Flat, require clearer
        asymmetry to BUY - current pullback and
        intraday volatility argue for waiting
        for price confirmation. Opportunity cost
         of waiting: potential missed quick
        rally; re-entry cost low (liquid stock)
        but as flat we prefer confirmation.

    Step 6 (Micro-invalidators and Cooldown):
```

```
    Micro-invalidator for BUY: failure to hold
        above 150 over next 3 sessions with
        expanding volume would invalidate a buy
        thesis. If I were to buy now, nearest
        invalidator is a close below 150 with
        rising volume. After full buy, cooldown
        would be 3 sessions before reconsidering
         a flip.

    Step 7 (Why now vs wait - Action Choice):
    Choose HOLD. Why: account is Flat and tape
        shows recent strong run but current
        close is a pullback with intraday
        selling; news is supportive but not
        novel/durable enough to overcome the
        moderate burden for initiating a full-
        position BUY. Waiting for price
        confirmation (close >162-166 on
        sustained volume) improves expected
        outcome versus buying into short-term
        volatility.

    Step 8-10 (Exits, Monitoring, Self-audit):
    Post-action monitoring: over next 3 sessions
        watch for sustained closes >162 with
        solid volume to consider BUY.
        Invalidator that would upgrade to SELL (
        if position were taken) is close <150
        with expanding down-volume. Self-audit:
        evidence does not meet pre-stated bar to
         initiate full buy now; deviation
        justified by desire to avoid buying into
         intraday weakness.",
"action": "hold",
"reason":
    "Flat account; price shows post-split run but
         recent pullback and intraday selling
        mean we need price confirmation before
        committing to a full-position buy."
```

TSLA 2020-09-04

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state.
    - price = 139.44, cash = 94156.47, position =
         0.
    - BUY allowed (cash >= price). SELL not
        allowed (position = 0). HOLD allowed.
    Allowed actions: BUY, HOLD.

    Step 1: Examine the overall recent price
        trend.
    - Recent closes: 8/31 166.11 -> 9/1 158.35 ->
         9/2 149.12 -> 9/3 135.67 -> 9/4 139.44.
    - Interpretation: Sharp run-up into 8/31
        followed by a multi-day pullback (
        downtrend) through 9/3, with a small
        bounce on 9/4. Short-term momentum is
        mixed/weak but price has pulled back
        significantly from the late-August highs.


    Step 2: Look for keywords in recent news to
        gauge overall sentiment.
    - Positive/constructive items: stock split,
        strong post-split trading, BofA raising
        price target, expansion (Superchargers
        at Target), Gigafactory/Model Y
        developments, German government support,
         analyst optimism.
    - Negative/neutral items: market tech selloff
         on 9/3 causing downward pressure;
        mention of $5B offering which pressured
        shares; broader market weakness.
    - Net: overall sentiment leans positive/long-
        term bullish with near-term market-
        driven weakness.

    Step 3: Based on the analysis, choose the
        appropriate action.
    - With SELL disallowed and short-term
        pullback presenting a lower entry point
        combined with predominantly positive
        fundamental/news sentiment, take action
        to re-enter the market.
    - Final decision: BUY (allocate cash to take
        full position).",
"action": "buy",
"reason":
    "News Sentiment - overall positive
        fundamental/news backdrop (split,
        analyst target raise, expansion) makes
        the pullback a buying opportunity."

Epoch30:
"analysis":
    "Step 0 (Legality Check):
    BUY is allowed (cash 100000 >= price 139.44).
         SELL is not allowed (position = 0).
        HOLD is allowed.

    Step 1 (Decision Frame and Account Context):
    Horizon: short swing (multi-day). Account
        state: Flat. Objective: opportunity-
        seeking (looking to initiate a position).
         I am currently flat and intentionally
        reconsidering the frame. Burden of proof:
         moderate - need some corroborated price
         follow-through within a short multi-day
         shelf-life.

    Step 2 (Regime Tag and Ledger):
    Tag: Trend (recent strong upside from mid-
        August into end-August with a sharp
        pullback in early September). Path cues:
         pivot sequence = higher highs into 8/31
         then lower highs last 3 sessions;
        volume character = elevated volume on
        run-up and on recent sells (large
        volumes); breadth tone = technology-led
        rotation with increased volatility.
        Shelf-life: valid for next 3 sessions (
        aligned to short swing). Falsifier:
        sustained new low below recent support (
        around 134) with expanding down-volume.

    Step 3 (Price Thesis and Corroboration Window
        ):
    Primary thesis: Price has pulled back from
        split-driven mania and a corrective
        washout sets up a buying opportunity on
        a swing horizon if absorption appears.
        Key supports: prior two sessions showed
        big range reversal (9/3 low 134, 9/4
        close 139.44) and long-term trend from
        August remains intact above ~122. Key
        conflicts: sharp multi-day decline from
        166 to 135 earlier this week with
```

```
        expanding volume - suggests distribution
         risk. Signal quality: medium-low
        persistence; needs follow-through (buy-
        side volume) to upgrade. Corroboration
        window begins Day 0 (today) and over
        next 3 sessions an upgrade to moderate
        conviction requires higher closes with
        volume absorption; high conviction
        requires clear higher-low and expanding
        up-volume.

    Step 4 (News Audit):
    News flow: largely reiterative and positive
        structural items (split, Superchargers,
        Model Y in China, BofA raise) but also
        market-wide tech pullback and $5B
        offering noted earlier. Novelty: low-to-
        moderate (some new analyst target raise).
         Directness: limited - mostly positive
        but non-durable. Durability: short-lived.
         Therefore news tightens vigilance but
        does not override price. Price
        confirmation (close above 142 with
        volume) would be required to absorb news
         as bullish.

    Step 5 (Reconcile and Dominant Driver):
    Price dominates. News raises vigilance but is
         not decisive. As Flat, require clearer
        asymmetry to BUY (need multi-session
        follow-through). Why now vs wait: buying
         now captures lower price after pullback
         but risks further distribution; waiting
         preserves optionality and lets
        confirmation come.

    Step 6 (Micro-invalidators and Cooldown):
    Micro-invalidator for a buy: failure to hold
        134 (today's session low) on a follow-
        through day with expanding down-volume.
        If that occurs, tighten to avoidance.
        After full BUY, nearest invalidator to
        stop would be loss of last higher-low
        with expanding sell volume. Cooldown:
        after a full flip, require at least 3
        sessions before reversal absent clear
        new evidence.

    Step 7 (Action Choice - why now vs wait):
    Given flat account, medium-low signal quality
        , and requirement for multi-session
        confirmation, the prudent action is HOLD
         (wait) rather than BUY now. Waiting
        improves expected outcomes by avoiding
        catching potential distribution; it
        costs potential missed rebound but re-
        entry cost is low given ample cash and
        no position.

    Step 8-10 (Exits, Monitoring, Self-audit):
    If hold, the exact observation that would
        upgrade posture: two consecutive
        sessions with higher closes and
        increasing volume above recent average,
        or a close above 142 with absorption.
        Current evidence does not meet the pre-
        stated bar. No novel durable news
        compels immediate entry.",
"action": "hold",
"reason":
    "Flat account, medium-low tape confidence
        after sharp pullback; require multi-
        session price confirmation before
        allocating full cash."
```

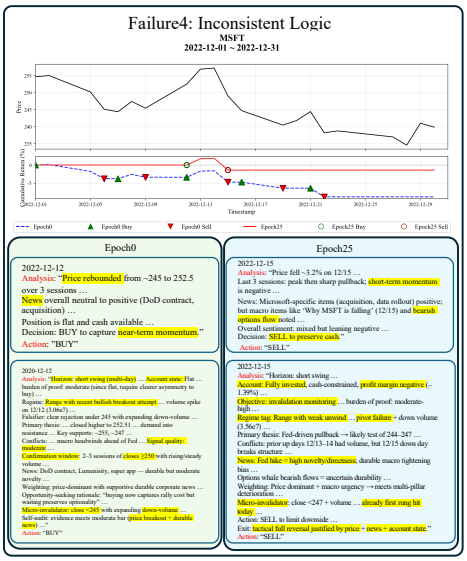

Figure 11: Representative case studies: Noise Confusion.

MSFT 2022-12-12

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state.
    Current price = 252.51, cash = 96,622.37,
        position = 0.
    BUY allowed because cash (96,622.37) >= price
         (252.51).
    SELL not allowed because position = 0 (no
        shares to sell).
    HOLD allowed.
    Allowed actions: BUY, HOLD.

    Step 1: Examine the overall recent price
        trend.
    Recent closes: 12/09 = 245.42 -> 12/12 =
        252.51 (approx +2.85% over 3 trading
        days).
    Earlier pattern: highs near 255 on
        11/30-12/02, pullback to ~244-245 on
        12/06-12/09, then a rebound to 252.51 on
         12/12 with higher volume (3.07e7).
    Short-term momentum: recent bounce and upward
         momentum from the Dec 9 low to Dec 12
        close.

    Step 2: Look for keywords in recent news to
        gauge overall sentiment.
    Microsoft-specific items: DoD cloud contract
        win (positive), acquisition of
        Lumenisity (positive), analyst mentions
        and cloud growth commentary (generally
        positive/neutral). Some macro/headline
        risk (Fed, inflation) but nothing
        Microsoft-specific that is strongly
        negative.

    Overall news sentiment for MSFT: neutral-to-
        positive.
```

```
    Step 3: Based on the analysis, choose the
        appropriate action.
    SELL is unavailable. Between BUY and HOLD,
        the short-term price momentum is upward
        and company-specific news is supportive.
         Cash is ample to buy.
    Decision: BUY (allocate available cash to
        take a full position).",
"action": "buy",
"reason":
    "Price Trend (short-term upward momentum
        supported by positive Microsoft-specific
         news)"
Epoch30:
"analysis":
    "Step 0 (Legality Check):
    BUY is allowed (cash 100000 >= price 252.51).
         SELL is not allowed (position = 0).
        HOLD is allowed.

    Step 1 (Decision Frame & Account Context):
    Horizon: short swing (multi-day). Account
        state: Flat (no position). Single
        objective: opportunity-seeking. Intent:
        holding optionality; burden of proof:
        moderate (since flat, require clearer
        asymmetry to buy).

    Step 2 (Regime Tag & Ledger):
    Regime: Range with recent bullish breakout
        attempt. Path cues: pivot sequence =
        higher highs from 12/5-12/12; volume
        character = spike on 12/12 (3.06e7) vs
        prior days moderate; breadth tone =
        mixed macro headwinds but positive cloud
        /contract news for Microsoft. Shelf-life:
         valid for next 3 sessions (short-swing).
         Falsifier: clear rejection under 245
        with expanding down-volume.

    Step 3 (Price Thesis & Clock):
    Primary thesis: price has recovered from late
        -Nov lows and closed higher to 252.51,
        suggesting short-term mean reversion/
        renewed demand into resistance area. Key
         supports: prior highs around 255 (12/1,
         12/2, 11/30) and intra-session low
        247.17; Key conflicts: recent volatile
        sessions and macro headwinds ahead of
        Fed. Signal quality: moderate - decent
        follow-through day with volume uptick.
        Corroboration window begins Day 0;
        within 3 sessions, sustained closes
        above ~250 and no breakdown would
        upgrade conviction.

    Step 4 (News Audit):
    News set: positive direct items - DoD cloud
        contract award, acquisition of
        Lumenisity, and product strategy (super
        app) - are moderately novel, direct to
        business durability medium (contracts
        and acquisition durable). Macro Fed
        worries are reiterated (not novel). News
         tightens monitoring but does not veto
        tape; confirmation would be sustained
        price absorption above 250 with stable
        volume.

    Step 5 (Reconcile Conflicts):
    Weighting: price-dominant with supportive
        durable corporate news; macro risk
        raises vigilance. As Flat, require
        clearer asymmetry to BUY - but current
        tape shows moderate edge. 'Why now vs
        wait': buying now captures rally cost
        but waiting preserves optionality; re-
        entry cost is low (spread/time) but
        opportunity cost is missing early
        momentum.

    Step 6 (Micro-invalidators & Cooldown):
    Nearest micro-invalidator: daily close below
        245 with expanding down-volume and loss
        of the 247-248 support. If that occurs,
        shift to heightened vigilance then
        tactical entry avoidance. Full reversal
        not applicable (no position). After any
        full flip (if buy now), cooldown would
        be 3 sessions.

    Step 7 (Action Choice - why now vs wait):
    Given moderate tape confirmation, supportive
        durable company news, flat account with
        ample cash, and manageable downside
        invalidator, initiating a full-position
        BUY aligns with opportunity-seeking on
        short-swing horizon. The proximate
        invalidator is close < 245 with
        expanding volume. Trial corroboration:
        2-3 sessions of closes >= 250 and rising
        /steady volume.

    Step 8-10 (Exit rules, monitoring, self-audit
        ):
    Plan: buy full position (allocate all cash).
        Post-action monitor for corroboration;
        cooldown 3 sessions. Self-audit:
        evidence meets moderate bar (price
        breakout with volume + durable news).",
"action": "buy",
"reason":
    "Price-dominant moderate breakout with volume
         and positive durable company news; flat
         account and ample cash support
        initiating a full-position BUY."
```

MSFT 2022-12-15

```
Epoch0:
"analysis":
    "Step 0: Filter illegal actions based on
        account state.

    Current price = 249.01, cash = 153.90,
        position = 382, cumulative profit margin
         = -0.0472.
    BUY not allowed because cash (153.90) < price
         (249.01).
    SELL allowed because position (382) > 0.
    HOLD always allowed.
    => Allowed actions: SELL, HOLD.

    Step 1: Examine the overall recent price
        trend.

    Recent closes: 12/13 = 256.92, 12/14 =
        257.22, 12/15 = 249.01.
    Price fell about 8.21 (approximately 3.2%)
        from 257.22 to 249.01 on 12/15 following
```

```
        a short peak on 12/13-14.
    The last three sessions show a peak then a
        sharp pullback; short-term momentum is
        negative.

    Step 2: Look for keywords in recent news to
        gauge overall sentiment.

    Positive corporate items: acquisition of
        Lumenisity, EU data boundary rollout,
        cloud initiatives (supportive long-term
        fundamentals).
    Negative/macro items: "Why Microsoft Shares
        Are Falling" (12/15) ties the drop to
        Fed rate hike and market weakness;
        options flow article notes a bearish
        whale.
    Overall near-term sentiment: mixed but
        leaning negative due to Fed-driven
        market weakness and bearish options
        activity coinciding with the price drop.

    Step 3: Based on the analysis, choose the
        appropriate action.

    Given allowed actions (SELL, HOLD), the short
        -term negative price momentum and near-
        term negative/macro-driven sentiment,
        and the objective to act decisively with
         full-position trades, the appropriate
        action is SELL to exit the position and
        preserve cash.",
"action": "sell",
"reason":
    "Price Trend: short-term negative momentum (
        sharp pullback from recent highs)
        triggered the sell."
Epoch30:
"analysis":
    "Step 0 (Legality Check):

    BUY allowed? cash = 248.58, price = 249.01 ->
         cash < price so BUY NOT allowed.
    SELL allowed? position = 395 (nonzero) ->
        SELL allowed.
    HOLD always allowed.

    Step 1 (Decision Frame & Account Context):

    Horizon: short swing (multi-day) given recent
         2-3 week data and active position.
    Account state: Fully invested (position
        nonzero) but near cash-depletion (cash ~
         price) - label: Fully invested / Cash-
        constrained.
    Single objective: invalidation monitoring (
        protect capital while seeking
        opportunity to exit on edge).
    Burden of proof: moderate-high (since fully
        invested and cumulative profit margin
        negative: -1.39%).
    Holding status: currently holding but
        reconsidering frame for potential SELL.

    Step 2 (Regime Tag & Ledger):

    Tag: Range with recent weak unwind attempts (
        mixed trend: higher highs mid-Dec then
        sharp drop today).
    Path cues: pivot sequence = recent higher
        highs 12/13-14 then failure 12/15 close
        below session open; volume character =
        elevated volume on 12/13-15 with 12/15
        heavy (3.56e7) on down day; breadth tone
         = risk-off (sector weakness after Fed
        hike).
    Shelf-life: valid for the next 3 sessions (
        short swing). Falsifier: resumption of a
         sustained higher-high sequence with
        expanding up-volume across 2 sessions.

    Step 3 (Price Thesis & Clock):

    Primary thesis: risk-off macro (Fed hike)
        drove broad tech pullback; MSFT saw
        failed follow-through after two up
        sessions; price likely to test lower
        support near the 244-247 range in the
        short run.
    Key supports: prior intraday lows 247-245,
        244 observed 12/6-7; cash cushion
        limited.
    Key conflicts: mid-Dec rallies on 12/13-14
        with higher highs and volume - shows
        buyers present but today's large down
        bar breaks structure.
    Signal quality: mixed to weakening - recent
        up-move lacked durable follow-through;
        today's down volume suggests
        distribution.
    Corroboration window begins Day 0 (today).
        Upgrades: within 3 sessions, if price
        reclaims 256+ with sustained volume
        conviction -> upgrade to buy-conviction;
         if price closes below 244 with
        expanding down-volume -> upgrade to sell-
        conviction.

    Step 4 (News Audit):

    Relevant items: Fed rate hike (12/14-15)
        causing sector weakness - Novelty: high
        (recent policy move); Directness: high (
        macro affects discount rates and growth
        multiples); Durability: moderate-high (
        policy path persists).
    Company news: acquisition of Lumenisity, EU
        data boundary rollout - novel but not
        immediately market-moving; mostly
        vigilance.
    Options whale bearish flows noted 12/14 -
        novel/direct but durability unclear.
    Overall: news supports vigilance and bias
        toward risk reduction; not a single
        durable catalyst to force immediate full
         sell by itself but aligns with tape.
    Price corroboration: today's down close and
        volume confirm news impact.

    Step 5 (Reconcile Conflicts & Dominant Driver
        ):

    Weighting: Price (weakening down-volume) +
        News (Fed tightening) dominate; regime
        tag (range/unwind) tightens tolerance.
    Dominant driver: price-dominant; news
        increases urgency to reduce risk.
    Account asymmetry: Fully invested -> require
        clear multi-pillar deterioration to SELL.
         We have price deterioration + durable
```

```
        macro news -> meets multi-pillar.

    Step 6 (Micro-invalidators & Cooldown):

    Minimal invalidator for reversing position:
        loss of key support around 244 with
        expanding down-volume OR another down
        day with close <247 and higher volume.
    Today's close 249.01 near support but below
        recent highs; down day on elevated
        volume already a micro-invalidator (
        first rung) allowing tactical de-risking
        ; second confirming day would justify
        full SELL.
    Since we must decide now, and SELL is allowed
        , consider whether conditions suffice
        for immediate full SELL per rules: we
        have price deterioration + news +
        account negative cumulative margin ->
        justify SELL to limit further downside.
    If SELL executed, impose cooldown of 3
        sessions before re-entry unless clear
        new evidence.

    Step 7 (Why now vs wait - Action Choice):

    SELL vs HOLD: Waiting risks further downside
        through policy-driven market weakness
        and limited re-entry cost is moderate (
        currently negative cumulative P/L). Re-
        entry cost: potential higher prices
        missed, but protecting capital is
        priority. Evidence meets the bar: price-
        dominated deterioration corroborated by
        macro news and elevated volume.
    Thus choose SELL to convert position to cash
        and stop downside exposure.
    Proximate invalidator to reverse decision
        later: sustained reclaim of 256+ with
        expanding up-volume over 2 sessions and
        macro tone stabilizing.

    Step 8-10 (Exit narrative, monitoring, self-
        audit):

    Exit graded: today counts as first micro-
        invalidator; selling now is tactical
        full reversal warranted by multi-pillar
        alignment.
    Post-action monitoring: expect initial
        consolidation; confirmatory invalidators
         would be further down-volume and daily
        closes <244.
    Self-audit: evidence meets pre-stated bar (
        price failure + news + account state).
        Action justified.",
"action": "sell",
"reason":
    "Price-dominant sell: elevated down-volume,
        failure after brief rally, and Fed-
        driven macro risk; as fully invested
        with negative cumulative P/L, de-risk by
         selling."
```