

# TRUSTTRADE: Human-Inspired Selective Consensus Reduces Decision Uncertainty in LLM Trading Agents

Minghan Li<sup>1</sup>, Rachel Gonsalves<sup>1,2</sup>, Weiyue Li<sup>1</sup>, Sunghoon Yoon<sup>3</sup>,  
Mengyu Wang<sup>1,4\*</sup>

<sup>1</sup>Harvard AI and Robotics Lab, Harvard University .

<sup>2</sup>Harvard Business School, Harvard University .

<sup>3</sup>Daegu Gyeongbuk Institute of Science and Technology .

<sup>4</sup>Kempner Institute for the Study of Natural and Artificial Intelligence,  
Harvard University.

\*Corresponding author(s). E-mail(s): [Mengyu\\_Wang@meei.harvard.edu](mailto:Mengyu_Wang@meei.harvard.edu);

Contributing authors: [mili4@meei.harvard.edu](mailto:mili4@meei.harvard.edu);

[rgonsalves@mba2027.hbs.edu](mailto:rgonsalves@mba2027.hbs.edu); [weiyueli@fas.harvard.edu](mailto:weiyueli@fas.harvard.edu);

[shyoon@dgist.ac.kr](mailto:shyoon@dgist.ac.kr);

## Abstract

Large language models (LLMs) are increasingly deployed as autonomous agents in financial trading. However, they often exhibit a hazardous behavioral bias that we term uniform trust, whereby retrieved information is implicitly assumed to be factual and heterogeneous sources are treated as equally informative. This assumption stands in sharp contrast to human decision-making, which relies on selective filtering, cross-validation, and experience-driven weighting of information sources. As a result, LLM-based trading systems are particularly vulnerable to multi-source noise and misinformation, amplifying factual hallucinations and leading to unstable risk–return performance. To bridge this behavioral gap, we introduce TRUSTTRADE (Trust-Rectified Unified Selective Trader), a multi-agent selective consensus framework inspired by human epistemic heuristics. TRUSTTRADE replaces uniform trust with cross-agent consistency by aggregating information from multiple independent LLM agents and dynamically weighting signals based on their semantic and numerical agreement. Consistent signals are prioritized, while divergent, weakly grounded, or temporally inconsistent inputs are selectively discounted. To further stabilize decision-making, TRUSTTRADE incorporates deterministic temporal signals as reproducible anchors and

a reflective memory mechanism that adapts risk preferences at test time without additional training. Together, these components suppress noise amplification and hallucination-driven volatility, yielding more stable and risk-aware trading behavior. Across controlled backtesting in high-noise market environments (2024 Q1 and 2026 Q1), the proposed TRUSTTRADE calibrates LLM trading behavior from extreme risk–return regimes toward a human-aligned, mid-risk/mid-return profile.

**Keywords:** large language models, autonomous trading agents, multi-agent consensus, factual hallucination, decision uncertainty

The rapid advancement of large language models (LLMs)<sup>1–3</sup> has accelerated their deployment as autonomous agents<sup>4,5</sup> in financial decision-making<sup>6</sup>. Owing to their capacity to ingest and synthesize heterogeneous information, including market data, corporate fundamentals, news, and social sentiment, LLMs offer an appealing foundation for automated trading systems<sup>7</sup> operating at unprecedented scale and speed<sup>5,8</sup>.

Recent research has consequently developed a range of advanced LLM-based trading frameworks<sup>9–14</sup>. These approaches often leverage LLMs to incorporate chain-of-thought<sup>15</sup>, external information retrieval, tool use, multi-agent collaboration<sup>16</sup> or reinforcement learning<sup>17</sup> for portfolio optimization<sup>10,18</sup>, market sentiment analysis<sup>19</sup>, cross-modal information integration<sup>20</sup>, and end-to-end trading decision generation<sup>11,21</sup>. Overall, these cutting-edge techniques extend model capabilities in both simulated and realistic market environments.

However, in quantitative trading settings<sup>22–24</sup>, the limitations of LLM-based approaches have become increasingly evident. Because LLM reasoning is inherently stochastic and highly sensitive to how information is formulated and combined, trading decisions can be overly responsive to multi-source noise and minor input perturbations, resulting in unstable and poorly reproducible outcomes. To mitigate uncertainty<sup>25</sup> and non-stationarity<sup>26</sup> in financial markets, recent approaches introduce structural constraints through multi-agent strategy search and evolutionary modeling<sup>27,28</sup>, or through logic-oriented representations of market narratives using semantic and temporal alignment<sup>29</sup>; however, these methods primarily focus on strategy generation and market structure characterization.

In contrast, we argue that decision uncertainty in LLM-based trading arises from a more fundamental limitation: a systematic bias in how information is trusted and integrated. Recent studies show that under open-ended retrieval and online search, LLM agents are prone to factual hallucinations and spurious correlations<sup>30–32</sup>. In trading settings, such failures extend beyond semantic errors and can directly influence portfolio decisions, amplifying volatility and drawdowns.

Crucially, this vulnerability is not driven by insufficient model capacity but by an implicit uniform-trust assumption in many existing approaches, which treat information retrieved or generated by LLM agents as factual and equally reliable<sup>20,33,34</sup>. This assumption ignores substantial variation in source quality, reliability, and temporal relevance. In high-noise, temporally dependent, and potentially adversarial financial

environments, even minor distortions can therefore be magnified into disproportionate risk exposure. Our empirical results indicate that this limitation reflects a behavioral mismatch between LLM agents and human annotators, rather than a purely technical shortcoming.

In our empirical study, human annotators follow a markedly different strategy in financial decision-making<sup>35</sup>. Rather than weighting information uniformly, human annotators rely on selective filtering<sup>36</sup>: they prioritize signals that are temporally coherent, repeatedly corroborated, and historically reliable, while down-weighting narrative-driven, emotional<sup>37</sup>, noisy, or weakly grounded inputs. This process is tightly coupled with memory<sup>38,39</sup> (such as past experiences, prior outcomes, and long-term context), which stabilizes decision-making and mitigates overreaction to transient market narratives.

To bridge this behavioral gap, we introduce TRUSTTRADE (Trust-Rectified Unified Selective Trader), a multi-agent selective consensus framework inspired by human epistemic heuristics. Instead of relying on a single LLM agent, our approach deploys multiple independent agents to collect and interpret information in parallel. The framework operationalizes the principle of invariance of truth: objective and reliable signals should remain consistent across independent reasoning paths. By quantifying cross-agent agreement in both semantic and numerical space, we derive dynamic credibility scores that govern information weighting. High-consensus signals are prioritized, while divergent, weakly grounded, or temporally inconsistent inputs, including those implicitly encoding look-ahead bias, are selectively discounted.

Beyond consensus filtering, TRUSTTRADE incorporates a deterministic temporal signal module that compresses raw price dynamics and market states into reproducible and auditable time-series indicators, including trend, momentum, volatility, drawdown, and risk exposure. These signals serve as stable anchors for downstream decision-making, constraining the influence of unreliable textual evidence and providing a consistent temporal reference across trading horizons. As a result, strategy updates become less sensitive to spurious information and less prone to abrupt fluctuations driven by factual hallucinations.

Finally, TRUSTTRADE integrates a memory bank with short- and long-term decision reflection mechanisms to enable test-time adaptation without additional training. The system continuously records historical trading contexts, including supporting evidence, consensus scores, temporal signals, executed actions, and confidence levels, and performs retrospective evaluations based on realized return–risk outcomes. Through this process, the model progressively calibrates its risk preferences and information-weighting strategy.

We evaluate TRUSTTRADE through controlled A/B experiments against standard LLM agents and human annotators in high-noise market simulations and real-world markets. The results demonstrate that TRUSTTRADE fundamentally reshapes LLM behavior: maximum drawdown is substantially reduced, risk–return trade-offs shift toward the human-preferred regime, and performance becomes more stable. Importantly, these improvements arise not from increased model capacity, but from behavioral alignment.

# Results

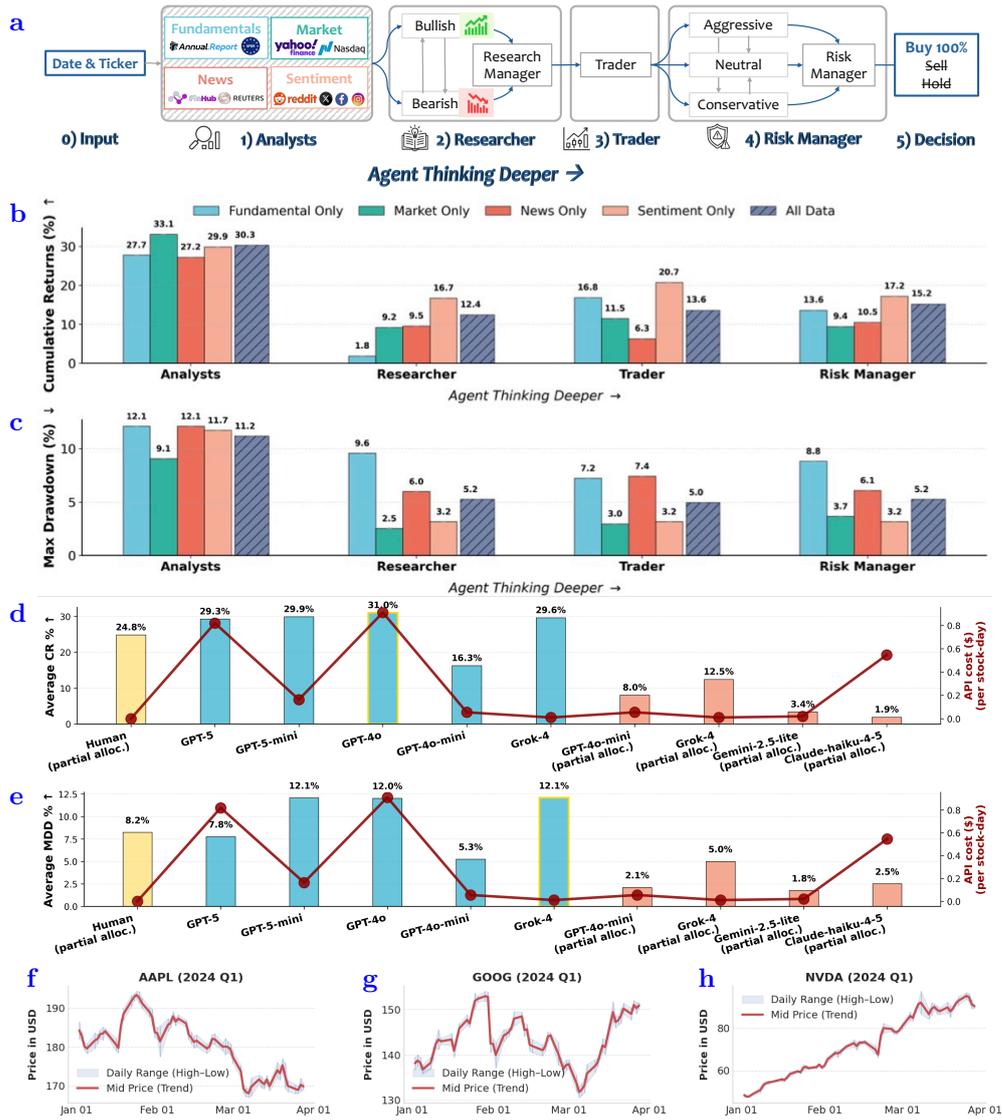
## Study I: Diagnosing Instability in LLM-Based Trading Agents

Across all three market regimes examined in 2024-Q1 (Fig. 1 f-h), we observe pronounced heterogeneity in trading behavior among LLM agents, whose structure is showed in Fig. 1 a. Even when exposed to identical market conditions, different agents frequently arrive at divergent trading decisions, leading to substantial variation in cumulative returns, drawdowns, and risk-adjusted performance (Fig. 1 b-e). This heterogeneity persists across reasoning depth, agent architectures, and the composition of accessible information sources, indicating that trading outcomes are not determined solely by market dynamics but are strongly shaped by model-specific decision processes.

**Reasoning depth of LLM agents.** We examined how trading performance changes as LLM agents operate with increasing reasoning depth (Fig. 1 b-c). The results reveal a clear stage-wise structure rather than a gradual or monotonic improvement. Advancing from the Analyst to the Trader stage leads to substantial gains in decision quality, reflected by markedly reduced maximum drawdowns. This transition captures nearly all performance benefits associated with deeper reasoning, indicating that the integration of analysis and execution constitutes the critical bottleneck for effective trading decisions. In contrast, adding a dedicated Risk Manager stage does not produce systematic benefits. Its behavior largely overlaps with that of the Trader stage and fails to consistently improve either returns or risk control. On the basis of these results, the Risk Manager stage is excluded from subsequent experiments, as it introduces additional complexity without measurable gains in decision quality.

**Non-additive effects of information modalities.** Information sources have non-additive effects (Fig. 1 b,c). At the Analyst stage, market-only signals perform best (33.1% CR; 9.1% MDD), while adding fundamentals/news/sentiment reduces returns and increases drawdown, indicating that naive multi-source aggregation injects noise. At the Trader stage, sentiment-only appears strong but is prone to leakage because it is retrieved via online search; combining sources still fails to improve performance (13.6% return; 5.0% MDD), suggesting limited ability to filter unreliable evidence.

**Systematic bias across LLM agent variants.** Fig. 1 d-e reveals clear heterogeneity in risk-return behavior across LLM agents and human annotators. Under full allocation, large LLMs (GPT-5, GPT-5-mini, GPT-4o, and Grok-4) achieve the highest cumulative returns ( $\approx 30\%$  on average), but consistently incur large drawdowns ( $\approx 12\%$  MDD), indicating aggressive, high-confidence trading with limited risk control. In contrast, GPT-4o-mini exhibits substantially lower returns but markedly reduced drawdowns, despite operating under the same allocation constraint, suggesting more conservative timing and weaker signal amplification. Under partial allocation, GPT-4o-mini and Grok-4 shift toward a risk-controlled regime, with returns decreasing from 29.6% to 12.5% and drawdowns reduced from 12.1% to 5.0%. Smaller models (Gemini-2.5-lite and Claude-Haiku-4.5) consistently produce low returns ( $<4\%$ ) with minimal drawdowns ( $<2.5\%$ ), reflecting strongly risk-averse behavior. Overall, these results show a systematic separation in behavior: larger LLMs favor high-reward strategies with elevated downside risk, whereas smaller models and human annotators prioritize



**Fig. 1: Diagnosing instability in LLM-based trading agents (with GPT-4o-mini).** This figure illustrates how data sources, reasoning depth, and allocation regimes jointly contribute to the instability and risk-return profiles of LLM trading agents. **a**, Schematic of a standard LLM-based trading pipeline<sup>20</sup> with increasing reasoning depth, progressing from Analysts to Researcher, Trader, and Risk Manager. **b-c**, Cumulative returns and maximum drawdown across LLM agents under different data-source and reasoning-depth ablations. Return and risk vary across data sources, transitioning from high-risk/high-return to more balanced trade-offs with increasing reasoning depth. **d-e**, Risk-return heterogeneity across human and LLM trading agents under different allocation regimes. Average cumulative return (CR) and maximum drawdown (MDD) across stocks, with API cost per stock-day shown on the right axis. Full-allocation agents earn higher returns but suffer larger drawdowns than partial-allocation agents and human annotators. **f-h**, Stock price dynamics during 2024-Q1. Note that all reported results are averaged over these three stocks (AAPL, GOOG, NVDA) during 2024-Q1.

drawdown control, highlighting a fundamental gap between LLM-driven and human behavior.

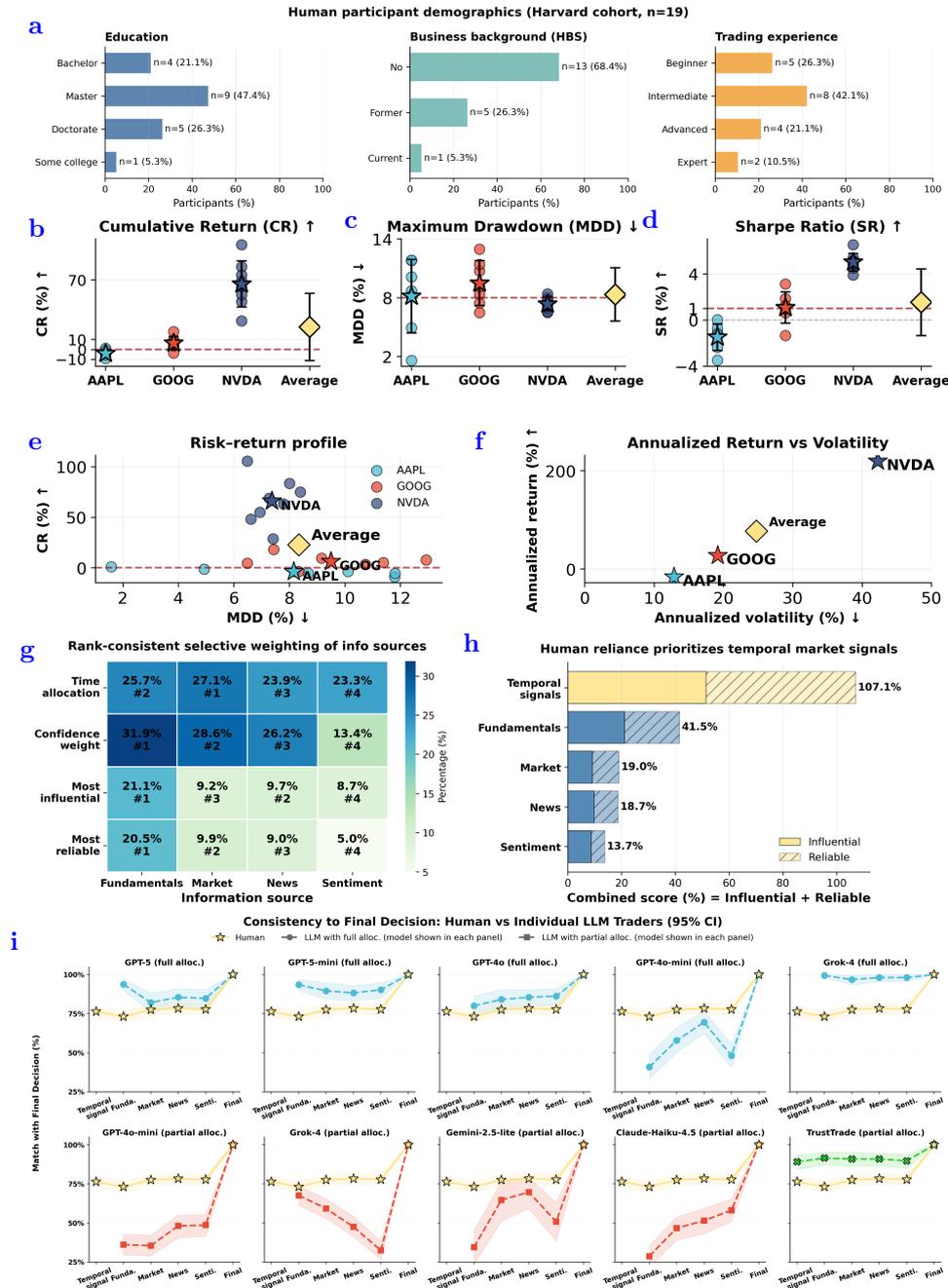
## Study II: Characterizing Human Stabilization Mechanisms

Human trading behaviour was studied using an online, stepwise simulation with structured logging. Participants progressed through a fixed sequence of information stages (temporal signals, fundamentals, market indicators, news and sentiment) and provided stage-wise actions, confidence ratings and rationales. We additionally recorded time spent on each information source and post-hoc judgments of influence and reliability. These signals yield reproducible behavioural traces that enable a direct comparison with sequential LLM decision pipelines (Fig. 2). Nineteen human annotators were recruited (students or faculty members at Harvard University), whose personas are shown in Fig. 2 a.

**Human trading exhibits stable risk–return integration.** Human annotators exhibit moderate returns with consistently controlled risk across assets in Fig. 2 b–d. While cumulative returns vary by asset, highest for NVDA and lowest for AAPL, performance remains stable and avoids extreme outcomes. Maximum drawdowns are relatively contained in all cases, indicating a strong preference for capital preservation over aggressive return maximization. In the risk–return space of Fig. 2 e, human strategies cluster within a mid-risk, mid-return regime, avoiding both high-volatility/high-return and low-risk/low-return extremes. Consistent with this behavior, the annualized return–volatility analysis in Fig. 2 f shows that human annotators systematically accept lower returns in exchange for reduced volatility, in contrast to strategies that aggressively exploit high-volatility opportunities.

**Empirical evidence of data filtering and temporal signals.** Human annotators exhibit a highly structured and selective pattern of information use. In Fig. 2 g, we separately quantify (i) the time that human annotators spend on each information source, (ii) their confidence weight assigned to that source (0-100), and (iii) post-hoc selections of the most influential and most reliable sources. The proportion of time allocation, confidence weighting, and post-hoc judgments of influence and reliability are strongly rank-consistent across information sources, with fundamentals and market signals consistently prioritized over news and sentiment. Together, these measures point to an implicit filtering strategy in human information use, in which attention and confidence co-vary with perceived utility rather than being allocated independently across sources. Fig. 2 h reveals a second, complementary signature of human stabilization: a marked preference for temporally anchored evidence. When influence and reliability are aggregated, temporal signals dominate (combined score, 107.1%), exceeding fundamentals, market signals, news and sentiment by a wide margin. Thus, even in the presence of narrative inputs, human judgements remain grounded in verifiable price dynamics and benchmark movements, with news and sentiment treated as ancillary and comparatively unreliable.

**Decision convergence distinguishes human and LLM trading decisions.** Decision convergence reveals fundamental behavioral differences between human annotators and LLM traders. Fig. 2 i shows the consistency between intermediate decisions and the final trading action across sequential information stages. Human annotators



**Fig. 2: Behavioral signatures of human trading.** **a**, Demographic profile of human annotators (n=19). **b-d**, Human trading outcomes across stocks, showing moderate cumulative returns, tightly controlled drawdowns, and near-neutral Sharpe ratios, with consistent risk exposure across assets. **e-f**, Risk-return and volatility profiles illustrate that human exhibits substantially greater dispersion across assets. **g-h**, Selective information weighting by human annotators: time allocation, confidence weighting and post-hoc influence/reliability are aligned, and the combined influence-reliability score emphasizes temporally grounded signals (price trends and market indices) over narrative-driven inputs (news and sentiment). **i**, Decision convergence across sequential information stages for human annotators and nine LLM-based traders under full and partial-allocation settings. Human annotators show consistently high convergence to the final action across stages, whereas LLM traders exhibit lower and more variable convergence.

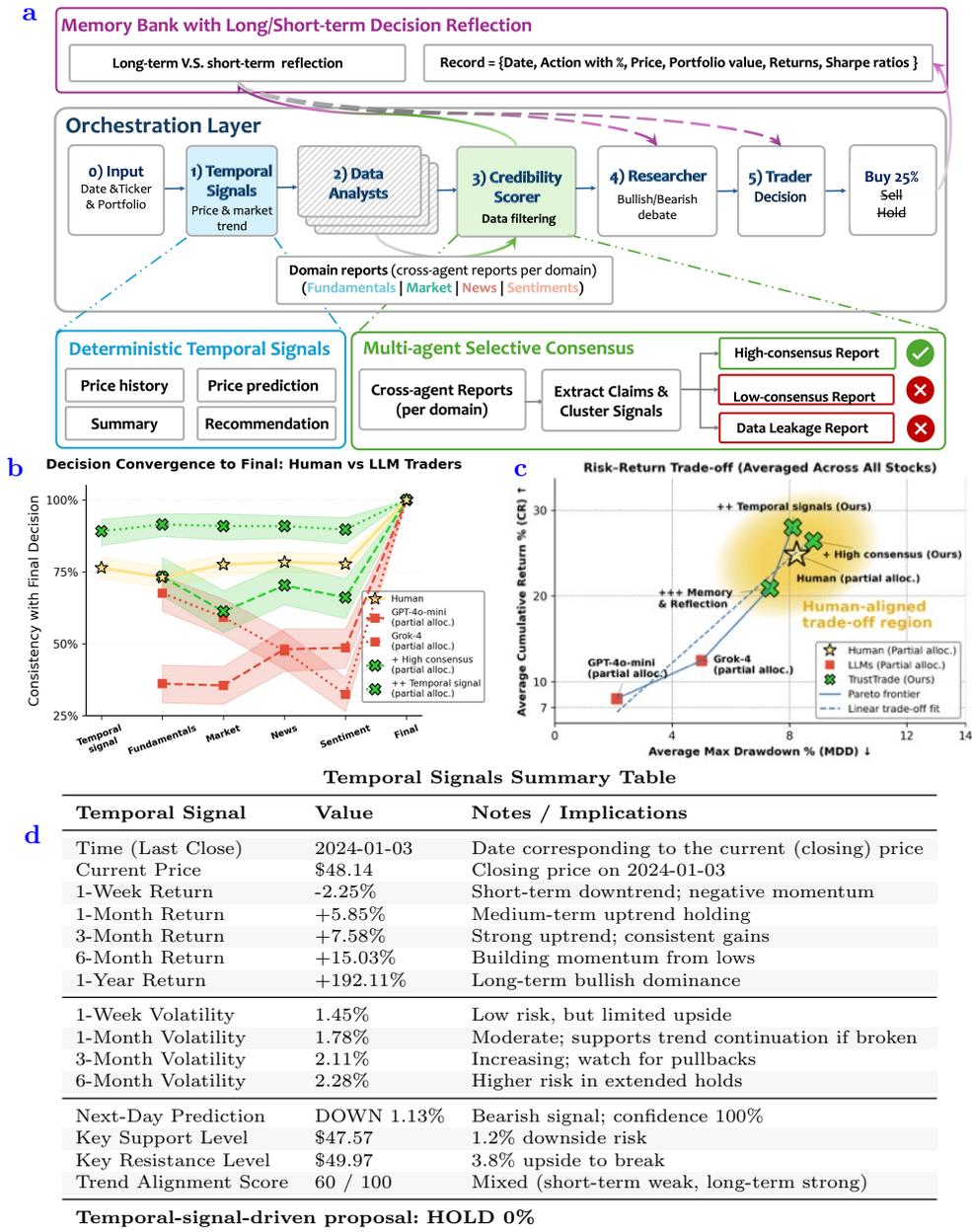
exhibit uniformly high convergence throughout the decision pipeline, indicating that early-stage judgments are largely preserved and incrementally refined rather than overwritten. This pattern suggests stable internal representations supported by selective data filtering and historical memory. In contrast, LLM traders show pronounced stage-wise inconsistency. Across models and allocation regimes, intermediate decisions often diverge sharply from the final action, with convergence occurring abruptly only at the final stage. This behavior indicates reactive decision revision rather than cumulative reasoning, consistent with limited temporal memory and weak commitment to earlier evidence. Partial-allocation strategies moderately improve convergence for some models, particularly GPT-4o-mini and Grok-4, but substantial instability remains across intermediate stages. Overall, these results demonstrate that human annotator behavior is characterized by early stabilization and memory-consistent integration, whereas current LLM traders rely on late-stage aggregation, highlighting a core behavioral gap in sequential decision-making.

### Study III: Aligning LLM Trading Behavior via Consensus, Temporal Signals and Memory

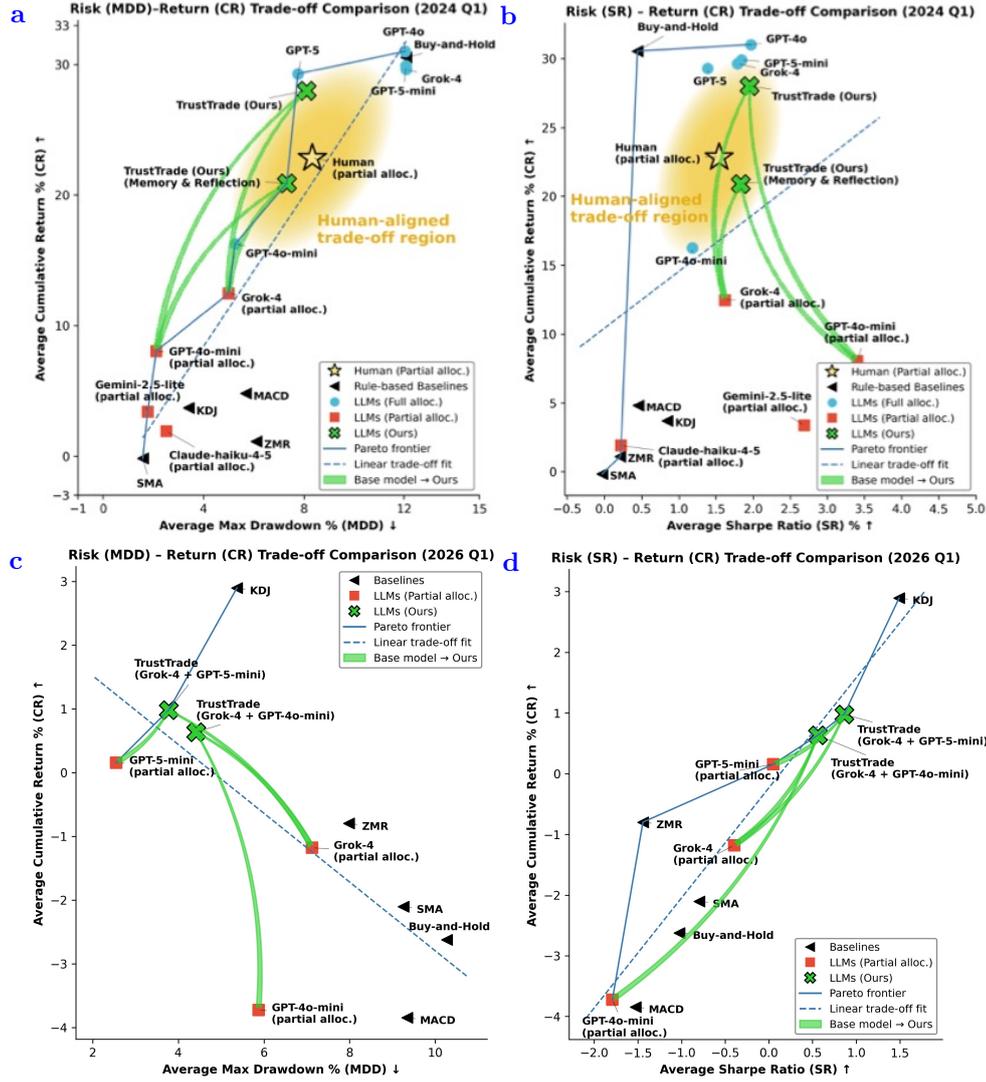
Motivated by the behavioural gap identified in Study II, we next introduce interventions that explicitly target stage-wise stabilization and human-aligned risk control. Fig. 3 summarizes how our design choices improve both the *stability* and the *human alignment* of LLM trading. We decompose the effect into three components: selective consensus, deterministic temporal signals, and a memory bank with long/short-term decision reflection, and report their individual contributions to convergence and risk–return behaviour.

**TrustTrade overview.** Fig. 3 a provides an overview of TrustTrade, which couples an orchestration layer with a memory bank updated via long/short reflection. For each information domain (fundamentals, market data, news, sentiment), we query multiple heterogeneous LLM agents (e.g., GPT-family and Gemini-family models) to collect independent evidence and form cross-agent domain reports. A credibility scorer then filters and ranks these reports to retain high-consensus information and suppress noisy or conflicting claims. A researcher consolidates the retained evidence before the trader produces the final action. The reflection module equips the memory bank with explicit strategy-evaluation signals (e.g., temporal return and Sharpe slopes over short and long horizons), allowing past decisions to be scored and summarized before being fed back into subsequent decision-making.

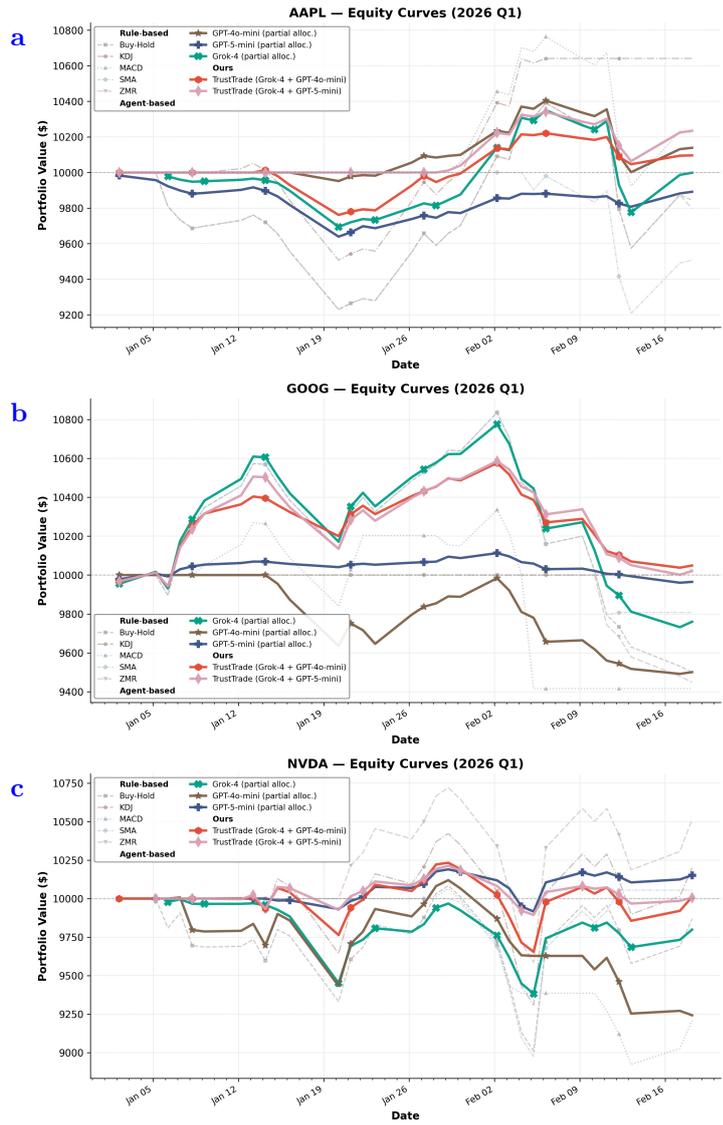
**Selective consensus improves decision stability.** Using diverse LLM agents (e.g., GPT-4o-mini and Grok-4) to generate domain reports without filtering leads to poor decision consistency across stages in Fig. 3 b. In contrast, introducing selective consensus substantially improves consistency with the final decision by retaining only high-consensus information: our full pipeline maintains high agreement with the final action across intermediate reasoning stages, whereas baseline LLM traders exhibit larger mid-stage reversals before converging at the final step. This stability gain is accompanied by a marked return improvement in the risk–return plane (Fig. 3 c),



**Fig. 3: Human-aligned trading behavior induced by multi-agent consensus filtering, temporal signals, and memory bank with long/short-term decision reflection.** **a**, Overview of the proposed TRUSTTRADE framework: multiple agents collect information from diverse sources, a credibility scorer filters for high-consensus evidence, and the resulting decision is used to update a memory bank with both short- and long-term reflections. **b**, Decision convergence across sequential stages, comparing human annotators, baseline LLM traders, and the proposed TRUSTTRADE. **c**, Risk–return trade-off averaged across stocks; the shaded ellipse marks the human-aligned preference region defined by the standard errors of human-annotator cumulative return (CR) and maximum drawdown (MDD). **d**, Newly introduced temporal-signal summary reporting deterministic, price-derived trends and indicators to reduce hallucination and improve decision reliability.



**Fig. 4:** Comprehensive backtesting comparison across all baselines over the 2024 Q1 and 2026 Q1 trading period respectively. **a & b**, Risk–return trade-off averaged across NVDA, AAPL, and GOOG in 2024 Q1. Each point represents an agent configuration, plotted by average cumulative return (CR) and average maximum drawdown (MDD) / average Sharpe Ratio (SR). Human annotators (yellow star) define a human-aligned risk–return preference region (shaded ellipse). Full-allocation LLM agents (blue dots) achieve high returns but incur substantial drawdowns, whereas partial-allocation LLMs (red squares) reduce risk at the cost of diminished returns. Our TRUSTTRADE builds on GPT-4o-mini and Grok-4 under partial allocation, and we compare two variants: without memory and reflection, the method achieves higher returns than human annotators at comparable maximum drawdown (approaching GPT-5 performance), while adding memory and reflection slightly reduces returns but further lowers risk. The Pareto frontier and linear trade-off fit are shown for reference. **c & d**, Day-by-day backtesting risk–return trade-off across NVDA, AAPL, and GOOG during 2026 Q1, comparing rule-based baselines, single-LLM traders, and our multi-LLM framework. Our TRUSTTRADE achieves a substantially improved return–risk balance, with higher returns and lower risk than the comparison methods.



**Fig. 5:** Daily real-time trading performance on AAPL, GOOG, and NVDA during 2026 Q1. This figure reports day-by-day results in a forward-time setting to reduce potential leakage from earlier market-period evaluation. Rule-based baselines show return swings that closely track price volatility, while pure LLM traders exhibit comparatively unstable behaviors and outcomes. By integrating multi-agent information with selective consensus, TRUSTTRADE improves return performance with more stable trading trajectories.

with average cumulative return increasing from roughly 10% to  $\sim 26\%$  under the high-consensus configuration, at the cost of an unavoidable increase in MDD from about 3% to  $\sim 8\%$ .

**Temporal signals provide deterministic grounding.** Fig. 3 d shows a temporal-signal summary table generated by our newly introduced analyst that reports deterministic, price-derived trends and indicators. Fig. 3 b further shows that introducing these reproducible temporal signals substantially improves decision consistency, bringing LLM decisions better than human stage-wise convergence. At the portfolio level (Fig. 3 c), adding temporal signals increases average cumulative return by about 1% while slightly reducing risk (MDD), confirming that deterministic temporal grounding is critical for stabilizing trading decisions.

**Memory bank regularizes sequential updates via long/short-term decision reflection.** Fig. 3 a outlines the memory bank with long/short reflection and its role in our orchestration layer. The reflection module evaluates whether a strategy is improving or degrading by tracking temporal performance slopes (e.g., return and Sharpe slopes over short and long horizons) and uses these signals to update the memory bank at test time. This feedback is immediately reused in subsequent trading steps, enabling test-time adaptation: the agent adjusts its allocation and risk posture by borrowing from similar past situations and down-weighting strategies whose recent slopes deteriorate. Compared with variants without memory and reflection, this component substantially reduces risk, with an acceptable trade-off of slightly lower returns, yielding more conservative and robust decisions and shifting the agent configuration toward the human-aligned preference region in Fig. 3 c.

## Study IV: Overall Benchmark Comparison

Fig. 4 and Fig. 5 consolidate all baselines into a unified return–risk comparison. Fig. 4 involves human annotators as a behavioural reference to contextualize where LLM agents and our method operate on the performance frontier. The plot shows average CR or SR versus average MDD, aggregated across NVDA, AAPL and GOOG. Each point corresponds to one agent configuration; lower MDD or higher SR indicate lower risk, and higher CR indicates higher return. Human annotators (partial allocation) define a human-aligned preference region (shaded ellipse), providing a behavioural reference for acceptable risk exposure.

**Risk-return comparison on 2024-Q1.** Fig. 4 a,b summarizes the overall benchmark comparison in 2024-Q1 and highlights three findings. High-return baselines (e.g., buy-and-hold and full-allocation LLM traders) largely lie in a high-drawdown regime, underscoring that raw return alone is misaligned with human risk preferences. Across LLM baselines, model capacity and allocation policy jointly shape a clear return–risk trade-off: LLM agents tend to operate in either a high-return/high-risk regime or a low-return/low-risk regime, with stronger models under full allocation attaining the largest gains but also the largest drawdowns, whereas partial allocation reduces risk at the expense of return, with weaker models clustering near the low-return end. Against this backdrop, our TRUSTTRADE based on Grok-4 and GPT-4o-mini model shifts the partial-allocation frontier by improving return at comparable drawdown (and/or reducing drawdown at comparable return), and adding memory and reflection further robustifies performance by lowering drawdown with only a modest return sacrifice, moving the operating point toward the human-aligned region.

**Real-time backtesting to mitigate data leakage.** When trading on historical market periods, a potential concern is data leakage: language models may implicitly encode future price movements through pretraining or contamination, which could artificially inflate performance and distort decision rationales. To address this risk, we conduct a real-time backtest from January 1, 2026 to February 18, 2026. Each trading day at 1:00 PM, the agent collects contemporaneous information available up to that timestamp, generates a trading decision, and executes the simulated trade, ensuring that future prices are not accessible at decision time. Fig. 5 plots the day-by-day return curves of all compared methods on AAPL, GOOG, and NVDA in 2026 Q1, while Fig. 4 c,d reports the corresponding stock-averaged risk–return trade-off comparison. Overall, rule-based strategies fluctuate strongly with market price swings: KDJ achieves a relatively strong return, whereas most other rule-based baselines remain in negative-return regimes. Single-LLM agents are also unstable across assets, with different models peaking on different stocks (e.g., GPT-4o-mini on AAPL, Grok-4 on GOOG, and GPT-5-mini on NVDA). In contrast, our multi-LLM consensus model delivers more stable performance across stocks, reducing risk while improving returns.

## Discussion

We investigated why LLM trading agents often fail to exhibit human-like risk control and decision stability, and proposed TRUSTTRADE to align LLM trading behavior with human annotators through three complementary components: (i) selective, credibility-scored multi-agent consensus, (ii) deterministic temporal signals that summarize market state, and (iii) a memory bank updated via long/short reflection for test-time adaptation. Across controlled studies in 2024-Q1, our analysis shows substantial heterogeneity across agents and reasoning pipelines, and reveals that instability is driven by noisy multi-source integration and weak commitment to intermediate evidence. In Study III, selective consensus improves stage-wise consistency with final actions, temporal signals provide reproducible grounding to reduce hallucination, and memory+reflection further regularize sequential updates by using temporal return/Sharpe slopes to evaluate and adapt strategies.

Overall benchmark results (Study IV) show that TRUSTTRADE improves the return–risk trade-off relative to strong partial-allocation baselines, and that memory with reflection yields a more conservative configuration by reducing drawdown at the cost of a small return decrease. To mitigate concerns about data leakage when evaluating on historical periods, we further introduce a real-time backtesting protocol (2026-Q1) that restricts each decision to contemporaneous information available at that timestamp.

**Limitations.** First, while deterministic temporal signals constrain downstream reasoning, the pipeline still relies on LLM-generated domain reports that may contain omissions, framing bias, or subtle errors. Second, the credibility scorer and consensus filtering can suppress minority-but-correct evidence, particularly in fast-changing or low-coverage market regimes. Third, to enable a controlled comparison to human annotator behavior, we focus on a small set of representative stocks to make the task quantifiable; due to limited API budgets, we also restrict the evaluation to limited time

windows. Future evaluations should extend to diversified multi-stock portfolios, which better reflect real-world trading settings and may change the observed return–risk trade-offs.

**Future work.** Promising directions include (i) real-time backtesting on a broader universe of stocks, including non-megacap and non-technology names, (ii) stronger anti-leakage evaluation protocols (e.g., strict time-stamped data feeds and controlled news snapshots), (iii) learning calibrated confidence and position sizing policies that explicitly target human-aligned drawdown constraints, and (iv) adaptive consensus mechanisms that incorporate uncertainty estimates to avoid discarding rare but informative signals. More broadly, integrating causal market features, transaction costs, and liquidity constraints, and extending the memory bank to support longer-horizon planning and regime detection, could further improve robustness and real-world applicability.

## Methods

This section is organized into three parts. We first describe the human trading simulation protocol and data collection pipeline. We then present TRUSTTRADE, which integrates selective multi-agent consensus, deterministic temporal-signal analytics, and a reflective memory bank for test-time adaptation. Finally, we summarize experimental details, including rule-based baselines, evaluation metrics, and the human-aligned preference region used for behavioral comparison.

### Human Trading Simulation and Result Collection

To study human decision-making under realistic market information, we designed a stepwise trading simulation with structured logging and persistent storage. The pipeline includes offline data preparation followed by an online interaction loop.

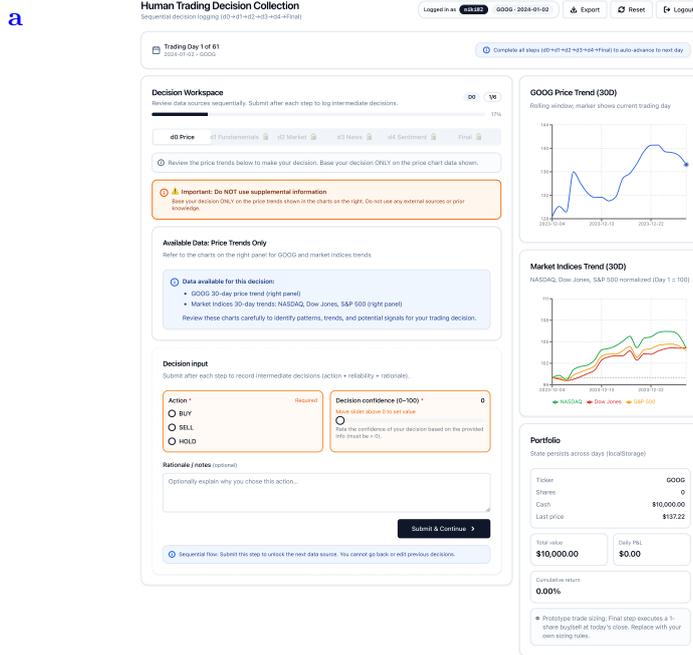
Before user interaction, we preprocess and standardize the raw market, fundamental, and news/sentiment data required by the simulation. To avoid priming participants, we use GPT-4o-mini to collect domain reports as the raw information shown in the interface, and remove any decision-related content (e.g., explicit buy/sell recommendations or position sizes). At runtime, the web interface authenticates participants via a user ID and collects basic demographic/background information (e.g., education level and business/finance experience), then restores prior progress if available or starts a new session by loading the prepared data. Participants iterate through trading days until completion, with automatic portfolio updates and cloud saves after each day. Each trading day is decomposed into six ordered stages:

- d0: Temporal signals (30-day price history with ticker and indices)
- d1: Fundamentals (e.g., market cap, P/E, EPS)
- d2: Market/technical indicators (e.g., 10/30-day moving averages)
- d3: News (title, source, summary)
- d4: Social sentiment (sentiment and analyst ratings)
- Final composite (aggregated view with execution)

At every stage, participants provide an **action** (BUY/SELL/HOLD), a reliability score (1–100), and a free-text rationale. For stages d1–d4, we additionally record a data-leakage flag indicating whether AI-generated decisions were inadvertently visible. In the final stage, participants also specify the most influential source, most reliable source, and trade size (25/50/75/100%), after which portfolio execution is applied.

To support recovery and analysis, results are written to Amazon Web Services (AWS) S3 under structured keys: full-session exports, portfolio state, and progress markers. This design enables interruption-safe sessions, reproducible behavioral traces, and standardized downstream analysis across participants.

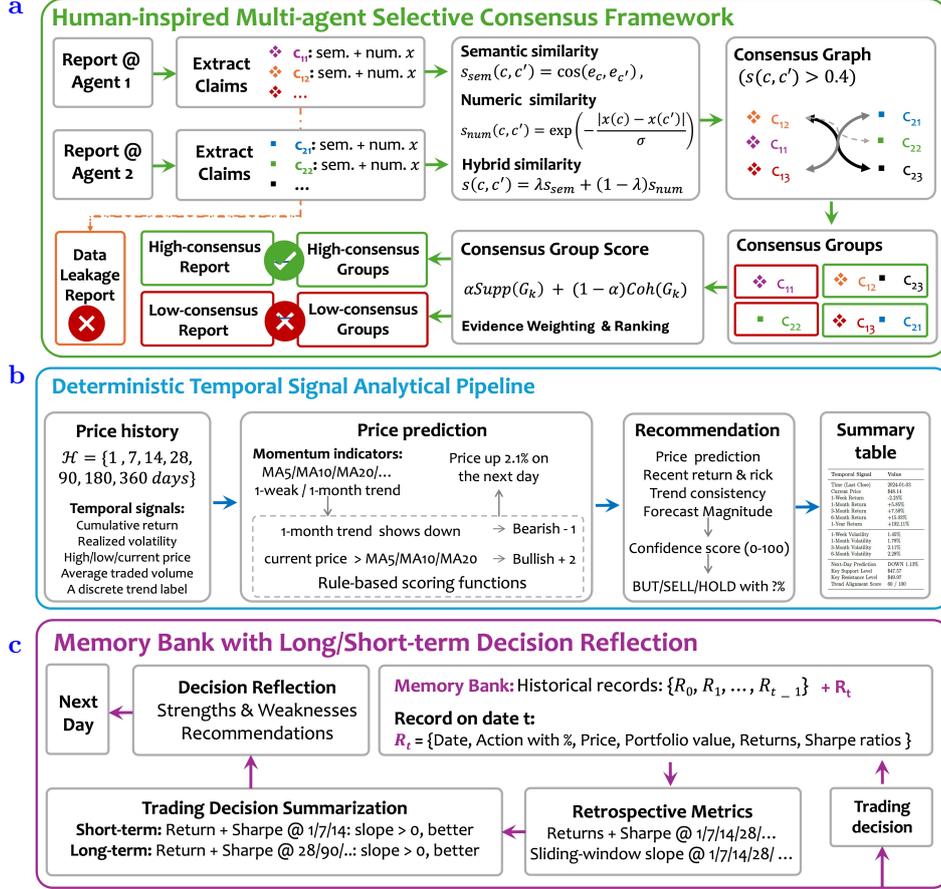
## TrustTrade: Trust-Rectified Unified Selective Trade



**Fig. 6:** User interface for the systematic collection of human decision annotations.

**Overview.** TRUSTTRADE, adapted from<sup>20</sup>, comprises three components: (i) multi-agent domain report collection and credibility-scored selective consensus, (ii) deterministic temporal signals with price-derived indicators, and (iii) a memory bank with long/short reflection that supports test-time adaptation.

**Human-inspired multi-agent selective consensus framework.** We introduce a selective consensus mechanism to quantify the reliability of retrieved financial data. The framework departs from the single-agent paradigm by using  $N$  independent agents to collect and parse domain reports in parallel. We then apply domain-wise credibility



**Fig. 7: Three main modules of the proposed TrustTrade pipeline. a**, Multi-agent selective consensus for domain report collection and credibility scoring. **b**, Deterministic price-based analysis module that produces temporal signals and rule-based forecasts. **c**, Memory bank with short/long-term decision reflection agents that enable test-time adaptation.

scoring (fundamentals, market, news, and sentiment) to aggregate these reports: the scorer outputs a structured evidence summary (market signal, source, cross-report consistency, and confidence level), promotes cross-report consistent facts and signals to a high-confidence set, flags conflicting or weakly supported claims as low confidence, and performs a data-leakage audit. The high-confidence summary is used as the default evidence for downstream decision making.

To make credibility scoring more reproducible, we additionally implement a deterministic consensus procedure over atomic claims. For each report  $r_i$ , we extract a set of point-by-point claims  $C_i = \{c_{ij} | j = 1, \dots, m_i\}$ , where each claim is normalized into a fixed schema (domain, subject, predicate, polarity, optional numeric value, and

timestamp). We embed each claim into a vector  $e_{c_{ij}} \in \mathbb{R}^d$  and compute the *semantic* similarity between claims from different reports using cosine similarity, denoted by  $s_{\text{sem}}(c, c')$ . For claims with numeric values, we define a numeric agreement term  $s_{\text{num}}(c, c')$ . These terms are combined into a final hybrid similarity score:

$$\begin{aligned} s_{\text{sem}}(c, c') &= \cos(e_c, e_{c'}), & s_{\text{num}}(c, c') &= \exp(-|x(c) - x(c')|/\sigma); \\ s(c, c') &= \lambda s_{\text{sem}}(c, c') + (1 - \lambda) s_{\text{num}}(c, c'). \end{aligned} \quad (1)$$

We build a cross-report similarity graph whose nodes are claims and whose edges connect claims from different reports when  $s(c, c') \geq \tau$ . Connected components define consensus groups  $G_k$ . For each group, we compute the support ratio and define the cohesion weight as the within-group average pairwise similarity,

$$\text{supp}(G_k) = \frac{|\{j : \exists c \in C_j \cap G_k\}|}{N}, \quad \text{coh}(G_k) = \frac{2}{N_k(N_k - 1)} \sum_{\substack{c_u, c_v \in G_k \\ u < v}} s(c_u, c_v). \quad (2)$$

where  $N_k = |G_k|$  is the number of claims in group  $G_k$ . The consensus score is assigned as  $S(G_k) = \alpha \text{supp}(G_k) + (1 - \alpha) \text{coh}(G_k)$ . Groups with high consensus score are retained as high-confidence evidence, while low-support or low-cohesion groups are treated as low-consensus evidence and down-weighted.

**Deterministic temporal signal analytical pipeline.** We implement a deterministic, price-only analytical module that transforms historical financial data into four structured outputs: price history, price prediction, trading recommendations, and a summary table. Given a ticker symbol and trade date, the system retrieves daily price data over a maximum lookback window determined by predefined horizons ( $\mathcal{H} = \{1, 7, 14, 28, 90, 180, 360 \text{ days}\}$ ). The resulting series is standardized into a date-indexed tabular format.

**Price history.** For each horizon  $h$ , the module computes a compact set of descriptive statistics, including cumulative return and realized volatility, high/low/current price levels, average traded volume, and a discrete trend label. Trend is estimated by fitting a polynomial regression to stock prices; the fitted start-to-end change and terminal slope are combined into a single scalar trend score (capturing both direction and strength). We then discretize this score into one of five labels by thresholding its magnitude: large positive values indicate a strong uptrend, small positive values indicate an uptrend, values near zero indicate sideways movement, small negative values indicate a downtrend, and large negative values indicate a strong downtrend.

**Price prediction.** The prediction module extracts short-term momentum indicators (MA5/MA10/MA20 alignment, 5-day rate of change, and recent return volatility) and combines them with cross-horizon trend consistency (1-week and 1-month trends) to produce a deterministic bullish/bearish estimate. To ensure reproducible evaluation, we implement a small set of rule-based scoring functions. For example, if the current price is above MA5/MA10/MA20, we increase the bullish confidence by +2; if it is below all three moving averages, we increase the bearish confidence by +2. The final bullish and bearish scores are then used to output a next-day direction label (**up**, **down**, or **uncertain**), an expected movement magnitude, and the confidence score itself, together with a brief textual rationale. To reduce unstable

extrapolation under high-noise conditions, the predicted magnitude is bounded to a conservative interval.

**Trading recommendations.** We construct a confidence score in  $[0, 100]$  by combining four signals: trend consistency across horizons, next-day forecast confidence, recent return strength, and forecasted magnitude. The final action is selected from  $\{\text{BUY}, \text{SELL}, \text{HOLD}\}$ , and the position size is determined by pre-defined confidence thresholds together with the current portfolio state (position present versus absent). This rule-based layer yields transparent and reproducible recommendations with explicit rationales.

**Summary table.** For downstream decision interfaces, the module compiles a concise summary table containing the recommended action and position size, an aggregate confidence score, the next-day forecast (direction and magnitude), and key horizon returns. The final output concatenates the four components above, balancing interpretability (narrative detail) with operational usability (tabular summary).

**Memory bank with long/short-term decision reflection.** To support interruption-safe evaluation and test-time adaptation, TRUSTTRADE maintains a memory bank that logs each executed trade together with performance metrics computed over multiple horizons. This memory is later summarized by decision reflection agents into short- and long-term feedback that conditions subsequent decision-making.

**Update memory bank and derived metrics.** For each trading date  $t$ , we store a comprehensive trade record

$$\mathcal{R}_t = \left\{ a_t, p_t, q_t, P_t^{\text{entry}}, V_t, \{R_{t,h}, v_{t,h}^{(R)}\}_{h \in \mathcal{H}}, \{\text{SR}_{t,h}, v_{t,h}^{(\text{SR})}\}_{h \in \mathcal{H}} \right\}, \quad (3)$$

where  $a_t \in \{\text{BUY}, \text{HOLD}, \text{SELL}\}$ ,  $p_t$  is trade percentage,  $q_t$  is shares changed,  $P_t^{\text{entry}}$  is entry price, and  $V_t$  is pre-trade portfolio value.  $R_{t,h}$  and  $\text{SR}_{t,h}$  denote the return and Sharpe ratio at horizon  $h$ , and  $v_{t,h}^{(R)}$  and  $v_{t,h}^{(\text{SR})}$  denote rolling-window trend slopes (computed over a recent window of trades) for returns and Sharpe ratio, respectively.

To prevent information leakage, we initialize  $R_{t,h}$ ,  $\text{SR}_{t,h}$ , and their corresponding slope features as missing/empty at trade time  $t$ . These metrics are backfilled retrospectively only once the future date  $t+h$  is observed; at that point, realized return and Sharpe ratio at horizon  $h$  (and their rolling-window slopes) are computed and appended to the stored record. We retrospectively update realized return (PnL) for each prior trade  $i < t$ :

$$R_{t,h}(\%) = 100 \cdot \frac{\Delta \text{PnL}_{t,h}}{V_t}, \quad \Delta \text{PnL}_{t,h} = (P_{t+h}^{\text{entry}} - P_t^{\text{entry}}) \cdot d_t \cdot q_t, \quad (4)$$

where  $h \in \mathcal{H}$ ,  $d_t \in \{+1, -1\}$  denotes buy/sell direction. For  $h \geq 7$ , we also compute annualized Sharpe from daily portfolio-level returns within the horizon:

$$\text{SR}_{t,h} = \sqrt{252} \cdot \frac{\mu_{t,h}}{\sigma_{t,h}}, \quad \mu_{t,h} = \frac{1}{h} \sum_{\tau=1}^h R_{t-\tau,1}, \quad \sigma_{t,h} = \sqrt{\frac{1}{h} \sum_{\tau=1}^h (R_{t-\tau,1} - \mu_{t,h})^2}. \quad (5)$$

To capture trend-of-performance (not single-point performance), we compute rolling-window ( $w$ ) linear slopes over recent trades for each horizon:

$$v_{t,h}^{(R)} = \text{slope}(\{R_{t-\tau,h}\}_{\tau < w}), \quad v_{t,h}^{(\text{SR})} = \text{slope}(\{\text{SR}_{t-\tau,h}\}_{\tau < w}). \quad (6)$$

Positive slopes indicate improving recent performance, whereas negative slopes suggest deterioration. We use these trend signals to judge whether the strategy is getting better or worse over time (for both returns and risk-adjusted returns).

**Short-term vs. long-term decision reflection.** TRUSTTRADE generates reflections through an agentic function call, which converts stored trading history into an LLM-readable prompt. The memory bank first converts trade records  $\{\mathcal{R}_t\}_t$  into a tabular dataframe that summarizes, for each trade date, the action/position and multi-horizon performance statistics (e.g.,  $r_{i,h}$  and  $SR_{i,h}$ ) together with their recent trend slopes (e.g.,  $s_h^{(r)}$  and  $s_h^{(S)}$ ). Conditioned on a selected horizon configuration, the function assembles a compact prompt containing a dataset overview, per-horizon performance/trend summaries, and explicit reflection instructions.

At each trading date  $t$ , we instantiate two specialized reflection agents that share this functional interface but differ only in horizon configuration: a short-term agent ( $\mathcal{H}_{\text{short}} = \{1, 7, 14\}$ ) and a long-term agent ( $\mathcal{H}_{\text{long}} = \{28, 90, 180, 360\}$ ). Formally, letting  $\Pi(\cdot)$  denote the generated prompt, we obtain

$$F_t^{\text{short}} = \text{LLM}_{\text{short}}(\{\mathcal{R}_t\}_t, \Pi(\mathcal{H}_{\text{short}})), \quad F_t^{\text{long}} = \text{LLM}_{\text{long}}(\{\mathcal{R}_t\}_t, \Pi(\mathcal{H}_{\text{long}})). \quad (7)$$

The resulting reflections are stored back into the memory bank  $\mathcal{R}_t$  and injected into the next decision cycle as historical performance context (used for confidence/risk/position sizing calibration), enabling test-time adaptation. Role-specific reflection memory (Bull/Bear/Trader) is updated separately using realized outcomes and stored for retrieval-augmented reasoning.

## Experimental Details

**Rule-based baselines.** We compare our proposed TRUSTTRADE against several baselines:

- **Buy and Hold:** Investing equal amounts in all selected stocks and holding them throughout the simulation period<sup>17</sup>.
- **MACD (Moving Average Convergence Divergence):** A trend-following momentum strategy that generates buy and sell signals based on the crossover points between the MACD line and signal line<sup>7</sup>.
- **KDJ and RSI (Relative Strength Index):** A momentum strategy combining KDJ (stochastic oscillator) and RSI (relative strength index) indicators to identify overbought and oversold conditions for trading signals<sup>40</sup>.
- **ZMR (Zero Mean Reversion):** A mean reversion trading strategy that generates signals based on price deviations from and subsequent reversions to a zero reference line<sup>26</sup>.
- **SMA (Simple Moving Average):** A trend-following strategy that generates trading signals based on crossovers between short-term and long-term moving averages<sup>41</sup>.

**Performance metrics and human-aligned evaluation. Evaluation metrics.** We evaluate TRUSTTRADE using standard portfolio metrics that capture profitability, risk, and safety relative to baselines. For clarity, we use a unified notation throughout:  $T$  is the total number of trading days;  $V_t$  is the portfolio (or asset) value at time  $t$  (with

$V_0$  the initial value);  $R_t := \frac{V_t}{V_{t-1}} - 1$  is the simple return at time  $t$  and  $\bar{R} := \frac{1}{T} \sum_{t=1}^T R_t$  is the average return. We denote the per-period risk-free rate by  $R_f$ , the benchmark return by  $b_t$ , the active (excess) return by  $a_t := R_t - b_t$ , and the per-period loss by  $L_t := -R_t$ . We report cumulative return,  $\text{CR} = \left( \frac{V_T - V_0}{V_0} \right) \times 100\%$ , and annualized return,  $\text{AR} = \left( \left( \frac{V_T}{V_0} \right)^{\frac{k}{T}} - 1 \right) \times 100\%$ , where  $k$  is the number of trading days per year (e.g.,  $k = 252$ ). We also report the Sharpe ratio,  $\text{SR} = \frac{\bar{R} - R_f}{\sigma}$ , with volatility defined as  $\sigma = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (R_t - \bar{R})^2}$ ; if reporting annualized Sharpe, we use  $\bar{R}_{\text{ann}} = \bar{R} \cdot k$  and  $\sigma_{\text{ann}} = \sigma \sqrt{k}$ . Finally, we measure maximum drawdown via the running peak  $M_t = \max_{0 \leq s \leq t} V_s$ , drawdown  $DD_t = \frac{V_t - M_t}{M_t} \leq 0$ , and  $\text{MDD} = \max_{0 \leq t \leq T} \left( \frac{M_t - V_t}{M_t} \right) = -\min_{0 \leq t \leq T} DD_t$ .

**Human-aligned preference region.** To assess alignment with human annotator behavior, we define a *human-aligned preference region* centered at the empirical human performance. Let  $(\mu_{\text{MDD}}^{(h)}, \mu_{\text{CR}}^{(h)})$  denote the mean MDD and CR of human annotators under partial allocation, and let  $(\sigma_{\text{MDD}}^{(h)}, \sigma_{\text{CR}}^{(h)})$  denote the corresponding standard errors across stocks. We define the preference region as an elliptical level set:

$$\mathcal{E} = \left\{ (x, y) \mid \left( \frac{x - \mu_{\text{MDD}}^{(h)}}{\sigma_{\text{MDD}}^{(h)}} \right)^2 + \left( \frac{y - \mu_{\text{CR}}^{(h)}}{\sigma_{\text{CR}}^{(h)}} \right)^2 \leq c \right\}, \quad (8)$$

where  $c > 0$  controls the size of the preference region. This formulation characterizes the statistical dispersion of human risk-return outcomes rather than imposing hard thresholds. Agents closer to the center of  $\mathcal{E}$  exhibit stronger alignment with typical human annotator behavior, while increasing distance indicates greater deviation in risk exposure, return realization, or both. For visualization,  $\mathcal{E}$  is rendered with graded shading to reflect degrees of alignment and serves as a soft behavioral reference rather than a strict decision boundary.

## Data availability

This study involved human annotators providing trading decisions in an experimental setting. All participants provided informed consent prior to participation. The links below correspond to the web interfaces used to collect human annotations during the trading simulation process: <https://tradingagents-human-aapl.netlify.app/>, <https://tradingagents-human-goog.netlify.app/>, <https://tradingagents-human-nvda.netlify.app/>. For the simulation itself, all multi-source information presented to participants (e.g., fundamentals, market signals, news summaries, and sentiment cues) was generated from language-model-produced reports. No human-authored analysis, recommendation, or manual curation was included in the decision-facing information stream.

## Code availability

The code used for data processing, model implementation, and analysis in this study is available at <https://github.com/Harvard-AI-and-Robotics-Lab>.

## References

- [1] Hurst, A., Lerer, A., Goucher, A.P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A., Welihinda, A., Hayes, A., Radford, A., et al.: Gpt-4o system card. arXiv preprint arXiv:2410.21276 (2024)
- [2] Team, G., Anil, R., Borgeaud, S., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A.M., Hauth, A., Millican, K., et al.: Gemini: a family of highly capable multimodal models. arXiv preprint arXiv:2312.11805 (2023)
- [3] Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., et al.: Deepseek-v3 technical report. arXiv preprint arXiv:2412.19437 (2024)
- [4] Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., et al.: A survey on large language model based autonomous agents. *Frontiers of Computer Science* **18**(6), 186345 (2024)
- [5] Ding, Y., Li, J., Wang, X., Chen, H.: Large language model agent in financial trading: A survey. arXiv preprint arXiv:2408.06361 (2024)
- [6] Ingersoll, J.E.: *Theory of Financial Decision Making* vol. 3. Bloomsbury Publishing PLC, ??? (1987)
- [7] Lopez-Lira, A., et al.: Can large language models trade? testing financial theories with llm agents in market simulations. arXiv preprint arXiv:2504.10789 (2025)
- [8] Bai, Y., Gao, S., Wan, J., Zhang, L., Song, H.: A review of reinforcement learning in financial applications. arXiv preprint arXiv:2411.12746 (2024)
- [9] Anonymous: LLM-Based Routing in Mixture of Experts: A Novel Framework for Trading. arXiv:2501.09636 (2025)
- [10] Anonymous: A Deep Reinforcement Learning Framework for Dynamic Portfolio Optimization: Evidence from China’s Stock Market. arXiv:2412.18563 (2024)
- [11] Anonymous: Decision-informed Neural Networks with Large Language Model Integration for Portfolio Optimization. arXiv:2502.00828 (2025)
- [12] Anonymous: A Multimodal Foundation Agent for Financial Trading: Tool-Augmented, Diversified, and Generalist. arXiv:2402.18485 (2024)
- [13] Zhang, C., Liu, X., Zhang, Z., et al.: When ai meets finance (stockagent): Large

- language model-based stock trading in simulated real-world environments. arXiv preprint arXiv:2407.18957 (2024)
- [14] Cao, L., Xi, Z., Liao, L., Yang, Z., Cao, Z.: Chain-of-alpha: Unleashing the power of large language models for alpha mining in quantitative trading. arXiv preprint arXiv:2508.06312 (2025)
  - [15] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D., *et al.*: Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems* **35**, 24824–24837 (2022)
  - [16] Tran, K.-T., Dao, D., Nguyen, M.-D., Pham, Q.-V., O’Sullivan, B., Nguyen, H.D.: Multi-agent collaboration mechanisms: A survey of llms. arXiv preprint arXiv:2501.06322 (2025)
  - [17] Sutton, R.S., Barto, A.G., *et al.*: Reinforcement Learning: An Introduction vol. 1. MIT press Cambridge, ??? (1998)
  - [18] Coriat, T., Benhamou, E.: Harlf: Hierarchical reinforcement learning and lightweight llm-driven sentiment integration for financial portfolio optimization. arXiv preprint arXiv:2507.18560 (2025)
  - [19] Unnikrishnan, A.: Financial news-driven llm reinforcement learning for portfolio management. arXiv preprint arXiv:2411.11059 (2024)
  - [20] Xiao, Y., Sun, E., Luo, D., Wang, W.: Tradingagents: Multi-agents llm financial trading framework. arXiv preprint arXiv:2412.20138 (2025)
  - [21] Li, Y., Yang, X., Yang, X., Xu, M., Wang, X., Liu, W., Bian, J.: R&d-agent-quant: a multi-agent framework for data-centric factors and model joint optimization. arXiv preprint arXiv:2505.15155 (2025)
  - [22] Mandelbrot, B., *et al.*: The variation of certain speculative prices. *Journal of business* **36**(4), 394 (1963)
  - [23] Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007 (1982)
  - [24] Diebold, F.X., Yilmaz, K.: On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of econometrics* **182**(1), 119–134 (2014)
  - [25] Brock, W.A., Hommes, C.H.: Rational expectations and rational learning in a simple asset pricing model. *Journal of Economic Dynamics and Control* **21**(8-9), 1115–1146 (1997)
  - [26] Brock, W.A., Hommes, C.H.: Heterogeneous beliefs and routes to chaos in a

- simple asset pricing model. *Journal of Economic dynamics and Control* **22**(8-9), 1235–1274 (1998)
- [27] Yun, J., Lee, H.J., Jeon, I.: Quantevolve: Automating quantitative strategy discovery through multi-agent evolutionary framework. *arXiv preprint arXiv:2510.18569* (2025)
- [28] Han, J., Zhang, S., Li, W., Yang, Z., Dong, Y., Hu, T., Yuan, J., Yu, X., Zhu, Y., Lou, F., et al.: Quantaalpha: An evolutionary framework for llm-driven alpha mining. *arXiv preprint arXiv:2602.07085* (2026)
- [29] Guo, T., Shen, H., Luo, J., Xing, Z., Lian, H., Huang, J., Chen, B., Liu, L., Ma, Y., Zhang, M.: Meme: Modeling the evolutionary modes of financial markets. *arXiv preprint arXiv:2602.11918* (2026)
- [30] Benhenda, E.: Finrl-deepseek: Llm-infused risk-sensitive reinforcement learning for trading agents. *arXiv preprint arXiv:2502.07393* (2025)
- [31] Lee, H., Seo, J., Park, S., Lee, J., Ahn, W., Choi, C., Lopez-Lira, A., Lee, Y.: Your ai, not your view: The bias of llms in investment analysis. *arXiv preprint arXiv:2507.20957* (2025)
- [32] Lin, X., Ning, Y., Zhang, J., Dong, Y., Liu, Y., Wu, Y., Qi, X., Sun, N., Shang, Y., Wang, K., et al.: Llm-based agents suffer from hallucinations: A survey of taxonomy, methods, and directions. *arXiv preprint arXiv:2509.18970* (2025)
- [33] Swanson, K., Wu, W., Nash, L.B., Pak, J.E., Zou, J.: The virtual lab: Ai agents design new sars-cov-2 nanobodies with experimental validation.  *biorxiv* (2024)
- [34] Piatti, G., Jin, Z., Kleiman-Weiner, M., Schölkopf, B., Sachan, M., Mihalcea, R.: Cooperate or collapse: Emergence of sustainability behaviors in a society of llm agents.  *CoRR* (2024)
- [35] Preis, T., Moat, H.S., Stanley, H.E.: Quantifying trading behavior in financial markets using google trends.  *Scientific reports* **3**(1), 1684 (2013)
- [36] Coval, J.D., Shumway, T.: Information and noise in financial markets: An experimental study.  *Journal of Finance* **56**, 1141–1179 (2001)
- [37] Tuckett, D., Taffler, R.: Emotions and financial markets.  *Handbook of Behavioural Economics and Smart Decision-Making* (2011)
- [38] Kahneman, D.:  *Thinking, Fast and Slow* vol. 0. Farrar, Straus and Giroux, ??? (2011)
- [39] Froot, K., Scharfstein, D.S., Stein, J.C.: Explaining investor beliefs.  *Journal of Finance* **47**, (1992)

- [40] Wilder, J.W.: *New Concepts in Technical Trading Systems*. Trend Research, Greensboro, NC (1978)
- [41] Brock, W., Lakonishok, J., LeBaron, B.: Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance* **47**(5), 1731–1764 (1992)