FINTRIAD: A MULTIMODAL, MULTI-AGENT SYSTEM WITH ADAPTIVE MEMORY FOR TRI-HORIZON PORTFOLIO CONSTRUCTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Professional investing requires integrating multimodal information across heterogeneous decision horizons, while experienced traders accumulate market experience that informs subsequent decisions. Although recent work has increasingly sought to replicate this decision-making process with LLM-based agents, existing systems remain limited in multimodal reasoning, portfolio-level decision-making, and, critically, the continual accumulation of experience from realized outcomes. We introduce **FinTriad**, a multimodal multi-agent framework with adaptive memory for concurrent tri-horizon portfolio construction. FinTriad integrates broadcast financial media, market-wide candidate selection, specialist analysis, and portfolio management across daily, weekly, and monthly horizons. Its verifiable-reward reflexive learning incorporates realized outcomes into temporally layered memory under a verifiable reward, admitting a correction only when the re-decision that produced it is confirmed against the realized market outcome. Across an 18-month evaluation window spanning multiple market regimes, FinTriad demonstrates competitive portfolio performance and robust adaptation to heterogeneous and evolving market conditions.

1 INTRODUCTION

Market information arrives as a continuous, multimodal stream (Soleymani et al., 2017): macroeconomic releases and central bank communications shape the broader market environment; corporate filings reveal firm-level fundamentals; news broadcasts provide factual content alongside vocal and visual cues; and price and volume series record the aggregate response of market participants. Translating these signals into financial decisions is therefore a sequential task that unfolds across multiple horizons. Financial institutions address this complexity through a division of labor, with market analysts, technical analysts, fundamental researchers, and portfolio managers processing different information sources and time horizons before synthesizing their views into final allocations (Radner, 1993; Sinkula, 1994). Automated trading systems therefore face a challenge beyond prediction: integrating heterogeneous information and coordinating decisions across multiple investment horizons.

Research on automated trading has progressed from rule-based and statistical methods to deep reinforcement learning (DRL), which enables agents to adapt decisions to evolving market conditions (Jiang et al., 2017; Li et al., 2023b; Sun et al., 2021). However, these approaches primarily operate on structured market observations and remain limited in integrating heterogeneous textual, visual, and numerical evidence across different time scales (Pippas et al., 2025). Large Language Models (LLMs) and their multimodal extensions (MLLMs) (OpenAI, 2023b;a) offer a complementary paradigm for financial decision-making, with demonstrated capabilities in reasoning over financial narratives and integrating qualitative and quantitative evidence (Wu et al., 2023; Xie et al., 2023; Lopez-Lira & Tang, 2025; Wang et al., 2024b; Cao et al., 2025). Building on these capabilities, LLM-based trading agents have emerged, initially focusing on single-asset trading (Yang et al., 2023; Yu et al., 2023; Zhang et al., 2024) and more recently extending to portfolio investment (Yu et al., 2024; Zhao et al., 2025; Luo et al., 2025; Paquette-Greenbaum & Yu, 2026; Caner et al., 2026). An extended discussion of related work is deferred to Appendix A.

These developments point to a broader challenge: how to construct portfolios from heterogeneous, evolving information while coordinating decisions across multiple investment horizons and con-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

tinuously adapting to new outcomes. Existing LLM-based financial agents leave four gaps: (1) **Multimodal reasoning.** Existing frameworks rely primarily on textual inputs, sometimes supplemented by static visual artifacts, whereas financial media can convey information through vocal and visual cues that are absent from transcripts, such as tone, hesitation, and gestures (Tan et al., 2025). (2) **Multi-horizon planning.** The same market information can affect prices over heterogeneous time horizons (Hong & Stein, 1999; Chan, 2003), while market narratives often evolve through incremental updates; existing agents, however, typically operate at a single decision cadence. (3) **Open universe selection.** Existing LLM-based systems generally operate within narrow, ex-ante selected asset lists. In contrast, professional portfolio construction also requires identifying which assets merit consideration. (4) **Verifiable-reward learning.** Prior reflection-based trading agents typically use discrete train/test partitions and hold learned state fixed during evaluation (Yu et al., 2023; 2024). Moreover, reflection converts observed failures directly into lessons that the agent itself grades (Shinn et al., 2023), without checking whether the resulting correction would have improved the decision it is meant to fix (Allard et al., 2026).

Motivated by these gaps, we present **FinTriad** (Figure 1), a multimodal, multi-agent architecture with adaptive memory for tri-horizon portfolio construction. FinTriad addresses these limitations through four core components. **(1) Direct multimodal financial media ingestion:** FinTriad processes Bloomberg Brief videos directly through a multimodal LLM, jointly extracting textual, vocal and visual cues rather than relying on transcripts alone. **(2) Tri-horizon financial decision-making:** driven by a unified point-in-time data stream, FinTriad makes portfolio decisions across daily, weekly and monthly horizons. **(3) Open universe selection:** rather than relying on a predefined shortlist, FinTriad autonomously generates investment candidates by combining the LLM's parametric knowledge with available point-in-time market information. **(4) Verifiable-Reward Reflexive Learning (VRRL):** rather than writing every reflection into memory, FinTriad grades each diagnosed failure with a reward computed from realized prices (Lambert et al., 2024; Guo et al., 2025). Experience retained in memory therefore reflects market outcomes rather than the agent's own assessment of its reflection. The graded re-decision is never executed and never alters reported returns; it determines only what is written to memory, which conditions decisions from the following day onward.

2 PRELIMINARIES

Trading as a partially observed decision problem. Investment decisions are sequential and risk-controlled, taken under partial observability with non-stationary dynamics and heterogeneous-horizon rewards (Lo, 2017): the canonical setting for a partially observable Markov decision process (POMDP) (Kaelbling et al., 1998). The market state is never directly observable (Grossman & Stiglitz, 1980); the transition law is non-stationary; and we assume sufficient liquidity that FinTriad acts as a price-taker. On trading day t, FinTriad observes the heterogeneous tuple $o_t = (n_t, p_t, f_t)$, multimodal news content, price history truncated at $t-1$, and the most recent fundamental panel, from which it emits an action $a_t = (a_t^{1\mathrm{d}}, a_t^{7\mathrm{d}}, a_t^{\mathrm{val}})$ across the horizons $\mathcal{H} = \{1\mathrm{d}, 7\mathrm{d}, \mathrm{val}\}$: a 1-day horizon (opened at the open, closed at the same session's close), a 7-day horizon, and a month-long value horizon. Each a_t^h is a weighted portfolio over a market-wide equity universe under per-position and gross-exposure constraints (Grinold & Kahn, 2000), realizing a mark-to-market return r_t^h; we write $r_t = (r_t^h)_{h\in\mathcal{H}}$. Evaluation runs over a finite window $[t_0, t_0 + W - 1]$ of W trading days. **Memory-conditioned decisions.** FinTriad trains no policy. A fixed multimodal LLM backbone maps observations, specialist-agent outputs, and retrieved experience to decisions, $a_t = \pi_{\theta_0}(o_t, \mathcal{B}_t, m_t)$, where θ_0 denotes the fixed LLM and prompt configuration, $\mathcal{B}_t$ the live book carried into day t, and m_t the structured experience distilled from prior decision outcomes. Adaptation is relocated from the parameters to the memory, which evolves under a reflexive update rule $m_{t+1} = \mathcal{U}(m_t, a_t, r_t)$, closing the loop: actions realize returns, returns drive memory updates, and updated memory conditions subsequent actions. **From verifiable rewards to VRRL.** We design $\mathcal{U}$ by analogy with reinforcement learning from verifiable rewards (RLVR; Lambert et al., 2024; Guo et al., 2025), where outcome-based feedback provides the reward signal. In our setting, this idea is adapted to three elements. *The verifier* is the market: the reward is computed from realized prices rather than from a checker supplied by the designer. *The update* is a write to m_t rather than a gradient step on θ. *Credit assignment* is outcome-based: a re-decision is credited only if it recovers, against the same realized prices, the shortfall that triggered it. The analogy is not exact, since an RLVR verifier is a deterministic checker supplied by the designer whereas here the market prices the outcome. What bounds what a re-decision can exploit is the channel: it is conditioned on a revised reading of the day's events, not on the

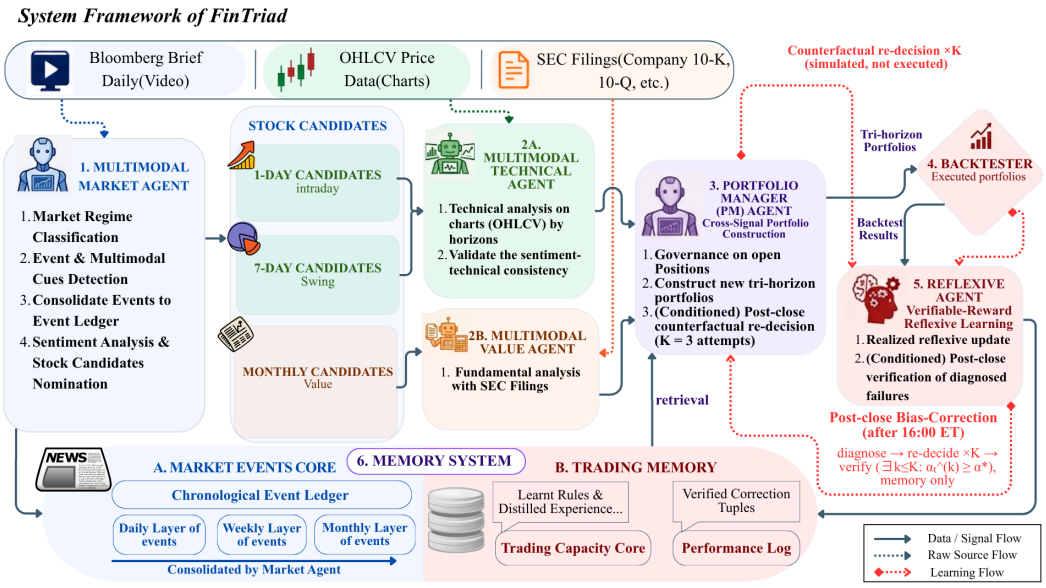


Figure 1: Inputs (left) flow through specialist agents to the Portfolio Manager (center); the Reflexive Agent distills backtested outcomes into memory (right). The red dotted post-close loop is Verifiable-Reward Reflexive Learning, which writes a correction to memory only when the re-decision that produced it is verified against the realized outcome; reported returns come solely from executed portfolios.

executed portfolio or its returns (Appendix B.6). We refer to this mechanism as *Verifiable-Reward Reflexive Learning* (VRRL).

3 Architecture

3.1 System Overview

Figure 1 illustrates the FinTriad pipeline, which operates on news-available trading days with horizon-specific decision cycles: a daily cycle for the 1-day horizon, an additional 7-day cycle on the last news-available trading day of each ISO week, and an additional value cycle on the first news-available trading day of each calendar month, each constructing the basket held over the following period. The pipeline comprises six components. (1) The **Multimodal Market Agent (MMA)** analyzes multimodal market news and constructs a temporally layered event ledger with horizon-specific stock candidates. (2A) The **Multimodal Technical Agent (MTA)** analyzes per-ticker price charts for the 1-day and 7-day horizons, while (2B) the **Multimodal Value Agent (MVA)** analyzes SEC filings for the value horizon. (3) The **Portfolio Manager (PM)** synthesizes specialist outputs with retrieved trading memory to produce the tri-horizon portfolio $a_t = (a_t^{1\mathrm{d}}, a_t^{7\mathrm{d}}, a_t^{\mathrm{val}})$, thereby instantiating the fixed decision rule π_{θ_0}. (4) The **Backtester** evaluates each portfolio against realized prices to obtain horizon-specific returns r_t^h. (5) The **Reflexive Agent** instantiates the update rule $\mathcal{U}$, performing end-of-period memory updates across horizons and an additional post-close diagnostic pass for the 1-day horizon. Because 1-day positions are fully closed at the session close, their realized outcome is available the same evening, so a diagnosis can be verified against that outcome and written to memory before the next session opens. This pass never alters executed trades or reported returns. (Appendix B.7). (6) The structured **Memory System** stores the resulting experience as the Market Events Core E_t and Trading Memory (C_t, D_t), which are retrieved by the PM at decision time. Together, these components form a closed loop in which multimodal market information is transformed into horizon-specific portfolios, evaluated against realized outcomes, and incorporated into memory to adaptively condition subsequent decisions.

3.2 MULTIMODAL MARKET AGENT

The MMA treats Bloomberg Brief videos, available on business days, as the multimodal news observation n_t and performs four operations before market open: it classifies the prevailing market regime; extracts market events and salient multimodal cues, linking each to prior ledger entries through two-stage semantic matching or creating a new entry when no match is found; organizes them into the daily, weekly and monthly ledger layers; and analyzes sentiment over the appropriate layer to nominate candidates on each horizon's cycle, routing the 1-day and 7-day baskets to the Technical Agent and the value basket to the Value Agent. The prompt template and matching procedure are detailed in Appendix B.1.

Our Multimodal Market Agent reflects two deliberate design choices. Direct video ingestion preserves cross-modal signals that transcription-and-classification pipelines may discard (Tan et al., 2025), while the chronologically layered ledger captures the gradual diffusion of news about a single development across trading periods (Hong & Stein, 1999; Chan, 2003). This prevents sequential updates about the same development from being treated as independent events and helps maintain a consistent estimate of its net market impact.

3.3 MULTIMODAL TECHNICAL & VALUE AGENTS

Multimodal Technical Agent. The MTA performs technical analysis on chart-based price signals to assess the sentiment and technical consistency of each ticker proposed by the MMA for the 1-day and 7-day horizons. For each ticker, it constructs horizon-calibrated charts from OHLCV data truncated at one trading day before t, using daily and 7-day candlesticks with predefined technical indicators (Appendix B.2) based only on point-in-time information. It returns a scored verdict that filters out rejected tickers and passes the rest to the PM (Section 3.4) with key signals and entry, stop-loss and take-profit levels.

Multimodal Value Agent. The MVA performs fundamental analysis of the monthly basket by grounding its assessment in regulatory filings and verified financial data (Graham, 1962; Buffett, 1984; Tan et al., 2026). For each ticker, it identifies relevant SEC EDGAR filings, including 10-K and 10-Q reports, and analyzes quantitative financial metrics and qualitative disclosures, such as management discussion, risk factors, and legal proceedings (Li, 2008; Loughran & Mcdonald, 2011; 2016; Tsai et al., 2025). To capture changes over time, it jointly analyzes the current and prior-period filings, allowing shifts in language and reported fundamentals to inform the assessment rather than treating each filing in isolation (Cohen et al., 2020). Industry-specific metrics are dynamically selected to reflect sector-relevant fundamentals, and reported figures are cross-validated against an external source using only information available as of the filing date. The MVA emits a directional verdict with a confidence score, price target and thesis; prompts, output fields and the exposure-dependent handling of *HOLD* are given in Appendix B.3.

3.4 ADAPTIVE MEMORY, PORTFOLIO MANAGER, AND VERIFIABLE-REWARD REFLEXIVE LEARNING

Adaptive Memory. FinTriad maintains a structured memory $m_t = (E_t, C_t, D_t)$ comprising *Market Events Core* E_t, *Trading Capacity Core* C_t, and *Performance Log* D_t. The Market Events Core E_t organizes evolving market developments across daily, weekly, and monthly layers. The Trading Capacity Core C_t stores reusable trading knowledge accumulated from realized outcomes, including event-to-market mappings, sector-specific playbooks, and horizon-specific feedback. The Performance Log D_t records previously diagnosed failures together with the alternative decisions that passed verification against the realized outcome. Together, these components form an evolving case-based prior that conditions future decisions.

Portfolio Manager. Given the current memory m_t, PM applies the static prompt, which defines the fixed decision rule π_{θ_0}, synthesizing upstream specialist outputs, the live book $\mathcal{B}_t$, and date-gated memory retrieval. For each horizon, it first reviews inherited positions against current analysis and assigns a governance action of *HOLD* or *EXIT*. It then constructs candidate portfolios from the specialist signals and reconciles them with retained positions through a rebalancer, producing the tri-horizon action $a_t = (a_t^{1\mathrm{d}}, a_t^{7\mathrm{d}}, a_t^{\mathrm{val}})$. Each position carries a backtest-ready specification including *entry timing*, *weight*, *stop-loss*, *take-profit*, and a one-sentence *thesis*, together with self-reported

Algorithm 1 FinTriad Daily Control Loop with Verifiable-Reward Reflexive Learning

Require: trading day t, threshold α^*, attempt budget K, $\mathcal{B}_t$, $o_t = (n_t, p_t, f_t)$, $m_t = (E_t, C_t, D_t)$, fixed π_{θ_0}, and $\mathcal{U} = (\mathcal{U}_{\text{diag}}, \mathcal{U}_{\text{update}})$
Ensure: $\mathcal{B}_{t+1}$, $m_{t+1} = (E_{t+1}, C_{t+1}, D_{t+1})$

Trading session (executed; sole source of reported returns)

Portfolio construction: $a_t^{(0)} \leftarrow \pi_{\theta_0}(o_t, \mathcal{B}_t, m_t) = (a_t^{1\text{d},(0)}, a_t^{7\text{d},(0)}, a_t^{\text{val},(0)})$
Outcome evaluation: $r_t^{h,(0)} \leftarrow \textsc{Backtester}(a_t^{h,(0)})\ \forall h \in \mathcal{H}$
$\mathcal{B}_{t+1} \leftarrow \textsc{UpdateLiveBook}(\mathcal{B}_t, a_t^{(0)}, r_t^{(0)})$
▷ $r_t^{(0)}$ is the only return recorded for evaluation

After market close (memory only; no trade executed or modified)

Post-close diagnosis and verification: $\Delta_t, k^\star, v_t \leftarrow \varnothing, \varnothing, 0$
if $\alpha_t^{1\text{d},(0)} < \alpha^*$ **then** ▷ same-day realized alpha
  $\Delta_t \leftarrow \mathcal{U}_{\text{diag}}(a_t^{1\text{d},(0)}, r_t^{1\text{d},(0)}, o_t)$
  if Δ_t contains a learnable failure **then**
    for $k = 1$ **to** K **do** ▷ $\delta_t \subset \Delta_t$ held fixed across attempts
      $a_t^{1\text{d},(k)} \leftarrow \pi_{\theta_0}(o_t, \mathcal{B}_t, m_t; \delta_t)$; $r_t^{1\text{d},(k)} \leftarrow \textsc{CounterfactualSim}(a_t^{1\text{d},(k)})$ ▷ never executed; not reported
      if $\alpha_t^{1\text{d},(k)} \geq \alpha^*$ **then**
        $v_t \leftarrow 1$; $k^\star \leftarrow k$; **break** ▷ verifiable reward: priced by the market
      end if
    end for
    if $v_t = 0$ **then**
      $k^\star \leftarrow \arg\max_{1 \leq k \leq K}\ \alpha_t^{1\text{d},(k)}$ ▷ best of the K attempts; the executed k=0 is excluded
    end if
  end if
end if
Reflexive update: $(E_{t+1}, C_{t+1}, D_{t+1}) \leftarrow \mathcal{U}_{\text{update}}\big(E_t, C_t, D_t, a_t^{(0)}, r_t^{(0)}, [a_t^{1\text{d},(k^\star)}, r_t^{1\text{d},(k^\star)}, \Delta_t, v_t]\big)$
▷ Closed-position outcomes update C_{t+1}; only verified corrections ($v_t = 1$) update D_{t+1}
▷ If $v_t = 0$, the best attempt still informs C_{t+1}, but no correction enters D_{t+1}
▷ m_{t+1} conditions decisions from day $t+1$ onward
return $\mathcal{B}_{t+1}$, m_{t+1}

conviction and weakest-link fields used for subsequent outcome attribution. Detailed retrieval, governance, and rebalancing procedures are provided in Appendix B.5.

Verifiable-Reward Reflexive Learning (VRRL). The Reflexive Agent realizes the update rule $\mathcal{U} = (\mathcal{U}_{\text{diag}}, \mathcal{U}_{\text{update}})$, which runs after the market close and modifies only memory, never executed trades. For completed positions on all three horizons, $\mathcal{U}_{\text{update}}$ distills realized action-return pairs into event-to-market mappings, sector-specific playbooks, event-type statistics, and horizon-specific feedback in C_{t+1}. The verifiable reward governs the corrective channel D_{t+1} on the 1-day horizon, the only horizon whose positions close within the session. Triggering is therefore 1-day only, but retrieval is not: D_t is read by the PM at all three horizons, so a verified correction conditions subsequent daily, weekly and value decisions alike. Throughout, alpha is measured against the S&P 500 (SPY) benchmark, $\alpha_t^{1\text{d},(k)} = r_t^{1\text{d},(k)} - r_t^{\text{SPY}}$, the 1-day sleeve's own same-day excess return, with $k = 0$ denoting the executed decision. If the executed alpha $\alpha_t^{1\text{d},(0)}$ falls below α^* (0.3 pp), $\mathcal{U}_{\text{diag}}$ diagnoses structured failure types from the executed decision and its outcome and re-analyses the day's event stream into a revised reading δ_t, under which the PM is re-invoked for up to $K = 3$ attempts $a_t^{1\text{d},(k)}$, none of which is executed. Only δ_t reaches the PM: it carries no record of the executed positions, no per-position returns and no direction, so the re-decision works from a revised account of the same pre-open evidence rather than from a report of what the executed portfolio got wrong. Each attempt is evaluated against realized prices, and the loop stops at the first whose alpha clears the threshold that triggered it, $v_t = \mathbf{1}\Big[\exists\, 1 \leq k \leq K :\ \alpha_t^{1\text{d},(k)} \geq \alpha^*\Big]$, so that a single constant serves as both trigger and success criterion. When $v_t = 1$ the successful re-decision enters D_{t+1} as a verified correction; otherwise the diagnosis is not credited, and only the best attempt is distilled into C_{t+1}, informing the situational prior without entering the corrective channel (Appendix B.6). All reported performance uses the executed returns $r_t^{(0)}$, and the updated memory m_{t+1} conditions decisions only from day $t+1$ onward. Algorithm 1 summarizes the procedure.

4 EXPERIMENTS

We conduct experiments to address the following Research Questions (RQs): **RQ1:** Does FinTriad achieve competitive portfolio performance against strategy and passive baselines? **RQ2:** Do the gains arise from open-universe asset selection, from portfolio construction, or from both? **RQ3:** Does multimodal financial-media information improve portfolio performance compared with audio-only or transcript-only input? **RQ4:** Does grading reflection with a verifiable reward improve the quality and performance of constructed portfolios?

4.1 EXPERIMENTAL SETUP

Fusion-Integrated Multimodal Dataset (FIMD). We introduce *FIMD*, which integrates heterogeneous sources into a consistent representation reflecting the information used by professional traders. **Bloomberg Brief videos**, published between roughly 07:20 and 08:10 ET on each trading day before the U.S. market open, provide daily market summaries together with vocal and visual cues such as speaker tone, pace and facial expressions; **1-minute bar data** from Massive (formerly Polygon.io) cover the stock candidate universe, supporting chart-based analysis and backtesting; and **SEC EDGAR filings** (10-K and 10-Q) provide standard and sector-specific fundamentals. All sources are gated to a point-in-time cut; Appendix B.7 gives the intraday timeline and the execution rule for each horizon.

Backbone Model & Evaluation Metrics. **Backbone model:** all FinTriad agents use Gemini 3 Flash with the default deployed decoding temperature of 1.0, a knowledge cutoff of January 2025 and web search disabled. **Warm-up and evaluation:** from the earliest available Bloomberg Brief video date, April 2, 2024, to January 31, 2025 (210 trading days), the system runs a *warm-up* period to initialize and populate the memory $m_t = (E_t, C_t, D_t)$; evaluation runs from February 3, 2025 to July 31, 2026 (375 trading days), after both the model's knowledge cutoff and the warm-up period, ensuring that evaluation-period information does not enter the initial memory. Following Yu et al. (2024), we run the full pipeline five times and report the median-CR trajectory. **Evaluation metrics:** Cumulative Return (CR) and Compound Annual Growth Rate (CAGR) for returns, Annualized Volatility (Vol) and Maximum Drawdown (MDD) for risk, and Sharpe Ratio (SR) for risk-adjusted performance. Further details on metric definitions and settings are given in Appendix C.

Baseline Models. For comprehensive evaluation, we aggregate FinTriad's tri-horizon sub-portfolios by equally weighting the horizons holding positions on each session (Eq. (6); Appendix C.1). Because the active set is known at market open, this aggregation is implementable ex ante. We then benchmark the resulting portfolio against six baselines: FinRL's Advantage Actor-Critic (A2C) and Proximal Policy Optimization (PPO) (Liu et al., 2022), the Minimum-Variance Portfolio (MVP) (Markowitz, 1952; 1959), the 12-1 cross-sectional momentum strategy (Jegadeesh & Titman, 1993; Asness et al., 2013), the Invesco S&P 500 Equal Weight ETF (RSP), and FinMem (Yu et al., 2023). Because FinTriad selects from an open universe while the non-passive baselines require a predefined basket, all non-passive baselines are evaluated on FinTriad's **core selected basket**: the tickers accounting for 90% of its total absolute exposure-days, measured as |position weight| $\times$ sessions held and summed across all orders in the evaluation window. This removes differences in asset availability while preserving each baseline's directional constraints, and the 90% cutoff excludes names held only briefly or at small weights, isolating portfolio construction from asset selection. Because the basket is defined over the evaluation window, the baselines carry an in-sample advantage in asset selection. Full configurations are given in Appendices D and E.

4.2 RESULTS

Baseline Comparison. To address RQ1, we report the performance of FinTriad and all baselines together with passive market references in Table 1 and Figure 2a. Over the 375 trading days in the evaluation window, FinTriad traded 372 distinct tickers, of which 49 form its core selected basket (see Appendix D). FinTriad achieves the highest raw return, with a cumulative return of 112.97% and CAGR of 66.20%, as well as the highest risk-adjusted return, with a Sharpe ratio of 2.23. It also records the lowest maximum drawdown of 7.46%, with annualized volatility of 22.14%. The three

Table 1: Performance of FinTriad and baselines, 2025-02-03 to 2026-07-31 (375 trading days). FinTriad is the median-CR run of five (§4.1); sub-portfolio metrics are computed on their own active days and the aggregate on all days. Among compared methods, the best value per metric is in bold and the second best is underlined.

Strategy	CR (%)	CAGR (%)	Vol (%)	MDD (%)	SR
FinTriad	**112.97**	**66.20**	22.14	**7.46**	**2.23**
1-Day Intraday	*33.84*	*24.93*	*13.54*	*6.79*	*1.43*
7-Day Swing	*29.68*	*22.48*	*21.17*	*19.46*	*0.88*
Value (Monthly)	*170.31*	*99.06*	*45.16*	*17.13*	*1.66*
FinRL-A2C	80.78	48.87	27.77	25.67	<u>1.43</u>
FinRL-PPO	79.71	47.98	27.99	26.26	1.40
MVP	39.44	25.03	**15.00**	<u>12.25</u>	1.31
Momentum	<u>100.13</u>	<u>59.40</u>	72.27	54.41	0.93
RSP	21.67	14.09	<u>15.69</u>	15.96	0.67
FinMem	56.43	35.08	20.16	20.15	1.40
FinTriad Core Basket Buy&Hold	79.84	48.35	28.01	26.17	1.41
S&P 500 Buy&Hold	26.25	16.96	17.81	18.76	0.75
NYSE Composite Buy&Hold	21.32	13.87	14.58	15.08	0.70

Table 2: Tri-horizon aggregate for the four runs other than the median-CR run reported in Table 1. Mean and SD (sample) are over all five runs.

Run	CR (%)	CAGR (%)	Vol (%)	MDD (%)	SR
1	104.76	61.87	21.49	9.51	2.17
2	120.88	70.32	25.22	8.42	2.08
3	102.10	60.45	21.57	13.46	2.12
4	118.07	68.86	26.48	12.31	1.96
Mean	111.76	65.54	23.38	10.23	2.11
SD	8.17	4.29	2.31	2.56	0.10

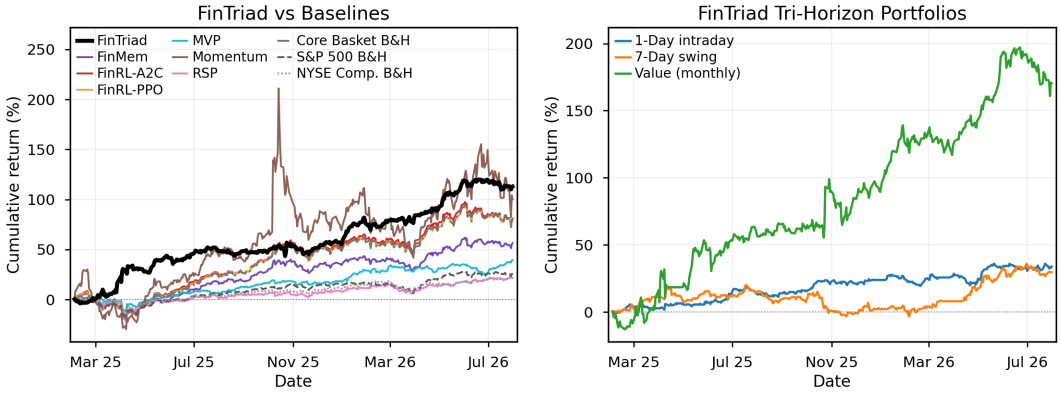
horizons generate weakly correlated return streams (pairwise correlations 0.198, -0.059 and 0.196; Appendix C.1), so the aggregate attains a Sharpe ratio above that of any single horizon.

FinTriad outperforms all strategy baselines on return, Sharpe ratio and drawdown. Momentum reaches a comparable CR (100.13%) only with far higher volatility and drawdown, while MVP achieves lower volatility at the cost of substantially lower return.

Run-to-run Dispersion. Across the five runs (Table 2), the lowest cumulative return (102.10%) and the lowest Sharpe ratio (1.96) still exceed the best baseline values in Table 1 (100.13% and 1.43). The Sharpe ratio is the most stable metric (2.11 ± 0.10) and maximum drawdown the least (10.23 ± 2.56); the weakest run draws down 13.46%, above MVP's 12.25% but below every other baseline.

Asset-Selection Control. To address RQ2, the **core selected basket** provides a favorable control for the baselines, as it gives them access to the assets selected by FinTriad while isolating portfolio construction from asset selection. The 49-stock basket spans 10 sectors, with Technology accounting for 29.8% of exposure-days and the top five names accounting for 31.2% (see Appendix D for full details). It achieves a 79.84% cumulative return under equal-weight buy-and-hold. Only Momentum substantially exceeds this passive benchmark, at the cost of a 54.41% drawdown, while FinRL-A2C essentially matches it (80.78% versus 79.84%) with a 25.67% drawdown; the other baselines underperform it. FinTriad reaches 112.97%, 33.13 points higher, but on a universe that is not identical: the basket covers 90% of its absolute exposure-days, so FinTriad also traded the 10% residual withheld from the baselines. The complement control (Appendix D) shows that residual to be market-like, returning 26.39% under equal-weight buy-and-hold against 26.25% for the S&P 500, so the gap does not come from a stronger opportunity set outside the cut. The result therefore reflects both the selected investment universe and the portfolio construction applied to it.

Cost Sensitivity. FinTriad's multi-horizon design entails different trading frequencies, particularly on the 1-Day horizon. With asset selection held fixed by the core basket, we evaluate FinTriad and the baselines at 5 and 10 bps applied to realized one-way daily turnover (Table 3); the three market baselines are buy-and-hold and unaffected. At 5 bps FinTriad retains the highest SR and the lowest



(a) Cumulative return of FinTriad (tri-horizon aggregate, solid black) against the baselines

(b) Per-horizon decomposition of FinTriad.

Figure 2: Cumulative-return analysis of FinTriad over the February 3, 2025 to July 31, 2026 evaluation window.

Table 3: Performance under 5 and 10 bps transaction costs. Costs are applied to realized one-way daily turnover. Best value per column in bold, second best underlined.

	5 bps			10 bps		
Strategy	CR (%)	SR	MDD (%)	CR (%)	SR	MDD (%)
FinTriad	<u>85.83</u>	**1.81**	**8.19**	62.14	<u>1.40</u>	**10.76**
FinMem	52.65	1.32	20.54	48.96	1.24	20.92
FinRL-A2C	80.01	<u>1.42</u>	25.69	<u>79.25</u>	**1.41**	25.72
FinRL-PPO	78.43	1.39	26.28	77.69	1.38	26.31
MVP	39.12	1.30	<u>12.26</u>	38.80	1.29	<u>12.27</u>
Momentum	**99.48**	0.93	54.46	**98.84**	0.93	54.51

MDD. At 10 bps its CR falls to 62.14%, behind Momentum and the RL baselines, but its SR remains 1.40 against a best of 1.41 and its MDD remains the lowest at 10.76%. Competitiveness under cost is therefore concentrated in risk-adjusted return and drawdown rather than raw return.

Ablation Study. To address RQ3 and RQ4, we ablate the input modality and the verifiable reward, which respectively determine the market memory E_t and the corrective memory D_t. Each configuration is run five times, and we report the median-CR result in Table 4, with cumulative-return trajectories provided in Appendix F. **Input modality.** Replacing the Bloomberg Brief video input with audio-only or transcript-only representations substantially reduces aggregate performance, with CR falling to 22.93% and 13.70% and SR to 0.71 and 0.37, respectively. Although audio-only input performs better on the 1-Day horizon, this advantage does not extend to the 7-Day and Value horizons. One possible explanation is that the video's visual cues (e.g., speakers' facial expressions) emphasize short-term, session-specific signals whose relevance may diminish for longer-horizon decisions (Kim & Kim, 2026). **Verifiable-reward reflexive learning.** Disabling VRRL removes the verifier while leaving the diagnostic machinery intact, so that every diagnosed failure is written to D_{t+1} regardless of whether the re-decision recovers the shortfall. This ablation therefore isolates the verifiable reward itself, comparing verified experience against the self-graded reflection adopted by prior reflection-based agents (Shinn et al., 2023; Yu et al., 2023). It reduces CR to 40.39% and SR to 1.19, while increasing MDD to 9.37% despite lower volatility. These results support the role of the full multimodal input in integrating information across horizons and indicate that verifying a correction before admitting it to memory contributes to higher return and lower drawdown in the evaluated setting. Across both the ablations and baseline strategies, performance deteriorates most sharply around the April 2025 tariff shock, whereas the full FinTriad configuration remains comparatively resilient (Appendix F).

Behaviour under an Extreme Market Shock. Following the April 2025 tariff announcement, baseline drawdowns ranged from 12% (MVP) to 46% (Momentum) and the core basket declined 26% from its February high; FinTriad's maximum drawdown was 6.5%, attributable in part to its capacity

Table 4: Ablation results, 2025-02-03 to 2026-07-31 (375 trading days). Left: tri-horizon aggregate; right: per-horizon CR / SR. CR, CAGR, Vol and MDD in %. Best value for each metric in bold; parentheses denote negative values.

	Tri-horizon aggregate					Per horizon: CR / SR		
Configuration	CR	CAGR	Vol	MDD	SR	1-Day	7-Day	Value
FinTriad (full)	**112.97**	**66.20**	22.14	**7.46**	**2.23**	33.84 / 1.43	**29.68 / 0.88**	**170.31 / 1.66**
Bloomberg Brief input modality								
Audio only	22.93	14.88	**15.76**	15.69	0.71	**39.97 / 1.67**	(22.79) / (0.85)	70.28 / 1.35
Transcript only	13.70	9.01	16.53	15.09	0.37	17.66 / 0.66	(10.49) / (0.39)	30.86 / 0.61
Reflexive learning								
w/o VRRL	40.39	25.61	17.14	9.37	1.19	(2.67) / (0.37)	23.05 / 0.51	71.67 / 1.39

to take short positions. The decisive contribution came from a well-timed short in Cleveland-Cliffs (CLF): a 20% position opened on the value horizon on February 3, held through the announcement and closed on April 14 for a 25.1% gain. The Portfolio Manager opened it against both upstream agents, which read the February tariffs as protection for domestic steel: the MMA called CLF a "national champion" with pricing power, and the MVA returned HOLD at 0.65 confidence with a price target above spot. Whether such a reversal is systematic or incidental is testable: holding the February 3 inputs fixed, we re-run the value-horizon decision 30 times, and the Portfolio Manager shorts CLF in all 30. A representative output is shown below.

Representative Portfolio Manager Output

Decision: "Overruling Market Agent LONG. CLF has negative real free cash flow and only $39M in cash. Relying on tariffs that Trump has recently called 'negotiating tools' (2025-01-24) is high-risk."
memory (D_t correction tuple, 2025-01-22): "CLF LONG 10%, labeled *wrong-thesis*: entered on tariff-protection headlines, but Trump's January 20 America First Trade Policy set the Section 232 steel review for April 1, so near-term policy support was absent while cash burn persisted. Verified alternative: SHORT CLF."
thesis: "CLF is a value trap with deteriorating liquidity. Without immediate tariff protection, its cost structure cannot compete with foreign dumped steel."

We then remove the dated memory entries cited above, including the 2025-01-22 correction tuple in D_t, and repeat the procedure; the Portfolio Manager now follows the upstream recommendation in 27 of 30 runs. Prompt, fundamentals and upstream signals are identical across conditions, so the reversal is a property of the accumulated memory rather than of the current day's evidence. The verified alternative that carried it was written on the 1-day horizon and reused here on the value horizon, so VRRL transfers a correction across horizons rather than replaying a stored trade.

5 Conclusion

We introduced **FinTriad**, a multimodal, multi-agent system with adaptive memory for tri-horizon portfolio construction. Across 375 trading days it achieves a 112.97% cumulative return, 66.20% CAGR and 2.23 Sharpe ratio at a 7.46% maximum drawdown, and remains competitive under transaction costs. Controlled comparisons on its selected basket indicate that the gains arise from portfolio construction as well as asset selection, and ablations show that removing multimodal inputs, or removing the verifier while leaving reflection in place, substantially degrades performance. These results suggest that LLM-based portfolio construction benefits from integrating heterogeneous signals, coordinating decisions across horizons, and retaining only market-verified experience in memory.
Limitations and future directions. FinTriad is evaluated as a backtest over a single market and period with one backbone, so the results speak to feasibility rather than to generality across regimes, asset classes or model families. The evaluation window is placed after the backbone's stated knowledge cutoff, which protects against memorized outcomes but bounds how much market history remains available to study. Grounding retained experience in realized prices rather than in the agent's own judgment is, in our view, the more durable part of the design; prospective evaluation, replication across markets and backbones, and admission criteria that aggregate evidence across several outcomes are the directions we consider most promising.

AI USE STATEMENT

Generative AI as the object of study. All LLM calls in the FinTriad pipeline (every agent, the event-matching arbiter, and every ablation) use a single fixed backbone, Gemini 3 Flash (decoding temperature 1.0, web search disabled). The only other learned models are auxiliary, non-generative components: a MiniLM sentence encoder used for event matching (Appendix B), the A2C and PPO policies trained for the FinRL baselines, and, for the FinMem baseline, gemini-embedding-001, paired with Gemini 3 Flash as its backbone so that the comparison is not confounded by model choice. Model versions, decoding settings and tool access are documented in Appendix B.

Generative AI as an aid to this work. Separately from the system and its evaluation, we used a general-purpose assistant (Claude) to polish parts of the manuscript and edit them for readability and grammar. No such assistant is called anywhere in the pipeline or experiments, and none was used to produce any reported number. Research ideas, hypotheses, the system design, and the experimental design are the authors' own, and no generative AI was used to create or clean the dataset. We reviewed all AI-assisted work and take responsibility for the final content of this paper, including all text, claims, and artifacts.

ETHICS STATEMENT

This work involves no human-subject experiments. Bloomberg Brief videos are publicly available on Bloomberg's official YouTube channel and were used solely for non-commercial research; minute-bar data are licensed from Massive, and SEC filings are public. We do not redistribute any videos or price data. The system analyzes presenters' vocal and visual cues only as aggregate market signals and does not identify or profile individuals. All results come from a historical backtest and are not investment advice; live deployment of LLM trading agents would require human oversight and independent risk controls.

REPRODUCIBILITY STATEMENT

The experimental setup is described in section 4, and implementation details, evaluation metrics and baseline configurations are provided in the appendices. The full codebase will be released as open source upon paper acceptance. Raw videos and licensed price data cannot be redistributed but can be obtained from their original sources.

REFERENCES

Marc-Antoine Allard, Arnaud Teinturier, Victor Xing, and Gautier Viaud. Experiential reflective learning for self-improving llm agents. In *ICLR 2026 Workshop on Memory-based Agents (MemAgents)*, 2026.

Andrew Ang and Joseph Chen. Downside risk. *Journal of Portfolio Management*, 29(4):103–112, 2003.

Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *Journal of Finance*, 68(3):929–985, 2013.

Zvi Bodie, Alex Kane, and Alan J. Marcus. *Investments*. McGraw-Hill Education, 10th edition, 2014.

Warren E Buffett. The superinvestors of graham-and-doddsville. *Hermes*, 4:15, 1984.

Mehmet Caner, Agostino Capponi, Nathan Sun, and Jonathan Y. Tan. Designing agentic ai-based screening for portfolio investment, 2026. URL https://arxiv.org/abs/2603.23300.

Yupeng Cao, Zhi Chen, Prashant Kumar, Qingyun Pei, Yangyang Yu, Haohang Li, Fabrizio Dimino, Lorenzo Ausiello, K. P. Subbalakshmi, and Papa Momar Ndiaye. Risklabs: Predicting financial risk using large language model based on multimodal and multi-sources data, 2025. URL https://arxiv.org/abs/2404.07452.

David R. Cariño. Combining attribution effects over time. *The Journal of Performance Measurement*, 3(4):5–14, 1999.

Wesley Chan. Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, 70:223–260, 02 2003. doi: 10.1016/S0304-405X(03)00146-6.

Marvin M. Chun and Marcia K. Johnson. Memory: enduring traces of perceptual and reflective attention. *Neuron*, 72(4):520–535, 2011.

John H. Cochrane. Volatility tests and efficient markets: A review essay. *Journal of Monetary Economics*, 22(3): 463–485, 1988.

Lauren Cohen, Christopher Malloy, and Quoc Nguyen. Lazy prices. *The Journal of Finance*, 75(3):1371–1415, 2020. doi: https://doi.org/10.1111/jofi.12885. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12885.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953, 2009. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/30226017.

Yujie Ding, Shuai Jia, Tianyi Ma, Bingcheng Mao, Xiuze Zhou, Liuliu Li, and Dongming Han. Integrating stock features and global information via large language models for enhanced stock return prediction, 2023.

Georgios Fatouros, Konstantinos Metaxas, John Soldatos, and Dimosthenis Kyriazis. Can large language models beat wall street? unveiling the potential of ai in stock selection, 2024.

B. Graham. *Security Analysis: Principles and Technique*. McGraw-Hill, 1962. URL https://books.google.com.tw/books?id=9Vo90AEACAAJ.

Richard C. Grinold and Ronald N. Kahn. *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*. McGraw-Hill, 2nd edition, 2000.

Sanford Grossman and Joseph Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 02 1980.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.

Orris C. Herfindahl. *Concentration in the Steel Industry*. PhD thesis, Columbia University, 1950.

Albert O. Hirschman. *National Power and the Structure of Foreign Trade*. University of California Press, Berkeley, 1945.

Harrison Hong and Jeremy C. Stein. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 54(6):2143–2184, 1999. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/797990.

John Hull. *Risk Management and Financial Institutions*. John Wiley & Sons, 2007.

Dennis Huynh, Garrett Audet, Nikolay Alabi, and Yuan Tian. Stock price prediction leveraging reddit: The role of trust filter and sliding window. In *2021 IEEE International Conference on Big Data (Big Data)*, pp. 1054–1060, 2021.

Paul Jaccard. The distribution of the flora in the alpine zone. *New Phytologist*, 11(2):37–50, 1912.

Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91, 1993. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2328882.

Zhengyao Jiang, Dixing Xu, and Jinjun Liang. A deep reinforcement learning framework for the financial portfolio management problem, 2017.

Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3):535–547, 2019.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1):99–134, 1998. ISSN 0004-3702. doi: https://doi.org/10.1016/S0004-3702(98)00023-X. URL https://www.sciencedirect.com/science/article/pii/S000437029800023X.

Sungsoo Kim and Ha Young Kim. Can multimodal large language models extract investment signals from financial charts? evidence from portfolio backtests. *Finance Research Letters*, 108:110453, October 2026. doi: 10.1016/j.frl.2026.110453.

Kemal Kirtac and Guido Germano. Sentiment trading with large language models. *Finance Research Letters*, 62:105227, 2024.

Kelvin J.L. Koa, Yunshan Ma, Ritchie Ng, and Tat-Seng Chua. Learning to generate explainable stock predictions using self-reflective large language models. In *Proceedings of the ACM on Web Conference 2024*, WWW '24, 2024.

Nathan Lambert, Jacob Morrison, Valentina Pyatkin, Shengyi Huang, Hamish Ivison, Faeze Brahman, Lester James V. Miranda, Alisa Liu, Nouha Dziri, Shane Lyu, Yuling Gu, Saumya Malik, Victoria Graf, Jena D. Hwang, Jiangjiang Yang, Ronan Le Bras, Oyvind Tafjord, Chris Wilhelm, Luca Soldaini, Noah A. Smith, Yizhong Wang, Pradeep Dasigi, and Hannaneh Hajishirzi. Tülu 3: Pushing frontiers in open language model post-training. *arXiv preprint arXiv:2411.15124*, 2024.

Feng Li. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2):221–247, 2008. ISSN 0165-4101. doi: https://doi.org/10.1016/j.jacceco.2008.02.003. URL https://www.sciencedirect.com/science/article/pii/S0165410108000141. Economic Consequences of Alternative Accounting Standards and Regulation.

Yang Li, Yangyang Yu, Haohang Li, Zhi Chen, and Khaldoun Khashanah. Tradinggpt: Multi-agent system with layered memory and distinct characters for enhanced financial trading performance, 2023a.

Yinheng Li, Junhao Wang, and Yijie Cao. A general framework on enhancing portfolio management with reinforcement learning, 2023b.

Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning, 2023. URL https://arxiv.org/abs/2304.08485.

Xiao-Yang Liu, Hongyang Yang, Qian Chen, Runjia Zhang, Liuqing Yang, Bowen Xiao, and Christina Dan Wang. Finrl: A deep reinforcement learning library for automated stock trading in quantitative finance, 2022. URL https://arxiv.org/abs/2011.09607.

A. Lo. *Adaptive Markets: Financial Evolution at the Speed of Thought*. Princeton University Press, 2017.

Alejandro Lopez-Lira and Yuehua Tang. Can chatgpt forecast stock price movements? return predictability and large language models, 2025. URL https://arxiv.org/abs/2304.07619.

Tim Loughran and Bill Mcdonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65, 2011. doi: https://doi.org/10.1111/j.1540-6261.2010.01625.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2010.01625.x.

TIM Loughran and BILL Mcdonald. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230, 2016. doi: https://doi.org/10.1111/1475-679X.12123. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-679X.12123.

Yichen Luo, Yebo Feng, Jiahua Xu, Paolo Tasca, and Yang Liu. LLM-powered multi-agent system for automated crypto portfolio management, 2025. URL https://arxiv.org/abs/2501.00826.

Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2975974.

Harry M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, 1959. ISBN 9780300013726. URL http://www.jstor.org/stable/j.ctt1bh4c8h.

Robert C. Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361, 1980.

Richard O. Michaud. The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42, 1989.

OpenAI. Gpt-4v(ision) system card. https://openai.com/research/gpt-4v-system-card, 2023a.

OpenAI. Gpt-4 technical report, 2023b.

Simon Paquette-Greenbaum and Jiangbo Yu. Llm agents for combinatorial efficient frontiers: Investment portfolio optimization, 2026. URL https://arxiv.org/abs/2601.00770.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior, 2023.

Nikolaos Pippas, Elliot A. Ludvig, and Cagatay Turkay. The evolution of reinforcement learning in quantitative finance: A survey. *ACM Computing Surveys*, 57(11):1–51, 2025. ISSN 1557-7341. doi: 10.1145/3733714. URL http://dx.doi.org/10.1145/3733714.

Roy Radner. The organization of decentralized information processing. *Econometrica*, 61(5):1109–1146, 1993. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2951495.

Gerard Salton, Anita Wong, and Chung-Shu Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.

Anubhav Sarkar, Swagata Chakraborty, Sohom Ghosh, and Sudip Kumar Naskar. Evaluating impact of social media posts by executives on stock prices. In *Proceedings of the 14th Annual Meeting of the Forum for Information Retrieval Evaluation*, FIRE '22, pp. 74–82, 2023.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017.

William F. Sharpe. The Sharpe ratio. *The Journal of Portfolio Management*, 21(1):49–58, 1994.

Noah Shinn, Federico Cassano, Edward Berman, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

James M. Sinkula. Market information processing and organizational learning. *Journal of Marketing*, 58(1): 35–45, 1994. ISSN 00222429. URL http://www.jstor.org/stable/1252249.

Mohammad Soleymani, David Garcia, Brendan Jou, Björn Schuller, Shih-Fu Chang, and Maja Pantic. A survey of multimodal sentiment analysis. *Image and Vision Computing*, 65:3–14, 2017.

Steve Strongin, Melanie Petsch, and Greg Sharenow. Beating benchmarks: A stockpicker's reality. *Journal of Portfolio Management*, 26(4):11–27, 2000.

Shuo Sun, Rundong Wang, and Bo An. Reinforcement learning for quantitative trading, 2021. URL https://arxiv.org/abs/2109.13851.

Su Tan, Chi Chiu So, Yueyue Sun, Jun-Min Wang, Wai Keung Anthony Loh, and Siu Pang Yung. Vision, voice, and text: Pioneering zero-shot multimodal llms for sentiment-driven investment. In *Proceedings of the 6th ACM International Conference on AI in Finance*, pp. 960–968, 2025.

Su Tan, Chi Chiu So, Yueyue Sun, Jun-Min Wang, Anthony Loh, and Siu Yung. Beyond buffett: Large language models for intelligent value investing. 01 2026.

Sheng-Hua Tsai, Hsin-min Lu, and Huan-Hsun Yen. Unified item segmentation for 10-q and 10-k filings using item-aware document-level auxiliary tasks. In *Proceedings of the 6th ACM International Conference on AI in Finance*, ICAIF '25, pp. 942–950, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400722202. doi: 10.1145/3768292.3770359. URL https://doi.org/10.1145/3768292.3770359.

Meiyun Wang, Kiyoshi Izumi, and Hiroki Sakaji. Llmfactor: Extracting profitable factors through prompts for explainable stock movement prediction, 2024a.

Qian Wang, Tianyu Wang, Zhenheng Tang, Qinbin Li, Nuo Chen, Jingsheng Liang, and Bingsheng He. Megaagent: A large-scale autonomous llm-based multi-agent system without predefined sops. pp. 4998–5036, 01 2025. doi: 10.18653/v1/2025.findings-acl.259.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

Xinlei Wang, Maike Feng, Jing Qiu, Jinjin Gu, and Junhua Zhao. From news to forecast: integrating event analysis in llm-based time series forecasting with reflection. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24, Red Hook, NY, USA, 2024b. Curran Associates Inc. ISBN 9798331314385.

Ruoxu Wu. Portfolio performance based on llm news scores and related economical analysis. SSRN, 2024. Available at SSRN: https://ssrn.com/abstract=4709617.

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

Qianqian Xie, Weiguang Han, Xiao Zhang, Yanzhao Lai, Min Peng, Alejandro Lopez-Lira, and Jimin Huang. Pixiu: A large language model, instruction data and evaluation benchmark for finance. In *Proceedings of NeurIPS Datasets and Benchmarks*, 2023.

Frank Xing. Designing heterogeneous llm agents for financial sentiment analysis, 2024.

Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang. Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031*, 2023.

Yangyang Yu, Haohang Li, Zhi Chen, Yuechen Jiang, Yang Li, Denghui Zhang, Rong Liu, Jordan W. Suchow, and Khaldoun Khashanah. Finmem: A performance-enhanced llm trading agent with layered memory and character design, 2023.

Yangyang Yu, Zhiyuan Yao, Haohang Li, Zhiyang Deng, Yupeng Cao, Zhi Chen, Jordan W. Suchow, Rong Liu, Zhenyu Cui, Zhaozhuo Xu, Denghui Zhang, Koduvayur Subbalakshmi, Guojun Xiong, Yueru He, Jimin Huang, Dong Li, and Qianqian Xie. Fincon: A synthesized llm multi-agent system with conceptual verbal reinforcement for enhanced financial decision making, 2024. URL https://arxiv.org/abs/2407.06567.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah D. Goodman. STaR: Bootstrapping reasoning with reasoning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Wentao Zhang, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, Longtao Zheng, Xinrun Wang, and Bo An. 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*, KDD '24, pp. 4314–4325, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704901. doi: 10.1145/3637528.3671801. URL https://doi.org/10.1145/3637528.3671801.

Xue Zhang, Hauke Fuehres, and Peter A. Gloor. Predicting stock market indicators through twitter "i hope it is not as bad as i fear". *Procedia - Social and Behavioral Sciences*, 26:55–62, 2011.

Tianjiao Zhao, Jingrao Lyu, Stokes Jones, Harrison Garber, Stefano Pasquali, and Dhagash Mehta. Alphaagents: Large language model based multi-agents for equity portfolio constructions, 2025. URL https://arxiv.org/abs/2508.11152.

A Related Work

Our work intersects multiple research streams spanning MLLMs, LLMs for finance, multi-agent systems, and adaptive investment strategies. We organize the related literature into three key areas.

A.1 Multimodal Learning for Financial Analysis

Multimodal learning has gained traction in finance as researchers recognize that market-moving information transcends text. Classical work in multimodal sentiment analysis (Soleymani et al., 2017) established foundational methods for combining text, audio, and visual signals. Recent applications have explored social media (Huynh et al., 2021; Sarkar et al., 2023; Zhang et al., 2011), where textual discussions on platforms such as Twitter, Reddit, and StockTwits provide real-time market sentiment signals.

The advent of Multimodal Large Language Models (MLLMs) such as GPT-4V (OpenAI, 2023a) and LLaVA (Liu et al., 2023) has enabled zero-shot processing across modalities, offering powerful tools for financial signal extraction. FinAgent (Zhang et al., 2024) represents an early effort to apply MLLMs to financial trading, incorporating K-line charts and trading visualizations alongside numerical and textual data, and demonstrating that visual integration can enhance trading performance. To date, however, MLLM applications in finance have largely been confined to isolated sentiment classification or single-stock trading tasks, rather than end-to-end multi-asset investment systems.

A.2 Multi-Agent Systems for Trading

Multi-agent architectures have emerged as a paradigm for enhancing LLM capabilities through specialized role assignment and collaborative reasoning (Wang et al., 2025). In financial trading, several recent works have explored agent-based approaches with varying degrees of sophistication.

News-Driven Agents. The most fundamental architectures leverage LLMs to analyze financial news and generate trading signals (Wu, 2024; Kirtac & Germano, 2024). LLMFactor (Wang et al., 2024a) uses LLMs to identify factors by analyzing relationships between historical news and stock price movements, then extracts these factors from daily news for prediction. MarketSenseAI (Fatouros et al., 2024) develops progressive summarization modules for daily news, fundamentals, and macroeconomics, storing them as memory for retrieval during trading decisions.

Reflection-Driven Agents. Building on generative agent architectures (Park et al., 2023), several works incorporate layered memory and reflection mechanisms. FinMem (Yu et al., 2023) introduces a trading agent with hierarchical memorization in which raw inputs (news, financial reports) are summarized into memories and aggregated into high-level reflections used for decision-making, with retrieval ranked by recency, relevancy, and importance. TradingGPT (Li et al., 2023a) extends this architecture with inter-agent debate, where agents discuss each other's actions and reflections to improve robustness. These designs draw inspiration from cognitive science (Chun & Johnson, 2011), mirroring how humans absorb feedback, generate memories, and apply learned lessons. Recognizing that multi-agent debate can enhance reasoning and factual validity, Xing (2024) proposes a heterogeneous debating framework in which LLM agents with different roles (mood, rhetoric, dependency) debate with each other to improve news sentiment classification.

Reinforcement Learning-Driven Agents. Several works combine LLMs with reinforcement learning for trading. SEP (Koa et al., 2024) leverages RL with memorization and reflection modules, using correct and incorrect predictions from financial market history to refine LLM-generated explanations and predictions. Ding et al. (2023) develop an RL-based framework in which LLMs generate embeddings from news headlines, integrated with stock features and trained via Proximal Policy Optimization (PPO) (Schulman et al., 2017). This line mirrors classical RL applications in trading (Jiang et al., 2017; Li et al., 2023b) while leveraging LLMs' text understanding capabilities.

A common limitation across these multi-agent trading systems is the absence of integrated multimodal information processing, with most architectures restricted to text-only or text-plus-numeric inputs.

A.3 Adaptive Learning in Trading Systems

Adaptive learning mechanisms are essential for trading systems operating in evolving markets. Traditional approaches employ reinforcement learning (Jiang et al., 2017) or online optimization, but these often lack interpretability and struggle to incorporate qualitative insights from performance analysis. Recent LLM-based trading agents have begun incorporating adaptive mechanisms; the reflection mechanisms in FinMem (Yu et al., 2023) and TradingGPT (Li et al., 2023a) enable agents to build high-level understanding from raw observations, though these systems lack explicit mechanisms to prevent overfitting to recent performance or to govern continuous self-revision under non-stationary regimes. In each case the lesson is graded by the agent that produced it, an arrangement that Reflexion (Shinn et al., 2023) makes explicit and that subsequent work has shown to admit unsupported corrections (Allard et al., 2026).

Verifiable rewards. A parallel development in LLM post-training replaces a learned reward model with a verifiable reward, in which an external program rather than a model checks the outcome and determines whether an attempt is credited. This signal underlies recent reasoning-oriented post-training (Lambert et al., 2024; Guo et al., 2025) and, earlier, rationale bootstrapping methods that retain only reasoning chains reaching a checkable answer (Zelikman et al., 2022). These methods update model parameters through gradient-based optimization. VRRL adopts the same principle in a setting where no parameters are updated: the backbone remains fixed and the verifier is applied to memory writes, admitting a correction only when one of a small number of sampled counterfactual re-decisions is priced by the market above the threshold that triggered the pass. We are not aware of prior work in which an outcome-based check governs what a financial agent retains from its own experience; prior reflection-based agents write the agent's own assessment of a failure directly to memory. The setting also differs from the post-training case: the market prices an outcome rather

than a program checking an answer, so credit is assigned empirically. This motivates evaluating a revised reading over several sampled attempts rather than a single draw, and restricting what that reading may carry (Appendix B.6).

B Implementation Details

B.1 Multimodal Market Agent: Structured Prompt & Event-matching

This appendix provides the structured prompt template invoked on each Bloomberg Brief video (§B.1.1) and the two-stage semantic procedure used to match each newly extracted event against the canonical event ledger (§B.1.2).

B.1.1 Video Analysis Prompt

Each Bloomberg Brief video is submitted to a multimodal LLM through a single call carrying both the video and the prompt template below. The prompt elicits four classes of structured output: a list of extracted events with rich classification fields, three lists of multimodal cues (visual, audio, and content), an overall day-level sentiment label, and a list of key moments with timestamps.

Video Analysis Prompts: Market Regime Synthesis (A) and Event Extraction (B)

Prompt A: Market Regime Synthesis

```
Analyze this entire video and synthesize the prevailing market
    regime
by jointly considering the spoken content, vocal cues (tone, stress,
hesitations), visual cues (body language, slide emphasis), and the
cumulative direction of the discussed events.

Assign one label from EACH of the two dimensions:
  - weekly_regime  in {bull_trend, bear_trend, sideways_volatile,
                       recovery, distribution}
  - monthly_regime in {expansion, contraction, stagflation,
                       rate_driven, geopolitical, earnings_driven}

Provide a one-sentence rationale citing the dominant multimodal
evidence (which modality and which cue most strongly informed the
assignment).

Output ONLY valid JSON:
{
  "market_regime": {
    "weekly_regime":    "bull_trend/bear_trend/sideways_volatile/
    recovery/distribution",
    "monthly_regime":   "expansion/contraction/stagflation/
    rate_driven/geopolitical/earnings_driven",
    "regime_rationale": "one sentence justifying the assignment"
  }
}
```

Prompt B: Event Extraction and Multimodal Cues

```
Analyze this entire video and extract all market-relevant events
together with the multimodal cues observed throughout.

1. Key Events -- For EACH event classify:
   - is_long_term_nature: true  -> policy change, fiscal reform,
        regulatory update, structural/strategic shift
   - is_long_term_nature: false -> daily market move, earnings beat/
    miss,
        short-term data release, immediate reaction
```

```
   - delta_assessment: how much NEW information this event adds vs
    prior
        knowledge (delta_significance = 1.0 for first-ever mention,
        lower for reiterations)

2. Multimodal Signals
   - Visual: body language, facial expressions, slide content
   - Audio:  tone, stress, hesitations, vocal emphasis
   - Content: what was said and what was emphasized

3. Overall Sentiment: bullish / bearish / neutral

4. Key Moments: notable timestamps or moments

Output ONLY valid JSON:
{
  "events": [{
    "description":         "...",
    "type":                "policy/fiscal/regulatory/data/earnings/
    announcement/other",
    "importance":          "high/medium/low",
    "sentiment":           "bullish/bearish/neutral",
    "is_long_term_nature": true,
    "delta_assessment": {
      "new_facts":          ["specific new information not
    previously known"],
      "changed_estimates":  ["any revised forecasts or expectations
    "],
      "confirmed_prior":    ["previously known facts now reconfirmed
    "],
      "delta_significance": 0.85
    }
  }],
  "visual_signals":     ["..."],
  "audio_signals":      ["..."],
  "content_highlights": ["..."],
  "sentiment":          "bullish/bearish/neutral",
  "key_moments":        [{"description": "..."}]
}
```

B.1.2 TWO-STAGE SEMANTIC EVENT MATCHER

The matcher reconciles each newly extracted event against the canonical event ledger so that paraphrases of the same underlying development collapse to a single ledger entry rather than fragmenting into spurious duplicates. Specifically, the procedure unfolds in two stages: a fast embedding-based hybrid scoring that handles the high-confidence and clearly novel cases, and a semantical LLM arbitration step reserved for the genuinely ambiguous middle band.

Stage 1: Hybrid embedding scoring. A FAISS index (Johnson et al., 2019) over normalized sentence embeddings produced by the MiniLM model (Wang et al., 2020) returns the top-K nearest neighbors of the new description d_{new}, restricted to entries whose `last_seen` predates the current trading day to enforce temporal causality. Each candidate d_{cand} is then re-scored by a hybrid similarity that combines semantic embedding agreement, measured by cosine similarity (Salton et al., 1975), with overlap on a financial-entity set $\mathcal{E}$. comprising tickers, curated financial keywords, and numeric-with-unit tokens extracted from each description, scored by the Jaccard coefficient (Jaccard, 1912):

$$\text{sim}(d_{\text{new}}, d_{\text{cand}}) \;=\; w_e \cdot \cos\big(\mathbf{e}_{\text{new}}, \mathbf{e}_{\text{cand}}\big) \;+\; w_j \cdot \text{Jaccard}\big(\mathcal{E}_{\text{new}}, \mathcal{E}_{\text{cand}}\big), \qquad (w_e, w_j) = (0.7, 0.3).$$

The 70/30 split prioritizes semantic similarity while letting entity overlap rescue paraphrased descriptions that share specific tickers, programs, or numeric magnitudes (e.g., two phrasings of a 25-basis-point Fed cut that diverge in surface form but coincide on the entity set $\{$`fed`, `25bps`,

cut}). The top-scoring candidate is forwarded to Stage 2; if no candidate exceeds the lower-band cutoff, the new event is taken as previously unseen.

Stage 2: Tiered match decision. The top-scoring candidate is dispatched into one of three bands defined on the hybrid similarity. High-confidence matches are auto-accepted and merged into the existing entry without further verification. Ambiguous matches fall in a calibrated middle band and are passed to an LLM arbiter that issues a binary same-event judgment with both descriptions presented side-by-side; an affirmative judgment triggers the same merge, while a negative judgment reroutes the event to the next band. Low-confidence matches fall below the lower cutoff and create a new ledger entry with a stable identifier that survives index rebuilds. Each merge increments the canonical entry's occurrence count, updates its `last_seen` date, and appends the day's `delta_assessment` as a new increment record, preserving the per-day evolution of long-running narratives. An idempotency guard suppresses double-counting under re-execution by skipping merges when an increment for the current date already exists in the entry.

B.2 Indicator Panel of Multimodal Technical Agent

Table 5 presents the indicator panel of MTA, which consists of a fast set for 1-day decisions and a classical set for 7-day swing positioning, with a shared volatility/liquidity block emitted at both horizons.

B.3 Value Agent: Implementation Details and Prompts

The extractor captures 17 standard XBRL fields spanning the income statement (revenue, net income, operating income, gross profit, diluted/basic EPS, and diluted/basic share counts), balance sheet (total assets, shareholders' equity, cash and equivalents, current assets, current liabilities, long-term debt, and short-term debt), and cash flow statement (operating cash flow and capital expenditures); **when more than $\tau_{\text{null}} = 3$ of these fields are missing**, Prompt A (§B.3) is invoked to identify the additional XBRL label for each null entry. **Qualitative sections** (Business Overview, Risk Factors, Legal Proceedings, and MD&A) are each truncated to 20,000 characters before insertion into Prompt B (§B.3), with prior-period sections passed as a separate block to enable trend comparison. **Should the LLM be unavailable or return unparseable output**, the system falls back to a rule-based signal that defaults to HOLD with confidence 0.5, flagging attractive valuation when P/E < 15 and valuation concern when P/E > 30, strong profitability when profit margin $> 15\%$ and weakness when $<5\%$, and strong liquidity when the current ratio > 2 and concern when <1.

System prompt.

System Prompt

```
You are a professional financial analyst and investment advisor.
Provide thorough, objective analysis based on financial data and SEC
filing information.  Always consider both risks and opportunities.
```

Prompt A: XBRL industry-metric selection. Curly-brace placeholders are substituted at runtime with the filing's extracted metrics and available XBRL labels.

Prompt A: XBRL Industry-Metric Selection

```
You are a value investing analyst.  We have extracted standard
financial metrics from a {FORM_TYPE} SEC filing for {TICKER}.
ALREADY EXTRACTED (standard metrics found):
{FOUND_METRICS}
STANDARD METRICS NOT FOUND (null - may not apply to this
company/industry):
{NULL_METRICS}
The null metrics often indicate the company is in a specialized
industry (e.g., banks, insurance, REITs) where conventional
accounting categories differ.
ALL AVAILABLE XBRL LABELS in this filing:
```

```
INCOME STATEMENT:
{INCOME_LABELS}
BALANCE SHEET:
{BALANCE_LABELS}
CASH FLOW STATEMENT:
{CASHFLOW_LABELS}
Task: Select 3-8 ADDITIONAL labels from the lists above that would
be most valuable for a comprehensive value investing analysis of
{TICKER} and its industry. These should complement (not duplicate)
what is already extracted. Focus on industry-specific metrics
that reveal profitability drivers, balance-sheet health, and risk
factors unique to this company's sector (e.g., Net Interest Income,
Provision for Credit Losses, Total Deposits for banks; Net Premiums
for insurers; Same-Store Sales for retailers).
Return ONLY valid JSON. Keys are descriptive snake_case metric
names; values are EXACT label strings from the lists above, or
null if the concept is important but the label is unavailable. No
markdown, no explanation:
{
"descriptive_metric_name": "Exact XBRL Label Or null",
...
}
```

Prompt B: Investment analysis. Square-bracket sections are populated with data from the upstream pipeline.

Prompt B: Investment Analysis

```
Please analyze the following financial data and SEC filing
information for {TICKER} as of {TARGET_DATE} and provide an
investment recommendation.
FINANCIAL METRICS
Valuation: P/E {pe_ratio}, P/B {pb_ratio}, P/S {ps_ratio}, Price
{stock_price}, Market Cap {market_cap}
Profitability: Revenue {revenue}, Net Income {net_income},
Operating Income {operating_income}, Gross Profit {gross_profit},
Profit/Operating/Gross Margin {margins}, ROE {roe}, ROA {roa}, EPS
{eps}, EBITDA {ebitda}
Liquidity: Current Ratio {current_ratio}, Cash {cash}, FCF {fcf},
OCF {ocf}, CapEx {capex}
Leverage: D/E {d_e}, Total Assets {total_assets}, LT Debt
{lt_debt}, ST Debt {st_debt}, Equity {equity}
Industry-Specific Metrics: [LLM-selected additional XBRL metrics
for this company's sector]
Data Source Discrepancies: [Metrics with >20% SEC vs. Yahoo
difference; both values shown]
SEC FILING INFORMATION (current period)
Business Overview [10-K Item 1] | Risk Factors [10-K Item 1A / 10-Q
Part II Item 1A] | Legal Proceedings [10-K Item 3 / 10-Q Part II
Item 1] | MD&A [10-K Item 7 / 10-Q Part I Item 2]
PRIOR PERIOD SEC FILING (for trend comparison)
Prior Business Overview | Prior Risk Factors | Prior Legal
Proceedings | Prior MD&A
ANALYSIS INSTRUCTIONS
Respond ONLY with valid JSON:
{
"trading_signal": "BUY" | "HOLD" | "SELL",
"confidence": 0.0-1.0,
"price_target": <float or null>,
"investment_thesis": "2-3 sentence summary",
"strengths": ["...", "...", "..."],
"risks": ["...", "...", "..."],
```

```
"financial_analysis": "Detailed analysis of metrics and trends",
"qualitative_analysis": "Analysis of SEC filing text",
"recommendation_reasoning": "Justification for signal and
confidence"
}
Focus on: (1) valuation vs. fundamentals; (2) financial health
and trends; (3) SEC-disclosed risks and opportunities; (4) MD&A
changes vs. prior period; (5) implications of large SEC vs. Yahoo
discrepancies; (6) industry-specific metric signals. Be specific
and quantitative. Consider both bull and bear cases.
```

B.4 Memory System Schema

Table 5: Indicator panel emitted by the technical agent. Inputs columns: O = open, H = high, L = low, C = close, V = volume.

Indicator	Parameters	Inputs	Role in the panel
Bollinger bands	20-period, 2σ	C	Volatility envelope around a moving mean; supplies mean-reversion and breakout reference levels.
ATR	period 14	H,L,C	Average true range; consumed numerically for volatility-adjusted stop-loss and take-profit sizing.
Relative volume	vs. 20-day average	V	Participation proxy; distinguishes conviction moves from low-liquidity drift.
Support / resistance	pivot-detected	H,L	Discrete reference prices from local extrema, used for entry placement and invalidation.

(a) Shared indicators (both horizons).

Indicator	1-day params	7-day params	Inputs	Role in the panel
EMA	9, 21	20, 50, 200	C	Trend direction and dynamic support/resistance; short windows track recent momentum, long windows anchor regime context.
VWAP (rolling)	20-day	n/a	H,L,C,V	Volume-weighted typical price $\frac{H+L+C}{3}$ over a rolling window; institutional fair-value benchmark, used only at the 1-day horizon.
RSI	period 9	period 14	C	Bounded momentum oscillator; the shorter window reacts to fresh price action, the classical 14 filters multi-day noise.
MACD	5:13:6	12:26:9	C	Momentum-of-momentum via EMA differential; fast triplet captures short-horizon turns, classical triplet identifies swing trend changes.
ADX	n/a	period 14	H,L,C	Wilder's directional index; gauges trend strength independent of direction at the swing horizon.
Stoch-RSI	n/a	period 14	C	Normalizes RSI within its own range to flag momentum extremes the level of RSI alone may obscure.

(b) Per-horizon indicators.

Table 6 presents the memory system schema, which is designed so that E_t supplies the situational context for a given trading day, C_t supplies the prior over how to act on that context, and D_t supplies the corrective signal that re-shapes both over time.

Table 6: Schema of the memory system m_t.

Component	Contents
Market Events Core E_t (event ledger)	
Daily layer	Per-trading-day event entries (announcement, data release, earnings, policy headline) with type, sentiment, and source.
Weekly layer	Weekly aggregation of daily entries; cross-day themes and re-mentions consolidated for swing-horizon retrieval.
Monthly layer	Monthly digest of structural and policy events; long-memory context for value-horizon decisions.
Trading Capacity Core C_t (case-based reasoning store)	
Event-to-market map	Dated event–outcome rules: "event e on day $\tau \Rightarrow$ realized market response."
Sector playbook	Per-sector triggers and confidence scores capturing which signals reliably move which sectors.
Event-type catalog	Per-category statistical impacts (mean return, hit rate, dispersion) aggregated across historical instances.
Feedback buckets	Horizon-specific records of per-basket returns and attributions (1-day, 7-day, value), used for parameter recalibration.
Performance Log D_t (corrective memory)	
Correction tuples	Each tuple records a position diagnosed as directionally incorrect, the structured failure label assigned to it (wrong-thesis, wrong-timing, wrong-sizing), and the corrected alternative produced by the verified re-decision.

B.5 PORTFOLIO MANAGER

Position constraints. Every parsed position is constrained to weight $\in [5\%, 20\%]$, with gross exposure across all positions in a basket capped at $\leq 100\%$. A position outside these bounds is removed and the remaining positions are re-normalized.

Governance action persistence. Governance decisions over inherited 7-day and value positions are persisted per date and reused across optimization iterations within the same day. Concretely, when the diagnostic loop re-runs the PM under the revised reading δ_t, the PM is given the same governance action set that was committed in the baseline run; it is allowed to revise position weights and construct new positions, but cannot retroactively re-decide whether a 7-day position should be EXIT_TODAY or HOLD. This separation is what prevents the verification pass from cascading exit decisions that were already committed to in the baseline run, and it keeps the counterfactual comparable to the executed decision.

Scenario-first retrieval rule. Within each horizon's prompt, retrieval enforces a scenario-first rule: if at least two stored entries share both the weekly and monthly regime label classified for t, the top three slots in the prompt are reserved for those entries and labeled SCENARIO_MATCH. The remaining slots come from a general relevance pool. The PM is explicitly instructed that retrieved analogies are starting hypotheses, not mandates, and that any divergence from a retrieved analogy must be recorded with reasoning in the portfolio's `warnings` field.

Rebalancer pass. The rebalancer receives the newly constructed proposal, the live book of same-horizon open positions, today's regime snapshot, and a mechanically computed *cross-basket freshness budget* that decays each basket's allocation by its age (a fresh proposal at age 0 receives the full freshness multiplier; a 7-day-old basket built on stale news receives a floor of 20%). The LLM is permitted to override the mechanical budget by at most ± 15 percentage points, citing specific same-day news to justify the swing. The rebalancer's output is a per-ticker action vector with vocabulary HOLD/TRIM/ADD_TO/CLOSE/NEW. Validity (action vocabulary, weight sum, forced-closure cap,

basket-budget consistency) is checked in Python before persistence; violations cause a fallback to the basket-builder's raw weights so a malformed LLM response cannot corrupt the live book.

Ticker-hygiene validation. A persistent failure mode of LLM-driven basket construction is hallucinated tickers, including sector codes, basket names, or non-existent symbols presented as positions. We run every parsed portfolio through three validators: a format check that each ticker conforms to the standard US-equity symbol pattern; an exit-parameter check that corrects sign conventions on stop-loss and take-profit levels and removes positions where both are zero; and a weight renormalizer that scales remaining positions to sum to 100% after any filtering. Cross-horizon conflicts, such as a long position on a ticker in the 1-day basket alongside a short on the same ticker in a 7-day basket on the same day, are flagged for review but not auto-resolved, since they can occasionally represent intentional hedges.

B.6 Reflexive Agent and the Verifiable-Reward Gate

Event-to-market extraction. For each new completed backtest, the agent constructs per-horizon date records pairing the news context with the realized horizon-h outcomes, then asks the LLM to extract reusable rules of the form "when event-type X occurs in scenario Y, sector Z tends to move direction D." Rules carry the source date, horizon tag, and scenario tag; one date can produce multiple rules (a Fed pivot and a tech earnings beat on the same day are separate entries). Newly extracted rules are merged with existing rules under a 50-entry cap that drops the lowest-quality entries when saturated.

Sector playbook update. A separate prompt aggregates all event-to-market entries into per-sector trigger rules with confidence scores ($\mathrm{count}/(\mathrm{count}+2)$ as a smoothed prior) and into per-event-category statistical impacts. This is the layer the PM queries when it needs to know what hawkish-Fed days have historically done to high-PE technology, independent of any specific date.

Trade-pattern analysis and core merge. The agent runs a coarser pattern analysis across all completed backtests in the window, identifying winning and losing signal combinations, systematic biases (e.g., shorts underperforming in low-VIX regimes), and calibrated thresholds (minimum technical score, RSI ranges, volume ratios) that the PM will use as soft guidance. The output is merged with the prior trading capacity core under an LLM-mediated update step that increments the core's version, weights new evidence at 40% and prior at 60%, and tags the core with the scenario observed during the run. We deliberately avoid hard rules of the form "never do X" in favor of calibrated thresholds and case-based lessons, because hard rules degrade catastrophically in regime shifts the system has not yet seen.

Diagnostic-loop output and two-stage verification. Reflexive agent receives the original positions with the thesis and signal context recorded for each, a binary flag marking which of them underperformed, the news context the PM had at construction time, the analogies the PM retrieved, and the portfolio's realized return together with the SPY return. It receives no per-position return figures. Its output Δ_t has two parts with separate destinations. For each flagged position it assigns one of the three failure types recorded in Table 6 (wrong-thesis, wrong-timing, wrong-sizing) together with the evidence for that label, stated in terms of signals that were available before the open; this part is written to memory and is never shown to the Portfolio Manager. The part that conditions the re-decision, $\delta_t \subset \Delta_t$, is a revised reading of the day's event stream: a re-analysis of the same pre-open evidence o_t, carrying no record of the executed positions, no return figures and no direction. A re-decision therefore operates over the same pre-open candidate set the specialist agents produced, must re-derive each direction from o_t, and cannot scan the universe for the day's largest movers. The revised reading δ_t is then held fixed while the PM is re-invoked up to $K = 3$ times. The reading is not revised between attempts, so the loop is a bounded re-decision under a single reading rather than iterative self-refinement. Scoring stops early at the first attempt whose same-day alpha reaches α^*, the threshold that triggered the loop; the same constant therefore serves as both the trigger and the success criterion, and no separate improvement margin is introduced. A verified correction is written to D_{t+1} as a correction tuple, whose corrected alternative is the direction the verified re-decision itself produced, and a single experience entry is distilled into the event-to-market map. If no attempt reaches α^*, the diagnosis is not credited: no correction tuple is written to D_{t+1}, and only the best of

the K attempts is distilled into the event-to-market map, so that the episode informs the situational prior without entering the corrective channel. The attempt budget addresses variability in the policy: a single sample at temperature 1.0 may fail to act on a sound revised reading, so K attempts test whether that reading can be realized at all. What bounds the consequence is that a credited correction never affects day t: it conditions decisions from day $t + 1$ onward, so its effect is borne entirely by executed returns that enter the reported results.

Bookkeeping and housekeeping. The reflexive agent maintains additional bookkeeping that does not appear in the main control loop: regime-feature timelines (VIX, 10-year yield drift, credit-spread proxy, S&P momentum) used as a numeric sidecar in PM prompts; an anti-pattern detector that surfaces repeat-failure signatures across correction tuples; a memory-distillation pass that prunes old correction tuples under a recency-decayed weighting; and a periodic housekeeping pass that deduplicates near-identical event mappings, resolves contradictory playbook entries, and recomputes catalog statistics. These run on predictable cadences (incremental updates per day, full rebuilds weekly) so that the memory store remains bounded and consistent without manual intervention.

B.7 Point-in-Time Protocol and Look-Ahead Controls

Every input consulted by a decision on trading day t is available before the day-t open, and every reported return is that of an executed position. Table 7 gives the intraday sequence; the gating rule for each source follows.

Bloomberg videos are broadcast between 07:20 and 08:10 ET, before the 09:30 open, and are the only same-day input. **SEC filings** are gated on the filing date, not the fiscal-period-end date. **Price history** is truncated at the previous close, so every indicator for day t uses data up to $t-1$. **Memory retrieval** returns an entry from $m_t = (E_t, C_t, D_t)$ only if it was recorded on a strictly earlier day, back to the start of the warm-up period (Section 4.1). Together these fix the portfolio for day t before the open, after which no further day-t information is consulted.

Execution enters 1-day positions at the day-t open and exits at the day-t close, with stop-loss and take-profit evaluated on 1-minute bars and the first level touched taking effect; the 7-day and value horizons carry no intraday barriers and are marked to the close. **The reflexive pass** runs after the close and writes only to m_{t+1}: the counterfactual re-decision it may produce is never executed and never enters a reported return, so it can influence day $t+1$ onward but not day t.

Table 7: Intraday point-in-time timeline for a decision on trading day t (times in ET).

Time (ET)	Stage	Information used
07:20–08:10	Bloomberg Brief published	The only same-day input
≈08:10	FinTriad ingests the brief and runs the agent chain	Prices and technical indicators to $t-1$; SEC filings filed on or before $t-1$; memory recorded strictly before t
before 09:30	Decision frozen	Portfolio for t fixed; no further day-t information consulted
09:30	Market open	1-day positions entered at the day-t open
09:30–16:00	Intraday	1-day stop/target evaluated on 1-minute bars
16:00	Market close	1-day positions closed; 7-day and value marked to the day-t close
after 16:00	Reflexive pass	Learning only, not scored

C Evaluation Metrics

Let V_t denote the portfolio value at time t and $r_t = (V_t - V_{t-1})/V_{t-1}$ the daily simple return. All metrics are computed on the series of days on which the corresponding sleeve is active, and n denotes the number of such days; for the tri-horizon aggregate n equals the full evaluation window, whereas an individual horizon is evaluated only over the days on which it holds a position. Days on which a sleeve is inactive are excluded rather than filled with zero, so that a horizon trading at a lower cadence is not penalized for holding cash.

Cumulative Return (Hull, 2007).

$$R_{\text{cum}} = \prod_{t=1}^{n} (1 + r_t) - 1 \tag{1}$$

Cumulative return compounds the daily simple returns over the evaluation period, giving the total growth of one unit of capital invested at the start of the window and reinvested throughout.

Compound Annual Growth Rate (CAGR) (Bodie et al., 2014).

$$\text{CAGR} = (1 + R_{\text{cum}})^{252/n} - 1 \tag{2}$$

where 252 is the standard number of trading days per year. CAGR rescales cumulative performance to a yearly basis via geometric compounding, enabling comparison across strategies evaluated over different horizons.

Annualized Volatility (Cochrane, 1988).

$$\sigma_{\text{ann}} = \sqrt{252} \cdot \sigma(r_1, \ldots, r_n) \tag{3}$$

where $\sigma(\cdot)$ denotes the sample standard deviation (Bessel-corrected) of the daily simple returns. Annualized volatility measures the dispersion of returns over a year, serving as a proxy for portfolio risk. Lower volatility is generally preferred, particularly for risk-averse investors.

Maximum Drawdown (Ang & Chen, 2003).

$$\text{MDD} = \max_{0 \le i \le j \le T} \frac{V_i - V_j}{V_i} \tag{4}$$

Maximum drawdown quantifies the largest peak-to-trough decline in portfolio value over the evaluation period, reflecting the worst-case capital loss an investor would have experienced. A smaller maximum drawdown indicates a more robust and capital-preserving strategy.

Sharpe Ratio (Sharpe, 1994).

$$\text{SR} = \frac{\overline{e}}{\sigma(e_1, \ldots, e_n)} \cdot \sqrt{252}, \qquad e_t = r_t - \frac{r_f}{252} \tag{5}$$

where e_t is the daily excess return, $\overline{e}$ its mean, and $r_f = 3.87\%$ the annualized risk-free rate, taken as the mean yield on the 13-week U.S. Treasury bill (`^IRX`) over the 375 sessions of the evaluation window. The same r_f is applied to FinTriad and to every baseline. Equivalently, $\text{SR} = (252\,\overline{r} - r_f)/\sigma_{\text{ann}}$: the numerator is the arithmetic annualization of the mean daily return, which is not the same quantity as the geometrically compounded CAGR reported above. The Sharpe ratio measures excess return per unit of volatility; higher values indicate more efficient risk-adjusted performance.

C.1 Reconciliation of the Tri-Horizon Aggregate

The aggregate is an equal weight over the horizons active that session: writing H_t for the set of horizons holding a position on day t,

$$r_t^{\text{FT}} = \frac{1}{|H_t|} \sum_{h \in H_t} r_t^h, \qquad w_{t,h} = \begin{cases} 1/|H_t| & h \in H_t \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

so a horizon holding nothing carries no weight rather than contributing a zero return. Horizon rows in Table 1 are computed on their own active days; the aggregate is active on all 375 sessions. Mean weights sum to one on every active session. The pairwise correlations in Table 8 instead set inactive days to zero, so that all three pairs are estimated on a common sample.

Table 9 attributes the cumulative return to the three horizons. Because cumulative return compounds, raw daily contributions sum to $\sum_t r_t^{\text{FT}} = 79.25\%$ rather than to the compounded 112.97%; we therefore apply the standard logarithmic linking coefficient (Cariño, 1999) $k_t = \log(1 + r_t^{\text{FT}})/r_t^{\text{FT}}$, normalized by $k = \log(1 + \text{CR})/\text{CR}$, so that the horizon contributions $\sum_t (k_t/k)\, w_{t,h}\, r_t^h$ sum exactly to the reported cumulative return.

The attribution accounts for where the return arises, but not for why the aggregate's Sharpe ratio of 2.23 exceeds that of every horizon taken individually (1.43, 0.88, 1.66). That follows from how weakly the horizons co-move.

Pairwise correlations. Table 8 reports the correlations quoted in §4.2 together with tests of $H_0 : \rho = 0$. Adjacent horizons are weakly but significantly correlated, while the 1-Day and Value horizons, the most widely separated pair, are not distinguishable from uncorrelated; we read that coefficient as zero rather than as evidence of a negative relationship. Lag-1 autocorrelations of the three daily series are -0.09, $+0.01$ and -0.13, so correcting the effective sample size for serial dependence leaves every p-value unchanged.

Table 8: Pairwise correlations of daily horizon returns over the 375-session evaluation window, inactive days treated as zero. t tests $H_0 : \rho = 0$ with 373 degrees of freedom; intervals are Fisher z.

Horizon pair	ρ	t	p	95% CI	Spearman ρ
1-Day, 7-Day	$+0.198$	3.90	< 0.001	$[+0.099, +0.293]$	$+0.177$
1-Day, Value	-0.059	-1.15	0.253	$[-0.160, +0.042]$	$+0.004$
7-Day, Value	$+0.196$	3.87	< 0.001	$[+0.097, +0.292]$	$+0.243$

Variance decomposition. Writing $c_{t,h} = w_{t,h}\, r_t^h$ for a horizon's weighted contribution on day t, so that $r_t^{\mathrm{FT}} = \sum_h c_{t,h}$, the aggregate's variance decomposes exactly into own-variance and covariance terms. Covariance accounts for only 9.9% of aggregate variance, so in the realized data the three streams are close to independent. Had the horizons moved together, the same contributions would have produced an annualized volatility of 30.96% against the observed 22.14%, a diversification ratio of 1.40; had they been mutually uncorrelated, the figure would have been 21.01%, so the design captures 88.6% of the diversification available at these contribution volatilities. Because cumulative return is additive in these contributions while volatility is not, the aggregate inherits the horizons' combined return at materially less than their combined risk, which is what lifts its Sharpe ratio above each of them.

Table 9: Attribution of FinTriad's cumulative return to the three horizons, 2025-02-03 to 2026-07-31 (375 trading days). Contributions are log-linked and sum to the reported cumulative return.

Horizon	Mean weight $\bar{w}_h$	Contribution to CR (%)
1-Day	0.315	16.19
7-Day	0.303	12.19
Value	0.382	84.59
Total	1.000	112.97

D Asset-Selection Control: the Core Selected Basket

Construction. FinTriad traded 372 distinct tickers over the evaluation window. For each ticker we sum exposure-days ($|\text{position weight}| \times$ sessions held) across all orders from the three horizons, rank tickers by this total, and retain the smallest set accounting for 90% of it: the 49 names in Table 11, which together carry 90.2% of all exposure-days. The remaining 323 tickers share the last 9.8% and are mostly single-session positions at small weight. Every strategy baseline in Table 1 trades this basket; the market indices (S&P 500, NYSE Composite, RSP) are passive buy-and-hold references and are not basket-dependent.

Composition and concentration. Sector labels follow Yahoo Finance. The basket spans 10 sectors; by exposure-days, Technology accounts for 29.8%, Industrials 18.3%, Energy 17.5%, Financial Services 10.3%, Basic Materials 9.0%, Communication Services 7.1%, Consumer Cyclical 4.0%, and Healthcare, Consumer Defensive and Utilities 3.9% combined. Within the basket, the five and ten most-traded names hold 31.2% and 49.8% of exposure-days; the Herfindahl index of exposure shares (Herfindahl, 1950; Hirschman, 1945) is 0.0385, an effective number of names of $1/\mathrm{HHI} = 26$ (Strongin et al., 2000).

Breadth. Summing sessions held over every order, the median name was held for 118 sessions and the mean for 153; all 49 were held for more than 20 sessions in aggregate, 40 for more than 60, and 47 were traded in at least two horizons (28 in all three). The basket therefore consists of names to which FinTriad returned repeatedly, not names it touched once.

Buy-and-hold contribution. The equal-weight, daily-rebalanced buy-and-hold of the basket returns 79.84% over the window (the "FinTriad Core Basket B&H" row of Table 1). The last column of Table 11 gives each name's own buy-and-hold return: the median is 52.9%, 39 of 49 are positive, and five AI-hardware names (MU +806%, INTC +364%, AMD +311%, DELL +300%, FIX +298%) dominate the upper tail. An equal-weight basket of the remaining 44 returns 56.36%; the five names alone return 462.7%. The basket's passive return is thus itself concentrated in a few in-sample winners, an advantage that accrues to every baseline trading it and not to FinTriad, which selected these names without foreknowledge of the window.

Complement control. To demonstrate the effectiveness of asset-selection control, we report the performance of the equal-weight buy-and-hold of the remaining 323 tickers that FinTriad traded but did not concentrate on (Table 10). The complement returns 26.39% over the window against 26.25% for the S&P 500, at a lower Sharpe ratio (0.65 against 0.75): the names outside the 90% cut performed like the market. The 53-point gap between the two halves of FinTriad's own traded universe is therefore attributable to where it concentrated capital rather than to the universe it drew from, and the advantage the baselines inherit from the core basket is correspondingly large.

Confining the baselines to a basket of this size is a requirement of their designs rather than a convenience. The minimum-variance programme estimates a sample covariance from 252 daily returns, which is rank-deficient at 372 names and leaves the programme ill-posed, whereas 49 names give $T/N \approx 5.1$. FinRL's environment carries an action dimension and a covariance state that both scale with the universe, and requires a rectangular panel over the training window that a larger universe would shorten further. FinMem instantiates one agent per ticker, so its cost scales linearly in the number of names. The 90% cut is accordingly the smallest set that preserves where FinTriad committed capital while remaining tractable for every baseline.

Table 10: Equal-weight, daily-rebalanced buy-and-hold of the core selected basket and of its complement, 2025-02-03 to 2026-07-31. The complement is the 323 tickers FinTriad traded that fall outside the 90% exposure-days cut.

Universe	CR (%)	CAGR (%)	Vol (%)	MDD (%)	SR
Core selected basket (49)	79.84	48.35	28.01	26.17	1.41
Complement (323)	26.39	17.04	21.82	23.30	0.65
S&P 500 Buy&Hold	26.25	16.96	17.81	18.76	0.75

E Baseline Implementation Details

This appendix details the implementation of every baseline in Section 4.2. All are evaluated over the same window as FinTriad (2025-02-03 to 2026-07-31, 375 sessions), on the same price data (1-min bar data from Massive), gross of transaction costs, and with the metric definitions of Appendix C; no baseline computes its own Sharpe ratio.

Universe. Every strategy baseline trades FinTriad's core selected basket of 49 tickers (Table 11). The market indices (S&P 500, NYSE Composite, RSP) are passive buy-and-hold references and are not basket-dependent. A single portfolio is formed over all 49 names; the baselines are not partitioned. With a 252-day estimation window this gives $T/N \approx 5.1$, so the sample covariance is well conditioned and a single quadratic programme is well posed (DeMiguel et al., 2009).

Minimum-Variance Portfolio (MVP). At each rebalance date r, MVP solves the long-only quadratic programme (Markowitz, 1952; 1959)

$$\min_{w} \; w^{\top}\Sigma w \quad \text{s.t.} \quad \mathbf{1}^{\top}w = 1, \;\; w \geq 0, \tag{7}$$

where Σ is the sample covariance of the strictly preceding 252 daily returns, with a ridge of $10^{-8}I$ for numerical stability. The programme is solved with CVXPY (OSQP backend) and a SciPy SLSQP fallback where OSQP fails to converge. Weights are re-solved on the first trading day of each calendar month (18 rebalances over the window) and drift multiplicatively with realized returns in between. Price history is taken from 2023-12-01 so the first rebalance already has a full estimation window. We

Table 11: FinTriad's core selected basket, in descending order of exposure-days (2025-02-03 to 2026-07-31). Share and Cum. are the name's share and cumulative share of the basket's exposure-days; Sess. is the number of sessions held summed over all orders; Hor. lists the horizons that traded the name (1d/7d/val); B&H is the name's own buy-and-hold return over the window.

#	Ticker	Sector	Share (%)	Cum. (%)	Sess.	Hor.	B&H (%)	#	Ticker	Sector	Share (%)	Cum. (%)	Sess.	Hor.	B&H (%)
1	CVX	Energy	10.2	10.2	399	1d/7d/val	40.7	26	BA	Indust.	1.3	85.7	122	1d/7d/val	22.4
2	HOOD	Financ.	6.3	16.5	228	1d/7d/val	66.6	27	MP	Materials	1.3	86.9	160	1d/7d/val	88.4
3	LMT	Indust.	5.4	21.9	572	1d/7d/val	31.1	28	CLF	Materials	1.2	88.2	94	1d/7d/val	12.5
4	TSM	Tech	5.2	27.1	229	1d/7d/val	96.7	29	INTC	Tech	1.0	89.2	164	1d/7d/val	364.2
5	NVDA	Tech	4.1	31.2	341	1d/7d/val	67.4	30	DAL	Indust.	1.0	90.2	94	1d/7d/val	32.6
6	MSFT	Tech	4.1	35.3	283	1d/7d/val	13.3	31	FIX	Indust.	0.9	91.1	56	val	297.8
7	ASML	Tech	4.0	39.3	123	1d/7d	123.5	32	ORCL	Tech	0.9	92.0	103	1d/7d	−22.3
8	GOOGL	Comm.	3.8	43.1	279	1d/7d/val	75.5	33	AAPL	Tech	0.8	92.9	94	1d/7d	31.7
9	XOM	Energy	3.6	46.7	360	1d/7d/val	52.9	34	SMCI	Tech	0.8	93.6	48	1d/7d	−0.4
10	MU	Tech	3.1	49.8	252	1d/7d/val	805.5	35	PLTR	Tech	0.8	94.4	71	1d/7d	49.2
11	HIMS	Health	3.1	53.0	95	1d/7d	−25.5	36	HAL	Energy	0.6	95.0	35	7d	28.8
12	CCJ	Energy	3.0	56.0	180	1d/val	75.1	37	VRT	Indust.	0.6	95.6	111	1d/7d/val	106.8
13	TMQ	Materials	3.0	59.0	186	7d/val	150.4	38	ADM	Cons. Def.	0.5	96.1	90	1d/7d	63.1
14	CAT	Indust.	3.0	62.0	212	1d/7d/val	123.2	39	GS	Financ.	0.5	96.6	88	1d/7d	63.8
15	NFLX	Comm.	3.0	65.0	118	1d/7d	−26.6	40	LULU	Cons. Cyc.	0.5	97.0	60	1d/7d/val	−71.3
16	GD	Indust.	2.6	67.6	199	1d/7d/val	53.5	41	AMZN	Cons. Cyc.	0.4	97.5	41	1d/7d	14.3
17	DELL	Tech	2.5	70.1	165	1d/7d/val	300.4	42	META	Comm.	0.4	97.9	70	1d/7d	−18.8
18	TSLA	Cons. Cyc.	2.5	72.5	189	1d/7d/val	−23.1	43	NEM	Materials	0.4	98.3	73	1d/7d	124.1
19	CSCO	Tech	2.2	74.7	152	1d/7d/val	97.9	44	BABA	Cons. Cyc.	0.4	98.6	61	1d/7d	26.9
20	UAL	Indust.	2.0	76.7	217	1d/7d/val	14.6	45	COIN	Financ.	0.3	98.9	38	1d/7d	−49.8
21	NUE	Materials	1.6	78.3	116	1d/7d/val	104.8	46	PVH	Cons. Cyc.	0.3	99.2	41	1d/7d/val	−3.0
22	RTX	Indust.	1.6	79.9	270	1d/7d/val	71.3	47	VST	Utilities	0.3	99.5	74	1d/7d	−11.1
23	FCX	Materials	1.5	81.5	168	1d/7d/val	78.0	48	AMD	Tech	0.3	99.7	36	1d/7d	310.7
24	JPM	Financ.	1.5	83.0	210	1d/7d/val	35.6	49	MS	Financ.	0.3	100.0	22	1d/7d	57.8
25	WFC	Financ.	1.5	84.4	105	1d/7d/val	13.3								

use the minimum-variance rather than the full mean–variance form of Markowitz (1952) because the sample mean $\hat{\mu}$ is a noisy estimator that significantly degrades out-of-sample performance (Merton, 1980; Michaud, 1989); MVP relies on Σ alone.

12-1 Cross-Sectional Momentum. At each monthly rebalance date r, each ticker's signal is its cumulative log-return from $r-252$ to $r-21$ trading days, that is the past 12 months excluding the most recent month (Jegadeesh & Titman, 1993; Asness et al., 2013), which avoids short-term reversal contamination. Names are ranked cross-sectionally over the basket, the top decile is held equal-weighted, and weights drift until the next rebalance. We leave the published parameters untuned, including $\mathrm{TOP_FRAC} = 0.10$. On a 49-name universe the top decile is $\mathrm{round}(49 \times 0.1) = 5$ names at 20% each, so this baseline is concentrated by construction rather than by malfunction; its volatility and drawdown in Table 1 should be read in that light.

FinRL-A2C and FinRL-PPO. Both agents use FinRL's `StockPortfolioEnv` for portfolio allocation: the action is a softmax weight vector over the 49 assets applied daily; the state is the 252-day return covariance matrix together with eight technical indicators per asset (MACD, RSI-30, CCI-30, DX-30, Bollinger upper and lower, 30- and 60-day SMA); the reward is the change in portfolio value scaled by 10^{-4}; initial capital is \$1M, $\mathrm{hmax} = 100$, and no transaction cost is charged. Policies are trained with Stable-Baselines3 through FinRL's `DRLAgent` for 80,000 timesteps at seed 42: A2C uses $\mathrm{n_steps} = 5$, $\mathrm{ent_coef} = 0.005$, $\mathrm{lr} = 7 \times 10^{-4}$; PPO uses $\mathrm{n_steps} = 2048$, $\mathrm{ent_coef} = 0.005$, $\mathrm{lr} = 1 \times 10^{-4}$, batch size 128. For assets without sufficient historical data to form the required training panel, we omit them from the training data while retaining the common evaluation universe.

We report both A2C and PPO because prior work finds inconsistent performance across policy-gradient DRL algorithms in financial environments (Liu et al., 2022). In our runs, both converge to near-uniform allocations: A2C's mean maximum weight is 2.8% versus $1/49 = 2.04\%$ under equal weighting, with a cross-asset standard deviation of 0.8%, and its daily returns correlate 0.994 with equal-weight buy-and-hold of the basket. Thus, their rows in Table 1 largely reflect equal-weight exposure to the basket rather than differentiated learned timing.

Market indices. The S&P 500 (SPY) and the NYSE Composite are held long over the full window as capitalization-weighted market references. RSP tracks the S&P 500 Equal Weight Index, which

holds each of the 500 constituents at $1/500$ and is rebalanced quarterly to equal weights by the index provider; it serves as a passive, no-information benchmark that controls for the market-cap concentration embedded in the standard S&P 500. None of the three is basket-dependent, and all three are buy-and-hold, so they incur no turnover and are unaffected by the cost analysis of Table 3.

FinMem. We use the authors' public implementation with their default memory configuration (top-$k = 3$, seven-day feedback look-back, and the published short/mid/long-layer decay and promotion thresholds), replacing the backbone and embedding model with Gemini 3 Flash and `gemini-embedding-001` so that the comparison is not confounded by model choice. One agent is instantiated per ticker in the 49-name basket, with the character profile templated from the company's sector and business description.

Following the original data pipeline, the short-term layer receives the day's news, here the Bloomberg Brief transcript of that day, reduced by one backbone call to the statements concerning that ticker plus the day's market-wide items (416 transcript days); the mid- and long-term layers receive LLM summaries of the MD&A section of each 10-Q and 10-K filed during the window, delivered on the filing date (288 summaries drawn from EDGAR). Memory is populated with FinMem's own train mode, which uses look-ahead reflection, over 2024-04-02 to 2025-01-31 (210 sessions); the evaluation window is then traded causally, without look-ahead, mirroring FinTriad's warm-up.

Each agent emits a buy, sell, or hold decision per session. The decision at t sets the position $d_t \in \{+1, -1, 0\}$ (long, short, or flat) held over $t \to t{+}1$, with ticker return $d_t\,(P_{t+1}/P_t - 1)$; the basket return is the equal-weight mean across the 49 assets. Following FinMem's position-based reward formulation, a hold retains the existing position rather than implying a zero return. Flat positions earn the daily risk-free rate $r_f/252$ ($r_f = 3.87\%$ annualized; Appendix C). Across the evaluation window, agents hold on average 41.5 long, 6.4 short, and 1.1 flat positions per session.

F Additional Ablation Results

Figures 3 and 4 give cumulative-return trajectories for the two ablations reported in Table 4 (§4.2), complementing the endpoint metrics with their path. Figure 3 contrasts the tri-horizon aggregate of FinTriad with the audio-only and transcript-only configurations; Figure 4 shows the analogous comparison with VRRL disabled. In both cases the full configuration leads from the first quarter onward and the ordering reported in Table 4 holds across the window rather than arising at its endpoint.

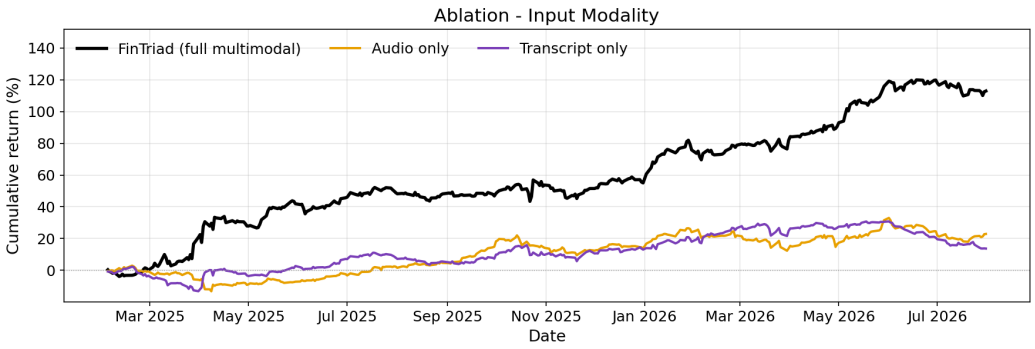


Figure 3: Cumulative-return trajectories for the input-modality ablation. The solid black line is FinTriad with the raw Bloomberg Brief video input; the orange and purple lines are the audio-only and transcript-only configurations.

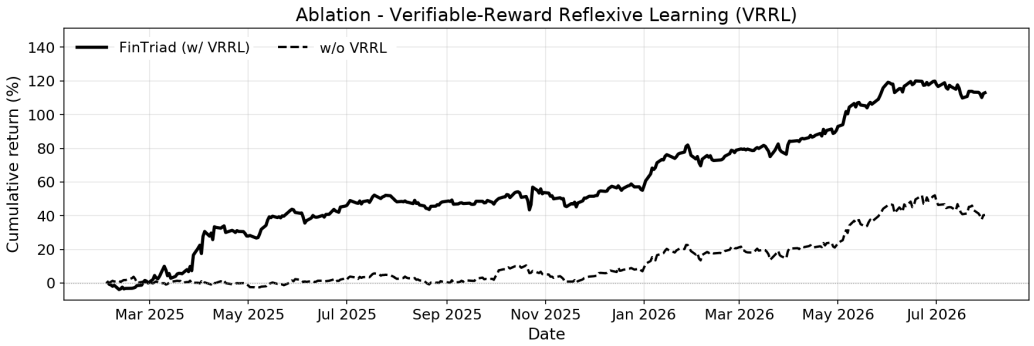


Figure 4: Cumulative-return trajectories for the VRRL ablation. The solid line is FinTriad with the verifiable-reward gate active; the dashed line is the configuration in which the gate is removed and every diagnosed failure is written to memory unverified.