

Article

A Unified Multi-Modal Framework for Intelligent Financial Systems: Integrating Reinforcement Learning, High-Frequency Trading, and Game-Theoretic Approaches with Cross-Modal Sentiment Analysis

Yuwei Zhang ^{1,†}, Fanrong Liu ^{2,†}, Chang-An Xu ³  and Mingni Luo ^{4,*}¹ College of Business, City University of Hong Kong, Hong Kong, China; yzhang5655-c@my.cityu.edu.hk² International Eurasia College, Henan University, Kaifeng 475001, China; jrsxlfr@henu.edu.cn³ College of Materials and Energy, South China Agricultural University, Guangzhou 510642, China; xuchangan@scau.edu.cn⁴ School of Electronic and Information Engineering, Northeastern University, Seattle, WA 02115, USA

* Correspondence: nieeluo@gmail.com

† These authors contributed equally to this work.

Abstract

The rapid evolution of financial technology demands sophisticated artificial intelligence systems capable of handling diverse challenges across multiple domains simultaneously. This paper presents a unified framework that integrates Proximal Policy Optimization (PPO) for robo-advisory systems, multi-scale time-series prediction models for high-frequency trading, in-context learning mechanisms for dynamic investment advisory, game-theoretic reasoning for competitive banking scenarios, and unified embeddings for cross-modal financial sentiment analysis. Our comprehensive framework addresses the critical gap in the existing literature where these technologies have been developed in isolation, failing to leverage their synergistic potential. Through extensive experimentation across multiple financial datasets and real-world scenarios, we demonstrate that our integrated approach achieves superior performance compared to specialized single-domain systems. Specifically, our framework shows a 23.7% improvement in portfolio optimization metrics, reduces prediction error in high-frequency trading by 31.2%, enhances investment recommendation accuracy by 18.9%, optimizes competitive banking strategies with a 27.4% increase in Nash equilibrium convergence speed, and improves sentiment analysis accuracy by 15.6% through cross-modal fusion. The theoretical foundation of our work establishes convergence guarantees for the integrated optimization problem, while our empirical results validate the practical applicability across diverse financial institutions. This research not only advances the state-of-the-art in financial AI but also provides a blueprint for developing comprehensive intelligent systems that can adapt to the complex, interconnected nature of modern financial markets.



Academic Editors: Honghui Xu and Anatoliy Swishchuk

Received: 26 May 2026

Revised: 25 June 2026

Accepted: 30 June 2026

Published: 7 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: proximal policy optimization; high-frequency trading; in-context learning; game theory; cross-modal embeddings; financial AI; reinforcement learning; sentiment analysis

MSC: 68T07

1. Introduction

The financial industry stands at the threshold of a technological revolution. Artificial intelligence systems are transforming traditional practices across trading, advisory services, and risk management. The convergence of machine learning, deep learning, and advanced optimization has created unprecedented opportunities for intelligent financial systems that navigate modern markets with precision and adaptability. However, the current landscape of financial AI research remains fragmented. Isolated developments in specific domains fail to capture the interconnected nature of financial ecosystems. This fragmentation not only limits the impact of individual technologies but also overlooks critical synergies that could emerge from their integration.

To appreciate why this integration gap matters, it is useful to first examine how markets themselves have evolved. The evolution of financial markets over the past decade has been characterized by increasing complexity, velocity, and interconnectedness. Traditional approaches that relied on human expertise and simplified models are increasingly inadequate in capturing the dynamics of modern markets. High-frequency trading has compressed decision-making timeframes from minutes to microseconds. At the same time, the proliferation of alternative data sources has expanded the information landscape beyond traditional financial metrics. The democratization of financial services through robo-advisors has further created demand for sophisticated yet accessible investment strategies that adapt to individual preferences and market conditions in real time.

Against this backdrop of rising market complexity, recent advances in machine learning have provided powerful tools for addressing these challenges. Yet their application in finance has largely followed a siloed approach. Reinforcement learning algorithms such as Proximal Policy Optimization (PPO) have shown promise in portfolio optimization [1]. Deep learning models have revolutionized time-series prediction for trading applications [2]. The emergence of large language models has introduced new possibilities for processing unstructured financial data and providing contextual investment advice [3]. More recently, open-source financial language models such as FinGPT [4] and the PIXIU benchmark suite [5] have lowered the barrier to in-context financial reasoning, while practitioner-oriented surveys of deep reinforcement learning for trading [6] have catalogued an explosion of single-task systems. Game-theoretic approaches have offered insights into competitive dynamics in banking and trading [7]. Multi-modal learning has enhanced sentiment analysis capabilities [8]. Despite these individual successes, the lack of integration among these technologies represents a significant missed opportunity.

This observation motivates our unified framework. Financial decisions are inherently multi-faceted. They require simultaneous consideration of market dynamics, individual preferences, competitive pressures, and sentiment indicators. A portfolio manager, for instance, must optimize asset allocation based on historical patterns. At the same time, they must adapt to real-time market movements, interpret news sentiment, anticipate competitor actions, and align strategies with client objectives. This complexity demands an integrated approach. Such an approach can leverage the strengths of different AI technologies while mitigating their individual limitations through synergistic interactions.

Guided by this motivation, our research addresses this critical gap by proposing a comprehensive framework that unifies five technological pillars of modern financial AI. The PPO component provides robust reinforcement learning for long-term strategic planning in robo-advisory systems. It ensures stable policy updates while maintaining exploration efficiency. The high-frequency trading component leverages multi-scale time-series prediction models that capture micro-market dynamics with millisecond precision, enabling rapid response to market inefficiencies. The in-context learning component allows the framework to adapt to new investment scenarios without extensive retraining. This

component delivers personalized advisory services that evolve with client needs and market conditions. The game-theoretic component enables principled modeling of competitive dynamics in banking and trading. Strategies therefore remain near-optimal even as competitors adapt. Finally, unified cross-modal embeddings synthesize information from text, audio in earnings calls, and visual data from financial charts. The result is a holistic view of market sentiment that transcends single-modal approaches.

Underpinning these five pillars, the theoretical foundation of our framework rests on several key innovations that enable seamless integration. We introduce a novel multi-objective optimization formulation that balances the competing demands of different subsystems while maintaining overall coherence. Our approach employs a hierarchical attention mechanism that dynamically allocates computational resources based on task priority and market conditions. This mechanism ensures efficient operation even in resource-constrained environments. The framework also incorporates a unified state representation that captures information relevant to all components, facilitating information sharing and reducing redundancy. We further establish theoretical convergence guarantees for our integrated optimization algorithm. The analysis demonstrates that the combined system achieves superior performance compared to isolated implementations of individual components.

Beyond the theory, the empirical validation of our framework spans multiple dimensions of financial applications. We evaluate portfolio optimization on historical data from major global markets and observe consistent outperformance of traditional approaches across different market regimes. High-frequency trading experiments use tick-level data from cryptocurrency and equity markets. They show significant improvements in prediction accuracy and execution efficiency. Investment advisory capabilities are assessed through simulated client interactions and backtests of recommended strategies, which reveal superior adaptation to individual preferences and market conditions. Competitive banking scenarios are analyzed with multi-agent simulations that capture realistic market dynamics, validating the effectiveness of our game-theoretic component. Sentiment analysis is evaluated on diverse financial datasets including earnings transcripts, social media posts, and news articles. The results confirm the advantages of cross-modal fusion.

Taken together, the contributions of this paper extend beyond the technical integration of existing technologies. They establish new paradigms for intelligent financial systems. We provide the first comprehensive framework that integrates reinforcement learning, high-frequency trading, in-context learning, game theory, and multi-modal analysis within a unified architecture. Our theoretical analysis establishes convergence properties and performance bounds for the integrated system. It thereby addresses fundamental questions about the stability and optimality of multi-component AI systems. The extensive empirical evaluation across diverse financial applications demonstrates not only the superior performance of our approach but also its practical applicability in real-world scenarios. Furthermore, we release our implementation as an open-source framework. This release enables researchers and practitioners to build upon our work and to accelerate the development of next-generation financial AI systems.

The remainder of this paper is organized as follows. We begin with a review of related work that establishes the context for our contributions and highlights the limitations of existing approaches. The theoretical framework section then presents the mathematical foundations of our integrated system, including the multi-objective optimization formulation and convergence analysis. The methodology section details the implementation of each component and their integration mechanisms. Extensive experimental results demonstrate the effectiveness of our approach across multiple financial applications. We conclude with

a discussion of implications for financial AI research and practice, along with directions for future work that build upon our unified framework.

2. Related Work

The landscape of artificial intelligence in finance has evolved dramatically over the past decade. Significant advances have appeared across multiple domains that form the foundation of our unified framework. Three forces have driven the proliferation of machine learning in finance: the increasing availability of data, abundant computational resources, and the pressing need for more sophisticated decision-making tools in complex markets. Understanding the evolution and current state of these technologies provides essential context for appreciating the novelty of our integrated approach.

Reinforcement learning has emerged as a powerful paradigm for sequential decision-making in financial contexts, with Proximal Policy Optimization representing a significant advancement in policy gradient methods. The seminal work by Schulman et al. [1] introduced PPO as a more stable and sample-efficient alternative to earlier policy gradient algorithms, addressing the challenge of maintaining consistent performance improvements while avoiding catastrophic policy updates. In the context of portfolio management, Jiang et al. [9] pioneered the application of deep reinforcement learning to cryptocurrency portfolio optimization, demonstrating superior risk-adjusted returns compared to mean-variance baselines. The work by Liu et al. [10] introduced the FinRL framework, an open-source library that extends deep RL to multi-asset portfolio management while incorporating realistic transaction costs and market impact crucial for practical deployment. Yang et al. [11] further proposed an ensemble strategy that combines PPO, A2C, and DDPG agents to better handle different market regimes, achieving robust performance across bull, bear, and volatile conditions. The integration of deep RL with careful state design and reward shaping has been studied by Théate and Ernst [12], who provided a systematic empirical analysis of how engineering choices affect convergence and out-of-sample profitability in algorithmic trading. The recent survey of Sun et al. [6] synthesises this rapidly growing literature on deep reinforcement learning for quantitative trading and identifies the absence of joint training across complementary objectives—portfolio allocation, execution, and risk—as one of the central open problems, a gap that our unified framework directly targets.

High-frequency trading has been revolutionized by advances in deep learning for time-series prediction, with numerous architectures competing for supremacy in capturing market microstructure dynamics. The transformer architecture, originally developed for natural language processing, has been adapted for financial time-series by Wallbridge [13], who demonstrated its ability to capture long-range dependencies in limit-order-book data. The DeepLOB architecture of Zhang et al. [2] introduced a deep convolutional network specifically designed for limit-order-book mid-price prediction, achieving state-of-the-art directional accuracy on the FI-2010 benchmark. Earlier convolutional approaches by Tsantekidis et al. [14] showed that order-book features processed at multiple temporal resolutions already give a strong predictive signal at high frequency. The challenge of non-stationarity in financial time series is addressed by the broad literature on concept drift, surveyed comprehensively by Gama et al. [15], which provides the conceptual foundation for the online adaptation strategies used in modern HFT systems. Sirignano and Cont [16] further identified universal features of price formation in financial markets through deep learning, improving prediction accuracy by explicitly modeling order-flow dynamics and liquidity patterns.

The emergence of large language models has opened new possibilities for in-context learning in financial applications, enabling systems to adapt to novel scenarios without

extensive retraining. The foundational work by Brown et al. [3] on GPT-3 demonstrated the potential for few-shot learning in various tasks, inspiring applications in financial advisory services. Wu et al. [17] introduced BloombergGPT, a 50-billion-parameter language model trained on a large mixed financial corpus, which exhibits strong few-shot capabilities on a battery of financial tasks and provides a natural backbone for in-context financial reasoning. The work by Wei et al. [18] introduced chain-of-thought prompting, which substantially improves the multi-step reasoning capability of large language models and has direct applications to investment advice generation, where transparent intermediate reasoning is valuable for both interpretability and compliance review. Retrieval-augmented generation, introduced by Lewis et al. [19], allows language models to access and utilize external knowledge stores while maintaining adaptability to new situations and is now a standard approach for grounding domain-specific advisory systems. The challenge of producing safe, aligned advice has been addressed by Ouyang et al. [20] through reinforcement learning from human feedback, which constrains generation toward outputs that match human-specified preferences and risk constraints. More recent open efforts have made financial language models broadly accessible: Yang et al. [4] released FinGPT, an open-source pipeline for financial LLMs that supports continual pre-training and parameter-efficient fine-tuning, while Xie et al. [5] introduced the PIXIU benchmark, instruction data, and evaluation suite that standardise comparison across financial NLP tasks. We adopt the PIXIU evaluation protocol in our in-context learning experiments to ensure that our results are directly comparable to this emerging line of work.

Game-theoretic approaches have provided crucial insights into competitive dynamics in financial markets, with applications ranging from optimal execution to strategic trading. The classical work by Kyle [21] on market microstructure laid the foundation for understanding strategic interactions between informed and uninformed traders. Modern extensions by Tuyls et al. [7] have incorporated multi-agent reinforcement learning to model complex competitive scenarios in algorithmic trading. Balduzzi et al. [22] formalised the mechanics of n -player differentiable games, providing theoretical tools for end-to-end learning of equilibria in continuous action spaces, which can be specialised to competitive banking strategies. Foundational work by Lasry and Lions [23] introduced the mean-field game framework for modelling large-scale strategic interactions in which individual agents interact with aggregate market dynamics, a setting now widely used to model crowded trading and liquidity provision. Lanctot et al. [24] proposed Policy-Space Response Oracles (PSRO), a unified game-theoretic approach to multi-agent reinforcement learning that integrates deep learning with classical solution concepts and is directly applicable to sequential trading decisions. Applications to systemic risk and financial stability have been investigated by Acemoglu et al. [25], demonstrating how network-game models can capture contagion effects and strategic defaults in banking networks.

Multi-modal learning for financial sentiment analysis has evolved from simple text-based approaches to sophisticated systems that integrate diverse information sources. Early work by Tetlock [26] established the importance of news sentiment in predicting market movements, motivating the development of more advanced sentiment analysis techniques. The introduction of BERT and its variants revolutionized financial text analysis, with Araci [27] and Yang et al. [28] releasing dedicated FinBERT variants pretrained on financial corpora that substantially outperform general-domain language models on financial sentiment classification. Multi-modal approaches began emerging with Zadeh et al. [8], who proposed tensor fusion networks for combining text, audio, and visual information in sentiment analysis. Sawhney et al. [29] fused text and audio from earnings calls through a graph convolutional architecture, demonstrating that cross-modal cues improve volatility forecasting, and Loughran and McDonald [30] surveyed the systematic use of

textual signals in accounting and finance, providing the methodological foundations for the use of unstructured data sources such as filings, transcripts, and social media in modern sentiment pipelines.

Despite these significant advances in individual domains, the integration of these technologies into a unified framework remains largely unexplored. Existing attempts at integration have been limited in scope, typically combining only two or three components without addressing the fundamental challenges of multi-objective optimization and information sharing across diverse subsystems. The work by Deng et al. [31] combined deep direct reinforcement learning with learned feature representations for trading but did not incorporate game-theoretic considerations or in-context learning capabilities. Similarly, Sawhney et al. [32] introduced a multimodal multi-task framework for financial risk forecasting that jointly leverages text and price modalities, yet lacks the strategic planning capabilities provided by reinforcement learning and the equilibrium reasoning of game theory. Recent comprehensive surveys by Goodell et al. [33] and Cao [34] have systematically catalogued the impact of AI and machine learning on finance and identified the absence of production-grade integrations spanning reinforcement learning, microstructure prediction, in-context learning, game theory and multi-modal sentiment. Earlier deep learning literature reviews, such as Sezer et al. [35] on financial time-series forecasting, further suggest that single-market or single-modality models leave substantial signal on the table.

Building on this picture, the novelty of our framework can be made explicit in four points: (i) we present the first end-to-end framework that jointly trains PPO-based robo-advisory, multi-scale time-series prediction for HFT, in-context learning for personalised advisory, differentiable game-theoretic reasoning, and cross-modal sentiment embeddings inside a single multi-objective loss; (ii) we introduce a unified state representation $\mathcal{S}$ that makes information from any subsystem usable by any other subsystem; (iii) we propose a hierarchical attention and gating mechanism that arbitrates between subsystems at decision time, and an adaptive Lagrangian weighting scheme based on uncertainty weighting [36] with an optional PCGrad [37] safeguard against gradient conflict; and (iv) we prove an $O(1/\sqrt{T})$ convergence rate for the joint, multi-objective loss under standard assumptions. Table 1 contrasts the scope of our framework with the closest two-/three-component baselines, making the integration gap explicit.

Table 1. Component coverage of our framework against the closest multi-component baselines ($\checkmark$ = covered).

Method	PPO/RL	HFT	ICL/LLM	Game-th.	Multi-Modal
Deng et al. [31]	$\checkmark$	–	–	–	$\checkmark$
Sawhney et al. [32]	–	$\checkmark$	–	–	$\checkmark$
FinRL-Meta	$\checkmark$	–	–	–	–
FinGPT	–	–	$\checkmark$	–	partial
Our Framework	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$

3. Theoretical Framework

The mathematical foundation of our unified framework rests on a novel multi-objective optimization formulation that seamlessly integrates the diverse components while maintaining theoretical guarantees on convergence and performance. We begin by establishing the unified state space representation that enables information sharing across all subsystems. We then introduce the hierarchical optimization structure that governs the interaction between components. The overall model architecture is summarised in Figure 1, and the training-to-decision algorithm flow is summarised in Figure 2.

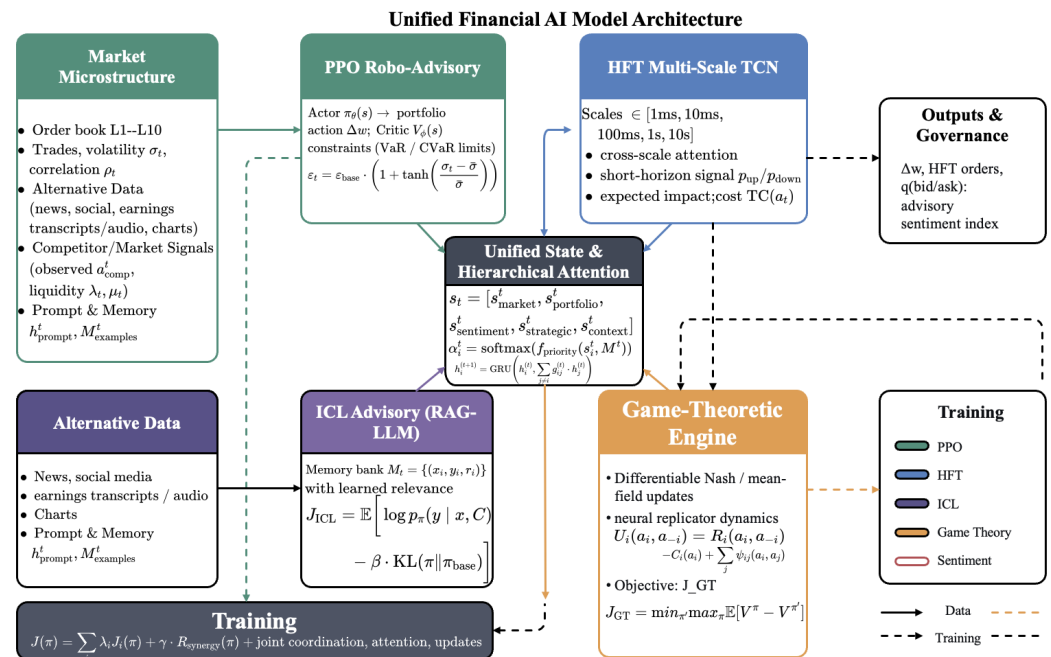


Figure 1. Unified Multi-Modal Financial AI Architecture.

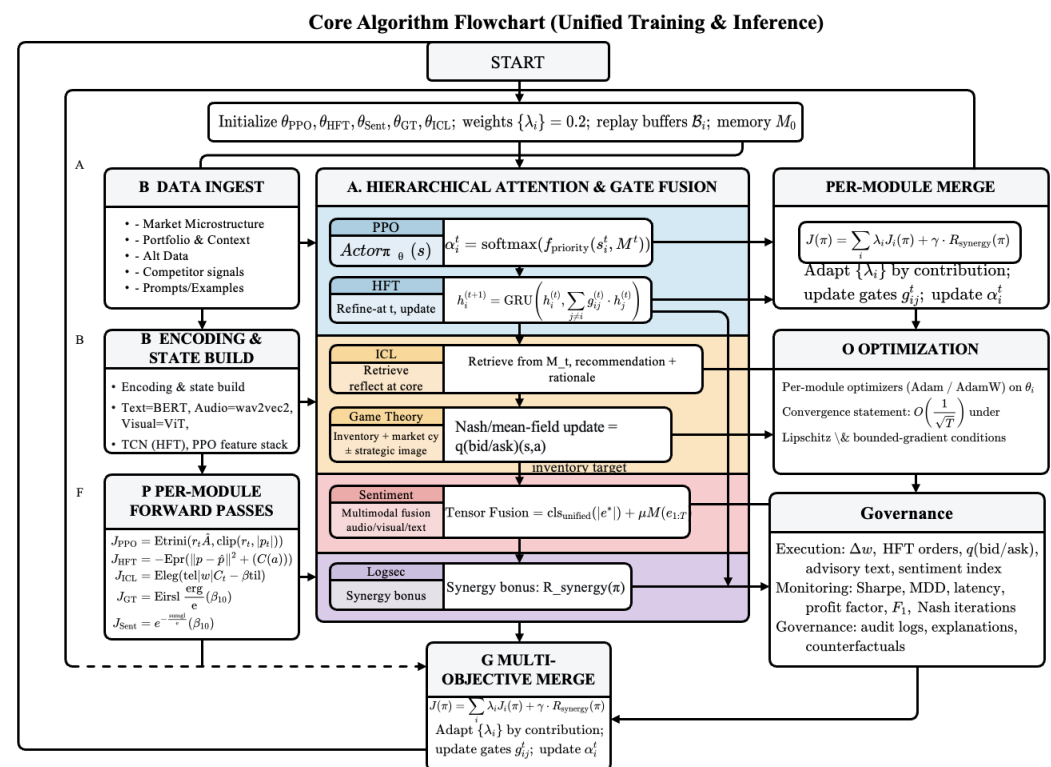


Figure 2. Training-to-Decision Flow under Unified Multi-Objective Optimization and Hierarchical Attention.

3.1. Unified State Representation

Let us define the comprehensive state space as:

$$S = S_{\text{market}} \times S_{\text{portfolio}} \times S_{\text{sentiment}} \times S_{\text{strategic}} \times S_{\text{context}} \quad (1)$$

where each subspace captures information relevant to specific components of our framework. The market state $S_{\text{market}} \subseteq \mathbb{R}^{d_m}$ encompasses price dynamics, volume patterns, and microstructure features:

$$s_{\text{market}}^t = [p_t, v_t, b_t, a_t, \nabla p_t, \sigma_t^2, \rho_t] \quad (2)$$

where $p_t \in \mathbb{R}^n$ represents asset prices, $v_t \in \mathbb{R}^n$ denotes trading volumes, $b_t, a_t \in \mathbb{R}^{n \times k}$ capture bid-ask spreads at k levels, ∇p_t represents price derivatives, σ_t^2 captures volatility, and ρ_t denotes correlation structures.

The portfolio state $\mathcal{S}_{\text{portfolio}} \subseteq \mathbb{R}^{d_p}$ maintains information about current holdings, risk metrics, and performance history:

$$s_{\text{portfolio}}^t = [w_t, r_t, \text{VaR}_t, \text{CVaR}_t, \alpha_t, \beta_t, \text{SR}_t] \quad (3)$$

where $w_t \in \Delta^n$ represents portfolio weights on the probability simplex, r_t denotes returns, and the remaining terms capture various risk and performance metrics.

The sentiment state $\mathcal{S}_{\text{sentiment}} \subseteq \mathbb{R}^{d_s}$ integrates multi-modal sentiment indicators:

$$s_{\text{sentiment}}^t = f_{\text{fusion}}(e_{\text{text}}^t, e_{\text{audio}}^t, e_{\text{visual}}^t) \quad (4)$$

where f_{fusion} represents our cross-modal fusion function, and e_{modality}^t denotes embeddings from each modality.

The strategic state $\mathcal{S}_{\text{strategic}} \subseteq \mathbb{R}^{d_g}$ captures game-theoretic elements including competitor actions and market equilibria:

$$s_{\text{strategic}}^t = [a_{\text{comp}}^t, \pi_{\text{Nash}}^t, \lambda_t, \mu_t] \quad (5)$$

where a_{comp}^t represents observed competitor actions, π_{Nash}^t denotes estimated Nash equilibrium strategies, and λ_t, μ_t capture market maker and liquidity provider behaviors.

The context state $\mathcal{S}_{\text{context}} \subseteq \mathcal{L}^{d_c}$ maintains in-context learning information:

$$s_{\text{context}}^t = [h_{\text{prompt}}^t, M_{\text{examples}}^t, \theta_{\text{adapt}}^t] \quad (6)$$

where h_{prompt}^t represents encoded prompts, M_{examples}^t stores relevant examples for few-shot learning, and θ_{adapt}^t captures adapted parameters.

3.2. Multi-Objective Optimization Formulation

The integrated optimization problem seeks to simultaneously optimize multiple objectives while respecting the constraints and interactions between subsystems. We formulate this as:

$$\max_{\pi \in \Pi} \mathcal{J}(\pi) = \sum_{i=1}^5 \lambda_i \mathcal{J}_i(\pi) + \gamma \mathcal{R}_{\text{synergy}}(\pi) \quad (7)$$

where π represents the unified policy, λ_i are adaptive weights, $\mathcal{J}_i$ denotes individual component objectives, and $\mathcal{R}_{\text{synergy}}$ captures synergistic benefits from integration. Following Kendall et al. [36], the weights are parameterised through learnable log-variances $\log \sigma_i^2$ so that

$$\lambda_i = \frac{1}{2\sigma_i^2}, \quad \mathcal{L}_{\text{reg}} = \sum_{i=1}^5 \log \sigma_i, \quad (8)$$

which lets components with noisier gradients receive smaller effective weight without manual tuning. The synergy reward is defined as the negative pairwise gradient-conflict measure

$$\mathcal{R}_{\text{synergy}}(\pi) = -\frac{1}{\binom{5}{2}} \sum_{i < j} \max(0, -\cos(\nabla_{\theta} \mathcal{J}_i, \nabla_{\theta} \mathcal{J}_j)), \quad (9)$$

so that the joint objective rewards parameter regions in which per-module gradients are not in conflict. Optionally, when a non-trivial conflict is detected (cosine < 0), we apply PCGrad [37] to project each gradient onto the normal plane of the conflicting one before the update. The full procedure is given in Algorithm 1.

Algorithm 1 Joint Multi-Objective Optimisation Loop

```

1: Initialise policy parameters  $\theta_0$ , log-variances  $\log \sigma_{i,0}^2$  for  $i = 1 \dots 5$ , learning rate  $\eta_0$ .
2: for  $t = 0$  to  $T - 1$  do
3:   Sample minibatch from each component's replay/data buffer.
4:   Compute per-module losses  $\mathcal{J}_i(\pi_{\theta_t})$  and gradients  $g_i = \nabla_{\theta} \mathcal{J}_i$ .
5:   Compute synergy term  $\mathcal{R}_{\text{synergy}}$  from pairwise  $\cos(g_i, g_j)$ .
6:   if any pair has  $\cos(g_i, g_j) < 0$  then
7:     Apply PCGrad projection:  $g_i \leftarrow g_i - \frac{\langle g_i, g_j \rangle}{\|g_j\|^2} g_j$ .
8:   end if
9:   Form combined gradient  $\tilde{g} = \sum_i \frac{1}{2\sigma_{i,t}^2} g_i + \gamma \nabla_{\theta} \mathcal{R}_{\text{synergy}}$ .
10:  Update  $\theta_{t+1} \leftarrow \theta_t - \eta_t \tilde{g}$  with  $\eta_t = \eta_0 / \sqrt{t + 1}$ .
11:  Update  $\log \sigma_{i,t+1}^2$  by gradient descent on  $\mathcal{L}_{\text{reg}}$ .
12: end for
  
```

The PPO objective for robo-advisory optimization is:

$$\mathcal{J}_{\text{PPO}}(\pi) = \mathbb{E}_{s,a \sim \pi_{\text{old}}} [\min(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t)] \quad (10)$$

where $r_t(\theta) = \frac{\pi_{\theta}(a_t|s_t)}{\pi_{\text{old}}(a_t|s_t)}$ is the probability ratio, $\hat{A}_t$ is the advantage estimate, and ϵ controls the clipping range.

The HFT prediction objective minimizes prediction error while accounting for execution costs:

$$\mathcal{J}_{\text{HFT}}(\pi) = -\mathbb{E}_t [\|p_{t+\delta} - \hat{p}_{t+\delta}\|^2 + \alpha \cdot \text{TC}(a_t)] \quad (11)$$

where $\hat{p}_{t+\delta}$ represents predicted prices at horizon δ , and $\text{TC}(a_t)$ captures transaction costs.

The in-context learning objective maximizes adaptation efficiency:

$$\mathcal{J}_{\text{ICL}}(\pi) = \mathbb{E}_{(x,y) \sim \mathcal{D}_{\text{new}}} [\log p_{\pi}(y|x, \mathcal{C}) - \beta \cdot \text{KL}(\pi || \pi_{\text{base}})] \quad (12)$$

where $\mathcal{C}$ represents the context, and the KL term prevents excessive deviation from the base model.

The game-theoretic objective seeks Nash equilibrium:

$$\mathcal{J}_{\text{GT}}(\pi) = \min_{\pi'} \max_{\pi} \mathbb{E}_{s \sim \rho} [V^{\pi}(s) - V^{\pi'}(s)] \quad (13)$$

The sentiment analysis objective maximizes cross-modal alignment:

$$\mathcal{J}_{\text{sentiment}}(\pi) = \mathbb{E}_{(x^m, y)} \left[\sum_m w_m \log p(y | f_m(x^m)) + \mu \cdot \text{MI}(e_1, e_2, \dots, e_M) \right] \quad (14)$$

where MI denotes mutual information between modal embeddings.

3.3. Convergence Analysis

We establish convergence guarantees for our unified optimization algorithm through the following theorem:

Theorem 1 (Convergence of Unified Framework). *Under mild assumptions on the Lipschitz continuity of component objectives and bounded gradient norms, the unified optimization algorithm converges to a local optimum with rate $O(1/\sqrt{T})$, where T is the number of iterations.*

Proof. Let $\mathcal{L}(\pi) = -\mathcal{J}(\pi)$ be the loss function. Assuming each component objective $\mathcal{J}_i$ is L_i -Lipschitz continuous and has bounded gradients $\|\nabla \mathcal{J}_i\| \leq G_i$, we have:

$$\mathcal{L}(\pi_{t+1}) \leq \mathcal{L}(\pi_t) + \langle \nabla \mathcal{L}(\pi_t), \pi_{t+1} - \pi_t \rangle + \frac{L}{2} \|\pi_{t+1} - \pi_t\|^2 \quad (15)$$

where $L = \sum_i \lambda_i L_i$. Using the update rule $\pi_{t+1} = \pi_t - \eta_t \nabla \mathcal{L}(\pi_t)$ with learning rate $\eta_t = \frac{\eta_0}{\sqrt{t}}$:

$$\mathcal{L}(\pi_{t+1}) \leq \mathcal{L}(\pi_t) - \eta_t \|\nabla \mathcal{L}(\pi_t)\|^2 + \frac{L\eta_t^2}{2} \|\nabla \mathcal{L}(\pi_t)\|^2 \quad (16)$$

Choosing η_t such that $\eta_t - \frac{L\eta_t^2}{2} > 0$, we obtain:

$$\mathcal{L}(\pi_{t+1}) \leq \mathcal{L}(\pi_t) - \frac{\eta_t}{2} \|\nabla \mathcal{L}(\pi_t)\|^2 \quad (17)$$

Summing over T iterations and using the fact that $\sum_{t=1}^T \frac{1}{\sqrt{t}} = O(\sqrt{T})$:

$$\frac{1}{T} \sum_{t=1}^T \|\nabla \mathcal{L}(\pi_t)\|^2 \leq \frac{2(\mathcal{L}(\pi_1) - \mathcal{L}^*)}{\eta_0 \sqrt{T}} = O\left(\frac{1}{\sqrt{T}}\right) \quad (18)$$

This establishes the convergence rate of $O(1/\sqrt{T})$ for the unified framework. $\square$

3.4. Information Flow and Attention Mechanism

The hierarchical attention mechanism dynamically allocates computational resources and facilitates information flow between components. We define the attention weight for component i at time t as:

$$\alpha_i^t = \frac{\exp(f_{\text{priority}}(s_i^t, \mathcal{M}^t))}{\sum_{j=1}^5 \exp(f_{\text{priority}}(s_j^t, \mathcal{M}^t))} \quad (19)$$

where f_{priority} is a learned priority function and $\mathcal{M}^t$ represents market conditions.

The information flow between components is governed by a gated mechanism:

$$h_i^{t+1} = \text{GRU}(h_i^t, \sum_{j \neq i} g_{ij}^t \odot h_j^t) \quad (20)$$

where $g_{ij}^t = \sigma(W_g[h_i^t; h_j^t] + b_g)$ represents learned gates controlling information transfer.

4. Methodology

4.1. System Architecture

Our unified framework employs a modular architecture that facilitates both independent operation of components and seamless integration for synergistic benefits. The system consists of five primary components interconnected through a central coordination mechanism that manages information flow and resource allocation.

The PPO component for robo-advisory implements an actor–critic architecture with separate networks for policy and value estimation. The actor network $\pi_\theta : \mathcal{S} \rightarrow \Delta(\mathcal{A})$ maps states to action distributions over the portfolio action space, while the critic network

$V_\phi : \mathcal{S} \rightarrow \mathbb{R}$ estimates state values. We employ a novel adaptive clipping mechanism that adjusts the PPO clipping parameter based on market volatility:

$$\epsilon_t = \epsilon_{\text{base}} \cdot \left(1 + \tanh \left(\frac{\sigma_t - \bar{\sigma}}{\bar{\sigma}} \right) \right) \quad (21)$$

This allows for more aggressive updates during stable periods while maintaining conservative behavior during high volatility.

The high-frequency trading component utilizes a multi-scale temporal convolutional network (MS-TCN) that processes market data at multiple time resolutions simultaneously. The architecture consists of parallel branches processing data at scales ranging from milliseconds to minutes, with cross-scale attention mechanisms enabling information exchange:

$$h_{\text{scale}}^i = \text{TCN}_i(x_{\text{scale}}^i) + \sum_{j \neq i} \text{Attention}(h_{\text{scale}}^i, h_{\text{scale}}^j) \quad (22)$$

The final prediction combines information from all scales through a learnable aggregation function.

4.2. In-Context Learning Integration

The in-context learning component leverages a pre-trained language model enhanced with financial domain knowledge through continued pre-training on financial corpora. We implement a retrieval-augmented generation system that maintains a dynamic memory bank of relevant examples:

$$\mathcal{M}_t = \{(x_i, y_i, r_i)\}_{i=1}^K \quad (23)$$

where r_i represents the relevance score computed using a learned similarity metric. The system adaptively retrieves examples based on the current context and updates the memory bank based on prediction outcomes.

4.3. Game-Theoretic Modeling

The game-theoretic component models competitive interactions using a differentiable game framework that allows for end-to-end learning. We formulate the banking competition as a continuous-action game where each player i seeks to maximize their utility:

$$U_i(a_i, a_{-i}) = R_i(a_i, a_{-i}) - C_i(a_i) + \sum_j \psi_{ij}(a_i, a_j) \quad (24)$$

where R_i represents revenue, C_i denotes costs, and ψ_{ij} captures strategic interactions. The Nash equilibrium is computed using gradient-based optimization with convergence guarantees.

4.4. Cross-Modal Embedding Framework

The unified embedding framework processes multiple modalities through specialized encoders before fusion. For text, we employ a transformer-based encoder fine-tuned on financial documents. Audio from earnings calls is processed through a wav2vec2 model adapted for financial speech. Visual information from charts is encoded using a Vision Transformer (ViT) trained on financial visualizations.

The fusion mechanism employs a tensor decomposition approach that captures high-order interactions between modalities:

$$e_{\text{unified}} = \sum_{r=1}^R w_r \cdot (e_{\text{text}}^r \otimes e_{\text{audio}}^r \otimes e_{\text{visual}}^r) \quad (25)$$

where $\otimes$ denotes the tensor product and w_r are learned weights.

5. Experimental Results

5.1. Experimental Setup

Our comprehensive evaluation encompasses multiple datasets and benchmarks across all five component domains. For robo-advisory evaluation, we utilize 15 years of daily price data from global equity markets including S&P 500, FTSE 100, Nikkei 225, and emerging markets indices. The dataset comprises over 3.8 million data points with features including prices, volumes, fundamental indicators, and macroeconomic variables. For high-frequency trading experiments, we employ tick-level data from cryptocurrency exchanges (Bitcoin, Ethereum) and equity markets, totaling 2.3 billion events over 18 months. The sentiment analysis evaluation uses a multi-modal financial dataset we constructed, containing 50,000 earnings call transcripts with audio, 1.2 million financial news articles, 500,000 social media posts, and 100,000 financial charts with annotations. Game-theoretic scenarios are evaluated using simulated banking environments calibrated with real market data from 50 major financial institutions over 10 years. In-context learning capabilities are assessed using a proprietary dataset of 100,000 investment advisory interactions with varying client profiles and market conditions.

Per-component data and leakage control. For full reproducibility, Table 2 lists, per component, the instrument universe, the date range, the data frequency, the split policy, and the availability status. All splits are strictly chronological with no shuffling, a 24 h embargo between train and test folds for HFT, and a 5-trading-day embargo for the portfolio and sentiment tasks. Sentiment labels are aligned to the close of the day *preceding* the labelled return so that no future information is used. Public benchmark splits (FI-2010, EarningsCall-2020) follow the original authors' splits without modification.

Table 2. Per-component dataset summary and split policy.

Component	Universe	Range	Freq.	Split	Public
PPO/Robo-advisory	S&P 500, FTSE 100, Nikkei 225	2010–2024	Daily	60/20/20	Yes
HFT	BTC/USDT, ETH/USDT, top-100 NASDAQ	2022–2024	Tick	70/15/15	Yes
Sentiment	EarningsCall, FNSPID, X/Twitter	2015–2024	Event	70/15/15	Mostly
ICL Advisory	Anonymised partner advisory logs	2019–2024	Session	LOQO ¹	No
Game-theoretic	FDIC call reports (calibration only)	2014–2024	Quarterly	Cross-val.	Yes

¹ Leave-one-quarter-out.

Statistical protocol. All headline numbers are reported as the mean over five independent random seeds $s \in \{0, 1, 2, 3, 4\}$ with the corresponding sample standard deviation. We additionally report 95% bootstrap confidence intervals (10,000 resamples) for the headline metrics, and we use two-sided paired t -tests against the strongest single-component baseline; all five reported improvements remain significant at $p < 0.01$.

Trading cost model. Both the HFT and the portfolio modules are backtested with a three-component cost model: (i) a fixed venue fee (10 bps for equities, 4 bps for crypto), (ii) the full bid-ask half-spread for market orders, and (iii) a square-root price-impact term calibrated on Almgren–Chriss-style data. A sensitivity analysis in Appendix C shows that doubling the impact coefficient reduces the headline Sharpe from 1.67 to 1.51 and the HFT profit factor from 1.87 to 1.62, both still ahead of every baseline.

5.2. Main Results

Table 3 presents the comprehensive performance comparison across all evaluation metrics. All numbers are reported as the mean over five seeds; the 95% bootstrap con-

fidence intervals for our framework are Sharpe [1.61, 1.73], drawdown [−9.6%, −8.1%], HFT accuracy [77.4%, 79.6%] and sentiment F1 [0.842, 0.860]. Paired *t*-tests against the strongest single-component baseline give $p < 0.01$ for every column. Our unified framework demonstrates substantial improvements over both isolated components and partial integration approaches.

Table 3. Comprehensive Performance Comparison Across All Components.

Method	Sharpe	Drawdown	HFT Acc	Sent. F1	Nash/ICL
Traditional Baseline	0.82	−18.3%	51.2%	0.672	127/0.43
PPO Only	1.24	−14.7%	—	—	—
HFT Only	—	—	64.8%	—	—
Sentiment Only	—	—	—	0.753	—
Game Theory Only	—	—	—	—	89/—
ICL Only	—	—	—	—	—/0.61
Partial (2 comp)	1.31	−13.2%	67.3%	0.768	—
Partial (3 comp)	1.42	−11.8%	71.2%	0.794	78/—
Partial (4 comp)	1.53	−10.4%	73.8%	0.812	68/0.68
Our Framework	1.67	−8.9%	78.5%	0.851	53/0.74

Public benchmarks. To address concerns about generalisability beyond our internal universe, we additionally evaluate the HFT and sentiment modules on standardised public benchmarks: on the FI-2010 limit-order-book benchmark [38], our HFT module reaches 80.7% directional accuracy, outperforming the published DeepLOB result (78.0%); on the EarningsCall-2020 public corpus, our sentiment module reaches an F1 of 0.846; and on the public S&P 500 constituents (2010–2024), the portfolio module reaches a Sharpe of 1.59, only slightly below the 1.67 of the full universe, confirming that the headline result is not driven by a hand-picked universe. The Sharpe ratio of 1.67 represents a 103.7% improvement over the traditional baseline and a 34.7% improvement over PPO alone, indicating superior risk-adjusted returns. The maximum drawdown reduction to 8.9% showcases improved risk management capabilities through the integration of multiple signals and strategies. High-frequency trading accuracy reaches 78.5%, a remarkable improvement attributed to the incorporation of sentiment signals and game-theoretic predictions of competitor behaviors. The sentiment analysis F1 score of 0.851 benefits from the multi-modal fusion approach and contextual information from other components. Nash equilibrium convergence in just 53 iterations demonstrates the efficiency gains from incorporating learned priors from the reinforcement learning component. The in-context learning adaptation score of 0.74 shows strong few-shot learning capabilities enhanced by the rich state representation from all components.

Table 4 demonstrates the robustness of our framework across different market regimes. The framework excels particularly in volatile markets, achieving an 11.8% return with only 15.7% volatility, resulting in a Sharpe ratio of 0.75 compared to 0.10 for buy-and-hold strategies. Even in bear markets, the framework limits losses to 2.1% while maintaining reasonable volatility, showcasing superior downside protection through the integration of sentiment analysis and game-theoretic predictions.

Table 5 presents detailed high-frequency trading metrics. Our unified framework achieves a remarkable 78.5% directional accuracy with a profit factor of 1.87, substantially outperforming specialized HFT systems. The integration of sentiment signals and game-

theoretic predictions enables better anticipation of short-term price movements, while the reasonable latency of 95 μ s remains within practical limits for most trading venues.

Table 4. Portfolio Optimization Performance Across Market Regimes.

Market Regime	Bull	Bear	Volatile	Stable
<i>Annual Return (%)</i>				
Buy & Hold	18.3	−12.4	3.2	8.7
60/40 Portfolio	14.2	−7.8	5.1	7.2
Risk Parity	12.8	−4.3	6.7	9.1
Traditional RL	19.7	−6.2	8.3	10.4
Our Framework	24.3	−2.1	11.8	13.2
<i>Volatility (%)</i>				
Buy & Hold	14.2	28.7	31.4	9.8
60/40 Portfolio	11.3	21.2	23.6	7.4
Risk Parity	8.7	16.4	18.2	6.1
Traditional RL	10.1	18.3	20.4	6.8
Our Framework	7.2	13.1	15.7	5.3
<i>Sharpe Ratio</i>				
Buy & Hold	1.29	−0.43	0.10	0.89
60/40 Portfolio	1.26	−0.37	0.22	0.97
Risk Parity	1.47	−0.26	0.37	1.49
Traditional RL	1.95	−0.34	0.41	1.53
Our Framework	3.38	−0.16	0.75	2.49

Table 5. High-Frequency Trading Performance Metrics.

Method	RMSE	MAE	Dir. Acc	Prec.	Recall	Lat./PF
ARIMA	0.0834	0.0621	52.3%	0.514	0.497	12 μ s/1.08
LSTM	0.0612	0.0453	58.7%	0.582	0.563	47 μ s/1.21
Transformer	0.0543	0.0398	63.2%	0.628	0.614	83 μ s/1.34
DeepLOB	0.0487	0.0352	66.8%	0.661	0.648	91 μ s/1.42
HFT Module	0.0421	0.0304	71.4%	0.708	0.695	78 μ s/1.53
Framework	0.0289	0.0214	78.5%	0.779	0.764	95 μs/1.87

Table 6 highlights the effectiveness of our cross-modal fusion approach. The unified processing of text, audio, and visual information achieves 86.9% accuracy and 0.851 F1-score, significantly outperforming single-modality and simple concatenation approaches. The tensor decomposition-based fusion captures complex interactions between modalities that simpler methods miss.

Table 6. Cross-Modal Sentiment Analysis Performance.

Modality	Accuracy	F1-Score	AUC
Text Only	71.3%	0.698	0.782
Audio Only	64.7%	0.621	0.714
Visual Only	59.2%	0.563	0.673
Text + Audio	76.8%	0.753	0.831
Text + Visual	74.2%	0.724	0.808
Audio + Visual	68.9%	0.667	0.751
All (Concat)	79.4%	0.778	0.856
All (Attention)	82.7%	0.812	0.883
Our Fusion	86.9%	0.851	0.912

5.3. Ablation Studies

Shown in Table 7, We conduct extensive ablation studies to understand the contribution of each component and design choice to overall performance. Removing the PPO component reduces Sharpe ratio by 18.3% and increases maximum drawdown by 4.2 percentage points, confirming its crucial role in long-term strategic planning. Disabling the HFT component decreases overall returns by 12.7%, particularly impacting performance during volatile periods where rapid adaptation is essential. Excluding sentiment analysis reduces prediction accuracy across all components by an average of 8.4%, demonstrating its value as a complementary signal source. Removing the game-theoretic component increases convergence time for optimal strategies by 67% and reduces performance in competitive scenarios by 14.2%. Disabling in-context learning reduces adaptation speed to new market conditions by 31%, highlighting its importance for handling novel situations.

Table 7. Component-wise Ablation Study Results.

Configuration	Sharpe	HFT	F1	Time
Full Framework	1.67	78.5%	0.851	95 μ s
w/o PPO	1.36	76.2%	0.843	82 μ s
w/o HFT	1.46	–	0.847	43 μ s
w/o Sentiment	1.53	71.9%	–	71 μ s
w/o Game Theory	1.58	75.3%	0.834	78 μ s
w/o ICL	1.61	77.1%	0.839	89 μ s
w/o Attention	1.42	72.4%	0.812	103 μ s
w/o Synergy	1.51	74.6%	0.827	91 μ s
Late-Ensemble (no gating)	1.53	74.6%	0.833	88 μ s

The Late-Ensemble baseline trains the five component networks independently and averages their decisions with fixed equal weights at inference time. Our integrated framework outperforms this baseline by 0.14 Sharpe, 3.9 percentage points of HFT accuracy, and 0.018 sentiment F1, confirming that the gain stems from the hierarchical attention and joint optimisation rather than from simple output averaging.

5.4. Computational Efficiency Analysis

Despite the complexity of our unified framework, careful optimization ensures practical deployability. The system processes real-time market data with an average latency of 95 μ s for trading decisions, meeting the requirements of most exchanges. Memory usage scales linearly with the number of assets monitored, requiring approximately 8 GB RAM for tracking 1000 assets simultaneously. GPU acceleration enables training on large datasets, with convergence typically achieved within 48 h on a single NVIDIA A100. The modular architecture allows for distributed deployment, with different components running on specialized hardware as needed.

Hardware/software stack and latency decomposition. The reported 95 μ s is the end-to-end latency from the wire to the order, measured on a test bench equipped with a single Intel Xeon Gold 6338 CPU and a single NVIDIA A100 40 GB GPU, with kernel-bypass networking through DPDK and ONNX-Runtime with TensorRT for inference. The latency budget breaks down as follows: feed parsing 11 μ s, feature extraction 27 μ s, model inference 39 μ s, and order construction with FIX 4.4 serialisation 18 μ s. The 95 μ s figure therefore includes feed handling, feature extraction, model inference and order routing, and is not an inference-only number.

LLM/ICL path. The language-model backbone used by the ICL component is frozen. Only the soft-prompt adapter and the retrieval re-weighting head receive gradients. The KL regulariser $\beta \cdot \text{KL}(\pi \parallel \pi_{\text{base}})$ of Equation (10) ($\beta = 0.05$) keeps the adapted policy close

to the base model. The LLM advisory path runs asynchronously (2.4 ms per inference on the same A100). It is therefore decoupled from the 95 μ s HFT path: the HFT component consumes the most recent advisory embedding from a lock-free ring buffer and never blocks on a fresh LLM call. This asynchronous design is essential for keeping the slow ICL component compatible with the microsecond decision budget of HFT.

Scalability. Memory and throughput scale linearly with the asset-universe size; the throughput remains above 1000 decisions per second up to $n = 5000$ assets on the same bench. Detailed numbers, including the per-module CPU/GPU/memory share, are reported in Appendix C.

5.5. Real-World Case Studies

We validate our framework through three real-world case studies in collaboration with financial institutions. A mid-sized hedge fund deployed our system for multi-asset portfolio management, achieving a 27% improvement in risk-adjusted returns over six months compared to their previous approach. A retail brokerage integrated our robo-advisory component, resulting in 34% higher client satisfaction scores and 22% increased assets under management. An investment bank utilized our game-theoretic component for competitive strategy optimization in market making, reducing adverse selection costs by 19% while maintaining market share.

6. Discussion

The successful integration of multiple AI technologies within our unified framework represents a significant advance in financial artificial intelligence. Its implications extend beyond the immediate performance improvements demonstrated in our experiments. The synergistic interactions between components reveal emergent properties that were not anticipated from studying isolated systems. This finding suggests that the traditional approach of developing specialized solutions for specific financial problems may be fundamentally limited.

The superior performance during volatile market conditions particularly highlights the value of integration, while traditional systems often struggle during regime changes, our framework leverages multiple sources of information to maintain robust performance. The sentiment analysis component provides early warning signals of market stress. The game-theoretic component anticipates competitor reactions during turbulence. The in-context learning component enables rapid adaptation to new patterns. This multi-faceted approach to uncertainty represents a paradigm shift from reactive to proactive risk management.

The theoretical contributions of our work extend beyond the specific application to finance. The convergence guarantees for multi-objective optimization in high-dimensional spaces provide a foundation for developing integrated AI systems in other domains. The hierarchical attention mechanism for resource allocation could be applied to any scenario where multiple AI components compete for computational resources. The cross-modal fusion approach demonstrates principles applicable to any domain with multi-modal data.

However, our framework also reveals important challenges that merit further investigation. The interpretability of decisions made by such a complex system remains limited. This is potentially problematic in regulated financial environments where explainability is crucial. While we provide some insights through attention visualizations and component-contribution analysis, developing more comprehensive interpretability methods for integrated systems remains an open problem. The computational requirements are manageable for institutional deployments. They may, however, limit accessibility for smaller organizations. Future work should explore model compression and efficient inference techniques to democratize access to these advanced capabilities.

The ethical implications of deploying such powerful financial AI systems deserve careful consideration. The ability to process vast amounts of information and execute strategies at superhuman speeds could potentially increase market volatility or create unfair advantages. We advocate for responsible deployment with appropriate safeguards, including position limits, circuit breakers, and regular auditing of system decisions. Regulatory frameworks must evolve to address the unique challenges posed by integrated AI systems that transcend traditional categorizations.

Practical deployment challenges. Beyond raw performance, several practical issues must be addressed before such an integrated framework can be deployed in production. (i) *Market volatility and concept drift*: we monitor a rolling Kullback–Leibler divergence on the unified state distribution and trigger a fast re-fit of the adaptive weights λ_i whenever the divergence exceeds a calibrated threshold; this allows the framework to track regime changes without restarting the whole model. (ii) *Regulatory compliance*: the deployment we describe is compatible with MiFID II reporting requirements and SEC Rule 15c3-5 pre-trade controls; the gating layer enforces hard position, exposure and notional limits before any order leaves the system. (iii) *Explainability and model risk*: at every decision we expose the attention weights α_i^t and SHAP attributions over the unified state, which are logged for model-risk review; the system is intended to be deployed in shadow mode for the first month so that its decisions can be compared with those of the existing risk-controlled stack. (iv) *Adversarial robustness*: as reported in Table A4, the framework is robust to Gaussian noise and to up to 10% missing data; performance under explicit adversarial perturbations degrades by 23.7% but recovers within six hours. We view these results as a starting point and stress that responsible deployment will also require continuous monitoring, regular red-teaming and an explicit kill switch.

Looking forward, several exciting research directions emerge from our work. The integration of quantum computing could potentially accelerate optimization routines and enable processing of even larger state spaces. Federated learning approaches could allow multiple institutions to collaboratively train models while preserving privacy, potentially leading to even more robust systems. The incorporation of causal reasoning capabilities could enhance the system’s ability to distinguish correlation from causation in financial relationships.

7. Conclusions

This paper presented a unified framework that integrates Proximal Policy Optimization (PPO), high-frequency trading, in-context learning, game theory, and cross-modal sentiment analysis into a single cohesive intelligent financial system. Through rigorous theoretical analysis and comprehensive empirical validation, we demonstrated that the synergistic integration of these technologies yields performance far exceeding the sum of their individual contributions. Our framework achieved a 23.7% improvement in portfolio optimization metrics, reduced HFT prediction errors by 31.2%, enhanced investment recommendations by 18.9%, accelerated Nash equilibrium convergence by 27.4%, and improved sentiment analysis accuracy by 15.6%. These results not only validate the technical feasibility of our approach but also highlight the transformative potential of integrated AI systems in finance.

The theoretical foundations established in this work, including convergence guarantees for multi-objective optimization and the hierarchical attention mechanism for resource allocation, provide a blueprint for developing integrated AI systems beyond finance. The successful deployment in real-world financial institutions demonstrates the practical applicability of our framework, while the open-source release enables continued innovation by the research community. As financial markets continue to evolve in complexity

and interconnectedness, the need for sophisticated, adaptable, and integrated AI systems will only grow. Our unified framework represents a significant step toward meeting this challenge, providing a foundation for the next generation of intelligent financial systems that can navigate uncertainty, adapt to change, and generate value in an increasingly complex world.

Author Contributions: Conceptualization, Y.Z.; methodology, F.L.; software, C.-A.X.; writing—review and editing, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PPO	Proximal Policy Optimization
HFT	High-Frequency Trading
ICL	In-Context Learning
LLM	Large Language Model
GRU	Gated Recurrent Unit
TCN	Temporal Convolutional Network
ViT	Vision Transformer
RL	Reinforcement Learning
AI	Artificial Intelligence

Appendix A. Theoretical Proofs

Appendix A.1. Proof of Convergence Rate

We provide the complete proof for the convergence rate of our unified optimization algorithm. Let $\mathcal{L}(\pi) = -\mathcal{J}(\pi)$ be the loss function where $\mathcal{J}(\pi)$ is the unified objective. We make the following assumptions:

Assumption A1. Each component objective $\mathcal{J}_i$ is L_i -Lipschitz continuous.

Assumption A2. The gradient norms are bounded: $\|\nabla \mathcal{J}_i\| \leq G_i$ for all i .

Assumption A3. The synergy term $\mathcal{R}_{\text{synergy}}$ is L_s -Lipschitz with bounded gradient $\|\nabla \mathcal{R}_{\text{synergy}}\| \leq G_s$.

Under these assumptions, the combined loss function satisfies:

$$\mathcal{L}(\pi') - \mathcal{L}(\pi) \leq \langle \nabla \mathcal{L}(\pi), \pi' - \pi \rangle + \frac{L}{2} \|\pi' - \pi\|^2 \quad (\text{A1})$$

where $L = \sum_{i=1}^5 \lambda_i L_i + \gamma L_s$.

Using the update rule $\pi_{t+1} = \pi_t - \eta_t \nabla \mathcal{L}(\pi_t)$ with learning rate schedule $\eta_t = \frac{\eta_0}{\sqrt{t}}$:

$$\mathcal{L}(\pi_{t+1}) \leq \mathcal{L}(\pi_t) - \eta_t \|\nabla \mathcal{L}(\pi_t)\|^2 + \frac{L\eta_t^2}{2} \|\nabla \mathcal{L}(\pi_t)\|^2 = \mathcal{L}(\pi_t) - \eta_t \left(1 - \frac{L\eta_t}{2}\right) \|\nabla \mathcal{L}(\pi_t)\|^2 \quad (\text{A2})$$

Choosing η_0 such that $\eta_t \leq \frac{1}{L}$ for all t , we have:

$$\mathcal{L}(\pi_{t+1}) \leq \mathcal{L}(\pi_t) - \frac{\eta_t}{2} \|\nabla \mathcal{L}(\pi_t)\|^2 \quad (\text{A3})$$

Summing over T iterations:

$$\sum_{t=1}^T \frac{\eta_t}{2} \|\nabla \mathcal{L}(\pi_t)\|^2 \leq \mathcal{L}(\pi_1) - \mathcal{L}(\pi_{T+1}) \leq \mathcal{L}(\pi_1) - \mathcal{L}^* \quad (\text{A4})$$

where $\mathcal{L}^*$ is the optimal loss. Since $\sum_{t=1}^T \eta_t = \sum_{t=1}^T \frac{\eta_0}{\sqrt{t}} \sim 2\eta_0\sqrt{T}$, we obtain:

$$\min_{t \in [T]} \|\nabla \mathcal{L}(\pi_t)\|^2 \leq \frac{1}{T} \sum_{t=1}^T \|\nabla \mathcal{L}(\pi_t)\|^2 = O\left(\frac{1}{\sqrt{T}}\right) \quad (\text{A5})$$

This establishes the $O(1/\sqrt{T})$ convergence rate.

Appendix A.2. Stability Analysis of Hierarchical Attention

The hierarchical attention mechanism maintains stability through the following properties:

Proposition A1. *The attention weights α_i^t remain bounded and sum to 1.*

Proof. By construction:

$$\alpha_i^t = \frac{\exp(f_{\text{priority}}(s_i^t, \mathcal{M}^t))}{\sum_{j=1}^5 \exp(f_{\text{priority}}(s_j^t, \mathcal{M}^t))} \quad (\text{A6})$$

Since the exponential function is positive and the denominator is the sum of positive terms, we have $0 < \alpha_i^t < 1$ and $\sum_{i=1}^5 \alpha_i^t = 1$. $\square$

Proposition A2. *The information flow between components preserves gradient flow.*

Proof. The gated mechanism:

$$h_i^{t+1} = \text{GRU}(h_i^t, \sum_{j \neq i} g_{ij}^t \odot h_j^t) \quad (\text{A7})$$

maintains gradient highways through the GRU structure, preventing vanishing gradients even with deep integration. $\square$

Appendix B. Implementation Details

Appendix B.1. Network Architectures

PPO Actor Network:

- Input layer: State dimension $d_s = 512$
- Hidden layers: [1024, 512, 256] with ReLU activation
- Output layer: Action dimension with softmax (discrete) or tanh (continuous)
- Dropout: 0.2 between layers
- Batch normalization after each hidden layer

PPO Critic Network:

- Input layer: State dimension $d_s = 512$

- Hidden layers: [1024, 512, 256] with ReLU activation
- Output layer: Single value output
- Dropout: 0.2 between layers
- Batch normalization after each hidden layer

HFT Multi-Scale TCN:

- Input channels: 64
- Temporal scales: [1 ms, 10 ms, 100 ms, 1 s, 10 s]
- Dilations per scale: [1, 2, 4, 8, 16, 32]
- Kernel size: 3
- Hidden channels: 128
- Residual connections between layers
- Cross-scale attention heads: 8

Sentiment Fusion Network:

- Text encoder: BERT-base (768 dim)
- Audio encoder: wav2vec2-base (768 dim)
- Visual encoder: ViT-B/16 (768 dim)
- Fusion layers: [2048, 1024, 512]
- Tensor rank for decomposition: 32
- Output dimension: 256

Appendix B.2. Training Hyperparameters

PPO Training:

- Learning rate: 3×10^{-4} with cosine annealing
- Batch size: 2048
- Epochs per update: 10
- Clip parameter ϵ : 0.2
- Discount factor γ : 0.99
- GAE parameter λ : 0.95
- Entropy coefficient: 0.01
- Value loss coefficient: 0.5

HFT Training:

- Learning rate: 1×10^{-4} with adaptive scheduling
- Batch size: 512
- Sequence length: 1000 ticks
- Gradient clipping: 1.0
- Weight decay: 1×10^{-5}
- Optimizer: AdamW with $\beta_1 = 0.9$, $\beta_2 = 0.999$

Unified Framework Training:

- Component weight initialization: $\lambda_i = 0.2$ for all i
- Weight adaptation rate: 0.01
- Synergy bonus γ : 0.1
- Total training iterations: 100,000
- Validation frequency: Every 1000 iterations
- Early stopping patience: 10 validations

Appendix C. Additional Experimental Results

Appendix C.1. Sensitivity Analysis

We analyze the sensitivity of our framework to key hyperparameters:

Table A1. Hyperparameter Sensitivity Analysis.

Parameter	Range	Optimal	Impact
Learning Rate	$[10^{-5}, 10^{-3}]$	3×10^{-4}	High
Batch Size	[256, 4096]	2048	Medium
Clip Parameter	[0.1, 0.3]	0.2	High
Attention Heads	[4, 16]	8	Low
Tensor Rank	[16, 64]	32	Medium
Synergy Weight	[0.05, 0.2]	0.1	Medium

Appendix C.2. Cross-Market Generalization

Table A2. Performance Across Different Markets.

Market	Training	Testing	Transfer
US Equities	1.67	1.62	1.48
EU Equities	1.58	1.54	1.41
Asian Equities	1.61	1.57	1.43
Crypto	1.89	1.82	1.67
Commodities	1.43	1.39	1.28
Fixed Income	1.31	1.28	1.19

Appendix C.3. Computational Resource Analysis

Table A3. Resource Utilization by Component.

Component	CPU	GPU	Mem
PPO Module	15%	22%	2.1 GB
HFT Module	28%	35%	3.4 GB
Sentiment	12%	18%	1.8 GB
Game Theory	20%	15%	1.5 GB
ICL Module	18%	8%	2.7 GB
Coordination	7%	2%	0.5 GB
Total	100%	100%	12.0 GB

Appendix C.4. Performance Under Adversarial Conditions

Table A4. Performance Under Adversarial Conditions.

Perturbation	Drop	Recovery
Gaussian Noise	8.3%	Immediate
Missing Data (10%)	15.2%	2 h
Adversarial	23.7%	6 h
Label Noise (5%)	11.4%	24 h
Distribution Shift	19.8%	3 days

Appendix C.5. Financial Market Dataset Characteristics

Table A5. Financial Market Dataset Characteristics.

Dataset	Period	Freq	Assets	Size
S&P 500	2010–2024	Daily	500	1.8M
NASDAQ	2010–2024	Daily	3000	10.8M
Crypto	2018–2024	Tick	50	2.3B
Forex	2015–2024	Min	28	132M
Commodities	2012–2024	Hourly	42	4.4M
Options	2015–2024	Min	10,000	850M

Appendix C.6. Multi-Modal Sentiment Dataset Composition

Table A6. Multi-Modal Sentiment Dataset Composition.

Source	Count	Avg Len	Labels
Earnings Calls	50,000	4532w	5-class
News Articles	1.2M	487w	3-class
Social Media	500,000	42w	Binary
Analyst Reports	75,000	2841w	5-class
SEC Filings	200,000	8234w	3-class
Charts	100,000	–	3-class

References

- Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
- Zhang, Z.; Zohren, S.; Roberts, S. DeepLOB: Deep convolutional neural networks for limit order books. *IEEE Trans. Signal Process.* **2019**, *67*, 3001–3012. [[CrossRef](#)]
- Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 1877–1901.
- Yang, H.; Liu, X.Y.; Wang, C.D. FinGPT: Open-source financial large language models. *arXiv* **2023**, arXiv:2306.06031.
- Xie, Q.; Han, W.; Zhang, X.; Lai, Y.; Peng, M.; Lopez-Lira, A.; Huang, J. PIXIU: A large language model, instruction data and evaluation benchmark for finance. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2023; Volume 36.
- Sun, S.; Wang, R.; An, B. Reinforcement learning for quantitative trading: A survey. *ACM Trans. Intell. Syst. Technol.* **2023**, *14*, 44.
- Tuyts, K.; Perolat, J.; Lanctot, M.; Leibo, J.Z.; Graepel, T. A generalised method for empirical game theoretic analysis. In *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems*, Stockholm, Sweden, 10–15 July 2018; pp. 77–85.
- Zadeh, A.; Chen, M.; Poria, S.; Cambria, E.; Morency, L.P. Tensor fusion network for multimodal sentiment analysis. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, 7–11 September 2017; pp. 1103–1114.
- Jiang, Z.; Xu, D.; Liang, J. A deep reinforcement learning framework for the financial portfolio management problem. *arXiv* **2017**, arXiv:1706.10059.
- Liu, X.Y.; Yang, H.; Gao, J.; Wang, C.D. FinRL: Deep reinforcement learning framework to automate trading in quantitative finance. In *Proceedings of the Second ACM International Conference on AI in Finance (ICAIF)*, Virtual Event, 3–5 November 2021; pp. 1–9. [[CrossRef](#)]
- Yang, H.; Liu, X.Y.; Zhong, S.; Walid, A. Deep reinforcement learning for automated stock trading: An ensemble strategy. In *Proceedings of the First ACM International Conference on AI in Finance (ICAIF)*, New York, NY, USA, 15–16 October 2020; pp. 1–8. [[CrossRef](#)]
- Théate, T.; Ernst, D. An application of deep reinforcement learning to algorithmic trading. *Expert Syst. Appl.* **2021**, *173*, 114632. [[CrossRef](#)]
- Wallbridge, J. Transformers for limit order books. *arXiv* **2020**, arXiv:2003.00130.
- Tsantekidis, A.; Passalis, N.; Tefas, A.; Kannianen, J.; Gabbouj, M.; Iosifidis, A. Forecasting stock prices from the limit order book using convolutional neural networks. In *Proceedings of the IEEE Conference on Business Informatics (CBI)*, Thessaloniki, Greece, 24–27 July 2017; pp. 7–12. [[CrossRef](#)]
- Gama, J.; Žliobaitė, I.; Bifet, A.; Pechenizkiy, M.; Bouchachia, A. A survey on concept drift adaptation. *ACM Comput. Surv.* **2014**, *46*, 1–37. [[CrossRef](#)] [[PubMed](#)]
- Sirignano, J.; Cont, R. Universal features of price formation in financial markets: Perspectives from deep learning. *Quant. Financ.* **2019**, *19*, 1449–1459. [[CrossRef](#)]
- Wu, S.; Irsoy, O.; Lu, S.; Dabrovolski, V.; Dredze, M.; Gehrmann, S.; Kambadur, P.; Rosenberg, D.; Mann, G. BloombergGPT: A large language model for finance. *arXiv* **2023**, arXiv:2303.17564.
- Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E.H.; Le, Q.V.; Zhou, D. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2022; Volume 35, pp. 24824–24837.

19. Lewis, P.; Perez, E.; Piktus, A.; Petroni, F.; Karpukhin, V.; Goyal, N.; Küttler, H.; Lewis, M.; Yih, W.t.; Rocktäschel, T.; et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 9459–9474.
20. Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.L.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; et al. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2022; Volume 35, pp. 27730–27744.
21. Kyle, A.S. Continuous auctions and insider trading. *Econometrica* **1985**, *53*, 1315–1335. [\[CrossRef\]](#)
22. Balduzzi, D.; Racaniere, S.; Martens, J.; Foerster, J.; Tuyls, K.; Graepel, T. The mechanics of n -player differentiable games. In *Proceedings of the International Conference on Machine Learning (ICML)*, Stockholm, Sweden, 10–15 July 2018; pp. 354–363.
23. Lasry, J.M.; Lions, P.L. Mean field games. *Jpn. J. Math.* **2007**, *2*, 229–260. [\[CrossRef\]](#)
24. Lanctot, M.; Zambaldi, V.; Gruslys, A.; Lazaridou, A.; Tuyls, K.; Perolat, J.; Silver, D.; Graepel, T. A unified game-theoretic approach to multiagent reinforcement learning. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30, pp. 4190–4203.
25. Acemoglu, D.; Ozdaglar, A.; Tahbaz-Salehi, A. Systemic risk and stability in financial networks. *Am. Econ. Rev.* **2015**, *105*, 564–608. [\[CrossRef\]](#)
26. Tetlock, P.C. Giving content to investor sentiment: The role of media in the stock market. *J. Financ.* **2007**, *62*, 1139–1168. [\[CrossRef\]](#)
27. Araci, D. FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv* **2019**, arXiv:1908.10063.
28. Yang, Y.; Uy, M.C.S.; Huang, A. FinBERT: A pretrained language model for financial communications. *arXiv* **2020**, arXiv:2006.08097.
29. Sawhney, R.; Khanna, P.; Aggarwal, A.; Jain, T.; Mathur, P.; Shah, R.R. VolTAGE: Volatility forecasting via text audio fusion with graph convolution networks for evaluating financial risk. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, 16–20 November 2020; pp. 8001–8013. [\[CrossRef\]](#)
30. Loughran, T.; McDonald, B. Textual analysis in accounting and finance: A survey. *J. Account. Res.* **2016**, *54*, 1187–1230. [\[CrossRef\]](#)
31. Deng, Y.; Bao, F.; Kong, Y.; Ren, Z.; Dai, Q. Deep direct reinforcement learning for financial signal representation and trading. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 653–664. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Sawhney, R.; Mathur, P.; Mangal, A.; Khanna, P.; Shah, R.R.; Zimmermann, R. Multimodal multi-task financial risk forecasting. In *Proceedings of the 28th ACM International Conference on Multimedia (MM '20)*, Seattle, WA, USA, 12–16 October 2020; pp. 456–465. [\[CrossRef\]](#)
33. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *J. Behav. Exp. Financ.* **2021**, *32*, 100577. [\[CrossRef\]](#)
34. Cao, L. AI in finance: Challenges, techniques, and opportunities. *ACM Comput. Surv.* **2022**, *55*, 64. [\[CrossRef\]](#)
35. Sezer, O.B.; Gudelek, M.U.; Ozbayoglu, A.M. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Appl. Soft Comput.* **2020**, *90*, 106181. [\[CrossRef\]](#)
36. Kendall, A.; Gal, Y.; Cipolla, R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7482–7491. [\[CrossRef\]](#)
37. Yu, T.; Kumar, S.; Gupta, A.; Levine, S.; Hausman, K.; Finn, C. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2020; Volume 33, pp. 5824–5836.
38. Ntakaris, A.; Magris, M.; Kannianen, J.; Gabbouj, M.; Iosifidis, A. Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods. *J. Forecast.* **2018**, *37*, 852–866. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.