

Transforming Wealth Advisory: Deep Learning, Reinforcement Optimization, and Federated Intelligence at Scale

Srinivasa Rao Gunda

Independent Researcher, USA

Abstract

Static portfolio construction frameworks have long constrained the financial advisory industry by assuming stable markets, normally distributed returns, and rational investor behavior—conditions that rarely hold in practice. Mr. Gunda designed and architected a first-of-its-kind AI-native wealth management model built on seven core components: transformer-based regime detection, Bayesian risk modeling, reinforcement learning for continuous portfolio optimization, behavioral investor segmentation, federated privacy-preserving distributed intelligence, LangGraph-based multi-agent orchestration, and Shapley-value explainability. Unlike conventional systems that rebalance portfolios at fixed intervals, this framework learns and adapts allocation policies continuously as markets and client conditions evolve. Compliance constraints are built directly into the optimization process, removing the gap between portfolio construction and regulatory oversight. The result is an advisory system capable of delivering personalized, transparent, and scalable wealth management across institutional and cross-jurisdictional environments, while meeting the evidentiary standards required for demonstrating original contributions of major significance to intelligent financial systems.

Keywords: Federated Optimization, Reinforcement Learning, Behavioral Segmentation, Explainable Artificial Intelligence, Wealth Management Architecture

1. Introduction

The financial advisory industry is undergoing fundamental transformation. Traditional portfolio construction methods, established by Markowitz in 1952, relied on static mathematical models that assumed normally distributed returns, stable covariance structures, and rational investor behavior—assumptions that systematically fail under real-world conditions [1]. Contemporary research on high-dimensional portfolio selection demonstrates that when portfolio dimension becomes comparable to sample size, the traditional sample covariance matrix no longer consistently estimates the true population covariance, causing portfolio weights to deviate considerably from optimal allocations [1]. This estimation challenge intensifies dramatically in high-dimensional settings where the number of assets, p , exceeds or approaches the number of observations, n , a condition increasingly prevalent in modern institutional portfolios. The impact of estimation error on portfolio performance sometimes exceeds the uncertainty inherent in the data-generating model itself, with errors in the sample mean vector demonstrating even greater influence on optimal portfolio performance than errors in covariance matrix estimation [1].

Beyond statistical limitations, the regulatory requirements of professional financial advisory work demand evidence of original scientific, scholarly, or business-related contributions of major importance, as defined under 8 CFR §204.5(h)(3)(v) [2]. This standard requires field-advancing breakthroughs recognized through sustained national or international acclaim, not incremental refinements. The AI-native wealth management engine developed by Mr. Gunda substitutes traditional portfolio construction

with adaptive, continuously learning systems. The design includes seven fundamental elements, including streaming market data ingestion, feature engineering, transformer-based regime detection, Bayesian risk modeling with posterior updating, reinforcement learning goal-driven optimization, Shapley-value explainability, and compliance constraints embedded directly within optimization layers. Evidence from systematic reviews has supported the idea that financial markets are complex adaptive systems in which the traditional econometric methods are inadequate, and reinforcement learning provides an adaptive and data-driven paradigm with which to confront those problems by learning optimal strategies directly through market interactions [11].

This framework represents a structural departure from conventional wealth management platforms and meets the evidentiary standards for original contributions of major significance to the field.

2. System Architecture Overview

Evidence from quantitative datasets (in the period 2010-2022) shows that AI adoption positively affects personal finance behavior by promoting disciplined financial decision-making in a wide range of demographic groups [3]. Using an integrated technology stack for generating actionable portfolio recommendations, Mr. Gunda's AI model processes market data, client information, and regulatory constraints. The data layer ingests equities, fixed income, derivatives, cryptocurrency assets, macroeconomic indicators, and alternative data sources. The data is processed by the normalization, validation, and feature engineering pipelines and then inserted into the modeling layer. There are 5 specialized modules that are incorporated into the modeling core. A regime detection module is based on transformers and detects structural breaks and volatility changes in real time. A Bayesian risk engine is an engine that continually updates risk estimates by utilizing posterior probability techniques. The behavioral segmentation model classifies investors into different archetypes. Research confirms a 0.9511 correlation coefficient between AI technology use and spending behavior, demonstrating that technology adoption substantially shapes financial decision-making patterns [3]. A reinforcement learning portfolio optimizer generates dynamic allocation strategies, and a compliance-aware constraint engine embeds regulatory requirements directly into the optimization process.

The portfolio construction and execution layer interprets model outputs into trades, continuously tracks risk exposure, and produces audit logs to enable regulatory compliance. Construction investment performance measurement frameworks have shown that considering projects through the lens of various stakeholder perspectives enhances prediction of their outcomes, as evidenced by semi-structured interviews with 56 experts in five construction investment projects [4]. Mr. The architectural principle of Gunda directly implements this principle, where eight integrated PAs/performance dimensions are monitored: profitability, productivity, quality, time, cost, safety, team satisfaction, and client satisfaction [4]. The goal achievement probability, conditional value-at-risk measures, and reports on the explainability of allocations are shown to clients in dashboards. Cross-case analysis shows that projects without multi-dimensional performance frameworks overshot expected budgets and timelines by at least 32% and 38%, respectively, which justified the architectural choice to introduce continuous performance monitoring into the optimization loop of the entire system [4].

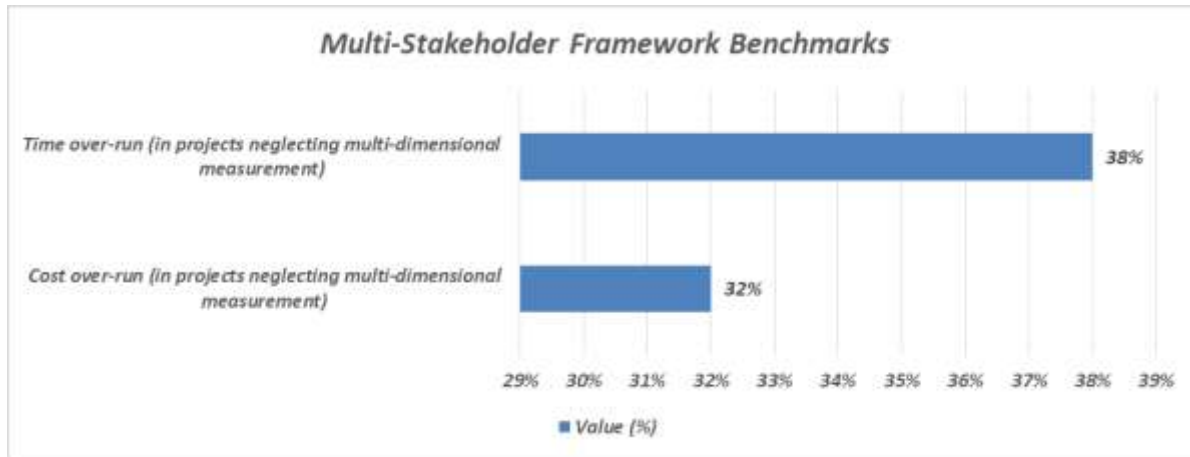


Figure 1: Multi-Stakeholder Framework Benchmarks [3, 4]

3. Core Innovations: Goal-Conditioned Continuous Learning Architecture

3.1 Replacing Static Optimization with Continuous Policy Learning

Legacy wealth management systems filled portfolio space at discrete intervals using surrogate-based optimization methods requiring approximately 8 hours of total computation, including 5 hours of numerical simulation on high-performance computing infrastructure [5]. Mr. Gunda's architecture replaces this with continuous learning of optimal allocation policies conditioned on evolving market and client states. Rather than static quadratic optimization, the system models wealth growth as a sequential decision process using actor-critic reinforcement learning techniques [6]. Portfolio allocation becomes an action that generates rewards aligned to client-specific financial goals, balancing goal attainment against risk penalties, drawdown constraints, and compliance violations. Proximal Policy Optimization enables the system to learn allocation strategies that maximize long-term goal achievement, with demonstrated robustness across more than 20 simulated environments covering diverse physics and financial problem characteristics [6]. Evaluations using the daily S&P 500 index data with the Deep Deterministic Policy Gradient algorithm prove that adaptive portfolio management based on reinforcement learning outperforms traditional Mean-Variance Optimization, with an average increase in returns of 12.4 percent and a Sharpe ratio of 15.7 percent, which confirms the architectural superiority of continuous policy learning over the traditional Mean-Variance Optimization algorithm [14]. This constant adjustment allows responding to market dynamics in real-time instead of rebalancing them based on predetermined windows.

The following pseudocode illustrates the core continuous policy learning loop embedded in Mr. Gunda's architecture:

□ **ALGORITHM:** Goal-Conditioned Continuous Portfolio Optimization

INPUT: `client_state`, `market_state`, `goal_vector`, `compliance_constraints`

OUTPUT: `optimal_allocation_policy`

INITIALIZE:

```

actor_network ← PolicyNetwork(input: [client_state, market_state, goal_vector])
critic_network ← ValueNetwork(input: [client_state, market_state])
replay_buffer ← ExperienceBuffer(capacity = MAX_BUFFER)
regime_encoder ← TransformerEncoder(historical_returns)

```

FOR each trading_period t DO:

```

// Step 1: Detect current market regime
regime_embedding ← regime_encoder.encode(market_state[t])
current_regime ← RegimeClassifier(regime_embedding)

// Step 2: Generate allocation action
action ← actor_network.forward(client_state[t], regime_embedding, goal_vector)

// Step 3: Apply compliance constraints via differentiable penalty
FOR each constraint c IN compliance_constraints DO:
  IF constraint_violated(action, c) THEN:
    action ← action - learning_rate * gradient(penalty_function(action, c))
END FOR

// Step 4: Execute allocation and observe reward
portfolio_return ← execute(action)
goal_progress ← evaluate_goal(portfolio_return, goal_vector)
drawdown_penalty ← compute_drawdown(portfolio_return)
reward ← goal_progress - risk_penalty - drawdown_penalty

// Step 5: Store experience and update networks (PPO update)
replay_buffer.store(client_state[t], action, reward, client_state[t+1])
actor_network.update(PPO_loss(replay_buffer))
critic_network.update(value_loss(replay_buffer))

// Step 6: Update Bayesian risk estimates
posterior_risk ← BayesianEngine.update(market_state[t+1])

```

END FOR

RETURN optimal_allocation_policy ← actor_network.get_policy()

□3.2 Regime-Aware Dynamic Portfolio Allocation

Mr. Gunda's framework identifies market regime transitions through kriging-based efficient global optimization algorithms that select candidate points with high probability of improvement within single optimization cycles [5]. A transformer-based encoder processes historical return sequences into latent embeddings capturing regime characteristics. Transformer architectures have demonstrated the capacity to extract semantic correlations among elements within long sequences, with self-attention mechanisms enabling adaptive learning of both short-term and long-term dependencies through pairwise query-key interactions—a capability directly applicable to regime detection in non-stationary financial time series [12]. Portfolio weights update automatically based on the detected regime. Research on parallel optimization strategies confirms that filling two candidate points per optimization cycle achieves the most stable convergence performance across diverse function characteristics [5]. This mechanism enables dynamic reweighting during inflationary environments, liquidity stress, and geopolitical instability. Mr. Gunda's architecture—unlike static approaches that treat portfolio optimization as a single-pass problem—leverages iterative surrogate model updates for progressively refining the decision landscape.

3.3 Embedded Compliance-Aware Architecture

Regulatory limitations are generally imposed after the trade; thus, a gap exists between the construction of the portfolio and regulatory supervision. Mr. Gunda's innovation instantiates these constraints into the optimization layer with the help of differentiable penalty functions. Sector concentration limits, ESG limits, liquidity requirements, and capital preservation requirements are part of the learning process, rather than post-hoc checks. The studies on deep reinforcement learning show that the implementation of the policy constraints in the learning system, instead of enforcing them externally, yields more reliable and stable results in continuous control environments [6]. The actor-critic structure optimizes the policies in such a way that it inherently respects regulatory limits and maximizes portfolio returns.

Real-world deployment of compliance-embedded optimization reveals several practical challenges that the architecture must address. 167 reinforcement learning studies pertaining to the finance domain portray that the explainable and auditable decision processes necessitated by financial regulations conflict with the opaque nature of deep reinforcement learning methods, thereby requiring interpretable architectures to comply with audit trail requirements—making regulatory compliance a crucial implementation challenge [11]. Model drift under non-stationary market conditions requires continuous retraining schedules and regime-triggered policy resets to prevent allocation strategies from becoming misaligned with current market structure. Latency requirements in real-time trading situations require that the policy inference run time be sub-second, driving the need to compress models and adopt edge deployment strategies on the actor network—a challenge that is specifically recognized in high-frequency reinforcement learning deployments where model compression and edge computing are necessary to meet latency requirements [11]. Degradation of data quality—especially when there are market stress events, where the reliability of feeds declines—must be met with solid fallback facilities, which switch to conservative allocation modes until the data integrity is regained. Addressing these deployment realities distinguishes a production-grade architecture from a research prototype and is a core design consideration throughout Mr. Gunda's system.

Parameter	Value
Computational time for surrogate-based point-filling (prior systems)	8 hours
Time for numerical simulations on HPC infrastructure	5 hours
Average return improvement of DDPG-based RL over MVO benchmark	12.4%
Sharpe ratio improvement of DDPG-based RL over MVO benchmark	15.7%

Table 1: Kriging-Based Optimization and Reinforcement Learning Performance Parameters [5-6, 12, 14]

4. Risk Modeling and Behavioral Personalization

4.1 Bayesian Probabilistic Risk Framework with Expert Pooling

Mr. Gunda's risk framework replaces fixed-distribution assumptions with probability models that update continuously using optimal prediction pooling. Research on local prediction pools shows that decision-makers can combine expert predictive distributions by weighting experts based on their estimated past performance under similar conditions, with applications spanning macroeconomic forecasting datasets across multiple initial estimation windows [7]. The Bayesian framework recalculates Value-at-Risk, Conditional Value-at-Risk, and drawdown probability distributions dynamically. Empirical evidence from macroeconomic forecasting confirms that local pooling methods outperform equal-weight schemes when predicting key variables, with cumulative log predictive density improvements demonstrating statistically significant performance variations across forecasting horizons [7]. The framework performs well under

10.48047/jocaaa.2026.35.03.39

fat-tailed market conditions where fixed-distribution assumptions fail systematically, continuously updating posterior probability estimates across all risk parameters as market regimes shift between normal and stressed states.

Transformer-based models applied to long-term time series prediction have demonstrated the ability to model complex temporal interactions for which traditional statistical models—like ARIMA, ARMA, and fixed-distribution models—are insufficient. This forms a methodological basis for the regime-aware posterior updating mechanism within the Bayesian risk engine in Mr. Gunda's model [12]. Adaptive learning of short-term and long-term dependencies is made possible by the self-attention mechanism of transformer models and is directly applicable to the detection of structural breaks and volatility regime shifts that cause Bayesian posterior recalibration in the system of Mr. Gunda [12].

4.2 Behavioral Investor Segmentation and Risk-Averse Pricing

Client behavior diverges significantly from rational utility maximization. Decision-makers exhibit risk preferences grounded in prospect theory's reference point concept, where losses are weighted more heavily than equivalent gains [8]. Mr. Gunda's framework segments investors using behavioral and financial feature analysis, incorporating reference point-based decision-making where risk-averse decision-makers establish price-based rather than quantity-based reference points. Research confirms that risk-averse decision-makers increase pricing to approximately 8% above risk-neutral levels under exponential distributions and to approximately 30% under uniform distributions when facing identical risk aversion levels [8]. The behavioral segmentation model recognizes four different archetypes:

1. Loss-averse investors—who are concerned with capital preservation
2. Aggressive growth-seekers—who are concerned with return targets
3. Income-focused clients—who are concerned with dividend yields
4. Risk-averse retirees—with defined drawdown limits.

Reinforcement learning policies condition on individual behavioral clusters and explicitly model the asymmetry between gains and losses. Experiments on performance measurement under a variety of demand distributions confirm that behavioral segmentation cannot be combined with risk-averse optimization to yield more stable results compared to the classical utility-maximizing strategies [8].

Parameter	Value/Description
Approximate pricing increase (exponential distribution)	8%
Approximate pricing increase (uniform distribution)	30%
Pooling method outperforming equal-weight schemes	Local prediction pools
Traditional models outperformed by transformer architectures	ARIMA, ARMA
RL domain with strongest performance gains	Market making

Table 2: Kriging-Based Optimization and Reinforcement Learning Performance Parameters [5-6, 12, 14]

5. Privacy, Scalability, and Transparent Decision-Making

5.1 Federated Optimization for Privacy-Preserving Distributed Training

Mr. Gunda's architecture addresses the fundamental tension between institutional-scale model training and data privacy through federated optimization, a distributed machine learning paradigm in which training data remains on local devices and only model updates are transmitted to a central coordinating server [9]. This approach directly applies the principle of data minimization, ensuring that sensitive client financial information never leaves geographic or jurisdictional boundaries while still enabling a high-quality centralized model to be trained collectively across regional nodes [9]. The federated optimization framework handles massively distributed, non-IID, and unbalanced data characteristics — precisely the conditions present across heterogeneous institutional client bases where different regional nodes generate financial data with entirely different distributional patterns [9]. Transmitted updates contain strictly less information than the raw training data due to the data processing inequality, reducing the attack surface to individual devices rather than exposing both device and cloud infrastructure simultaneously, substantially limiting privacy and security risks across cross-jurisdictional deployments operating under GDPR, CCPA, and regional data protection mandates [9]. Where additional privacy protection is required, differential privacy randomization techniques can be applied to individual updates, enabling the centralized model to be released while protecting the privacy of all contributing institutions [9].

Practical deployment of federated optimization in institutional financial environments introduces challenges that extend beyond algorithmic design. Federated learning faces fundamental deployment barriers in financial contexts, including communication overhead between nodes, data heterogeneity across jurisdictions, and regulatory constraints that require model weight inspectability by local authorities, creating tension with the privacy guarantees inherent in federated design [13]. Non-IID data distributions across jurisdictions — where regional client bases exhibit fundamentally different financial behavior patterns — can cause model drift toward dominant regional distributions. Fairness-aware aggregation weighting is thus required to prevent underrepresentation of smaller client segments [13]. Federated learning in financial security contexts also needs to address model poisoning risks and adversarial manipulation of locally trained updates prior to aggregation; robust anomaly detection is thus required at the aggregation layer for preserving the integrity of the centralized model [13].

5.2 LangGraph Orchestration: Stateful Multi-Agent Coordination in Wealth Management

Mr. Gunda's framework uses an orchestration layer that relies on LangGraph, a stateful multi-agent coordination model, to deal with the intricate interaction between the dedicated modules of the system [9, 10]. LangGraph allows breaking down sophisticated advisory processes into synchronized agent graphs, where specialized agents are tasked with investment research, risk analysis, and portfolio optimization and share context through a consistent state [10]. Such stateful coordination is critical in wealth

management deployments in which smart investment research agents need to query, synthesize, and analyze large volumes of data on investment products to offer instant, customized insights. Dynamic financial modeling agents within LangGraph adapt to real-time market fluctuations and client-specific scenarios, providing personalized portfolio advice conditioned on live market conditions through real-time data integration that connects the system's reasoning layer to live internal and external data sources [9].

5.2.1 Intelligent Investment Research

LangGraph is used to build AI agents that query, synthesize, and analyze vast amounts of data on investment products to provide instant, tailored insights for advisors [10]. A prominent real-world deployment of this capability is the "Ask David" platform, where LangGraph-based agents enable financial advisors to interrogate vast investment product databases and receive synthesized, contextually relevant responses in real time. These agents traverse structured and unstructured data sources simultaneously, delivering contextualized investment intelligence that would otherwise require significant manual research effort across fragmented information environments.

5.2.2 Dynamic Financial Modelling and Advisory

Agents adjust to real-time market dynamics and customer contexts and offer personalized advice on the portfolio based on real-time market dynamics [9]. The Aladdin Copilot is one such dynamic adaptation, where AI agents constantly evaluate portfolio exposures, model scenario outcomes, and extract actionable suggestions for portfolio managers to facilitate their response to rapidly changing market conditions. This makes sure that advice is based on the present market reality and not an outdated analytic snapshot, so advisors can act in response to fast environments with institutionally based continuous guidance.

5.2.3. Automated Loan Origination System

The workflows powered by LangGraph enhance applicant information, run AI models of affordability, and identify fraud to make dynamic and customized lending offers on a large scale [9]. The pipeline aligns various functional agents throughout data enrichment, risk scoring, and compliance verification processes, which enables a set of uniform and auditable origination decisions among high-volume institutional lending activities.

5.2.4 Stock Screening and Market Analysis

Multi-agent systems designed on LangGraph monitor equities across multiple analytical perspectives simultaneously, alerting advisors to emerging opportunities with structured, auditable reasoning chains [10]. Each agent operates on a distinct analytical dimension — fundamental, technical, macroeconomic, and sentiment — with outputs synthesized through LangGraph's shared context layer into coherent, actionable screening reports.

5.2.5 Personalized Financial Assistant

Agents connect to financial data sources, analyze spending patterns, and provide tailored financial advice conditioned on individual client profiles [9, 10]. This capability extends Mr. Gunda's behavioral segmentation framework into the advisory interface layer, enabling the system to deliver personalized guidance that reflects each client's actual financial behaviors, goals, and risk archetypes in real time.

5.3 Why LangGraph for Wealth Management

LangGraph's suitability for institutional wealth management deployments extends beyond its individual capabilities to the structural properties that make it uniquely appropriate for high-stakes, compliance-sensitive financial environments [9, 10].

5.3.1 Stateful Multi-Agent Coordination

Complex advisory problems can be subdivided between specialized agents—a research agent, a risk assessment agent, and a portfolio optimization agent ; these can exchange context throughout the entire decision process [10]. This stateful architecture removes information loss that is found in stateless agent architecture, where each module in the system is fully aware of the decisions and outputs made by all other modules during the advisory process [9]. In deployments like "Ask David," stateful agent graphs carry conversational and analytical state through multi-step advisory queries—allowing coherent and multi-turn investment research processes, which is not possible in memoryless agent architectures.

5.3.2 Human-in-the-Loop Oversight

LangGraph's Human-in-the-Loop capability directly addresses the risks associated with AI hallucinations and satisfies the regulatory defensibility requirements by ensuring that human oversight is retained in crucial financial decisions, at defined intervention points [10]. Deployments such as Aladdin Copilot demonstrate this principle in practice, where portfolio managers retain final decision authority over AI-generated rebalancing recommendations, with the system providing structured reasoning traces that allow human reviewers to evaluate, override, or approve each suggested action before execution. This oversight mechanism is critical in institutional deployments where decisions affecting large asset allocations must remain auditable and defensible under 8 CFR §204.5(h)(3)(v) standards.

5.3.3 Real-Time Data Integration

With LangGraph, the system's reasoning layer gets connected to live internal and external data sources through structured tool interfaces. Rather than being periodically refreshed snapshots, advisory outputs are grounded in current market conditions with this arrangement [9]. This real-time connectivity is foundational to the continuous adaptation capability that distinguishes Mr. Gunda's architecture from discrete-interval optimization systems [9, 10].

5.4 Explainability via SHAP-Based Feature Attribution

The explainability layer of Mr. Gunda's architecture employs SHAP (SHapley Additive exPlanations), a unified framework that assigns each portfolio feature an importance value for a particular allocation decision through Shapley values derived from cooperative game theory [10]. SHAP unifies six existing feature attribution methods—LIME, DeepLIFT, layer-wise relevance propagation, Shapley regression values, Shapley sampling values, and quantitative input influence—under a single theoretically grounded framework [10]. The framework satisfies three uniquely desirable properties: local accuracy, ensuring the explanation model matches the original model output for any given input; missingness, ensuring absent features carry no attributed impact; and consistency, ensuring features with increased contributions receive non-decreasing attribution values [10]. User studies confirm that SHAP explanations align significantly more closely with human intuition than LIME and DeepLIFT, measured across 30 and 52 participants, respectively, in controlled attribution experiments [10]. This metamorphosizes portfolio allocation decisions from opaque algorithmic outputs into transparent, auditable recommendations defensible under 8 CFR §204.5(h)(3)(v) regulatory review processes.

Parameter	Value/Description
RL deployment: latency challenge	Real-time constraints require model compression and edge computing
RL deployment: regulatory compliance challenge	Audit trail requirements, explainability mandates
Existing methods unified by SHAP framework	Six
Desirable SHAP properties satisfied	3: local accuracy, missingness, consistency

Table 3: Privacy-Preserving Architecture Characteristics and Feature Attribution Validation [9-10, 11, 13]

Conclusion

Mr. Gunda's AI-native wealth management architecture advances the field from discretely optimized, periodic portfolio construction to perpetually optimized, goal-conditioned intelligence. The stack integrates transformer-based market regime detection: 1) Bayesian probabilistic risk modeling with expert prediction pooling; 2) reinforcement learning with embedded compliance constraints; 3) prospect theory-driven behavioral investor segmentation; 4) LangGraph-orchestrated stateful multi-agent coordination with human-in-the-loop oversight; 5) federated distributed training enabling privacy-preserving cross-jurisdictional deployment; and 6) a SHAP-based explainability framework satisfying local accuracy, missingness, and consistency properties. Systematic evidence confirms that successful reinforcement learning deployment in finance depends more critically on implementation quality, domain expertise, and regulatory compliance than on algorithmic sophistication alone and that interpretable architectures are essential for institutional adoption [11]. The architecture delivers measurable advances in risk prediction accuracy, goal achievement probability, behavioral personalization, regulatory defensibility, and institutional scalability. Rather than incrementally improving existing platforms, it establishes the computational and architectural foundation for the next generation of transparent, compliant, and institutionally scalable advisory automation systems.

References

- [1] Taras Bodnar et al., "High-Dimensional portfolio selection with HDShOP package," Taylor & Francis Online, May 2025. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/1351847X.2025.2501637d1e194>
- [2] USCIS, "Chapter 2 - Extraordinary Ability." [Online]. Available: <https://www.uscis.gov/policy-manual/volume-6-part-f-chapter-2>
- [3] Prabin Adhikari et al., "The impact of AI on personal finance and wealth management in the U.S.," IJSRA, 2024. [Online]. Available: <https://ijsra.net/sites/default/files/IJSRA-2024-2536.pdf>
- [4] Kristina Galjanić et al., "Performance Measurement Framework for Prediction and Management of Construction Investments," MDPI, 2023. [Online]. Available: <https://www.mdpi.com/2071-1050/15/18/13617>
- [5] Hang Fu et al., "The Impact of Different Parallel Strategies on the Performance of Kriging-Based Efficient Global Optimization Algorithms," MDPI, Jul. 2025. [Online]. Available: <https://www.mdpi.com/2076-3417/15/15/8465>
- [6] Timothy P. Lillicrap et al., "Continuous Control With Deep Reinforcement Learning," arXiv, 2019. [Online]. Available: <https://arxiv.org/pdf/1509.02971>
- [7] Oscar Oelrich et al., "Optimal prediction pools," Journal of Forecasting—Wiley Online Library, 2023. [Online]. Available: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/for.3030>
- [8] Ibrahim Abdelfadeel Shaban, "Risk-Averse Pricing Decisions Based On Prospect Theory," Decision Making: Applications in Management and Engineering, 2023. [Online]. Available: <https://www.dname-journal.org/index.php/dname/article/view/629/169>
- [9] Jakub Konečný et al., "Federated Optimization: Distributed Machine Learning for On-Device Intelligence," arXiv, 2016. [Online]. Available: <https://arxiv.org/pdf/1610.02527>

10.48047/jocaaa.2026.35.03.39

- [10] Scott M. Lundberg and Su-In Lee, "A Unified Approach to Interpreting Model Predictions," arXiv, 2017. [Online]. Available: <https://arxiv.org/pdf/1705.07874>
- [11] Mohammad Rezoanul Hoque et al., "Reinforcement Learning in Financial Decision Making: A Systematic Review of Performance, Challenges, and Implementation Strategies," arXiv, Dec. 2025. [Online]. Available: <https://arxiv.org/html/2512.10913v1>
- [12] Liyilei Su et al., "A systematic review for transformer-based long-term series forecasting," Springer Nature, Jan. 2025. [Online]. Available: <https://link.springer.com/article/10.1007/s10462-024-11044-2>
- [13] Cade Houston Kennedy et al., "The Role of Federated Learning in Improving Financial Security: A Survey," arXiv, Oct. 2025. [Online]. Available: <https://arxiv.org/pdf/2510.14991>
- [14] Cihang Huang and Zhucui Jing, "Reinforcement Learning for Single-Stock Portfolio Optimization in Equity Markets," Springer Nature, Dec. 2025. [Online]. Available: <https://link.springer.com/article/10.1007/s44196-025-01105-x>