The American Journal of Interdisciplinary Innovations and Research
ISSN 2642-7478 | Open Access

Type Original Research
PAGE NO. 118-126

Adaptive Portfolio Management Using Cloud Driven Deep Reinforcement Learning Systems

Edward J. Belmont

Department of Computer Science and Quantitative Finance University of Melbourne, Australia

OPEN ACCESS

SUBMITTED 01 October 2025
ACCEPTED 15 October 2025
PUBLISHED 31 October 2025
VOLUME Vol.07 Issue 10 2025

CITATION

Edward J. Belmont. (2025). Adaptive Portfolio Management Using Cloud Driven Deep Reinforcement Learning Systems. The American Journal of Interdisciplinary Innovations and Research, 7(10), 118–126. Retrieved from https://theamericanjournals.com/index.php/tajiir/article/view/7379

Abstract- The rapid digital transformation of global financial markets has introduced unprecedented levels of complexity, uncertainty, and interconnectedness, challenging the foundations of classical portfolio theory and conventional risk management paradigms. Traditional optimization frameworks, while mathematically elegant, are increasingly unable to adapt to the nonstationary, nonlinear, and highly volatile environments that characterize modern asset markets. Within this evolving context, intelligent cloud-based deep reinforcement learning has emerged as a promising computational paradigm capable of learning, adapting, and optimizing financial decision-making under uncertainty. This study develops a comprehensive theoretical and methodological framework for intelligent cloud-driven dynamic portfolio risk prediction and optimization by synthesizing insights from classical portfolio theory, stochastic control, hierarchical risk models, and contemporary deep reinforcement learning systems.

Building upon foundational theories such as mean variance optimization, dynamic programming, information theoretic portfolio selection, and continuous-time asset allocation, this research situates deep reinforcement learning within a broader lineage of financial decision science. At the same time, it integrates recent advances in computational intelligence, including policy-based and value-based reinforcement learning, graph-based representation learning, and hierarchical

risk parity architectures, to address limitations inherent in static and parametric financial models. Central to the conceptual architecture developed in this study is the role of intelligent cloud infrastructures that enable distributed learning, real-time data ingestion, adaptive model updating, and scalable simulation of market environments.

The results demonstrate that deep reinforcement learning, when deployed within intelligent cloud frameworks, offers a fundamentally different epistemology of portfolio risk than classical models. Risk is no longer treated as a static distributional property of returns but as a dynamic, learned, and context-sensitive construct that evolves as agents interact with markets. By comparing hierarchical risk parity methods, graph-based learning architectures, and meta-policy reinforcement learning systems, this study shows how intelligent cloud-based models can internalize diversification, downside protection, and long-term growth objectives within their learning dynamics.

Keywords: Deep reinforcement learning, intelligent cloud computing, portfolio risk prediction, financial optimization, dynamic asset allocation, hierarchical risk models

Introduction

Using The intellectual history of portfolio management is inseparable from the broader evolution of financial economics, computational methods, and theories of uncertainty. Since the mid twentieth century, scholars have sought to formalize the problem of how rational investors should allocate wealth among risky assets in a way that balances expected return against exposure to uncertainty. The emergence of mean variance theory, dynamic stochastic optimization, and information theoretic approaches represented successive milestones in this ongoing project, yet each of these frameworks was built upon simplifying assumptions about market structure, distributional stability, and investor rationality that have become increasingly strained in contemporary financial environments (Markowitz, 1952; Samuelson, 1969; Kelly, 1956). The global financial system of the twenty first century is characterized by high frequency trading, algorithmic decision making, cross market contagion, and abrupt regime shifts, all of which challenge the viability of static or weakly adaptive portfolio models.

The growing recognition of these challenges has stimulated a search for computational paradigms capable of operating under conditions of deep uncertainty, partial observability, and continuous change. Reinforcement learning, originally developed within the field of artificial intelligence as a framework for sequential decision making, has emerged as one of the most influential of these paradigms. By framing portfolio management as a problem of learning through interaction with a stochastic environment, reinforcement learning offers a way to move beyond the closed form optimization of fixed distributions toward adaptive strategies that evolve in response to market feedback (Jang and Seong, 2023). In this view, the investor is modeled not as a passive optimizer of known parameters but as an active agent that continually updates its beliefs and actions as new information arrives.

At the same time, the rise of cloud computing has transformed the infrastructure within which financial analytics are conducted. Intelligent cloud platforms now enable the aggregation of massive data streams, the deployment of distributed learning agents, and the real time updating of predictive models at scales that were previously impossible. The convergence of deep reinforcement learning with intelligent cloud architectures has therefore created a new technological and theoretical frontier in portfolio risk management. Rather than treating learning and optimization as offline processes, cloud based reinforcement learning systems allow portfolios to be adjusted continuously in response to evolving market conditions, with risk predictions and control policies being recalibrated dynamically as new data becomes available.

A particularly influential articulation of this paradigm is provided by the intelligent cloud framework for dynamic portfolio risk prediction proposed by Mirza and colleagues, who demonstrated how deep reinforcement learning agents can be embedded within cloud infrastructures to generate adaptive and predictive risk management systems (Mirza et al., 2025). Their work represents a significant step toward operationalizing reinforcement learning in real world financial contexts, integrating data ingestion, model training, and decision execution within a unified cloud based architecture. However, while their contribution is substantial, it remains situated within a specific implementation context, leaving open broader theoretical questions about how such systems relate to the foundational principles of portfolio theory, risk measurement, and financial economics.

The present study seeks to address this gap by developing a comprehensive theoretical and methodological synthesis of intelligent cloud based deep reinforcement learning for dynamic portfolio risk prediction and optimization. Rather than focusing on a single algorithm or dataset, this research situates reinforcement learning within the long history of portfolio theory, tracing continuities and discontinuities between classical models and contemporary learning based approaches. By drawing on a wide range of scholarly contributions, including hierarchical risk parity methods, graph based learning architectures, and meta policy reinforcement learning systems, the study aims to articulate a unified conceptual framework that explains how intelligent cloud systems can coordinate learning, risk control, and portfolio allocation in complex financial environments (Millea and Edalat, 2022; Sun et al., 2024; Niu et al., 2022).

The theoretical foundation of this inquiry is rooted in the recognition that risk in financial markets is not merely a statistical property of returns but a dynamic and relational phenomenon that emerges from interactions among agents, assets, and institutions. Classical models such as mean variance optimization treat risk as variance or covariance, assuming that these quantities can be estimated reliably from historical data and remain stable over time (Markowitz, 1952). While such models have provided invaluable insights, they struggle to account for nonlinear dependencies, fat tailed distributions, and structural breaks that characterize real markets. Similarly, dynamic stochastic programming models, while more flexible, often rely on parametric assumptions that limit their adaptability to novel market regimes (Samuelson, 1969; Merton, 1975).

Deep reinforcement learning offers an alternative epistemology in which risk is learned implicitly through experience rather than specified ex ante. By interacting with simulated or real market environments, learning agents can discover patterns of reward and loss that reflect not only average outcomes but also rare and extreme events. This capability is particularly important in portfolio management, where tail risks and drawdowns often dominate investor outcomes. Hierarchical risk parity methods, for example, have shown how reinforcement learning can internalize diversification principles by structuring the learning problem in a way that reflects the hierarchical organization of assets and risk factors (Sen, 2023; Millea and Edalat, 2022).

Graph based reinforcement learning further extends this approach by representing financial markets as networks of interconnected assets, allowing agents to learn how shocks propagate through the system and how correlations evolve over time (Sun et al., 2024). Such models challenge the assumption of stable covariance structures that underlies much of classical portfolio theory, instead emphasizing the dynamic and context dependent nature of financial relationships. Meta policy learning, as exemplified by frameworks such as MetaTrader, adds another layer of adaptability by allowing multiple reinforcement learning policies to be combined and selected dynamically based on market conditions (Niu et al., 2022).

Within this rich and evolving landscape, intelligent cloud frameworks play a crucial enabling role. By providing scalable computational resources, distributed data storage, and real time connectivity, cloud platforms allow reinforcement learning systems to operate at the temporal and spatial scales required by modern financial markets. The work of Mirza and colleagues illustrates how such infrastructures can be used to integrate risk prediction, portfolio optimization, and deep learning into a cohesive operational system (Mirza et al., 2025). Yet their contribution also raises important questions about governance, interpretability, and systemic risk, as cloud based learning systems become increasingly embedded in the fabric of global finance.

The literature gap addressed by this study therefore lies not in the absence of algorithmic innovation but in the lack of an integrated theoretical account of how intelligent cloud based deep reinforcement learning reshapes the conceptual foundations of portfolio risk management. While numerous studies have compared reinforcement learning to traditional optimization models in empirical settings, fewer have examined how these approaches transform the meaning of risk, diversification, and optimality in a world of adaptive and self learning systems (Ngo et al., 2023). By synthesizing insights from financial economics, machine learning, and cloud computing, this research aims to provide such an account.

In pursuing this objective, the study adopts a deliberately expansive and interpretive approach. Rather than privileging any single metric of performance, it examines how different models conceptualize and operationalize risk, how learning dynamics interact with market structure, and how cloud infrastructures mediate the relationship between data, computation, and decision making. This allows for a nuanced assessment of both the promise and the limitations of intelligent cloud based reinforcement learning in portfolio management, grounded in a deep engagement with the scholarly literature.

Methodology

The methodological orientation of this research is grounded in qualitative analytical synthesis rather than empirical experimentation or algorithmic benchmarking. This choice reflects the conceptual ambition of the study, which seeks to integrate diverse strands of theory and evidence into a coherent framework for understanding intelligent cloud based deep reinforcement learning in portfolio risk management. By drawing on a broad corpus of peer reviewed research in financial economics, reinforcement learning, and computational intelligence, the methodology aims to construct a theoretically robust and critically informed account of how these domains intersect and coevolve.

At the core of this methodological approach is the principle of theoretical triangulation. Rather than relying on a single disciplinary perspective, the study juxtaposes classical portfolio theory, stochastic control, information theory, and contemporary machine learning to illuminate both continuities and ruptures in the evolution of portfolio optimization. This allows the analysis to move beyond surface level comparisons of model performance toward a deeper understanding of the epistemological and ontological assumptions that underlie different approaches to risk and decision making (Markowitz, 1952; Samuelson, 1969; Kelly, 1956).

The selection of references was guided by their relevance to three interrelated dimensions of the research problem. The first dimension concerns the foundational theories of portfolio selection and risk management, represented by works on mean variance optimization, dynamic stochastic programming, and continuous time portfolio choice (Merton, 1975; Black and Litterman, 1991). These provide the baseline against which newer approaches are evaluated. The second dimension involves computational and algorithmic methods for portfolio optimization, including simulated annealing, genetic algorithms, and conditional value at risk optimization, which illustrate the gradual incorporation of computational intelligence into financial decision making (Kirkpatrick et al., 1983; Arnone et al., 1993; Rockafellar and Uryasev, 2000). The third dimension encompasses contemporary deep reinforcement learning and intelligent cloud frameworks, which represent the current frontier of research and practice in this domain (Jang and Seong, 2023; Sun et al., 2024; Mirza et al., 2025).

The analytical procedure involves a close reading and interpretive comparison of these sources, focusing on how they conceptualize risk, learning, and optimization. Particular attention is paid to the reward structures, state representations, and policy update mechanisms used in reinforcement learning models, as these elements determine how risk and return are encoded within the learning process (Niu et al., 2022). By examining how different studies operationalize these components, the methodology seeks to identify common patterns and divergences that can inform a unified theoretical framework.

A central methodological pillar of this study is the integration of intelligent cloud computing into the analysis of reinforcement learning. Rather than treating cloud infrastructure as a neutral backdrop, the methodology conceptualizes it as an active component of the learning system that shapes data availability, computational capacity, and model deployment. This perspective is informed by the framework proposed by Mirza and colleagues, who emphasize the role of cloud based orchestration in enabling dynamic risk prediction and adaptive portfolio control (Mirza et al., 2025). By extending this insight, the study analyzes how cloud platforms mediate the relationship between local learning agents and global market information, creating new forms of collective intelligence in financial systems.

The limitations of this methodological approach are acknowledged as part of the analysis. Because the study does not involve original empirical data or experimental validation, its conclusions are necessarily interpretive rather than predictive. However, this is consistent with the aim of developing a theoretical synthesis that can guide future empirical research. By articulating clear conceptual linkages and identifying areas of tension and uncertainty, the methodology provides a foundation for subsequent quantitative studies that can test and refine the proposed framework (Ngo et al., 2023).

Another limitation concerns the rapidly evolving nature of both financial markets and machine learning technologies. Any theoretical account risks becoming outdated as new algorithms, data sources, and regulatory regimes emerge. To mitigate this, the methodology emphasizes underlying principles rather than specific implementations, focusing on how learning, risk, and computation interact at a fundamental level. This allows the analysis to retain relevance even as particular technologies change.

In sum, the methodological approach of this study is designed to balance depth and breadth, rigor and flexibility. By synthesizing insights from a wide range of scholarly traditions and situating them within the emerging paradigm of intelligent cloud based reinforcement learning, it aims to provide a comprehensive and critically informed account of dynamic portfolio risk prediction and optimization.

Results

The analytical synthesis undertaken in this study yields a set of interrelated findings that illuminate how intelligent cloud based deep reinforcement learning reshapes the practice and theory of portfolio risk management. These findings are not presented as numerical outcomes but as interpretive insights grounded in the comparative analysis of existing scholarly work. Each of these insights reflects a convergence of theoretical reasoning and empirical observation reported in the literature, particularly in studies that have contrasted reinforcement learning with traditional portfolio optimization methods (Ngo et al., 2023; Jang and Seong, 2023).

One of the most salient results is the recognition that deep reinforcement learning fundamentally alters the temporal structure of portfolio decision making. Classical models, whether static or dynamic, typically assume that the investor solves an optimization problem over a predefined horizon based on estimated parameters of return and risk (Markowitz, 1952; Merton, 1975). Even when dynamic programming is employed, the underlying stochastic processes are often assumed to be stationary or to follow known dynamics. In contrast, reinforcement learning agents operating within intelligent cloud frameworks treat the market as an evolving environment whose dynamics must be learned through ongoing interaction (Mirza et al., 2025). This means that portfolio weights are not the solution to a single optimization problem but the outcome of a continuous learning process in which policies are updated as new data arrives.

This temporal reconfiguration has profound implications for risk prediction. Rather than relying on historical estimates of volatility or covariance, reinforcement learning systems infer risk from the realized rewards and penalties associated with their actions. Hierarchical risk parity models, for example, structure this learning process in a way that encourages diversification across different levels of asset aggregation, allowing agents to internalize the tradeoff between concentration and stability (Millea and Edalat, 2022; Sen, 2023). The result is a form of risk awareness that is embedded in the policy itself rather than imposed externally through constraints.

Another key finding concerns the role of representation learning in capturing complex market relationships. Traditional portfolio theory often assumes that asset returns are jointly normally distributed or can be adequately summarized by a covariance matrix (Markowitz, 1952). However, empirical evidence suggests that financial markets exhibit nonlinear dependencies, regime shifts, and network effects that are poorly captured by such representations. Graph based reinforcement learning models address this limitation by encoding assets as nodes in a network whose edges represent dynamic relationships such as correlation, co movement, or exposure to common risk factors (Sun et al., 2024). By learning over these graph structures, reinforcement learning agents can detect changes in market topology that signal emerging risks or opportunities.

The integration of such models within intelligent cloud frameworks further amplifies their effectiveness. Cloud platforms allow for the continuous updating of graph structures as new data streams in, enabling the learning agent to adapt its representation of the market in near real time (Mirza et al., 2025). This dynamic representation learning contrasts sharply with the static factor models that underpin many traditional risk management systems, offering a more responsive and context sensitive approach to portfolio control.

A third result pertains to the comparative performance of reinforcement learning relative to traditional and deep learning based portfolio optimization models. Studies that have examined this question across different market regimes and asset classes generally find that reinforcement learning systems exhibit superior adaptability, particularly in volatile or rapidly changing environments (Ngo et al., 2023). While deep learning models can capture complex nonlinear patterns, they typically operate in a supervised learning paradigm that assumes a fixed relationship between inputs and outputs. Reinforcement learning, by contrast, optimizes directly for long term reward, allowing it to adjust its strategy when the underlying data generating process changes.

This adaptability is further enhanced by meta policy frameworks that integrate multiple reinforcement learning strategies and select among them based on market conditions (Niu et al., 2022). Such systems can be understood as a form of portfolio management at the level of policies rather than assets, dynamically reallocating computational and decision making resources in response to shifts in volatility, liquidity, or trend structure. Within intelligent cloud architectures, these meta policies can be trained and deployed at scale, leveraging distributed computing resources to explore a wide range of strategic possibilities (Mirza et al., 2025).

The results also highlight the importance of reward design in shaping portfolio behavior. In classical finance, the objective function is typically specified in terms of expected return and variance or some related risk measure such as conditional value at risk (Rockafellar and Uryasev, 2000). In reinforcement learning, however, the reward function can be crafted to reflect a broader set of objectives, including drawdown control, transaction costs, and long term growth. Hierarchical risk parity and other risk aware reward structures allow agents to internalize these considerations, producing strategies that are more aligned with real world investor preferences (Millea and Edalat, 2022; Sen, 2023).

Within intelligent cloud frameworks, reward functions can be updated and customized for different users or regulatory contexts, enabling a level of personalization and compliance that is difficult to achieve with static optimization models (Mirza et al., 2025). This flexibility represents a significant advance in the practical applicability of reinforcement learning to portfolio management.

Taken together, these findings suggest that intelligent cloud based deep reinforcement learning constitutes not merely an incremental improvement over existing methods but a qualitatively different approach to portfolio risk prediction and optimization. By reconceptualizing risk as a learned and dynamic property of agent environment interaction, and by leveraging cloud infrastructure to support continuous adaptation, these systems offer a powerful alternative to the static and parametric models that have long dominated financial theory.

Discussion

The results of this study invite a deeper theoretical reflection on the nature of portfolio risk, learning, and optimization in an era of intelligent cloud based computation. At the heart of this reflection lies a fundamental tension between two epistemological paradigms: one that views risk as an objective statistical property of asset returns, and another that treats risk as an emergent and context dependent phenomenon that arises through the interaction of agents with their environment. Classical portfolio theory is firmly rooted in the former paradigm, while deep reinforcement learning, particularly when embedded within intelligent cloud frameworks, exemplifies the latter (Markowitz, 1952; Mirza et al., 2025).

From the perspective of traditional financial economics, risk is typically quantified through measures such as variance, covariance, or tail risk, which are assumed to be estimable from historical data and stable enough to inform future decisions. This view underpins not only mean variance optimization but also more sophisticated models such as the Black Litterman framework, which combines investor views with market equilibrium to produce optimal asset weights (Black and Litterman, 1991). While these models have been enormously influential, they presuppose a level of structural stability and informational completeness that is increasingly at odds with the realities of global financial markets.

Reinforcement learning challenges this presupposition by shifting the focus from parameter estimation to policy learning. Rather than attempting to model the full joint distribution of asset returns, a reinforcement learning agent seeks to discover a mapping from observed market states to portfolio actions that maximizes long term reward. This mapping is learned through trial and error, guided by feedback from the environment. In this framework, risk is not something that is measured independently of action but something that is experienced through the consequences of decisions. Losses, drawdowns, and volatility become signals that shape the evolution of the policy itself (Jang and Seong, 2023).

The integration of this learning paradigm within intelligent cloud infrastructures further transforms the epistemology of portfolio management. Cloud platforms allow reinforcement learning agents to operate in parallel across multiple simulated or real market environments, aggregating experience and updating policies at a scale that far exceeds what any single agent could achieve in isolation (Mirza et al., 2025). This creates a form of collective learning in which knowledge about risk and return is distributed across a network of computational nodes, each contributing to the refinement of the overall policy.

Such collective intelligence raises both opportunities and challenges. On the one hand, it allows for a richer and more robust understanding of market dynamics, as agents can explore a wide range of scenarios and share insights through centralized or federated learning mechanisms. On the other hand, it introduces new forms of systemic risk, as correlated learning behaviors could lead to herding or instability if many agents converge on similar strategies. This concern echoes longstanding debates in financial economics about the role of homogeneous expectations and correlated trading in amplifying market volatility (Merton, 1975).

Hierarchical risk parity and graph based reinforcement learning offer partial responses to this challenge by embedding diversification principles within the learning architecture itself (Millea and Edalat, 2022; Sun et al., 2024). By structuring the learning problem in a way that reflects the multi level organization of financial markets, these models encourage agents to distribute risk across different asset classes, sectors, and network communities. In doing so, they operationalize a form of diversification that is more adaptive and context sensitive than the static diversification implied by covariance based optimization.

The meta policy frameworks developed in recent research further extend this idea by allowing multiple reinforcement learning strategies to coexist and compete within a single system (Niu et al., 2022). From a theoretical perspective, this can be seen as an application of evolutionary principles to portfolio management, where different strategies are selected based on their performance in different market environments. Intelligent cloud infrastructures provide the computational substrate for this evolutionary process, enabling large scale experimentation and rapid adaptation.

However, the very flexibility and power of these systems also raises important questions about interpretability and governance. Classical portfolio models, for all their limitations, offer a degree of transparency: investors can see how expected returns, variances, and covariances contribute to the optimal allocation. Reinforcement learning policies, particularly those implemented through deep neural networks, are often opaque, making it difficult to understand why a particular portfolio decision was made. This opacity poses challenges for regulatory compliance, risk oversight, and investor trust, especially in contexts where accountability is paramount (Ngo et al., 2023).

The intelligent cloud framework proposed by Mirza and colleagues attempts to address some of these concerns by integrating monitoring, reporting, and control mechanisms into the learning architecture (Mirza et al., 2025). By centralizing data and computation within a managed cloud environment, it becomes possible to audit model behavior, enforce risk limits, and intervene when necessary. Nevertheless, the tension between adaptability and transparency remains a central challenge for the deployment of deep reinforcement learning in finance.

Another critical dimension of the discussion concerns the relationship between reinforcement learning and classical financial theory. While these approaches are often presented as competitors, the analysis suggests that they can also be understood as complementary. Reinforcement learning does not render mean variance theory or stochastic control obsolete; rather, it provides a way to operationalize their insights in environments where parameters are unknown or unstable. For example, the concept of a reward function in reinforcement learning can be seen as a generalization of the utility functions used in economic theory, encompassing not only expected return and variance but also a wide range of behavioral and institutional constraints (Kelly, 1956; Rockafellar and Uryasev, 2000).

Similarly, the dynamic programming foundations of reinforcement learning resonate with the stochastic optimization models developed by Samuelson and Merton, even as they relax some of their restrictive assumptions (Samuelson, 1969; Merton, 1975). By learning value functions and policies directly from data, reinforcement learning agents approximate the solutions to complex dynamic optimization problems that would be intractable using traditional analytical methods.

The broader implication of these developments is that portfolio management is evolving from a problem of solving equations to a problem of designing learning systems. Intelligent cloud based deep reinforcement learning represents a shift in the locus of financial intelligence from static models to adaptive computational architectures. This shift has profound consequences not only for how portfolios are managed but also for how financial risk is understood, regulated, and distributed across society.

Future research must therefore grapple with a range of unresolved questions. How can reinforcement learning systems be made more interpretable and accountable without sacrificing their adaptability? How can intelligent cloud infrastructures be designed to promote systemic stability rather than amplify correlated behavior? How should regulatory frameworks evolve to accommodate learning based financial agents that operate across borders and markets? These questions point to a rich agenda for interdisciplinary inquiry that spans finance, computer science, economics, and public policy.

Conclusion

This study has developed a comprehensive theoretical and methodological synthesis of intelligent cloud based deep reinforcement learning for dynamic portfolio risk prediction and optimization. By situating contemporary learning architectures within the long history of portfolio theory and computational finance, it has shown how these systems represent both a continuation and a transformation of established ideas about risk, return, and rational decision making. Drawing on a wide range of scholarly contributions, including the intelligent cloud framework articulated by Mirza and colleagues, the analysis has highlighted the distinctive epistemology of reinforcement learning, in which risk is learned through interaction rather than specified in advance (Mirza et al., 2025).

The central conclusion is that intelligent cloud based deep reinforcement learning offers a powerful and flexible approach to portfolio management in complex and volatile financial environments. By leveraging distributed computation, adaptive learning, and rich market representations, these systems can respond to changing conditions in ways that static and parametric models cannot. At the same time, their deployment raises important challenges related to interpretability, governance, and systemic risk that must be addressed through careful design and regulation.

As financial markets continue to evolve, the integration of learning based intelligence and cloud computing will likely become an increasingly central feature of portfolio management. The theoretical framework developed in this study provides a foundation for understanding and guiding this evolution, offering insights that are relevant not only to researchers and practitioners but also to policymakers and regulators concerned with the stability and fairness of the global financial system.

References

1. Ngo V M, Nguyen H H, Van Nguyen P. Does reinforcement learning outperform deep learning and traditional portfolio optimization models in frontier and developed financial markets? Research in International Business and Finance, 2023, 65, 101936.

2. Black F, Litterman R. Global portfolio optimization. Financial Analysts Journal, 1992, 48, 28–43.
3. Arnone S, Loraschi A, Tettamanzi A. A genetic approach to portfolio selection. Neural Networks World, 1993, 3, 597–604.
4. Sun Q, Wei X, Yang X. GraphSAGE with deep reinforcement learning for financial portfolio optimization. Expert Systems with Applications, 2024, 238, 122027.
5. Rockafellar R T, Uryasev S. Optimization of conditional value at risk. Journal of Risk, 2000, 2, 21–42.
6. Niu H, Li S, Li J. MetaTrader: An reinforcement learning approach integrating diverse policies for portfolio optimization. Proceedings of the 31st ACM International Conference on Information and Knowledge Management, 2022, 1573–1583.
7. Samuelson P A. Lifetime portfolio selection by dynamic stochastic programming. Review of Economics and Statistics, 1969, 51, 239.
8. Millea A, Edalat A. Using deep reinforcement learning with hierarchical risk parity for portfolio optimization. International Journal of Financial Studies, 2022, 11, 10.
9. Black F, Litterman R. Global asset allocation with equities, bonds, and currencies. Fixed Income Research, 1991, 2, 1–44.
10. Jang J, Seong N Y. Deep reinforcement learning for stock portfolio optimization by connecting with modern portfolio theory. Expert Systems with Applications, 2023, 218, 119556.
11. Markowitz H M. Portfolio selection. Journal of Finance, 1952, 7, 77.
12. Merton R C. Optimum consumption and portfolio rules in a continuous time model. In Stochastic Optimization Models in Finance, Elsevier, 1975, 621–661.
13. Kelly J L. A new interpretation of information rate. Bell System Technical Journal, 1956, 35, 917–926.
14. Sen J. Portfolio optimization using reinforcement learning and hierarchical risk parity approach. In Data Analytics and Computational Intelligence: Novel Models, Algorithms and Applications, Springer Nature Switzerland, 2023, 509–554.
15. Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing. Science, 1983, 220, 671–680.