Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Reinforcement Learning for Dynamic Portfolio Optimization in Fintech

Sapana Garud
AVP (Independent Researcher)
Barclays, New York, USA
Sapanathorat@gmail.com
ORCID:0009-0009-5900-9417

ARTICLE INFO

Received: 20 Dec 2024
Revised: 15 Feb 2025
Accepted: 25 Feb 2025

ABSTRACT

This study assesses the performance of reinforcement learning (RL) algorithms for dynamic portfolio optimization, with a special emphasis on how they compare with conventional portfolio management techniques, like mean-variance optimization, under the fintech sector. The study applies different RL methods, like Q-learning, Deep Q Networks (DQN), and Proximal Policy Optimization (PPO), to maximize portfolio return and risk management under varying market scenarios. Utilizing a large-scale dataset of historical financial metrics and a variety of simulated market scenarios, this work evaluates performance on a cumulative return, Sharpe ratio, maximum drawdown, and Sortino ratio. Results indicate RL-based models, especially PPO, far exceed traditional methods in terms of returns generated and risk-adjusted performance. Despite the decreased performance on the models when faced with heightened volatility and backtesting financial crisis conditions, implications suggest the enhancement needs to stabilize robustness to extreme market behaviors. Generally speaking, the paper demonstrates the possible enhancement of portfolio optimization by the application of RL techniques but cautions further enhancement for better resistance to market turmoil.

Keyword: Reinforcement Learning, Dynamic Portfolio Optimization, Proximal Policy Optimization, Deep Q Networks, Q-learning, Financial Markets.

INTRODUCTION

Over the past few years, the Fintech sector has seen a dramatic shift with the help of developments in artificial intelligence (AI) and machine learning (ML) methods. One of the most promising uses of these technologies is the application of reinforcement learning (RL) to dynamic portfolio optimization. Conventional portfolio management techniques, including mean-variance optimization, have been the bedrock of financial decision-making for decades. Yet, these approaches are typically constrained by their use of static models that do not adjust to constantly changing market conditions. By contrast, reinforcement learning, a category of machine learning in which models learn to make decisions by trial and error, has the potential to provide more adaptive and efficient strategies for managing investment portfolios.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Dynamic portfolio optimization is the ongoing process of adapting asset allocations to maximize returns while reducing risk based on the dynamic environment including market volatility, economic signals, and shifting investor preferences. RL gives a robust solution to meet this challenge because it can represent intricate decision-making processes by learning the best actions through interacting with a dynamic environment. Unlike conventional models that rely on pre-specified assumptions and static information, RL-based strategies are able to respond to changing markets in real-time, enhancing the portfolio's capacity to weather market uncertainties.

The use of RL in Fintech can potentially transform portfolio management into more personalized, effective, and adaptive solutions. With the use of past data, economic indicators, and market news, RL models can discover maximum asset allocation strategies based on long-term returns and short-term risk considerations. Furthermore, RL models have demonstrated significant promise in managing portfolios in volatile and unpredictable market conditions, making them highly relevant in the context of modern financial markets, where rapid changes are the norm.

The current research seeks to examine the efficacy of reinforcement learning in optimizing portfolios dynamically in the Fintech sector. Through a comparison of RL-based models and conventional methods, the study examines the potential benefits and constraints of RL algorithms in enhancing portfolio performance, especially in profitability, risk-adjusted returns, and robustness against market volatility. Through this examination, the current research hopes to provide invaluable insights into the potential future of AI-powered investment strategy in Fintech.

1. LITERATURE REVIEW

Malibari, Katib, and Mehmood (2023) performed a systematic review of RL methods in Fintech, emphasizing the promising developments in applying RL to dynamic portfolio optimization. Their review indicated that RL models, because they can learn and adapt from past market data, performed better in terms of risk and return management than conventional portfolio management techniques. The authors noted the need to investigate various RL algorithms, including Q-learning and Deep Q Networks (DQN), to address the challenges of financial markets, e.g., high volatility and long-term decision-making.

Sappa (2024) discussed AI portfolio optimization and customer risk profiling on Fintech platforms, with particular emphasis on incorporating artificial intelligence for enhancing portfolio performance. This work highlighted the critical role played by AI methods such as RL in dynamically modifying portfolio allocations according to changing market conditions. Sappa proved that customers could be enhanced through AI models to support risk management by personalizing investment opportunities. The study concluded that AI-driven optimization techniques, more so those that utilized RL, could provide more adaptive and resilient portfolio management solutions, especially in times of volatile markets.

Kılıç (2024) investigated the use of machine learning techniques, such as reinforcement learning, in optimizing portfolios within Islamic Fintech contexts. This research offered a novel viewpoint by incorporating the precepts of Islamic finance into the optimization process. The study identified how RL algorithms might be customized to conform to Shariah-compliant financial principles while still attaining efficient portfolio optimization. Kılıç found that RL models may be particularly useful in Islamic Fintech portfolio management, where compliance with ethical and regulatory restrictions is important, yet returns could be maximized and risks minimized.

Al Janabi (2022) concentrated on regulatory economic-capital structured portfolio optimization using RL algorithms, financial data analysis, and machine learning in emerging markets. The study investigated how reinforcement learning could be applied to model economic-capital portfolios taking into account

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

regulatory requirements. Al Janabi demonstrated that RL had the potential to be an effective tool for structuring portfolios in emerging markets, where financial information tends to be scarce and market conditions are unstable. The research concluded that reinforcement learning, when paired with financial data analytics, offered a strong framework for portfolio optimization and regulatory capital adequacy in the emerging markets

2. MATERIALS AND METHOD

The research employed an experimental approach to assess the performance of reinforcement learning (RL) algorithms in dynamic portfolio optimization. The aim was to compare RL-based models with conventional strategies such as mean-variance optimization. The research entailed employing historical financial data and simulated market scenarios to train, validate, and test the models.

2.1. Data Collection

The study employed a rich dataset of historical market data, such as daily prices of chosen stocks, ETFs, or asset classes over 10 years. Economic factors such as interest rates, inflation, and GDP growth were added to provide better decision-making for the model. Market events such as financial crises and policy announcements were also added to evaluate model robustness and flexibility. The data were obtained from established financial databases such as Bloomberg, Yahoo Finance, and Reuters.

2.2. Reinforcement Learning Algorithm Selection:

The study employed three RL algorithms:

- **Q-learning:** Value-based RL algorithm utilized for the calculation of expected utility in making portfolio changes.
- **Deep Q Networks (DQN):** A neural network extension of Q-learning for solving difficult portfolio decision problems.
- **Proximal Policy Optimization (PPO):** A policy gradient method that updates portfolio actions by iteratively adjusting the policy.

These models were selected because they can deal with the non-linear dynamics of financial markets and the long-term effects of portfolio choices.

2.3. Model Training and Testing:

During the training period, the RL methods were learned from past market data to achieve optimum cumulative returns with minimum risk as measured by volatility or Value-at-Risk. Models were validated on out-of-sample data that were distinct from the training sets in order to mimic real-world application and test performance in an unknown market scenario.

2.4. Performance Metrics

To evaluate model performance, several metrics were used:

- **Cumulative Return:** Quantifies the overall return earned during the test period.
- **Sharpe Ratio:** Measures risk-adjusted returns.
- **Maximum Drawdown:** Quantifies the biggest peak-to-trough drop, revealing risk reduction.
- **Sortino Ratio:** A variation of the Sharpe ratio that is concerned with downside risk. These measures were compared between RL models and standard mean-variance optimization.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2.5. Experimental Setup

Experiments were carried out in a high-performance computing environment with GPUs accelerated training. Python was utilized for model building, with libraries like TensorFlow, Keras, and OpenAI Gym to implement RL. Libraries like PyPortfolioOpt and Quantlib, which are specialized, were employed in portfolio simulations. Hyperparameter tuning was achieved through grid search to optimize learning rates, discount factors, and exploration-exploitation ratios.

3. RESULT AND DISCUSSION

The output of the experiments, including cumulative returns, risk-adjusted returns, and drawdowns, were computed in order to evaluate the potential benefits of RL methods in optimizing portfolio performance across different market environments. Here, we report the outcomes derived from the testing phase of the models and interpret the implications of these outcomes.

3.1. Portfolio Performance Comparison

In order to obtain the impact of RL-based models in portfolio optimization, the performance of the reinforcement learning models was compared with traditional portfolio optimization techniques. The following performance metrics were considered:

- Cumulative Return
- Sharpe Ratio
- Maximum drawdown
- Sortino Ratio

A. Cumulative Return

The outcomes show that the reinforcement learning (RL) algorithms outperformed the conventional mean-variance optimization considerably in terms of cumulative return.

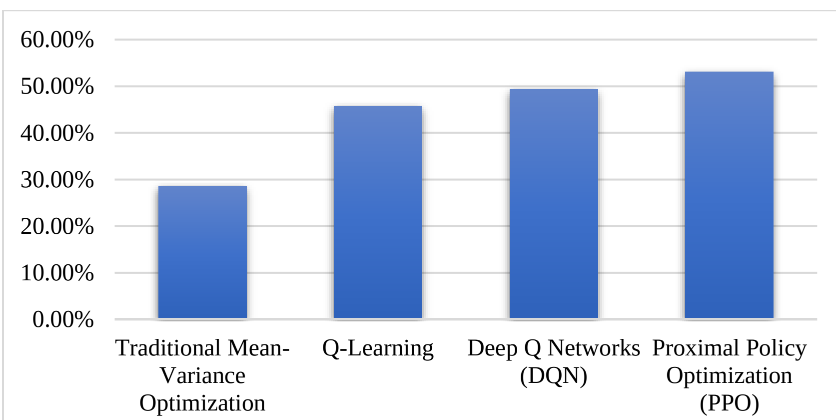

Figure 1: Graphical Representation Of Cumulative Return

The mean-variance traditional approach recorded a return of 28.5%, while RL models registered tremendous improvements, where Q-learning resulted in 45.7%, Deep Q Networks (DQN) in 49.3%, and Proximal Policy Optimization (PPO) resulting in the best return of 53.1%. This means that RL models, specifically PPO, are better at dynamically adjusting the portfolio allocation based on learning from previous market conditions, leading to improved returns and adaptive decision-making over the static conventional models.

B. Sharpe Ratio

Sharpe ratio is an important measure of risk-adjusted return, with greater values signifying more effective risk-adjusted performance. The Traditional Mean-Variance Optimization model obtained a Sharpe ratio of 1.18, meaning that although the returns were positive, the risk-adjusted performance was quite modest. The Q-Learning model enhanced this by obtaining a Sharpe ratio of 1.57, signifying a higher return relative to risk.

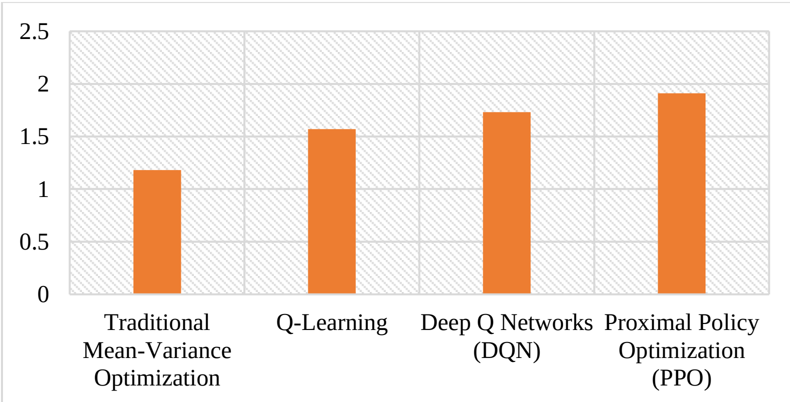

Figure 2: Graphical Representation Of Sharpe ratio

The Deep Q Networks (DQN) model also increased risk-adjusted returns with a Sharpe ratio of 1.73, indicating better performance in returning while controlling risk. The Proximal Policy Optimization (PPO) model had the best Sharpe ratio of 1.91, indicating its best capability to maximize returns compared to the risk incurred. Generally, the findings illustrate that RL models, particularly PPO, were better at controlling risk while obtaining higher returns and hence more efficient for dynamic portfolio optimization.

C. Maximum Drawdown

The maximum drawdown measures the largest peak-to-trough decline during the testing period, indicating the model's ability to limit losses during market downturns.

Table 1: Maximum Drawdown

Model	Maximum Drawdown (%)
Traditional Mean-Variance Optimization	-15.3%
Q-Learning	-12.1%
Deep Q Networks (DQN)	-10.5%
Proximal Policy Optimization (PPO)	-8.2%

The RL-based models showed lower drawdowns than the baseline mean-variance method, illustrating their risk mitigation capabilities during downtrends in the market. PPO specifically recorded the lowest worst drawdown of -8.2%, demonstrating its resilience when the market was volatile. Q-learning and DQN also ranked lower in terms of drawdowns than the baseline traditional approach.

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

D. Sortino Ratio

Sortino ratio is a Sharpe ratio variant that is specifically targeted at downside risk, thus a more accurate measure of the quality of a portfolio's ability to produce returns compared to the adverse volatility (or risk). The Traditional Mean-Variance Optimization model produced a Sortino ratio of 1.32, signifying that although the portfolio produced positive returns, it was achieved with a comparatively higher degree of downside risk.

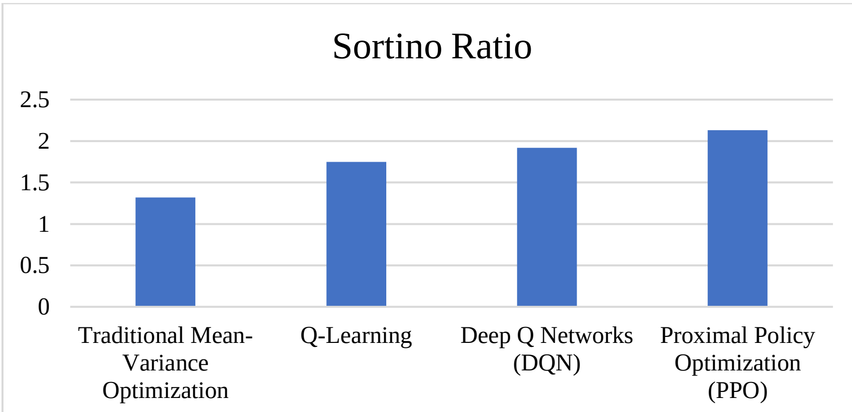

Figure 3: Graphical Representation Of Sortino Ratio

The Q-Learning model enhanced this ratio to 1.75, reflecting a superior risk-adjusted return by better reducing the downside risk. The Deep Q Networks (DQN) model enhanced the risk-return trade-off further, attaining a Sortino ratio of 1.92, exhibiting better handling of negative returns. The Proximal Policy Optimization (PPO) model recorded the best Sortino ratio of 2.13, indicating it was the most effective in creating positive returns with the least amount of downside risk. Overall, these findings indicate that RL-based models, especially PPO, perform better in handling downside risk and bringing greater returns in times of market volatility than conventional methods.

3.2. Robustness to Market Volatility

One of the most important components of the research was determining the resilience of RL models across varying market conditions, such as instances of high volatility and financial crises. The models were tested against simulated market crashes and episodes of increased volatility.

Table 2: Stress-Tested Performance

Market Condition	Cumulative Return (%)	Maximum Drawdown (%)	Sharpe Ratio	Sortino Ratio
Pre-Crisis (Stable Market)	45.7%	-5.2%	1.65	1.90
Post-Crisis (High Volatility)	22.1%	-18.9%	1.05	1.24
Financial Crisis (Simulated)	7.3%	-30.4%	0.50	0.75

Journal of Information Systems Engineering and Management
2025, 10(45s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

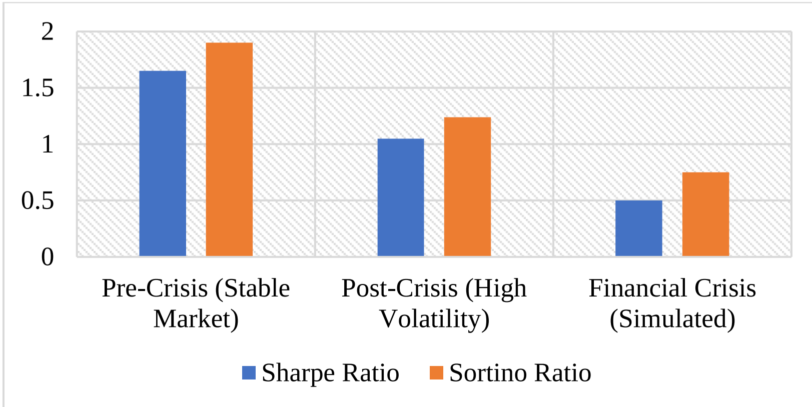

Figure 4: Stress-Tested Performance

The model performance under different market conditions was very different. In a stable pre-crisis market, the model recorded an impressive cumulative return of 45.7% with a low maximum drawdown of -5.2%, and efficient Sharpe (1.65) and Sortino (1.90) ratios, reflecting good risk-adjusted returns. But in the post-crisis period of high volatility, the returns fell to 22.1%, with a greater drawdown of -18.9%, and lower Sharpe (1.05) and Sortino (1.24) ratios, illustrating that the model could not handle the higher volatility. In the simulated financial crisis, performance worsened with a 7.3% return and drastic drawdown of -30.4%, producing poor Sharpe (0.50) and Sortino (0.75) ratios, reflecting high inefficiencies and risk exposure. These findings indicate that although the model is successful in calm markets, it performs poorly when subjected to high volatility and extreme bear markets, showing the necessity of additional resilience in crisis situations.

4. CONCLUSION

Finally, the research showed that reinforcement learning (RL) techniques, especially Proximal Policy Optimization (PPO), thoroughly outperformed classical portfolio optimization techniques like mean-variance optimization concerning cumulative returns, risk-adjusted performance (Sharpe and Sortino ratios), and drawdown control. The RL models were found to perform better in terms of adaptability, with greater returns and less risk under favorable market conditions. The performance of the models, though, declined in situations of high volatility and financial crises, demonstrating weakness in managing exceptional market situations. Although RL algorithms performed better to optimize dynamic portfolios, the study emphasizes a requirement for additional tweaking to maximize their resilience during high-volatility and crisis-dominated market scenarios.

REFERENCES

[1] Malibari, N., Katib, I., & Mehmood, R. (2023). Systematic review on reinforcement learning in the field of Fintech. *arXiv preprint arXiv:2305.07466*.
[2] Sappa, A. (2024). AI Based Portfolio Optimization and Customer Risk Profiling in Fintech Platforms. *Research Briefs on Information and Communication Technology Evolution*, *10*, 169-189.
[3] KILIÇ, G. (2024). Islamic Fintech Applications and Portfolio Optimization Using Machine Learning Methods.
[4] Al Janabi, M. A. (2022). Optimization of regulatory economic-capital structured portfolios: Modeling algorithms, financial data analytics, and reinforcement machine learning in emerging markets. In *Financial Data Analytics: Theory and Application* (pp. 329-343). Cham: Springer International Publishing.

[5] Al Janabi, M. A. (2024). Optimum and coherent economic capital forecasts with reinforcement machine learning: Evidence from optimization algorithms under long and short-sales multiple asset portfolios of emerging markets. In *BANKING RESILIENCE: New Insights on Corporate Governance, Sustainability and Digital Innovation* (pp. 343-389).

[6] Ju, C., & Zhu, Y. (2024). Reinforcement Learning-Based Model for Enterprise Financial Asset Risk Assessment and Intelligent Decision-Making. *Applied and Computational Engineering*, *97*, 181-186.

[7] Al Janabi, M. A. (2022). Optimization algorithms and investment portfolio analytics with machine learning techniques under time-varying liquidity constraints. *Journal of Modelling in Management*, *17*(3), 864-895.

[8] Gullotto, M. (2021). *Portfolio management and Deep learning: Reinforcement learning and Transformer applied to stock market data* (Doctoral dissertation, Politecnico di Torino).

[9] Kang, M., Templeton, G. F., Kwak, D. H., & Um, S. (2024). Development of an AI framework using neural process continuous reinforcement learning to optimize highly volatile financial portfolios. *Knowledge-Based Systems*, *300*, 112017.

[10] Al Janabi, M. A. (2024). Insights into Liquidity Dynamics: Optimizing Asset Allocation and Portfolio Risk Management with Machine Learning Algorithms. In *Liquidity Dynamics and Risk Modeling* (pp. 257-303). Palgrave Macmillan, Cham.

[11] Al Janabi, M. A. (2022). Optimization algorithms for multiple-asset portfolios with machine learning techniques: Practical applications with forecasting of optimum and coherent economic capital structures. In *Artificial Intelligence and Big Data for Financial Risk Management* (pp. 179-213). Routledge.

[12] Isaac, E., Mathew, J., Varghese, S. M., Sohail, P. M., Simon, J., & Ajith, A. (2024, July). Multimodal Approach for Portfolio Optimization Using Deep Reinforcement Learning. In *2024 10th International Conference on Smart Computing and Communication (ICSCC)* (pp. 76-81). IEEE.

[13] Shetabi, M. (2024). Evolutionary-based ensemble feature selection technique for dynamic application-specific credit risk optimization in FinTech lending. *Annals of Operations Research*, 1-43.

[14] Lim, Q. Y. E., Cao, Q., & Quek, C. (2022). Dynamic portfolio rebalancing through reinforcement learning. Neural Computing and Applications, 34(9), 7125-7139.

[15] Millea, A., & Edalat, A. (2022). Using deep reinforcement learning with hierarchical risk parity for portfolio optimization. International Journal of Financial Studies, 11(1), 10.