Comparing Deep Reinforcement Learning and Traditional Methods for Adaptive Portfolio Management

Blaine Swieder
Department of Computer Science
College of Engineering
Tennessee Technological University
Cookeville, Tennessee 38505
Email: bcswieder42@tntech.edu

***Abstract*—**

***Index Terms*—Algorithmic Trading, Machine Learning, Quantitative Finance, Risk Management**

1. Problem Statement and Significance

1.1. Problem

When it comes to stock trading, modern financial markets can present unique challenges for algorithmic trading systems where the effective implementation of machine learning involves balancing three competing objectives:

1) Maximizing risk-adjusted returns (Sharpe Ratio $>$ 1.5)
2) Maintaining sub-10 ms execution latency
3) Ensuring model stability across market regimes

The main objective of this project is to address the following question: ***Which reinforcement learning algorithm, compared to traditional machine learning algorithms, provides the most optimal performance in different trading frequencies and market volatility regimes?*** It is essential to address this question, as financial markets are inherently dynamic, thus showing sudden shifts in volatility, liquidity, and price behavior. These objectives are difficult to balance, since models optimized for high returns often suffer from latency issues, while those designed for low latency may compromise predictive accuracy. Algorithmic trading models that perform well under specific conditions may fail dramatically when exposed to different market environments will lead to significant financial losses. This will lose both the user and the investor's confidence in the system.

1.2. Significance

This study is of great practical importance for both institutional and retail traders, etc. Every year, institutional investors lose approximately $2.1 billion annually due to suboptimal algorithm selection, which highlights the financial impact of poorly performing models [8]. In addition, around 68% of retail trading platforms lack robust model evaluation frameworks, which result in increased exposure to market risks and inconsistent trading outcomes. From a legal and regulatory perspective, compliance with policies such as SEC Rule 15c3-5 mandates the implementation of explainable trading strategies that ensure transparency and prevent the market from being manipulation.

The aim of this project is picking which machine learning algorithm—amongst Proximal Policy Optimization (PPO), Advantage Actor-Critic (A2C), Deep Q-Networks (DQN), Gradient Boosting Machines (GBM), and Deep Neural Networks (DNN) – consistently has superior performance across various financial conditions. Every one of these machine learning algorithms has its proven strengths. For example, deep reinforcement learning models such as Proximal Policy Optimization (PPO) and Advantage Actor-Critic (A2C) usually excel in dynamic environments with continuous learning capabilities. Moreover, Gradient Boosting Machines (GBM) offer strong predictive power for structured data while Deep Neural Networks (DNN) capture complex nonlinear patterns. While prior literature has demonstrated the effectiveness of individual models, relatively few offers a comprehensive comparative analysis across multiple market conditions. If we evaluate these algorithms across diverse market regimes, such as bullish, bearish, and sideways; then, we will be able to identify not only top performer overall, but we also identify how each model effectiveness varies with market volatility, execution constraints, and risk profiles. By completing this study, we will be able to provide critical insights into algorithm selection for developing resilient, high-performance trading strategies for investors and users.

2. Background and Related Work

In the literature, deep reinforcement learning has emerged as a transformational approach in algorithmic trading due to its ability to be able to learn optimal trading strategies directly from dynamic market environments. Deep reinforcement learning methods, such as Proximal Policy Optimization (PPO), Advantage Actor-Critic (A2C), and

Deep Q-Networks (DQN) have been proven to show significant promise in handling non-stationary data, adapting to changing market conditions, and maximizing long-term returns. As opposed to traditional machine learning models that rely heavily on historical data for predictive accuracy, deep reinforcement learning models continuously learn from the interactions with the market environment hence allowing them to adapt to real-time fluctuations [1], [2].

Moreover, the literature also reveals that deep reinforcement learning outperforms conventional models in scenarios characterized by high volatility and complex market dynamics. For example, Zhang et al. [3] demonstrated the effectiveness of deep reinforcement learning algorithms in volatile markets, attributing their success to the ability of reinforcement learning to optimize decision-making under uncertainty. Similarly, Patel et al. [4] demonstrated that integrating deep neural networks with deep reinforcement learning frameworks enhances predictive accuracy, especially in high-frequency trading applications. In addition, traditional models such as Gradient Boosting Machines (GBM) and Deep Neural Networks (DNN), despite being robust in structured data analysis, often lacks the adaptability that is required in rapidly changing financial environments [7].

More recent work suggests that hybrid models that combine deep reinforcement learning with traditional techniques can mitigate challenges such as overfitting, computational inefficiency, and sensitivity to hyper parameters [6]. This research aims to build upon existing work by systematically comparing the performance of deep reinforcement learning algorithms with traditional models using different market regimes and by aiming to provide a comprehensive understanding of their strengths, limitations, as well as practical applications in algorithmic trading.

3. Data Aquistion

For this project, we will utilize a dataset that is comprised of historical price data that is sourced from reputable financial databases like Bloomberg, Yahoo Finance, and Quandl. In addition, macroeconomic indicators, which include interest rate fluctuations and GDP growth rates that are obtained from reliable sources such as the Federal Reserve Economic Data (FRED) and World Bank datasets; moreover, technical indicators like Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), and Bollinger Bonders are incorporated to capture market trends and volatility.

Our data preprocessing will involve normalization, feature engineering, and noise reduction techniques to enhance model performance. We could employ a rolling-window approach to simulate real-time trading conditions, to ensure robustness against data leakage and providing a realistic evaluation of model adaptability. The DRL algorithms are trained alongside traditional models such as GBM and DNNs to establish performance benchmarks, as done shown by Lee et al. and Prado [2], [7].

The model performance will be evaluated using key financial metrics, which includes the Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and annualized returns. This methodological framework is intended to provide a rigorous, empirical basis to evaluate the adaptability, robustness, and practical viability of machine learning algorithms in diverse financial market conditions.

4. Methodology

The aim of this project is to conduct a comparative analysis of Deep Reinforcement Learning (DRL) and traditional machine learning algorithms for adaptive portfolio management. Our data will be sourced from datasets that consist of historical price data from Bloomberg, Yahoo Finance, and Quandl, coupled with technical indicators such as Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI), and macroeconomic variables like interest rate fluctuations that are sourced from the sourced from the Federal Reserve Economic Data (FRED) and the World Banks will be used to simulate a realistic trading environment. Next, we will used data preprocessing techniques such as normalization and feature engineering will be applied to enhance the model's performance and ensure data consistency. Thirdly, we will implement Proximal Policy Optimization (PPO), Advantage Actor Critic (A2C), and Deep Q-Networks (DQN) to test efficiency of DRL strategies. Moreover, these will be benchmarked against traditional machine learning models like Gradient Boosting Machines (GBMs) and Deep Neural Networks (DNNs). Also, we will adopt a rolling-window approach to prevent data leakage and to ensure real-world applicability. As a result, these models will be iteratively trained and tested on sequential subsets of the data.

This approach is intended to replicate real-time trading conditions, providing a realistic evaluation of model adaptability and generalization across various market scenarios. In addition, we will use financial metrics like the Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and annualized returns as well as overall trading strategy effectiveness. Additionally, cost analysis, incorporating transaction costs will be conducted to evaluate real-world trading viability and to assess the impact of market volatility on algorithmic performance. By conducting a thorough comparative analysis across these dimensions, this project aims to give empirical insights into the advantages and disadvantages of each algorithm, which will guide the development of resilient, data-driven trading strategies applicable in dynamic financial environments.

5. Timeline

We will follow the following proposed timeline to conduct this experiment and its methodology. The project is split into four distinct phases across a duration of approximately 13 weeks.

5.1. Phase I: Data Collection and Preparation

This first phase will be focused on data collection and preparation. This phase will involve the identification, access, and collection of historical price data, technical indicators, and macroeconomic variables from Bloomberg, Yahoo Finance, Quandl, FRED, and the World Bank. During this phase, data cleaning, preprocessing, feature engineering, and normalization will be utilized to prepare the data for model training.

5.2. Phase II: Model Development and Implementation

The model development and implementation phase will last around 3 weeks, and will be dedicated to implementing deep reinforcement learning models Actor Critic, Deep Q Network, and Proximal Policy Optimization, alongside traditional machine learning models such as Gradient Boosting Machines (GBMs) and Deep Neural Networks (DNNs). The implementation will be following by fine-tuning hyperparameters for each model, using techniques such as grid search, Bayesian optimization, or other optimization algorithms to optimize performance. Moreover, this will also involve the development of the simulated trading environment, which will replicate real-world market conditions.

5.3. Phase III: Training, Backtesting, and Evaluation

Training, backtesting, and evaluation will take approximately 3 weeks. This thrid phase will concentrate on the training of each model based on historical data by using a rolling-window approach. Next, backtesting will be conducted across different market regimes (i.e., bullish, bearish, and sideways) to assess model performance under varying market conditions. We will use financial metrics such as the Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and annualized returns to evaluate performance. In addition, cost and risk analysis will also be integrated to assess the impact of transaction costs and market volatility.

5.4. Phase IV: Analysis, Reporting, and Conclusion

This phase will last about two weeks. This fourth phase will involve a comparative analysis of the performance of each algorithm, documenting the findings within a comprehensive report. This comprehensive report will include the methodology, results, conclusions, and recommendations based on the empirical insight gained throughout the project. This final week will be devoted to finalizing and revising the report, drawing conclusions, and preparing a presentation based on the findings.

6. Conclusion

In conclusion, this project will aim to provide a comparative analysis of Deep Reinforcement Algorithms such as Proximal Policy Optimization (PPO), Advantage Actor Critic (A2C), and Deep Q-Networks (DQN) for optimizing algorithmic trading strategies. Our goal will be to identify which DRL architecture will provide the superior risk-adjusted returns in dynamic financial markets. To benchmark performance, we will use tradition machine learning models such as Gradient Boosting Machines (GBMs) and Deep Neural Networks (DNNs) as a comparison for the evaluation.

For our data, we will use historical price data from Bloomberg, Yahoo Finance, and Quandl, as well as technical indicators (MACD, RSI) and macroeconomic variables (i.e., interest rate fluctuations from FRED and World Bank). We will also test across diverse market regimes such as bullish, bearish, and sideways to determine and assess adaptability and robustness. Some key performance metrics such as the Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and annualized returns.

Finally, the expected outcome will be to generate empirical insights into the advantages and disadvantages of each algorithm. We will aim to guide the development of robust, data-driven trading strategies that will adapt effectively to dynamic market conditions, which will contribute to more resilient and efficient trading systems.

References

[1] M. Smith and L. Johnson, "Deep Reinforcement Learning Robots for Algorithmic Trading: Considering Stock Market Conditions and U.S. Interest Rates," *Journal of Financial Data Science*, vol. 3, no. 1, pp. 15–27, 2021.

[2] R. Lee and S. Kim, "Algorithmic Trading on Financial Time Series Using Deep Reinforcement Learning," *IEEE Access*, vol. 8, pp. 212450–212460, 2020.

[3] X. Zhang, Y. Liu, and Q. Wang, "A Novel Deep Reinforcement Learning-Based Automatic Stock Trading Method and a Case Study," *Applied Soft Computing*, vol. 94, pp. 106–119, 2020.

[4] D. Patel and A. Banerjee, "High-Frequency Trading Strategy Based on Deep Neural Networks," *Quantitative Finance*, vol. 22, no. 4, pp. 567–578, 2022.

[5] J. Thompson, "Deep Reinforcement Learning in Quantitative Algorithmic Trading: A Review," *Finance and AI Review*, vol. 5, no. 2, pp. 34–48, 2021.

[6] K. Wang and H. Zhao, "Gradient Boosting Reinforcement Learning," *Proceedings of the International Conference on Machine Learning*, 2020.

[7] M. Lopez de Prado, *Advances in Financial Machine Learning*. Hoboken, NJ: Wiley, 2018.

[8] J.P. Morgan, "Annual Report on Institutional Trading Strategies," J.P. Morgan, 2023.

[9] Bloomberg, "Bloomberg Markets - Financial Data & Analytics," *Bloomberg*, 2024. [Online]. Available: https://www.bloomberg.com/. [Accessed: Feb. 3, 2025].

[10] Yahoo Finance, "Stock Market Data - Historical and Real-Time Prices," *Yahoo Finance*, 2024. [Online]. Available: https://finance.yahoo.com/. [Accessed: Feb. 3, 2025].

[11] Quandl, "Financial and Economic Data API," *Quandl*, 2024. [Online]. Available: https://www.quandl.com/. [Accessed: Feb. 3, 2025].

[12] Federal Reserve Bank of St. Louis, "FRED Economic Data," *Federal Reserve Economic Data (FRED)*, 2024. [Online]. Available: https://fred.stlouisfed.org/. [Accessed: Feb. 3, 2025].

[13] The World Bank, "World Bank Open Data," *The World Bank*, 2024. [Online]. Available: https://data.worldbank.org/. [Accessed: Feb. 3, 2025].