Comparing Deep Reinforcement Learning and Traditional Methods for Adaptive Portfolio Management

Blaine Swieder[1]

[1]Tennessee Technological University, Department of Computer Science
bcswieder42@tntech.edu

Abstract

In this project, we will aim to provide a comparative analysis of Deep Reinforcement Learning algorithms for optimizing algorithmic trading strategies; more specifically, Proximal Policy Optimization, Advantage Actor-Critic, and Deep Q-Networks. Our principal goal is to identify which deep learning architecture delivers superior risk-adjusted returns when applied to dynamic financial markets. Moreover, to establish a performance benchmark, we will utilize traditional machine learning models such as gradient boosting machines and deep neural networks to provide further evaluation. We will utilize historical price data, technical indicators (that is,is, MACD, RSI), and macroeconomic variables, as well as interest fluctuations, to create a robust trading environment. The models will be trained and tested across multiple different market regimes, bullish, bearish, and sideways, to allow a comprehensive assessment of adaptability and robustness. In addition, we will measure performance using key financial metrics such as the Sharpe Ratio, Sortino Ratio, Maximum Drawdown, and annualized returns. In addition, this project will incorporate cost analysis to assess real-world trading viability and evaluate the impact of market volatility on algorithmic performance. The expected outcome of this project is to generate empirical insights into the comparative strengths and limitations of each algorithm, which will offer practical guidance for developing robust data-driven trading strategies in dynamic financial environments.

Introduction

In the modern financial landscape, algorithmic trading systems have become increasingly dependent on predictive modeling to identify profitable trading opportunities. Traditional machine learning (ML) methods, such as Gradient Boosting Machines (GBM) and Deep Neural Networks (DNN), have demonstrated success in forecasting stock prices based on historical data. These models learn from patterns embedded in the data, optimizing for prediction accuracy. However, one of their primary limitations is the lack of adaptability to changing market conditions. Their predictive nature means that they make static forecasts based on past trends, without dynamically adjusting to evolving market regimes.

In contrast, Deep Reinforcement Learning (DRL) offers a paradigm shift by enabling agents to learn strategies through direct interaction with the environment, guided by reward signals. DRL models such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Advantage Actor-Critic (A2C) have the potential to adapt in real-time to shifting market dynamics. These models treat trading as a sequential decision-making process, enabling them to optimize strategies not just for accuracy, but for long-term returns.

This project aims to investigate whether DRL-based strategies can outperform traditional ML methods in both predictive accuracy and adaptability. By applying both model types to the same dataset—comprising daily stock prices from 2018 to 2023—we conduct a systematic comparison across both general performance metrics and regime-specific performance. We further introduce a custom trading environment to simulate realistic trading actions and assess agent performance over time. The broader goal is to evaluate the extent to which DRL can complement or even replace traditional models in algorithmic trading.

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 (Smith and Johnson 2021; Lee and Kim 2020).

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. (Zhang, Liu, and Wang 2020) 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. (Patel and Banerjee 2022) 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 (Lopez de Prado 2018).

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 (Wang and Zhao 2020). 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.

Methodology

To achieve our objective of comparing traditional machine learning models against deep reinforcement learning (DRL) agents, a multi-stage experimental pipeline was established. Traditional models focused on regression-based prediction tasks, while the Deep Reinforcement Learning (DRL) agents operated within a custom-built trading environment that simulated dynamic financial market behavior. We selected traditional machine learning models such as Gradient Boosting Machines (GBMs) using the XGBoost library and Deep Neural Networks (DNN) using Tensorflow/Keras. Our GBM was optimized with hyperparameters including 100 estimators, a learning rate of 0.1, and regularization to mitigate overfitting. The Deep Neural Network (DNN) architecture employed dense layers, batch normalization, dropout, and ReLU activations to improve generalization.

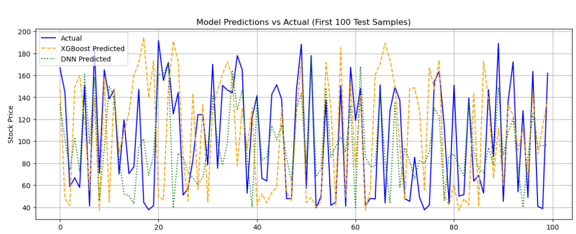

Figure 1: DNN vs XGBoost Comparison

With respect to the Deep Reinforcement Learning (DRL) side, three popular agents were implemented by using the stable-baselines3 library: Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Advantage Actor-Critic (A2C). A custom OpenAI Gym-compatible environment called `StockTradingEnv` was developed to allow sequential interaction between the agents and the market. This environment supported three discrete actions (hold, buy, sell) and rewarded agents based on their portfolio value changes.

To introduce realism, the dataset was partitioned into three market regimes: bullish, bearish, and volatile. Each model was trained independently within each regime, allowing for evaluation of adaptability across different financial conditions. Traditional models were evaluated utilizing Mean Squared Error (MSE) and R-squared (R^2) scores. Deep Reinforcement Learning (DRL) agents were evaluated on cumulative reward, Sharpe Ratio, Sortino Ratio, and maximum drawdown to capture profitability and risk management.

We trained the hyperparameters for the DRL models included 10,000 timesteps, Adam optimizers, learning rates ranging from 10^{-4} to 10^{-5}, and early stopping based on validation loss. The models were initialized with random seeds, and experiments were repeated across regimes to isolate environment-specific effects.

Experiments

Our data was sourced from the FRED. The experimental dataset consisted of historical stock price data that was collected from 2018 to 2023. This data included standard OHLCV features (Open, High, Low, Close, Volume) as well as technical indicators such as Moving Average Convergence Divergence (MACD) and Relative Strength Index (RSI). In addition, all numeric features were normalized between 0 and 1 by using the MinMaxScaler to ensure that scale differences would not impact model convergence.

For traditional ML models, a target variable was constructed by shifting the ”Close” price by one day to formulate a next-day prediction task. We split our dataset into training (80%) and testing (20%) partitions. Our models were trained on the training set and evaluated on the test set by using Mean Standard Error (MSE) and R^2 metrics.

For the Deep Reinforcement Learning (DRL) models, the `StockTradingEnv` provided a simulation of daily trading based on input features. Our agents observed market states and executed one of three actions: hold, buy, or sell. The reward at each step was computed based on daily portfolio value changes. At episode termination, the final portfolio value was recorded.

Market regimes were segmented based on 30-day moving averages and price volatility thresholds. Each DRL agent was retrained separately for bullish, bearish, and sideways conditions to assess robustness.

Results and Discussion

The traditional machine learning models demonstrated strong predictive accuracy. The Gradient Boosting Machines (GBMs) achieved the best results with a Mean Squared Error (MSE) of 5.60 and an R^2 score of 0.998, which indicates highly precise one-step-ahead predictions. In addition, Deep Neural Networks (DNNs) performed reasonably well but with slightly higher variance, showing an MSE of 485.60 and an R^2 of 0.819.

The Deep Reinforcement Learning (DRL) agents showed greater variability but higher adaptability to different market

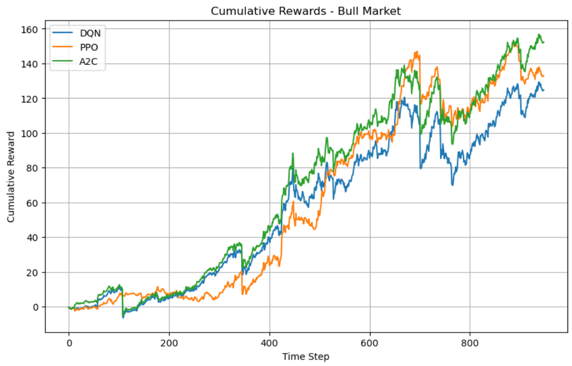

Figure 2: Cumulative Rewards in Bullish Market

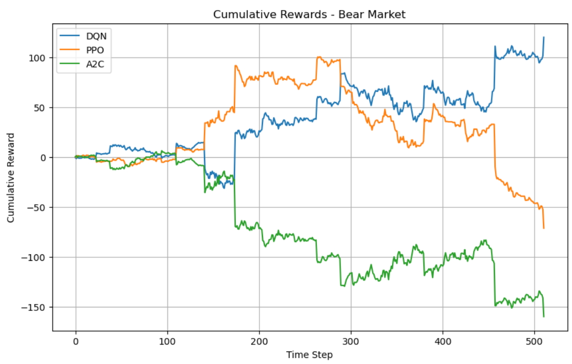

Figure 3: Cumulative Rewards in Bearish Market

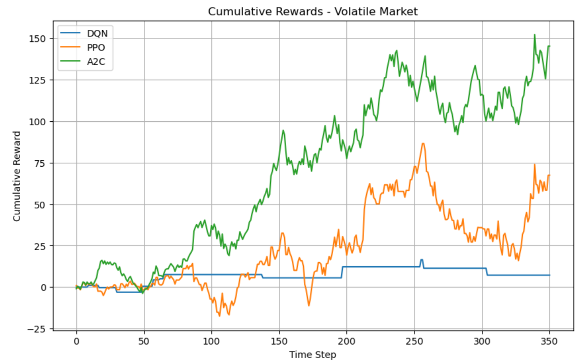

Figure 4: Cumulative Rewards in Volatile Market

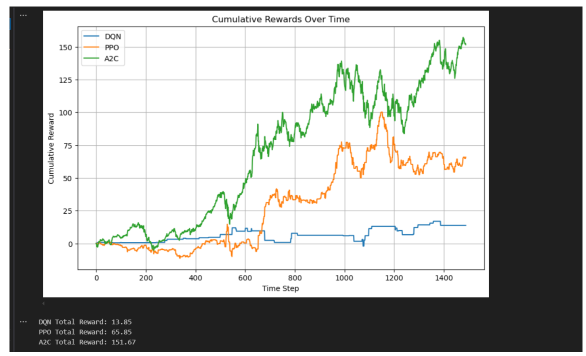

Figure 5: Overall Cumulative Rewards Comparison (A2C, PPO, DQN)

regimes. First, the Advantage Actor-Critic (A2C) agent consistently delivered the highest cumulative reward (151.67) across all regimes. Next, the Proximal Policy Optimization (PPO) agents followed with a cumulative reward of 65.85. Finally, the Deep Q-Network (DQN), while stable in bear markets, underperformed in bullish and volatile conditions with a cumulative reward of only 13.85.

When analyzing regime-specific behavior, Advantage Actor-Critic (A2C) agents achieved superior Sharpe and Sortino ratios and demonstrated the lowest maximum drawdown. The Proximal Policy Optimization (PPO) agents managed risk well but exhibited slightly lower reward accumulation. The Deep Q-Network (DQN) agents exhibited inconsistency, suffering from poor exploration strategies and failing to adapt effectively during regime changes.

These results reinforce our hypothesis that while traditional machine learning models provide strong short-term forecasting, deep reinforcement learning (DRL) agents offer superior dynamic decision-making ability when facing market uncertainty. The primary limitation observed was the inconsistency in DRL results across multiple training runs, which is a recognized challenge in reinforcement learning applications.

Further Research

While our study demonstrates the comparative strengths of traditional machine learning and deep reinforcement learning (DRL) models in portfolio management, there is still potential for further research with this project. One potential research direction could involve developing hybrid systems that leverage both predictive accuracy from traditional models and adaptive policy learning from DRL. For example, a GBM-based predictor could inform DRL agent states, potentially enhancing decision-making efficiency and mitigating poor exploration behavior observed in models such as Deep Q-Networks (DQN).

Another potential research direction could be the inclusion of transaction costs, slippage, and latency, which are essential in real-world trading. By incorporating these elements into the reward functions, it would lead to the DRL environment being more realistic and improve the robustness of learned policies. In addition, modeling order book dynamics and allowing agents to execute limits or market orders that could provide a more granular simulation of trading conditions.

The stochasticity of DRL performance implies a need for more rigorous reproducibility testing. Future work should be run on multi-seed experiments and perform statistical significance testing across runs. In addition, exploring fine-tuning strategies such as reward shaping, curriculum learning, and meta-learning could improve DRL training stability and convergence.

Finally, expanding the dataset to include high-frequency intra-day data or multi-asset portfolios (e.g., ETFs, commodities, crypto) would provide richer environments for agent learning. This would test the scalability and generalization of the models by offering deeper insights into their potential usage in complex portfolio optimization problems.

By addressing these limitations, future research can be effective in more practical, robust, and intelligent trading systems that bridge the gap between static prediction and dynamic market adaptation.

Conclusion

The objective of this research project was to compare the performance of traditional machine learning models and deep reinforcement learning agents for adaptive portfolio management. Gradient Boosting Machines proved to excel at next-day stock price prediction by achieving high predictive accuracy on the test data. On the contrary, their static nature made them less capable of changing financial conditions.

Deep Reinforcement Learning models, more specifically Advantage Actor-Critic (A2C), demonstrated robust adaptability and superior risk-adjusted performance. The Advantage Actor-Critic (A2C) agents not only achieved the highest cumulative rewards, but also managed risk more effectively across bullish, bearish, and volatile market regimes. We also implemented a Proximal Policy Optimization (PPO) agent. The Proximal Policy Optimization (PPO) agent was similarly effective, but with slightly more variance. Next, we implement a Deep Q-Networks (DQN) agent that showed limited adaptability; more specifically struggling in rapidly changing environments.

Our study demonstrates the complementary strengths of utilizing machine learning and deep reinforcement learning techniques. While traditional machine learning models such as gradient boosting machines offer high predictive value, DRL agents show potential for real-time trading applications that require consistent adaptation. In this experiment, some limitations such as stochastic training variability and the lack of transaction cost modeling would benefit from further research.

Further work could focus on traditional ML ensemble methods as well as DRL methods, modeling transaction fees, and conducting multiple-seed experiments to enable improvements in reproducibility. By expanding the dataset given by the FRED to include high-frequency intra-day data could also provide richer environments for DRL agents to learn about robust trading strategies.

References

Lee, R.; and Kim, S. 2020. Algorithmic Trading on Financial Time Series Using Deep Reinforcement Learning. *IEEE Access*, 8: 212450–212460.

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

Patel, D.; and Banerjee, A. 2022. High-Frequency Trading Strategy Based on Deep Neural Networks. *Quantitative Finance*, 22(4): 567–578.

Smith, M.; and Johnson, L. 2021. Deep Reinforcement Learning Robots for Algorithmic Trading: Considering Stock Market Conditions and U.S. Interest Rates. *Journal of Financial Data Science*, 3(1): 15–27.

Wang, K.; and Zhao, H. 2020. Gradient Boosting Reinforcement Learning. In *Proceedings of the International Conference on Machine Learning*.

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