Artificial Intelligence Research and Development
I. Sanz et al. (Eds.)

doi:10.3233/FAIA230667

Navigating Black Swan Events in Algorithmic Trading: A Reinforcement Learning Perspective

Fernando VILLAMARÍN DÍAZ [a] Carlos GUERRERO-MOSQUERA [a]
[a] *HER - Human-Environment Research Group, La Salle - Universitat Ramon Llull, Barcelona, Spain, Quatre Camins, 30, 08022 Barcelona, Spain*
ORCiD ID: Fernando Villamarín Díaz https://orcid.org/0009-0000-1877-5902, Carlos Guerrero-Mosquera https://orcid.org/0000-0001-8265-3651

Abstract. This study evaluates Reinforcement Learning (RL) techniques for financial trading during unpredictable market conditions, such as black swan events. Three experiments were conducted: one where the algorithms were trained and tested over the same period; another where they were trained and tested over different periods; and the final one where they were trained over a certain period and then tested during a period that included a black swan event (the market crash of March 2020). Results show that RL methods outperform traditional strategies in the in-sample period, but struggle to adapt during the black swan event. The results show the potential of RL techniques in financial trading with the right approach.

Keywords. AI applications, reinforcement learning, financial trading, black swan events, stock market, algorithmic trading, portfolio management

1. Introduction

Financial markets, often assumed to be efficient, have experienced significant bubbles and crashes, revealing their vulnerability to rare, high-impact "black swan" events. Over the past two decades, the financial industry has undergone a technological transformation, with algorithmic trading becoming increasingly prevalent. Simultaneously, advancements in **Machine Learning**, particularly **RL**, have shown promising results across various problem domains.

Despite growing interest in applying RL algorithms to trading, their performance during black swan events remains uncertain. This study aims to investigate the application of RL algorithms to trading by capturing the state of the market through technical indicators and exploring their performance during unexpected market downturns. Additionally, the study compares the performance of **RL-based trading strategies** with traditional man-made trading strategies to evaluate their efficacy and potential advantages.

2. Literature Review

Reinforcement learning (RL) is a learning paradigm where an agent learns to map situations to actions to maximize cumulative reward through trial and error [9] . Despite its increasing prominence, research on RL's applications in finance remains limited.

Few studies have investigated RL in financial settings, such as the application of inverse RL in order book dynamics [6] and the development of granular trading simulators for RL agents [1]. Deng et al. [3] stand out for applying **Deep Reinforcement Learning (DRL)** to financial trading, demonstrating the effectiveness of their proposed system in summarizing market conditions and learning optimal actions.

New approaches include extending the DRL framework to handle **multiple assets** and learn portfolio management strategies, as well as developing a method for intelligently selecting appropriate training periods to address the **non-stationary** nature of financial markets [3].

3. Methodology and Results

This study utilizes **daily stock price data** for JPMorgan Chase & Co. (ticker: JPM) from January 1, 2000, to January 1, 2023. This specific time period was selected due to its availability on Yahoo Finance [12]. The data is transformed into **technical indicators**—price-EMA ratio, %B, and momentum—and then discretized into n bins to capture the market's state. This state information is subsequently fed into the models, which use it to determine their actions: either buy, sell, or do nothing. The market simulator iteratively repeats this process for all the training days in the sample, drawing from historical data.

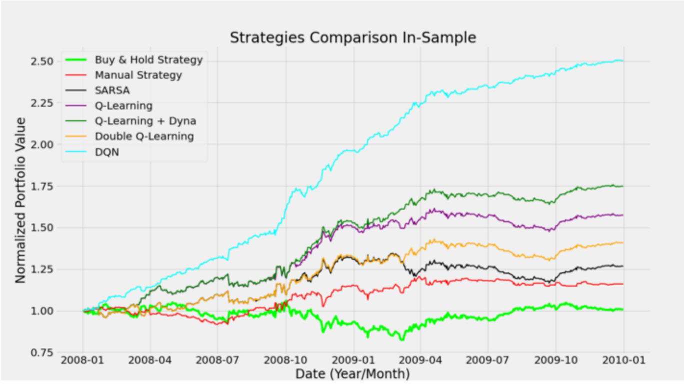

Figure 1. In-sample performance comparison of RL methods, manual, and buy-and-hold strategies. The DQN strategy outperforms all, with a Sharpe ratio of 4.75 and a 150% portfolio value increase. Other RL strategies also outperform the manual (Sharpe ratio: 0.604) and buy-and-hold (Sharpe ratio: 0.112) strategies. SARSA yields the lowest Sharpe ratio (0.810) among RL methods.

The authors have developed a **market simulator** that reads financial data from a CSV file to simulate trading over a specified period. The simulator calculates the daily portfolio value, which is the sum of stock holdings and cash. It also incorporates transaction costs such as a set commission fee per trade and accounts for slippage, representing the adverse price movement a trader might experience in a real market. The simulator restricts the range of stock holdings to between -1000 and 1000 shares. The reinforcement learning model integrates with the market simulator by transforming state elements (price-EMA ratio, %B, momentum, and holdings) into a single value using a **discretization formula**. Values have been previously discretized into n bins, for this article, $n = 10$ was used.

$$\text{State} = \text{Portfolio} \times n^3 + \%B \times n^2 + \text{Momentum} \times n + \frac{\text{Price}}{\text{EMA}} \tag{1}$$

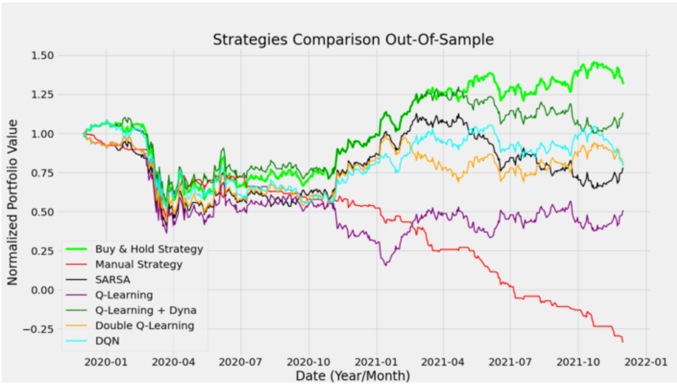

Figure 2. Out-of-sample performance comparison. DQN underperforms with a Sharpe ratio of -1.471, suggesting possible overfitting. Other RL methods continue to outperform manual and buy-and-hold strategies, though the hypothesis that RL models will outperform these strategies in the out-of-sample period is not fully validated.

The models used are **SARSA** [5], **Q-Learning** [11], **Dyna Q-Learning** [8], **Double Q-Learning** [10], and **Deep Q Learning (DQN)** [4].

The simulations use a straightforward **reward function** based on daily returns, allocating a positive reward of +1 when the algorithm generates a profit and -1 when it incurs a loss. This function guides the trading agent's learning process and decision-making within the simulated market environment. To evaluate the performance of the reinforcement learning methods, two benchmark strategies are employed: a **buy-and-hold strategy**, where the trader buys at the beginning of the time period and maintains the position, and a **manual strategy**, which combines technical indicators with simple logic to generate trading signals. These serve as comparison points for assessing algorithmic trading performance.

These experiments were designed to comprehensively evaluate the performance of the algorithms under a variety of conditions in the financial market. The **in-sample experiment** (Experiment 1) aimed to verify the learning capabilities of the RL algorithms in a controlled environment using known data. **Out-of-sample testing** (Experiment 2) was conducted to assess the generalizability of these methods to unseen data and their ability to avoid overfitting. Finally, the **black swan event experiment** (Experiment 3) was crafted to test the resilience and adaptability of these models during extreme market upheavals, such as those characterized by the **COVID-19 pandemic**. The timeframe for this last experiment is more restricted, focusing solely on the market crash of March 2020. To ensure a fair comparison, we optimized the hyperparameters of all models using a grid search approach, focusing solely on the in-sample data. Each experiment was run 100 times to account for variability. The results were then ranked based on the final Sharpe ratio. Graphs were created using data from the models that landed in the median rank for a more robust representation.

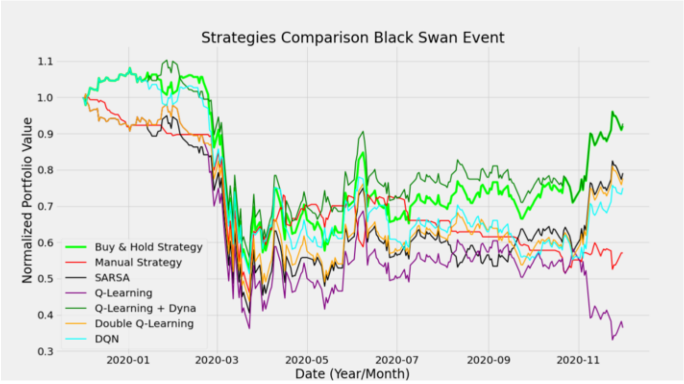

Figure 3. Performance comparison during the black swan event (COVID-19 pandemic). No RL methods outperform manual and buy-and-hold strategies. The buy-and-hold strategy tops with a Sharpe ratio of 0.224. Q-Learning performs the worst among RL methods (Sharpe ratio: -0.580). All strategies yield negative returns, indicating inability to predict or mitigate a black swan event.

4. Discussion and Conclusion

The purpose of this study was to evaluate the **effectiveness of RL techniques** for financial trading and to determine if these methods perform better than traditional strategies during unstable and unpredictable market conditions, such as black swan events. The study focused on experimenting with several value-based RL methods and compared them to a basic ”buy-and-hold” strategy and a manual strategy using the same technical indicators.

In the second experiment, RL methods performed significantly better than the benchmark strategies during the in-sample period. However, in the third experiment, the **Deep Q Network (DQN)** method suffered from overfitting, leading to a decrease in performance. The other RL methods, namely the tabular methods, performed much better, indicating a certain level of generalizability.

In the fourth experiment, **none of the RL methods** were able to adapt their behavior or take preventive measures during the black swan event, the market crash in March 2020 due to the COVID-19 pandemic. Despite mixed outcomes, the results demonstrate the potential of these algorithms, showing promise and effectiveness with the right approach.

Despite the promising results of this study, there are several areas of exploration for future research. Extensive model tuning could potentially have enhanced the performance of the RL methods used in this study, but the complexity of financial trading makes this a challenging task. Particularly for the DQN model, future work could focus on implementing techniques to combat overfitting, such as L2 regularization [2] and neural network dropout [7], as proposed in prior literature.

References

[1] David Byrd, Maria Hybinette, and Tucker Hybinette Balch. ABIDES: Towards High-Fidelity Market Simulation for AI Research. 4 2019.

[2] Corinna Cortes, Mehryar Mohri, and Afshin Rostamizadeh. L 2 Regularization for Learning Kernels. Technical report, 2012.

[3] Yue Deng, Feng Bao, Youyong Kong, Zhiquan Ren, and Qionghai Dai. Deep Direct Reinforcement Learning for Financial Signal Representation and Trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3):653–664, 3 2017.
[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with Deep Reinforcement Learning. 12 2013.
[5] Mahesan Niranjan. On-Line Q-Learning Using Connectionist Systems Machine Learning in Fixed Income Markets View project Deep Learning based visual object Tracking View project. Technical report, 1994.
[6] Jacobo Roa-Vicens, Cyrine Chtourou, Angelos Filos, Francisco Rullan, Yarin Gal, and Ricardo Silva. Towards Inverse Reinforcement Learning for Limit Order Book Dynamics. 6 2019.
[7] Nitish Srivastava. Improving Neural Networks with Dropout. Technical report, 2013.
[8] Richard S. Sutton. Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bulletin*, 2(4):160–163, 7 1991.
[9] Richard S Sutton and Andrew G Barto. Reinforcement Learning An Introduction second edition. Technical report, 2018.
[10] Hado Van Hasselt. Double Q-learning. Technical report, 2010.
[11] Christopher J C H Watkins and Peter Dayan. Technical Note Q,-Learning. Technical report, 1992.
[12] Yahoo Finance. JPMorgan Chase & Co. (JPM) Stock Price, News, Quote & History, 2023.