Reproducibility and Analysis of "Practical Deep Reinforcement Learning Approach for Stock Trading" "Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy"

Dixin Li
Siyuan Chen
Weiyue Cai

School of Computer Science
McGill University

April 17, 2024

Overview

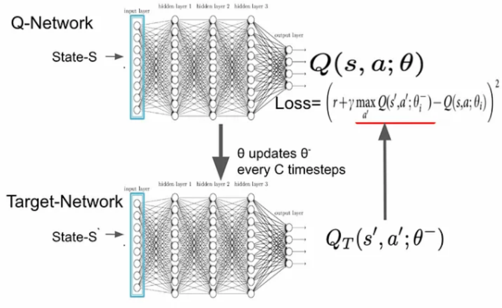
Article Summary

These two papers explore the application of reinforcement learning algorithms to develop adaptive trading strategies for the stock market. The first article focuses on the Deep Deterministic Policy Gradient (DDPG) algorithm, while the second introduces an ensemble strategy that integrates three algorithms: Proximal Policy Optimization (PPO), Advantage Actor Critic (A2C), and DDPG.

Project Objectives

- Explore actor-critic reinforcement learning algorithms to gain a deeper understanding of their mechanisms and applications.
- Learn how to define and construct states and actions within a trading environment using the OpenAI API.
- Discuss the limitations identified in the original study and explore potential opportunities for further improvements.

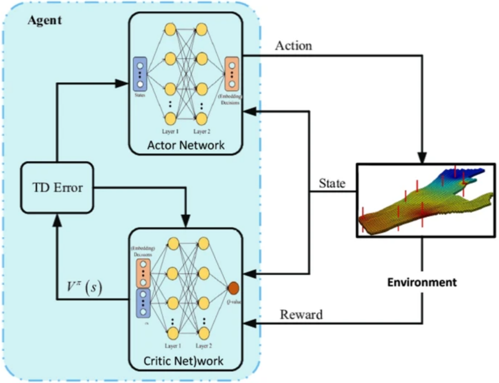
DQN



The idea of the Deep Q-learning (DQN) is to use a Q-value function to learn the optimal action-selection policy that maximizes the expected future reward given the current state.

Reference: `https://arshren.medium.com/deep-q-learning-a-deep-reinforcement-learning-algorithm-f1366cf1b53d`

Actor-Critic: A2C, DDPG, PPO, TD3, SAC



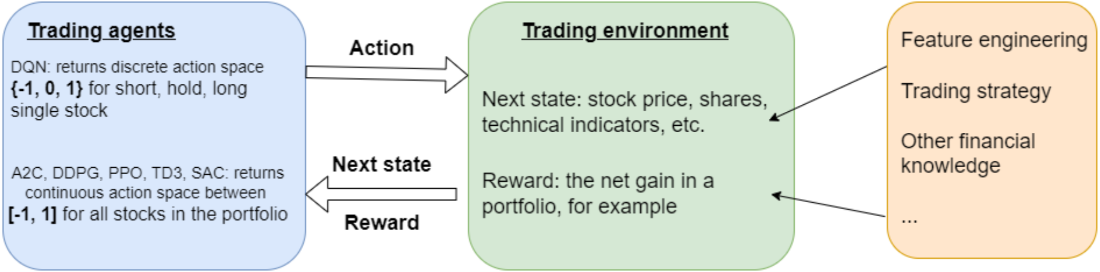
Actor-critic architecture and its interaction with the environment

In the actor-critic method, we have two main components working together: the actor and the critic. The actor network is responsible for representing and updating the policy, which dictates the actions to take. On the other hand, the critic network evaluates how good those actions are by estimating the value function.

Reference: `https://link.springer.com/article/10.1007/s00521-023-08537-6`

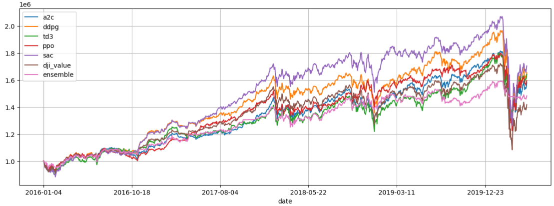
Articles Summary

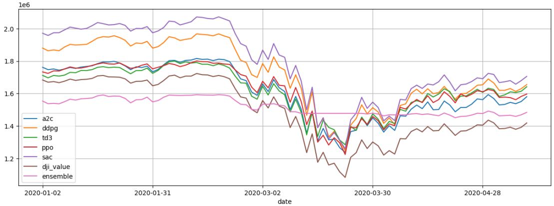
The action space is represented by a 29-dimensional vector. Each element of this vector corresponds to the number of shares we can buy or sell for each stock.
The state space is a multidimensional vector that includes several key components: portfolio value, stock close prices, the number of stock shares, and various technical indicators. This structured approach allows for a comprehensive view of the trading environment.



Results (1): Ensemble strategy as the most resilient during the stock market downturn

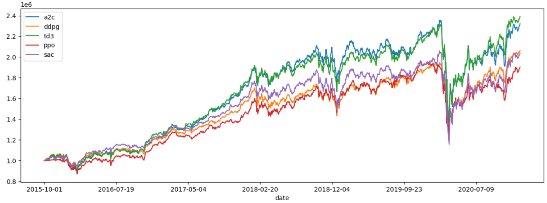
To ensure uniform environments for both methods, we adopt the default settings. The ensemble strategy is the most resilient during the stock market crash.

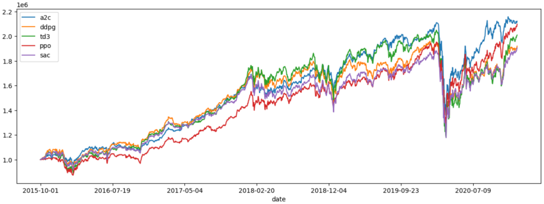


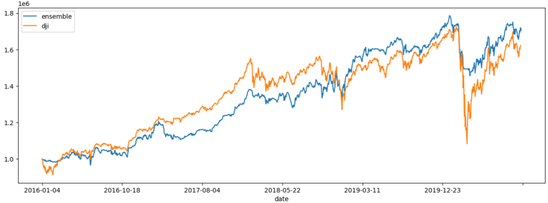


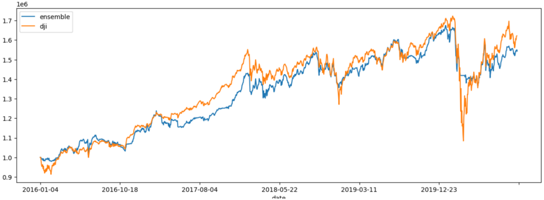
Results (2): Non-deterministic outcomes with same settings

first run and second run for single DLR and ensemble strategies: training from 2009-1-1 to 2015-9-30, trading from 2015-10-1 to 2020-5-8



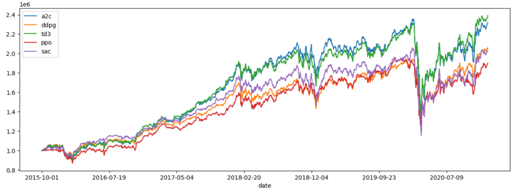


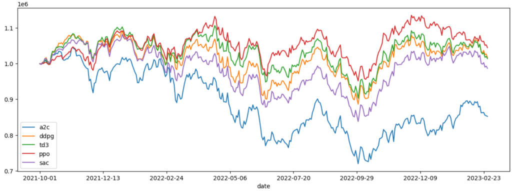


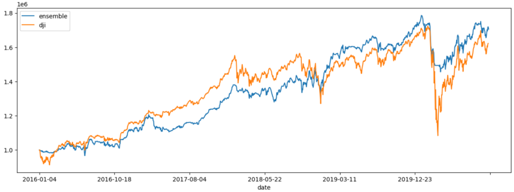


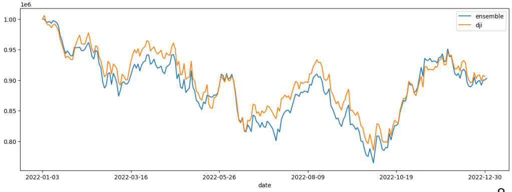
Results (3): Impact of Training Data on Results

First Period: training data: from January 1, 2009, to September 30, 2015, test data: from October 1, 2015, to May 8, 2020.
Second Period: training data: from January 1, 2010, to October 1, 2021, test data: from October 1, 2021, to March 1, 2023.









Conclusion

- The design of the state and action spaces is critical when applying Deep Reinforcement Learning to stock trading, potentially even more crucial than the choice of the algorithm itself.
- The current simplistic modeling of trading environment may not fully capture the complexities of real-world trading environments.
- Enhancing the sophistication of both state and action representations could significantly boost the system's effectiveness and adaptability in dynamic market conditions.
- It's crucial to develop a method that allows for multiple runs in a real-time trading environment to ensure consistently stable results.