

# Portfolio Management Using RL

Di He, Congyun Jin

# Portfolio Management

**Portfolio** is a collection of financial assets (cash, stocks, bonds..)

**Portfolio Management** is a decision making process of continuously re-allocating an amount of fund into a number of financial assets

- Goal: Maximize the return while restraining the risk
- Decisions/Actions: The investment weights for the M assets at every decision point t.

$$w_t = [w_{1,t}, w_{2,t}, \dots, w_{M,t}]^T \in R^M \quad s.t. \sum_{i=1}^M w_{i,t} = 1$$

**Previous Works** has did some RL experiments on single stock trading but less works focus on stocks portfolio management. We planed to do more experiments on stock portfolio management and explore the impact of observation window size.

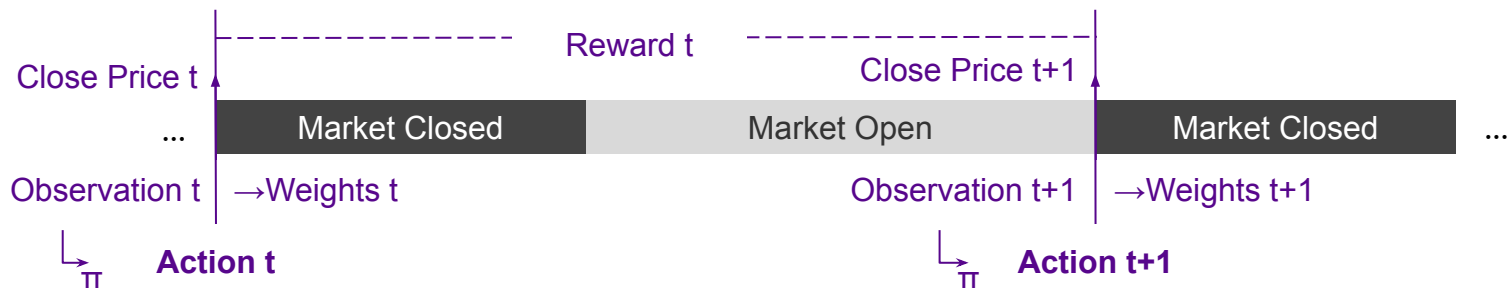
# Translate into a RL Problem

## Observation:

- Historical  $n$ -days prices for candidate assets ( $n$  depends on window size).
- Latest portfolio weights.

**Actions:** The portfolio weights. An asset with an increased target weight will be bought, and that with decreased weight will be sold.

**Rewards:** The average of logarithmic daily portfolio returns which also consider transition costs.



# Model and Data

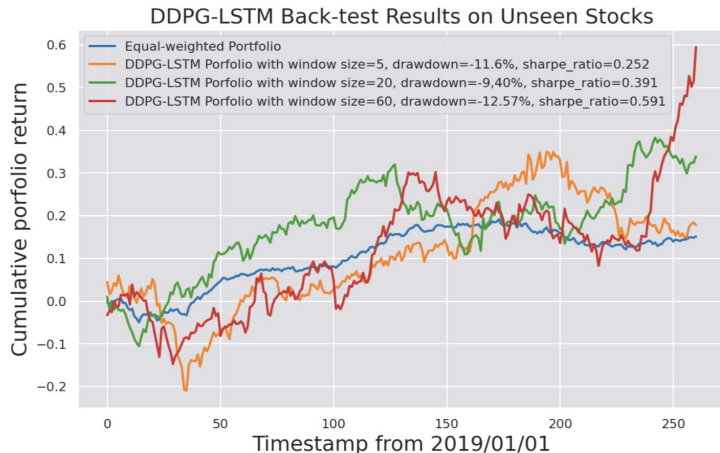
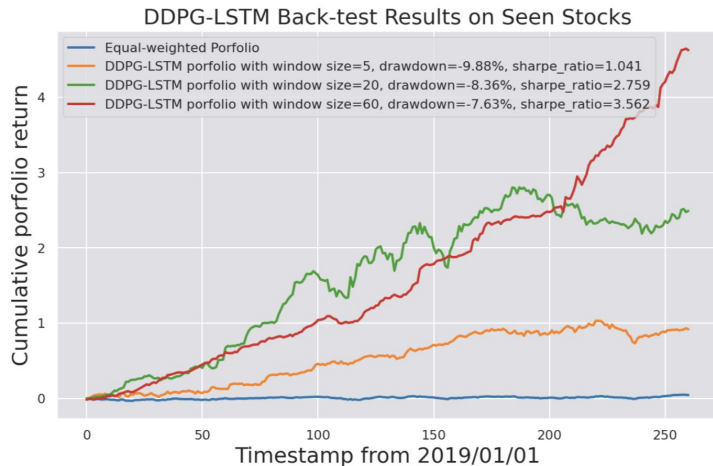
## Model

- Because the reward can be explicitly calculated and the reward of one timestamp only depends on the action of this timestamp (independent of actions of other timestamps), we want to do more exploitation rather than exploration, and finally choose to use **DDPG** (Deep Deterministic Policy Gradient).
- Considering our data is time series, we experimented DDPG with **CNN** and **LSTM**.
- We did some experiments on different **window sizes** which define our observation.

## Data

- To balance risk: we choose 17 S&P 500 stocks from 5 industries from 2015-01-02 to 2020-01-01. The train and test split point is at 2019-01-01.
- To check the generalization of our strategy: we also choose another unseen 17 stocks to test the agent.

# Results and Future Work



- DDPG with LSTM has a better performance than CNN
- **Longer window size** has a better performance
- Our RL agent on other unseen stocks is not as profitable as seen stocks
- In the future, we plan to explore the generalization of our method on a longer economic cycle. We also plan to take risk into the reward calculation.

**Thanks**