

Homework2: Exchange Rate Investment

CE6020 Artificial Intelligence (2023 Fall)

vinming852@gmail.com

111525018@cc.ncu.edu.tw

111522155@cc.ncu.edu.tw

Description

- **Task**

- 利用強化學習 (Reinforcement Learning) 進行外匯投資決策
- **Environment:** Gymnasium [宣網](#)

- **Data**

- **檔案名稱** - train.csv

- **檔案名稱** - test.csv

- **Date** : 日期

- **現鈔買入** : 現鈔買入價格

- **現鈔賣出** : 現鈔賣出價格

- **即期買入** : 即期買入價格

- **即期賣出** : 即期賣出價格

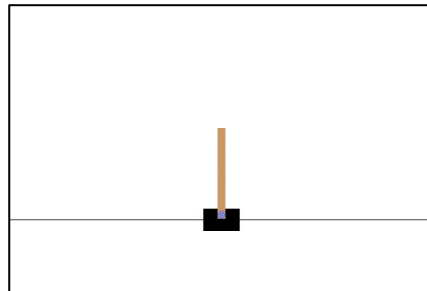
Date	AUD現鈔買入	AUD現鈔賣出	AUD即期買入	AUD即期賣出
2023/9/13	20.11	20.89	20.4	20.6
2023/9/12	20.19	20.97	20.48	20.68
2023/9/11	20.19	20.97	20.48	20.68
2023/9/8	20.08	20.86	20.37	20.57
2023/9/7	20.05	20.83	20.34	20.54
2023/9/6	20.01	20.79	20.3	20.5
2023/9/5	19.94	20.72	20.23	20.43
2023/9/4	20.22	21	20.51	20.71
2023/9/1	20.24	21.02	20.53	20.73

Gymnasium

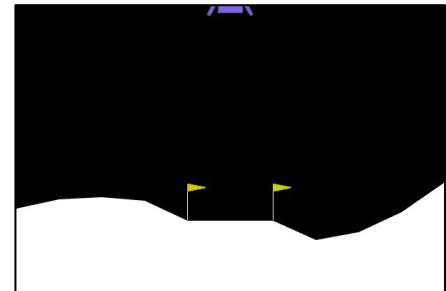
- 開源的強化學習 (Reinforcement Learning) 套件
 - 提供許多環境 (Environment), 藉以處理環境模擬及獎勵等過程
- <https://www.gymnasium.dev/>



Assault



Cart Pole

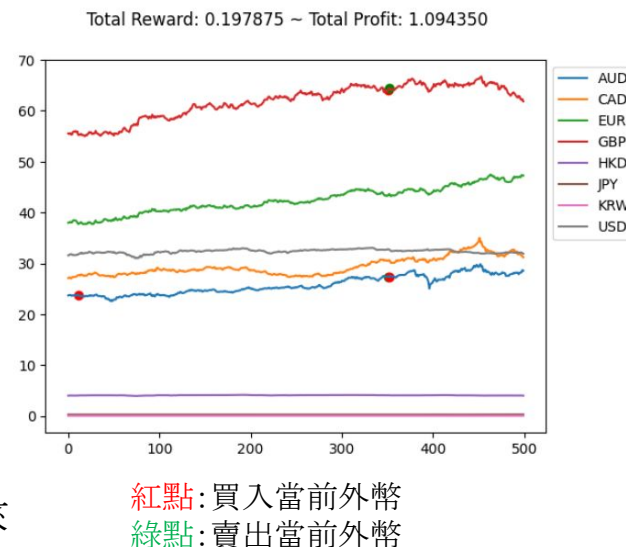


Lunar Lander

Gymnasium

• Environment 規則

- 有 8 種外幣+台幣一共 9 種
- 某個時間點 i 只能同時擁有一種幣
- 初始本金 = 1 元台幣
- **Observation:** 第 $i - 10$ 日 ~ 第 i 日的外匯資料 ($i \geq 10$)
- **Action:** 買入其中一種外幣 or 買入台幣總共 9 種 actions
 - 買入新貨幣時會先賣出原貨幣再買入新貨幣(都是以現鈔來看)
 - 買入貨幣會以"現鈔賣出"資料, 賣出貨幣會以"現鈔買入"資料做計算, 中間價差即為銀行手續費。(因為現鈔賣出、現鈔買入皆為銀行面相)
 - 若買入的幣別和持有的幣別相同, 則不做任何動作
- **Position:** 目前持有的幣別, 一共 9 種
- **Total Reward:** 持有台幣時為 0, 持有其他外幣, 則計算此外幣匯率第 i 日和 $i - 1$ 日的之間的漲跌幅, 每天累加後的總合
- **Total Profit:** 資金成長比 = 最終本金



$$\Rightarrow \frac{price_i - price_{i-1}}{price_{i-1}}$$

Submission

- 請將檔案請包成一個 zip 壓縮檔, 並繳交至 ee-class 平台
 - 壓縮檔名格式為 學號.zip (e.g. 112525000.zip)
 - 在你的 學號.zip 內至少要有下列 5 類檔案
 - report.pdf 你的報告
 - train.ipynb 訓練主程式
 - model.ckpt 模型 / 參數檔案(只能繳交一個模型讓助教評分)
 - README.md 說明要如何執行訓練主程式
 - requirements.txt 套件列表
 - Python 版本指定 3.10.x
- **Deadline:** 2023/11/16 23:59
 - 遲交: 本次作業成績 = 原始成績 $\times (1 - (\text{遲交天數} \times 0.1))$
 - 檔名錯誤: 本次作業成績 = 原始成績 - 5分

Grading Policy(Update)

• Programming (60%)

$$\bullet \text{ Score}_{\text{Testing}}(20\%) = \begin{cases} 20 \% & \text{助教能正確執行你的程式} \\ \leq 10 \% & \text{助教不能正確執行你的程式} \end{cases}$$

$$\bullet \text{ Score}_{\text{Ranking}}(40\%) = \begin{cases} 40 \% & \text{Total Profit Top 1 ~ 15} \\ 30 \% & \text{Total Profit Top 16 ~ 30} \\ 20 \% & \text{Total Profit Top 31 ~ 50} \\ 10 \% & \text{else} \end{cases}$$

$$\bullet \text{ Score}_{\text{base line}} = \begin{cases} 30 \% & \text{Total Profit} > 1.05 \\ 20 \% & \text{Total Profit} > 1.0 \end{cases}$$

base line和排名擇
優給分

$$\text{ProgrammingScore}(60\%) = \text{Score}_{\text{Testing}}(20\%) + \max(\text{Score}_{\text{Ranking}}(40\%), \text{Score}_{\text{base line}})$$

• Report (40%)

• Template Link:

<https://docs.google.com/document/d/18veUn-fNpalBUrHUafDAX0E6lhQZgLDqC82BWlXuLls>

• 請標明學號與姓名，並根據題號回答問題。

• 請輸出 A4 大小的 PDF 檔，檔名為 report.pdf (**Report.pdf** 是不正確的)

Report

1.(10%) Policy Gradient 方法

1. 請閱讀及跑過範例程式, 並試著改進reward 計算的方式。
2. 請說明你如何改進 reward 的算法, 而不同的算法又如何影響訓練結果?

2.(15%) 試著修改與比較至少三項超參數(神經網路大小、一個batch 中的回合數等), 並說明你觀察到什麼。

3.(15%) 請同學們從 Q Learning、Actor-Critic、PPO、DDPG、TD3 等眾多 RL 方法中擇一實作, 並說明你的實作細節。

4.(10%) 請具體比較(數據、作圖等)你實作的方法與Policy Gradient 方法有何差異, 並說明其各自的優缺點為何。

Grading Policy (cont.)

- Programming 評分細節

- 助教將從外匯交易資料集抽樣進行評分
- 從資料集抽樣短、中、長期各 3 筆資料, 以 **Total Profit** 作為該筆資料的評估標準
 - 短期=100天、中期=300天、長期=800天(抽樣的區間將在 **Deadline** 後擇日公布)
 - 考慮到每次執行結果不同, 同筆資料將執行 10 次並取平均
- 助教在批改程式部分時, 會執行以下指令:
執行 `train.ipynb` 中, `score` 以下的部分
 - 執行環境: Colab (請確認在 colab 上是否可正常執行)
 - 測試階段, 會讀取各位的 `model.ckpt` 來運行程式

Overview

- 利用強化學習 (Reinforcement Learning) 進行外匯投資決策
- 必須在Gymnasium 這個環境進行外匯投資
- 允許同學以範例程式碼為基礎進行修改

- code example:

https://colab.research.google.com/drive/1ZX94acJFBHYV-VNAixQ7_S4GGshvMqkb#scrollTo=VowxXC2zy-Ht&uniqifier=1

- **Deadline:** 2023/11/16 23:59

Reference / Resource

- [Gym Documentation](#)
- [Github: AminHP/gym-anytrading](#)
- [Github: datawhalechina / easy-rl](#)