

機器學習實務與應用

Homework #12

Due 2019 June 10 9:00AM

使用 OpenAI Gym 做為強化學習的環境，並自行撰寫 Q-learning、及 Policy gradient 的演算法進行訓練。

以下是使用的 OpenAI Gym 的環境及需要實作的強化學習方法

Part 1: FrozenLake8x8-v0

<https://gym.openai.com/envs/FrozenLake8x8-v0/>

使用 Q-learning

Part 2: Pong-v0

<https://gym.openai.com/envs/Pong-v0/>

使用 Policy gradient

作業需求：

如果有更動環境所給的 Reward 也寫下改變的方法

畫出訓練過程中的 Total Reward 的變化。X 軸為 Episode，Y 軸為 Reward

儲存訓練完後的模型參數

計算訓練模型所需要花的 Episode

嘗試設定不同超參數所造成的變化並記錄下來。

評分標準：

使用訓練完後的模型，執行 50 次 Episode 計算平均的 Total Reward 作為依據。

需畫出執行過程中的 Total Reward 的變化。X 軸為 Episode，Y 軸為

Reward，及顯示最後的平均值。

批改方式：

請繳交壓縮檔，檔名為 ML_HW7_<學號>.zip

其中需包含的檔案及資料夾為：

1. 資料夾 Part1

ML_HW7_part1_<學號>.ipynb

ML_HW7_part1_<學號>_test.py

2. 資料夾 Part2

ML_HW7_part2_<學號>.ipynb

ML_HW7_part2_<學號>_test.py

批改作業時不會執行 .ipynb 的檔案，**也不會重新訓練模型**，這個檔案內容須包含模型產生的方法，..._test.py 則存放訓練完畢的模型，為實際用來進行測試的檔案。