

# HW5

## Basic

Q learning in basic part, 每次iteration都會進行以下幾個步驟

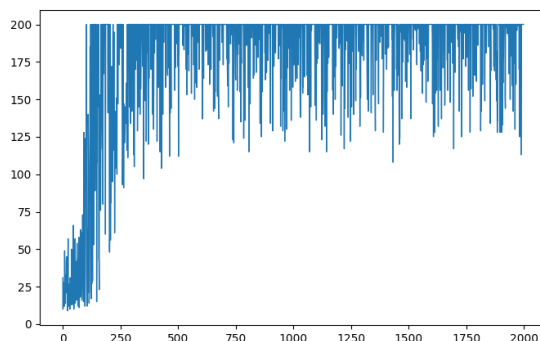
- 利用Q table, 決定本次iteration所要進行的action
- 把決定好的action套入environment中, 並取得reward與observation
- 和之前已取得的reward累加, 計算出截至目前為止取得了多少reward
- 利用action套入environment後取得的reward更新Q table
- 根據此state做出的action所取得的結果更新state, 並重複以上步驟

Hyperparameter

- N\_EPISODES = 2000
- EPISODE\_LENGTH = 2000
- position of cart = 5
- velocity of cart = 1
- angle of pole = 8
- rotation of pole = 8

Result

- 根據result結果發現其實N\_EPISODES設太多次了, N\_EPISODES設定在250~500即可



# Advanced

Deep Q learning in basic part, 每次iteration都會進行以下幾個步驟

- 根據DQN和STATE選擇接下來的action
- 將選到的action套入environment並取得next\_state和reward
- 將經驗(state action reward next\_state)存起來
- 當memory存夠多經驗時, agent就能以Q learning學習
- 將state轉換成next\_state, 並重複iteration

Hyperparameter

- n\_hidden = 200
- batch\_size = 10
- lr = 0.01
- epsilon = 0.2
- gamma = 0.99
- target\_replace\_iter = 500
- memory\_capacity = 1000

Result

