**DRL學習筆記\_DQN篇**

強化學習(Reinforcement Learning, RL)的原理是由代理人(agent)藉由與環境的互動(**回饋值**)，學習如何因應**狀態**的改變來採取行動，如下圍棋，打電玩等。若這類神經網路模型加上深度學習的做法，則統稱為深度強化學習(Deep Reinforcement Learning, DRL)演算法。

**DQN的概念**

強化學習是一種Markov決策過程(MDP)，重要的元素是執行環境輸出的回饋值(reward)及狀態(state)，學習模型中的代理人依據這兩個元素組合產生決策，並決定行動(action)，例如:迷宮遊戲中，決定向左或向右前進一步的行動。

但隨著多種組合與順序的不同，模型計算將會非常複雜，所以產生了Q-Learning價值迭代法(value iteration)的演算法，即為每次採取行動時，依據結果來更新這個行動的價值，然後利用這個更新的價值去決定下一個策略。

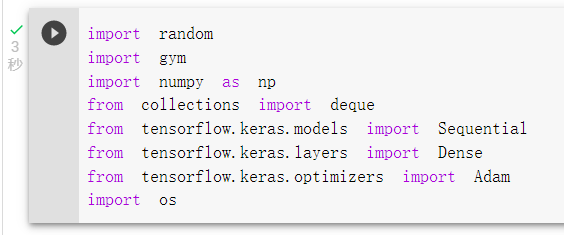
DQN(Deep Q Network)即為一種結合神經網路與Q-Learning的演算法。

**DQN設計實作範例**

使用Cola開放平台(可參考文末說明)練習此次的課題，訓練在木棒平衡車環境中行動，遊戲規則維保持木棒於滑車保持平衡，不會掉下來。我們套用Python OpenAI Gym函式庫中的遊戲環境來實作。

實作流程如下:

#### 步驟一， 載入相關模組。



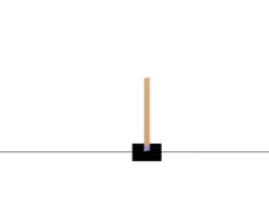
關鍵套件:gym。

#### 步驟二，設定參數



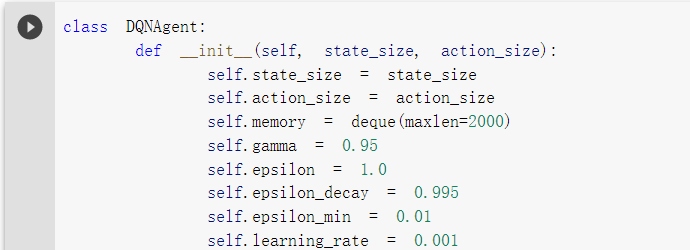
1.Open AI Gym的make()指定代理人執行的環境為木棒平衡滑車第0版(CarPole-v0)。

2.state\_size:狀態資訊的維度，此例中為4維(滑車位置，滑車速度，木棒角度，木棒角速度)



3.action\_size:可選的動作數，我們設計滑車僅能向左或式向右。

#### 步驟三，定義DQN代理人



初始化參數

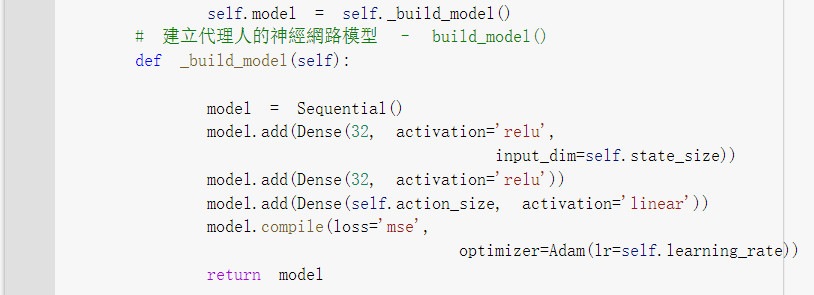
1.memory:儲存每一時步的資訊，我們設為2000筆，所以若加入第2001筆時，會將第1筆去掉。

2.gamma:折扣係數，用來將得到的回饋值打折，例如:0.95即為打95折。

3.epsilon:探索率，指【代理人不依循在遊戲累積的經驗來行動，而是隨機行動】的機率。

4. epsilon\_min: 探索率的下限。

5.learning\_rate:修正權重參數時的學習率。



模型架構

1.第一，二層隱藏層為DNN，由32個神經元組成，使用ReLU激活函數。

2.輸出層使用linear線性激活函數，回傳向左或是右的Q值。

3.本案例為回歸模型，使用MSE為損失函數，優化器則為Adam。



記憶遊戲的資訊:

1.代理人狀態(state) 。

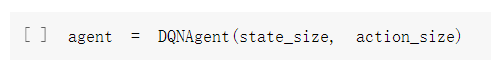
2.採取的行動(action) 。

3.環境針對行動的回饋(reward) 。

4.環境依據代理人的行動而反應出新狀態。

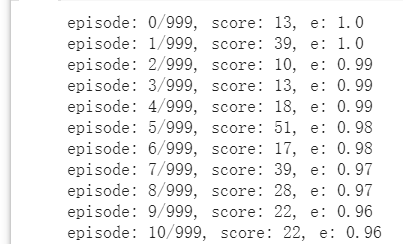
5.是否為最後一步。

#### 步驟四，以函數式API方式，建構鑑別器神經網路

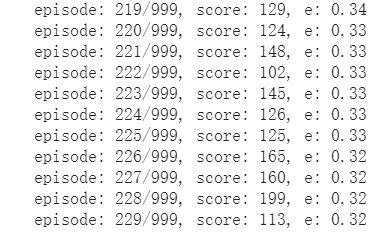




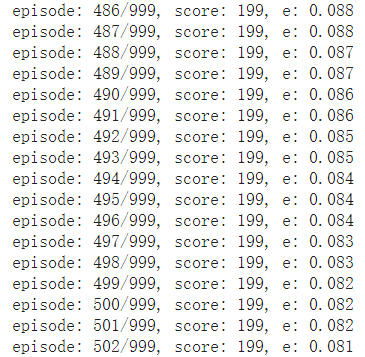
#### 步驟五，檢視訓練結果



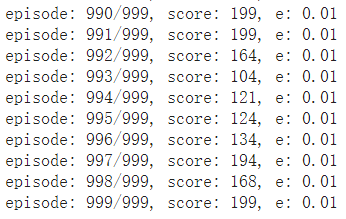
前10回合分數不高，最高僅有39分，探索率在第10回合為0.96，代表僅有4%的機率在使用學來的經驗值。



第219回合分數開始穩定破100分，探索率在第219回合為0.34，代表有66%的機率在使用學來的經驗值。

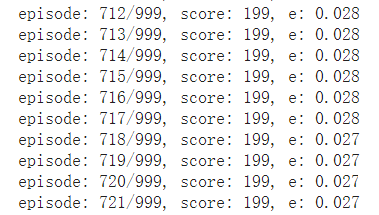


第486~502回合分數都有199分，探索率在第486回合為0.88，代表有12%的機率在使用學來的經驗值。



最後10回合分數有的有199分，但也有突然走鐘，探索率在第990回合為0.01，代表有99%的機率在使用學來的經驗值，代表代理人有些不穩定，當發生這種情況，可以使用agent.load()還原模型至表現較好的參數。

例如:

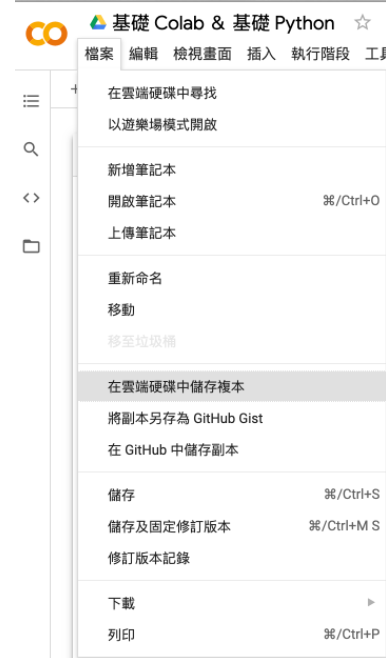


#### **附錄:前置作業Colab環境準備**

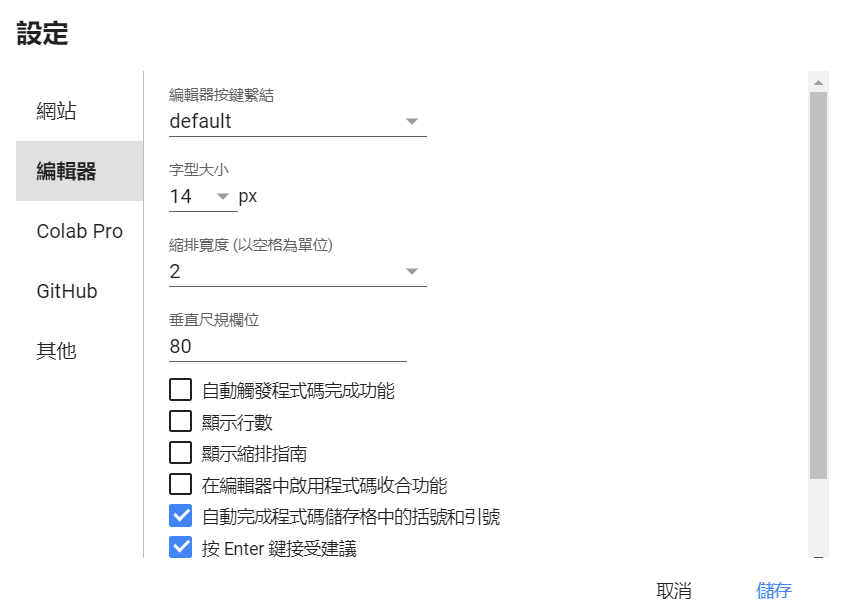
1.必須有自己的google帳號:使用自己的個人的 Google 帳號才可以執行自己 Colab上的程式。

2.先下載 Chrome瀏覽器，並登入 Google 帳號。

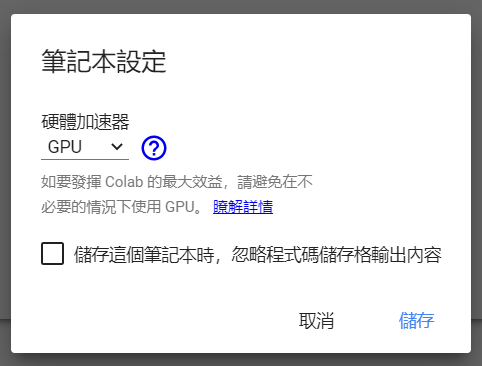
3.建立副本至個人目錄，否則不能存檔。



4.在⼯具 > 設定>編輯器中,反勾選「⾃動觸發程式碼完成功能」。



5.(必須做)設定執行階段類型。



6.療癒模式(非必須)，在⼯具 > 設定>其他, 勾選 「柯基⽝模式」和「貓咪 模式」，就有可愛的貓 貓狗狗來陪你寫程式了。



