

# ADLxMLDS HW3 Report

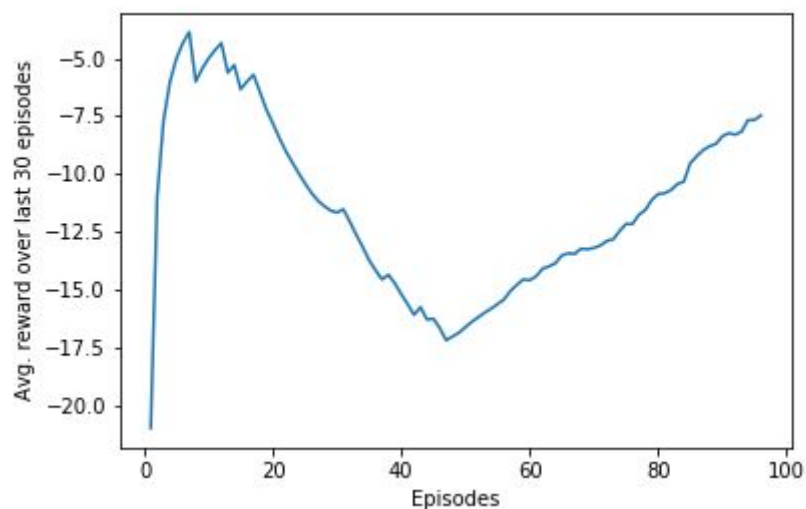
B03902086 資工四 李鈺昇

## Model Descriptions

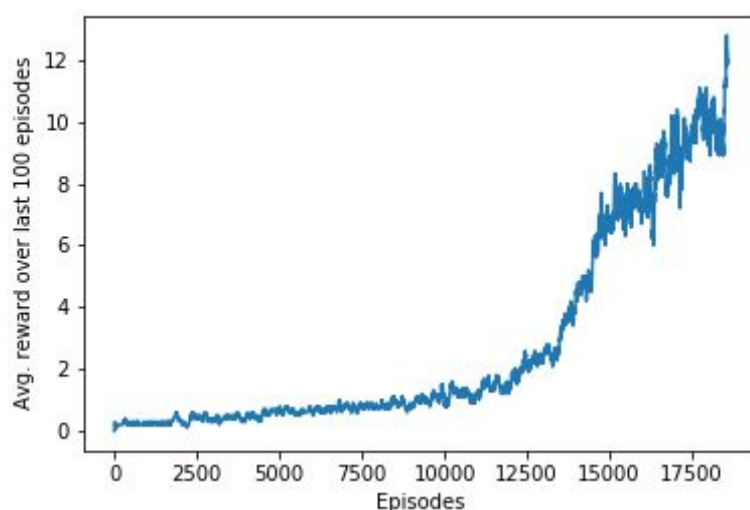
- PQ :  
Conv2d(in\_channels=1, out\_channels=16, kernel\_size=8, stride=4)  
ReLU  
Conv2d(in\_channels=16, out\_channels=32, kernel\_size=4, stride=2)  
ReLU  
Flatten  
Linear(2048, 128)  
ReLU  
Linear(128, 6)  
Softmax
- DQN :  
Conv2d(in\_channels=4, out\_channels=32, kernel\_size=8, stride=4)  
ReLU  
Conv2d(in\_channels=32, out\_channels=64, kernel\_size=4, stride=2)  
ReLU  
Conv2d(in\_channels=64, out\_channels=64, kernel\_size=3, stride=1)  
ReLU  
Flatten  
Linear(3136, 512)  
ReLU  
Linear(512, 4)

## Learning Curves

- PG

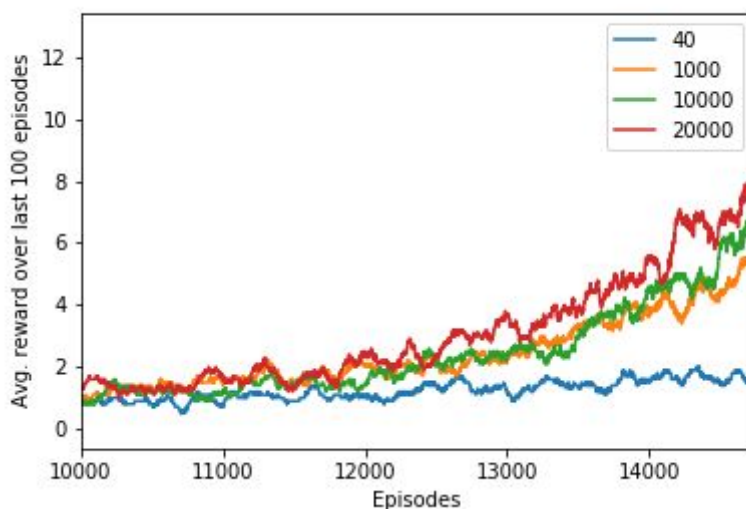


- DQN



## Experiments with Hyperparameter

我覺得 experience replay 的 memory 大小應該會影響結果，因此除了原本的 10000 之外分別試了 40、1000、20000 三種大小，結果如下圖所示：



由圖可見，除了 40 看起來幾乎學不太到東西之外，另外三種都有明顯的進步。而到了後期三者的 average reward 是跟 memory size 呈正相關，並且差距有增大的趨勢。

我覺得讓 memory size 變大的好處在於 model 可以看到分布比較廣的 data，不會只看到最近的（比較好的），而是連過去的（即使比較爛、reward 較低）也會看得到，這樣應該可以學地比較全面。