

# Hw4 code review report

111062503 吳冠志

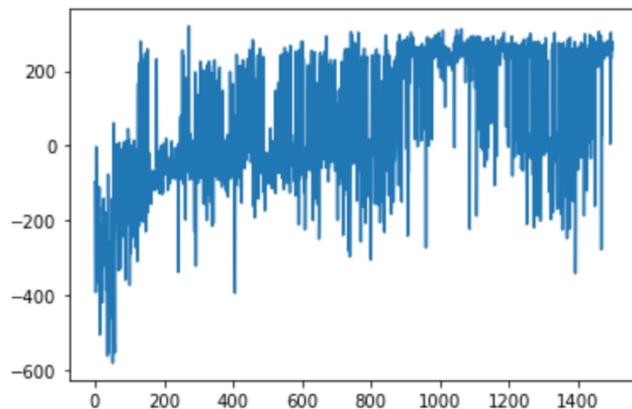
## 1. Model

```
class network(nn.Module):
    def __init__(self , num_state , num_action):
        super().__init__()
        self.layers=nn.Sequential(
            nn.Linear(num_state , 64 ),
            nn.LayerNorm(64),
            nn.ReLU(),
            nn.Linear(64 , 128)
            nn.LayerNorm(128),
            nn.ReLU(),
            nn.Linear(128 , 32)
            nn.LayerNorm(32),
            nn.ReLU(),
            nn.Linear(32 , num_action ),
            nn.LayerNorm(num_action),
            nn.Softmax()
        )
    def forward(self , x):
        return self.layers(x)
```

本次共訓練 1501 個 episode

## 2. Reward

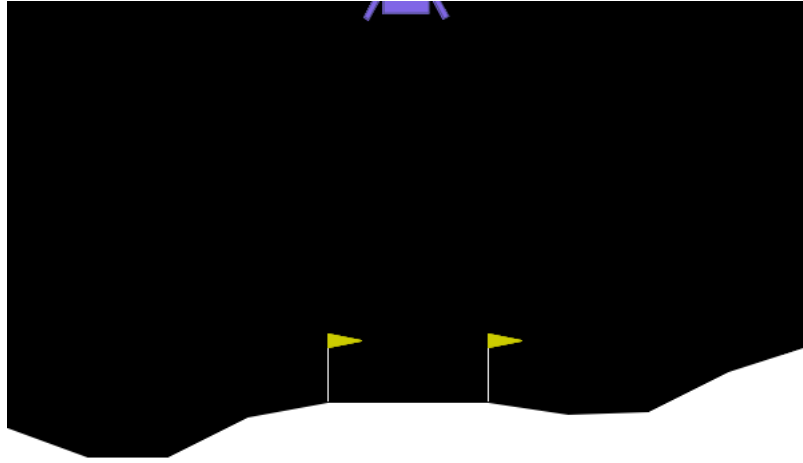
```
Out[29]: [<matplotlib.lines.Line2D at 0x7fc046c68c10>]
```



雖然各 episode 之 reward 不穩定，但明顯看出隨著訓練次數增加，其 reward 有逐漸提升，從一開始的-600~0 到最後其 reward 約落在-200~+200 間，可看出其學習成效

### 3. 各 Episode

0



1500

