

# HW1 検討

# 參考作法-把正確率提升策略

改變optimizer (SGD  $\rightarrow$  Adam )不過效果不彰  
將 learning rate 往下調整(ex. 0.01  $\rightarrow$  0.001  $\rightarrow$  0.0001)  
Random Seed 一個一個慢慢try...

# 評分標準

- Report (50%)
  - Method description – What is your strategy for parameter selection?--10
  - Experimental results--20
    - Two figures of accuracy and loss curves
    - Testing accuracy
  - Discussion--10
  - Problem and difficulties--10
- Source Code (50%)

# Q&A

## 1. Q: Colab跑出結果崩壞

- A:我們實測也是一樣，所以不會影響作業分數。

## 2. Q:修改模型架構後，仍無法更有效提升，卡在86-87% accuracy，像是加”dropout”和”batch normalization”.

- A: 1. 兩個一起使用會有衝突的產生可能  
2. Ensemble 的可能性，將不同模型的結果利用投票方式選出最大可能。  
ex. 五個模型(A,B,...,E)  
 $A \approx 1, B \approx 1, C \approx 1, D \approx 3, E \approx 1$ ，投票結果為1

## 3. Q:Training Loss跑出nan

- A:因為”lr”過大導致”梯度爆炸”，代表網路不能正常的收斂
- A:其他可能原因”資料過髒”、“網路架構問題”