

## DeepQ HW3 - Report

資料科學碩一 R07946007 陳庭安

- a. 做了那些比較的實驗
- b. 最後選擇此值 / 此 scheduling 的原因

### Resnet50

	Seed	Lr	batch norm	batch size	dropout	epoch	Train_loss Val_loss	Train_acc Val_acc
1	123	1e-2	0.9	32	0.2	8	0.0131 0.3338	0.9992 0.9060
2	123	1e-2	0.9	32	0.5	8	0.0193 0.3513	0.9998 0.8980
3	123	9e-3	0.9	32	0.2	18	0.0043 0.3517	1.0000 0.9060

Round 1. 發現 round 1 只跑到 epoch 8，training accuracy 遠大於 validating accuracy，有 overfitting 的現象。

Round 2. 故在 round 2 調高 dropout rate，多刪減沒有表現力的 nodes，期望降低參數數目。結果 overfitting 的現象沒有改善，表現也不如預期，甚至比 round 1 差。

Round 3. 再將 dropout 調回 0.2，改調降 learning rate，讓它走慢點，不要走過頭，多走幾個 epoch，別錯過 minimum，果然 performance 就有比較好。只是仍存在 overfitting 的風險。

### Mobilenetv2

	Seed	Lr	batch norm	batch size	dropout	epoch	Train_loss Val_loss	Train_acc Val_acc
1	123	1e-3	0.9	20	0.3	10	1.1664 0.8507	0.5768 0.6860
2	123	1e-2	0.9	20	0.3	22	0.1963 0.4178	0.9320 0.8780

3	123	3e-2	0.9	20	0.3	8	0.0043 0.3517	0.9352 0.8800
4	123	3e-2	0.9	20	0.5	8	0.2611 0.4376	0.9064 0.8800
5	123	5e-2	0.9	20	0.5	11	0.0832 0.5562	0.9704 0.8840
6	123	7e-2	0.9	20	0.5	9	0.1061 0.5371	0.9708 0.8440

Round 1-3. 發現 accuracies 上升很慢，故慢慢調升 learning rate

Round 4. 發現有 overfitting 的風險，調升 dropout rate，出來的結果確實改善許多，因此維持 dropout rate 再 0.5

Round 5-6. 希望表現能夠更好，所以就調降一點 learning rate、多跑幾個 epoch。結果 loss 下降，accuracies 也有上升。

### **Kaggle 表現 / Testing accuracies**

Public: 0.88583

Private: 0.89562

上述調參結果在 public 的表現雖還不到 strong baseline，甚至有些距離，但在 private 過了 strong baseline，原因可能跟調整 dropout 降低 overfitting 風險的策略有關。