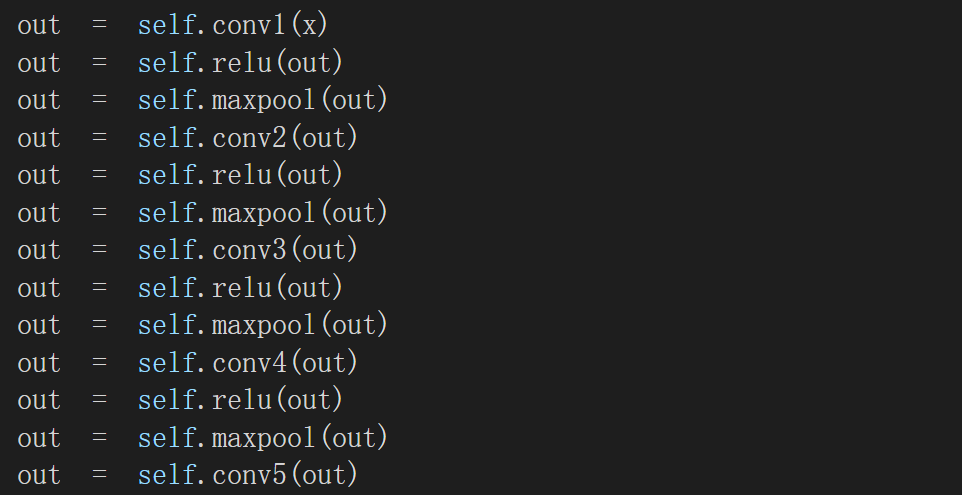
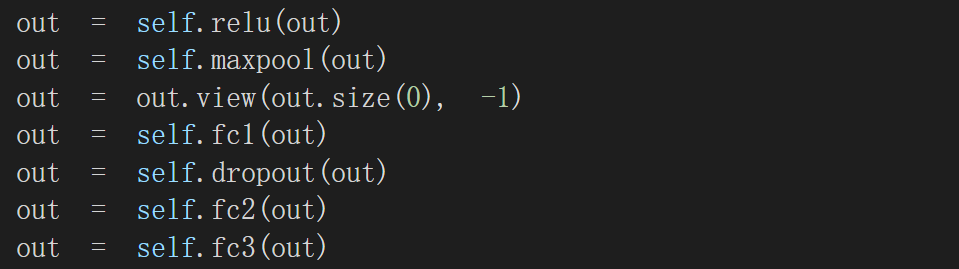
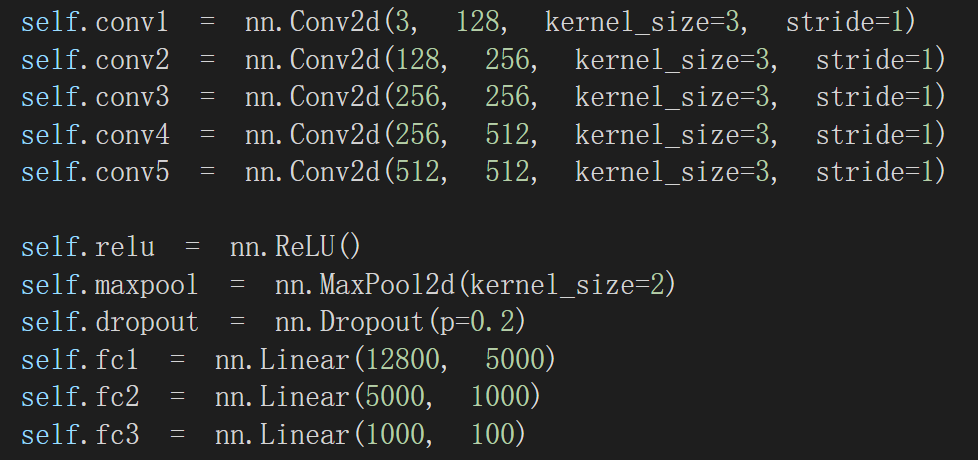
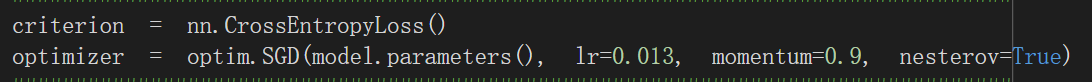
Model 1

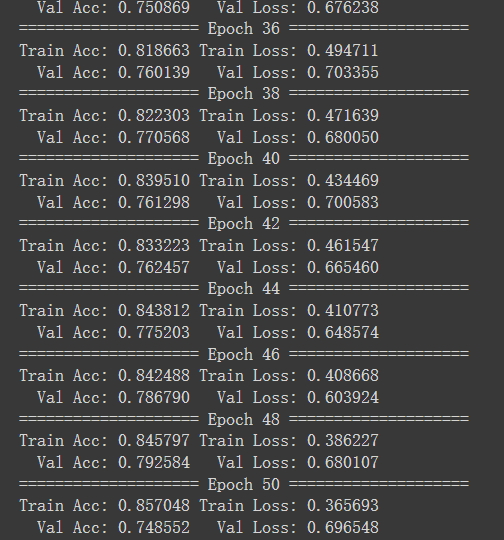
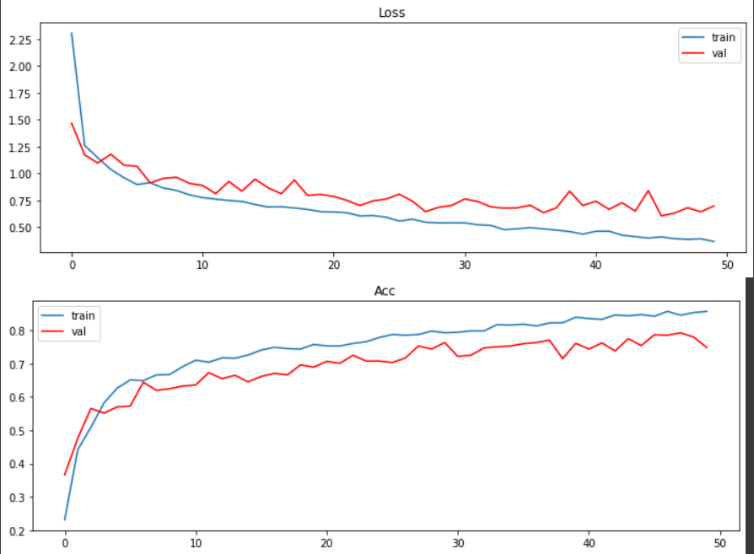
* Model layer : 5
* Full connect layer : 3



* Activation function : ReLU
* Optimizer function : SGD
* Nesterov 加速算法
* Learning rate : 0.013

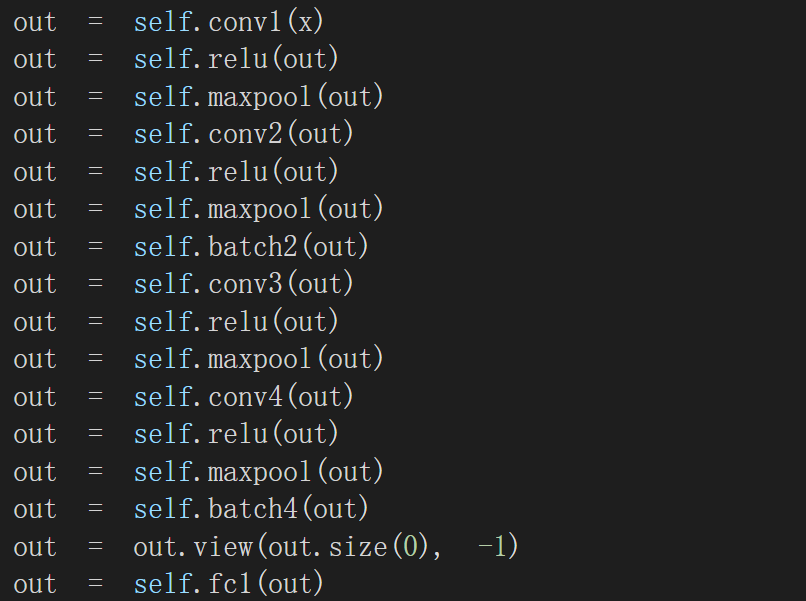
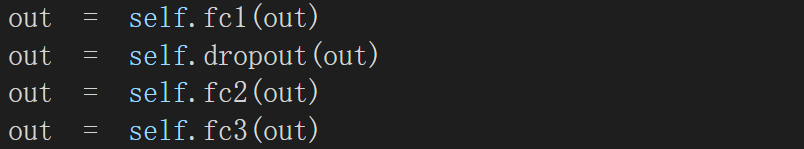


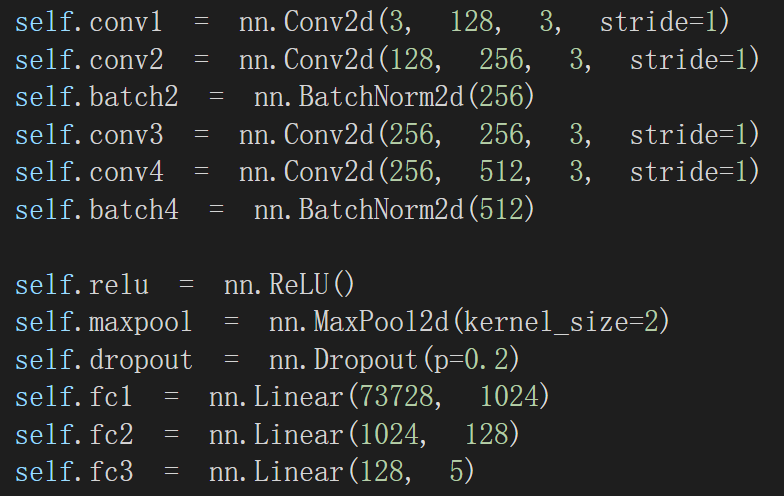
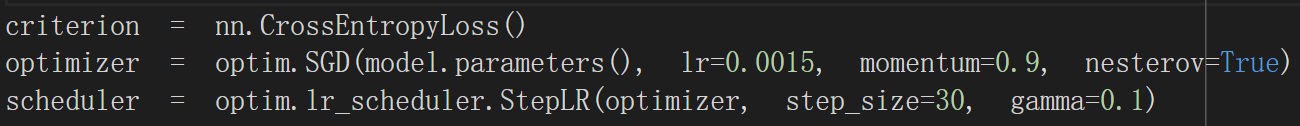
執行結果 :

* Epoch : 50
* Kaggle 測試結果 : 0.77076

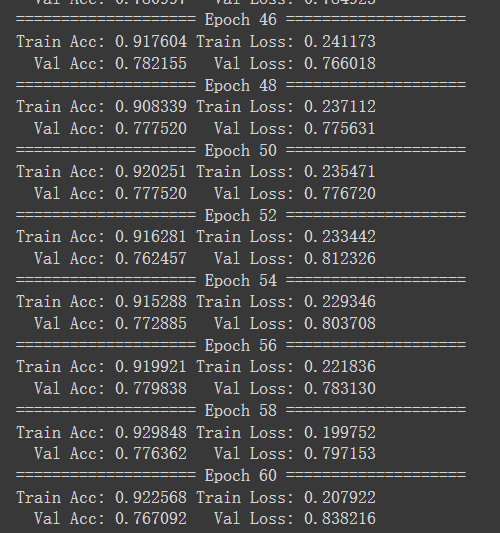
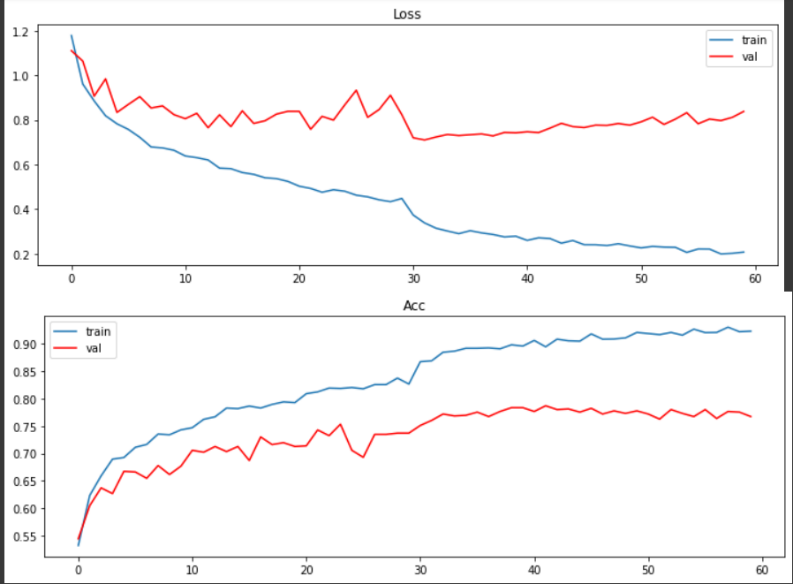
Model 2

* Model layer : 4
* Full connect layer : 3



* Activation function : ReLU
* Optimizer function : SGD + Schedular
* Nesterov 加速算法
* Learning rate : 0.0015

執行結果 :

* Epoch : 60
* Kaggle 測試結果 : 0.80730

Model 1 與 Model 2 的差異

1. Model Layer 減少 (5 → 4)
2. 加入2層batch，更新參數
3. Optimizer 搭配 schedular
4. Learning rate 調小 (0.013 → 0.0015)