Deep Learning Homework2

StudentID: 107062650 Name: 曾昱榮

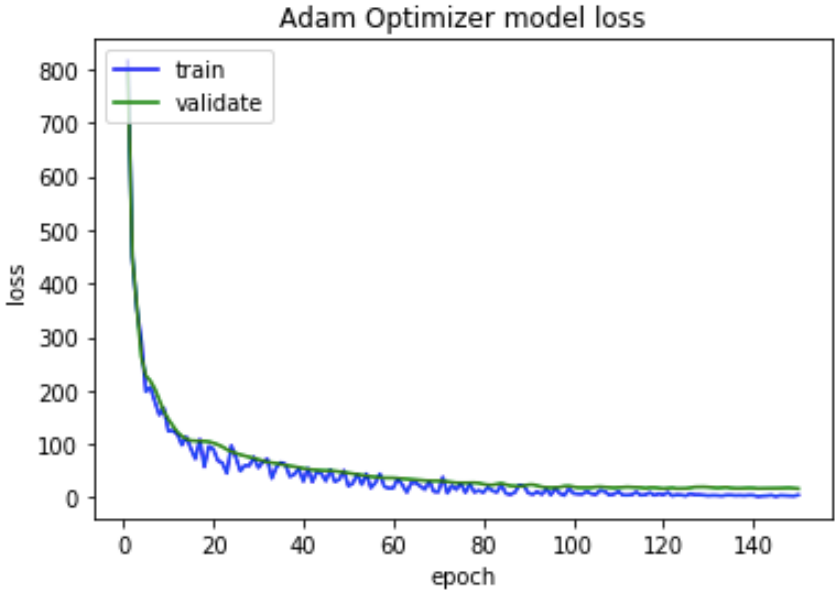
1.

(i)Construct a DNN for multi-class classification.

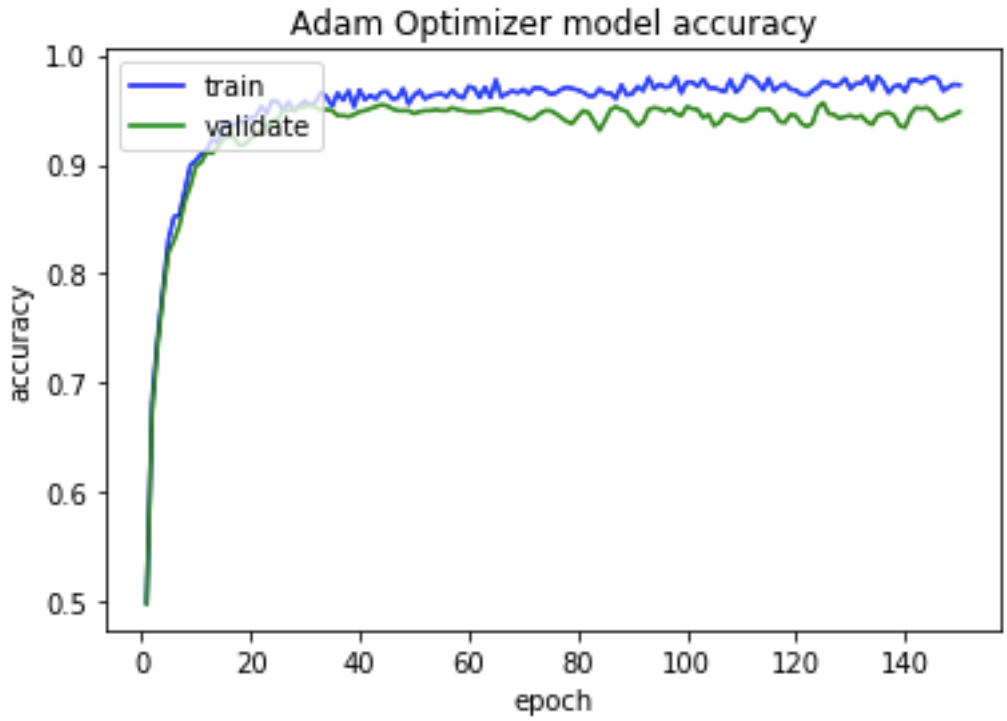
我建置的 DNN model 為有兩層分別皆有 256 個 neurons 的 hidden layer 的 CNN，learning_rate 為 0.01，訓練的 epoch 數目為 150，batch_size 為 1500。下圖為訓練過程中的 training loss 和 training accuracy。而最終 model 的 testing accuracy 為 0.962。

```
Adam Optimizer start
Step 1, Minibatch Loss= 1709.9467, Training Accuracy= 0.296
Step 10, Minibatch Loss= 238.7294, Training Accuracy= 0.853
Step 20, Minibatch Loss= 193.9726, Training Accuracy= 0.929
Step 30, Minibatch Loss= 78.3385, Training Accuracy= 0.955
Step 40, Minibatch Loss= 48.7880, Training Accuracy= 0.959
Step 50, Minibatch Loss= 37.5486, Training Accuracy= 0.963
Step 60, Minibatch Loss= 31.5224, Training Accuracy= 0.964
Step 70, Minibatch Loss= 22.2587, Training Accuracy= 0.963
Step 80, Minibatch Loss= 30.1856, Training Accuracy= 0.971
Step 90, Minibatch Loss= 24.1903, Training Accuracy= 0.970
Step 100, Minibatch Loss= 12.2513, Training Accuracy= 0.961
Step 110, Minibatch Loss= 9.6873, Training Accuracy= 0.969
Step 120, Minibatch Loss= 10.4345, Training Accuracy= 0.970
Step 130, Minibatch Loss= 4.3485, Training Accuracy= 0.977
Step 140, Minibatch Loss= 6.1443, Training Accuracy= 0.972
Step 150, Minibatch Loss= 5.0283, Training Accuracy= 0.969
Optimization Finished!
```

下圖為 training loss、testing loss 在每個 epoch 的折線圖。



下圖為 training accuracy、testing accuracy 在每個 epoch 的折線圖。



(ii) record precision, recall and F1-score as well as the averages of those criteria

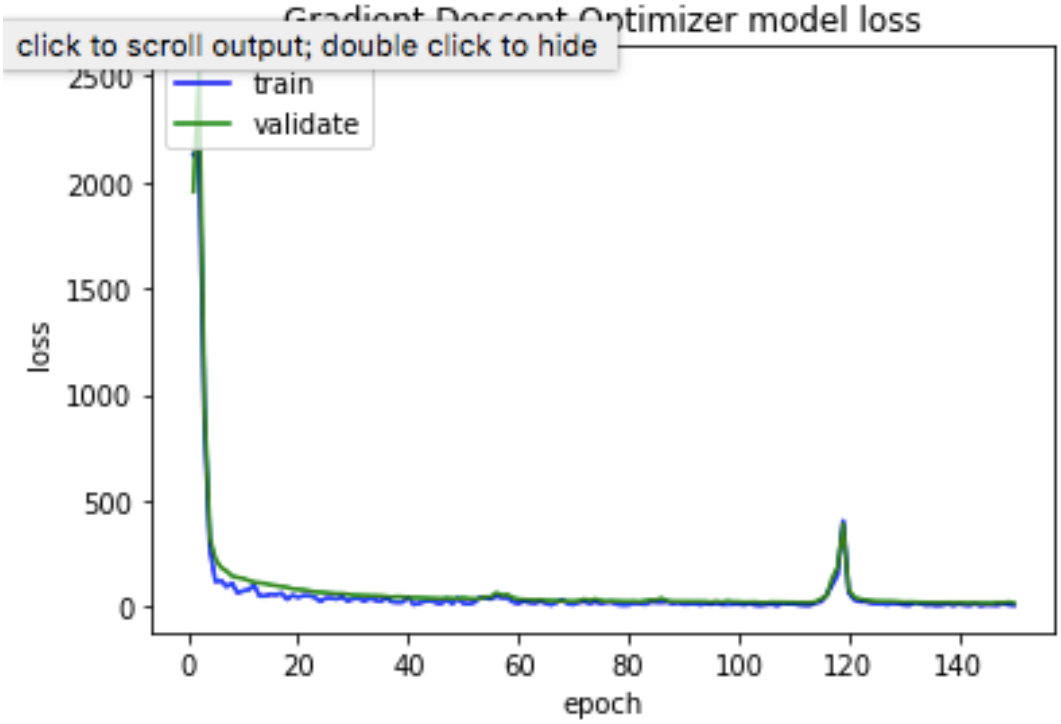
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Micro	Macro
Precision	0.900	0.991	0.946	0.883	0.945	0.996	0.956	0.944
Recall	0.675	0.834	0.731	0.757	0.896	0.844	0.815	0.790
F-score	0.771	0.906	0.825	0.815	0.920	0.914	0.880	0.860

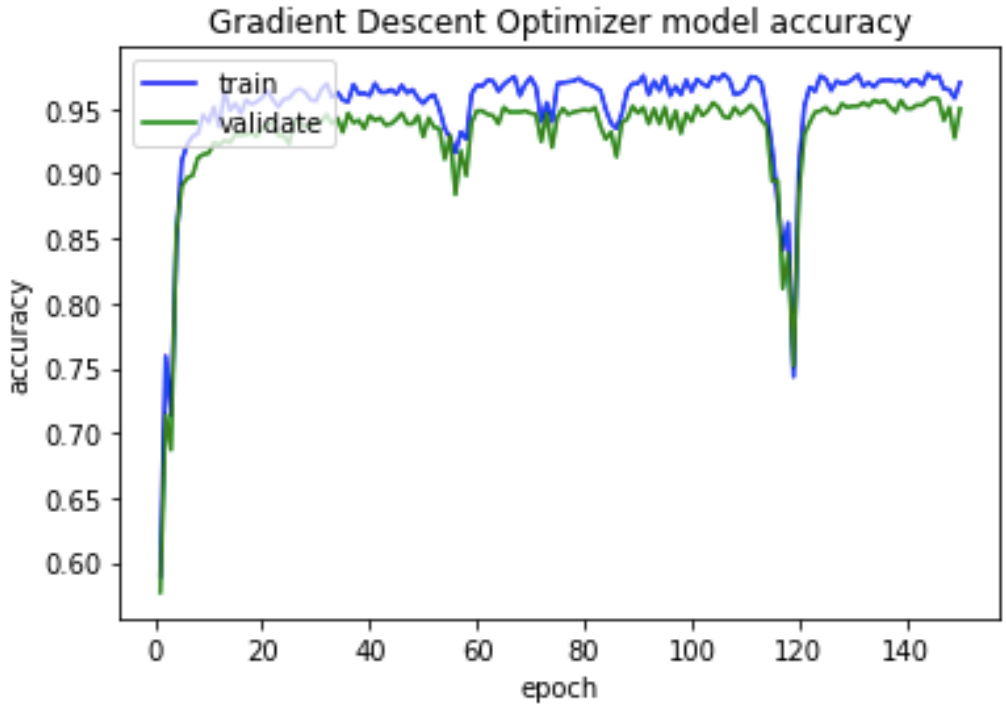
(iii) Repeat (i), please compare with different optimizers (Gradient descent, Ada-grad). Explain what you observe.

這部分Gradient descent和Ada-grad model的batch size、learning rate皆與(i)相同。

根據觀察，Gradient descent一開始的training loss下降速率為三者最快，但在訓練過程後半段其training loss和training accuracy會有較大幅度的動盪，可得知Gradient descent為較不穩定的optimizers，其最終testing accuracy為0.959。訓練過程如下圖。

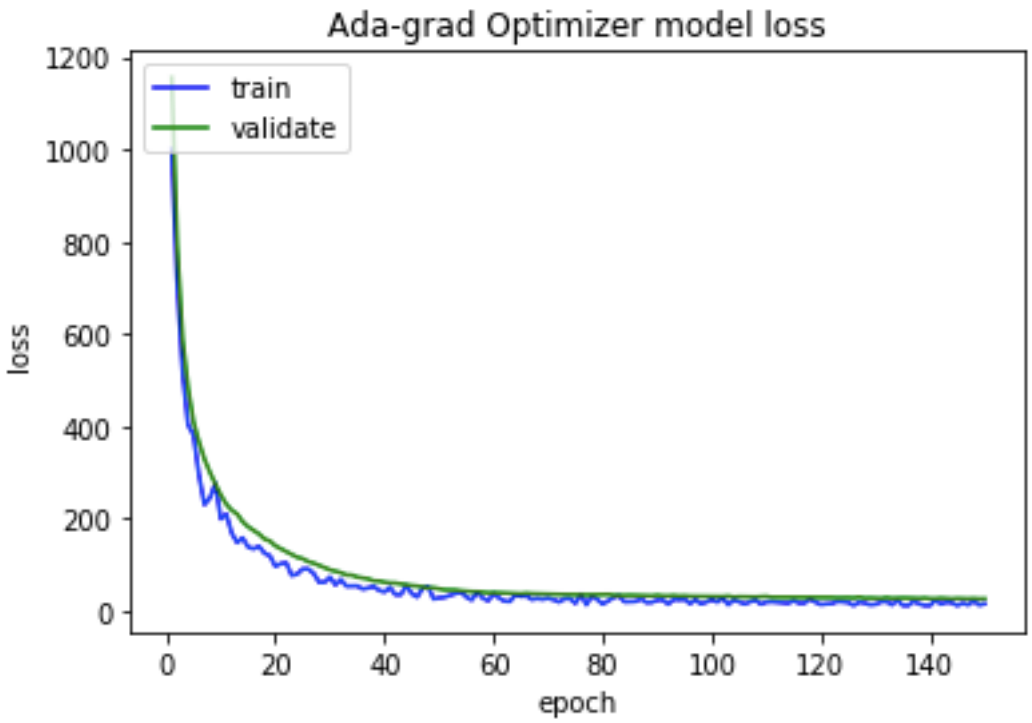
```
Gradient Descent Optimizer start
Step 1, Minibatch Loss= 2127.1929, Training Accuracy= 0.589
Step 10, Minibatch Loss= 74.7339, Training Accuracy= 0.939
Step 20, Minibatch Loss= 44.2047, Training Accuracy= 0.959
Step 30, Minibatch Loss= 37.6674, Training Accuracy= 0.956
Step 40, Minibatch Loss= 46.5174, Training Accuracy= 0.960
Step 50, Minibatch Loss= 17.1376, Training Accuracy= 0.954
Step 60, Minibatch Loss= 15.9458, Training Accuracy= 0.968
Step 70, Minibatch Loss= 10.7349, Training Accuracy= 0.974
Step 80, Minibatch Loss= 9.7558, Training Accuracy= 0.970
Step 90, Minibatch Loss= 12.4154, Training Accuracy= 0.969
Step 100, Minibatch Loss= 13.4725, Training Accuracy= 0.963
Step 110, Minibatch Loss= 13.9458, Training Accuracy= 0.965
Step 120, Minibatch Loss= 68.3201, Training Accuracy= 0.907
Step 130, Minibatch Loss= 10.0755, Training Accuracy= 0.973
Step 140, Minibatch Loss= 14.0125, Training Accuracy= 0.967
Step 150, Minibatch Loss= 6.8217, Training Accuracy= 0.970
Optimization Finished!
```

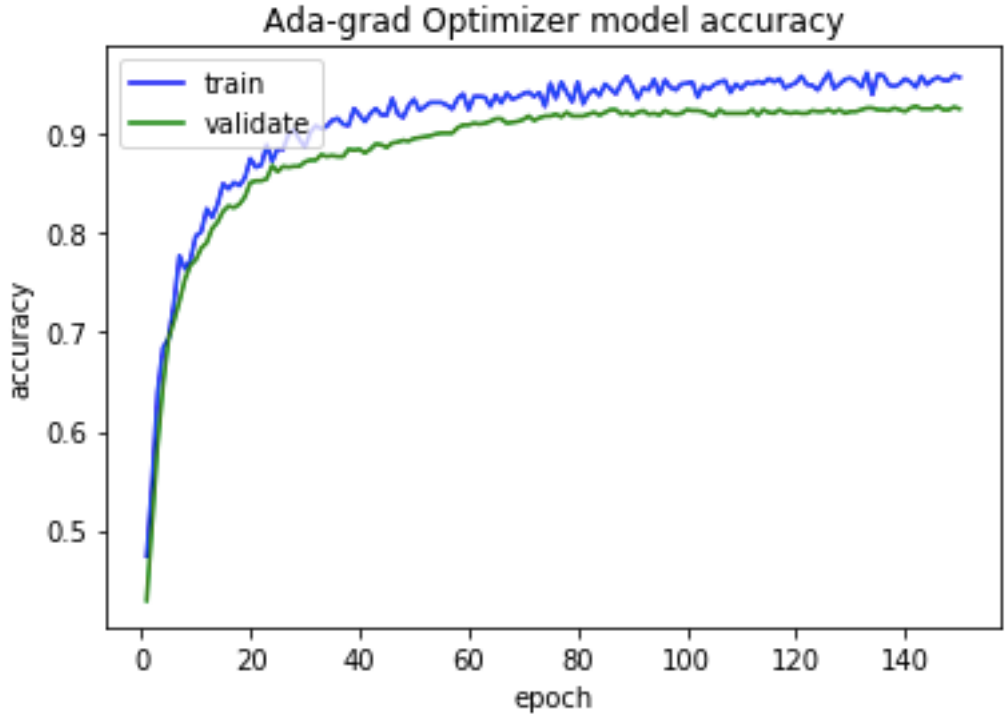




而Ada-grad在訓練時，training loss下降的速率為最慢，但其優點為其training loss和training accuracy為三者中最穩定，其需要較多的epoch才能將training loss下降到較低的值，其最終的training loss為0.943。

```
Ada-grad Optimizer start
Step 1, Minibatch Loss= 1000.0541, Training Accuracy= 0.474
Step 10, Minibatch Loss= 199.6039, Training Accuracy= 0.797
Step 20, Minibatch Loss= 97.3798, Training Accuracy= 0.875
Step 30, Minibatch Loss= 73.9579, Training Accuracy= 0.887
Step 40, Minibatch Loss= 41.8197, Training Accuracy= 0.919
Step 50, Minibatch Loss= 28.9931, Training Accuracy= 0.935
Step 60, Minibatch Loss= 38.3981, Training Accuracy= 0.925
Step 70, Minibatch Loss= 28.1582, Training Accuracy= 0.941
Step 80, Minibatch Loss= 15.5096, Training Accuracy= 0.953
Step 90, Minibatch Loss= 24.1603, Training Accuracy= 0.948
Step 100, Minibatch Loss= 25.5107, Training Accuracy= 0.951
Step 110, Minibatch Loss= 19.9208, Training Accuracy= 0.946
Step 120, Minibatch Loss= 14.6154, Training Accuracy= 0.952
Step 130, Minibatch Loss= 15.0393, Training Accuracy= 0.957
Step 140, Minibatch Loss= 12.4783, Training Accuracy= 0.956
Step 150, Minibatch Loss= 16.2204, Training Accuracy= 0.957
Optimization Finished!
```



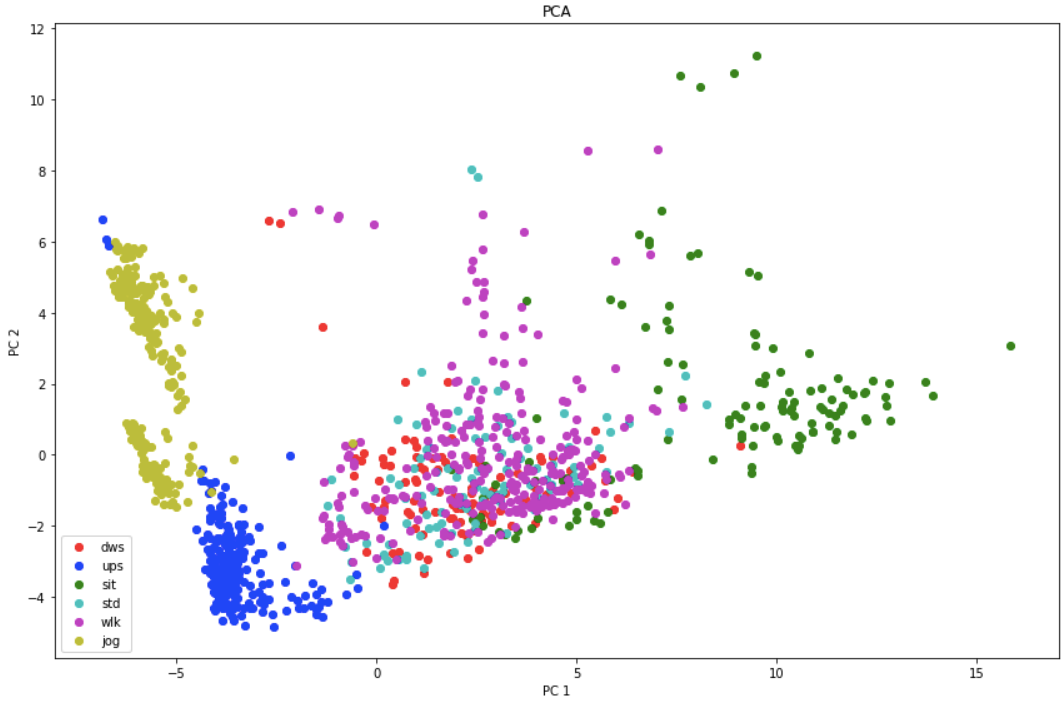


最後三者optimizer的training accuracy比較結果為Adam > gradient descent > Ada-grad。

(vi) Please explain what you observe from PCA(iv) and t-SNE(v).

PCA和t-SNE所plot出的figure如下方兩圖所示，可看出兩者在jogging、sitting和upstairs的分類在兩種方法皆可明顯和其他類別區分出來，但downstairs、walk、stand三個類別在PCA中幾乎全混在一起，而t-SNE可較優秀的將downstairs、walk、stand三個類別區分出來。但PCA的運行時間大幅短於t-SNE的運行時間。

PCA:



t-SNE