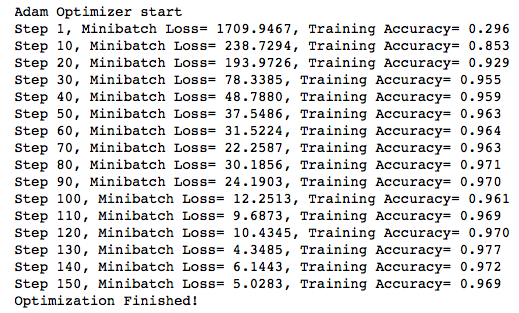
Deep Learning Homework2

StudentID: 107062650 Name: 曾昱榮

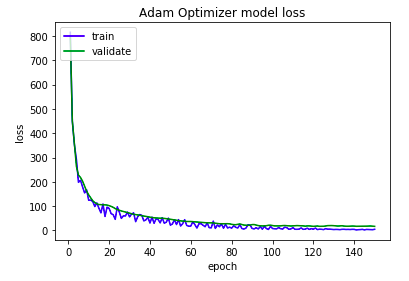
1.

(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。



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。



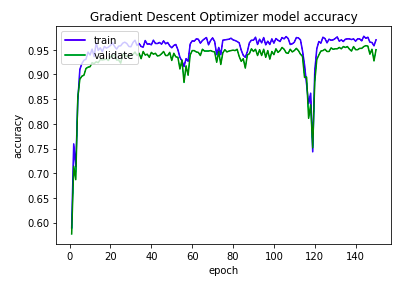
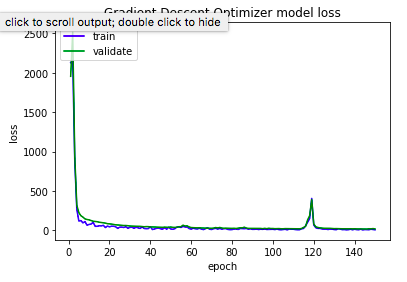
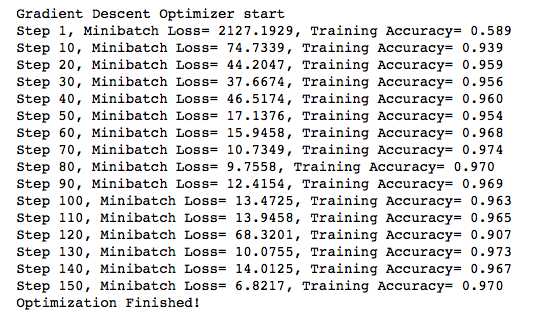
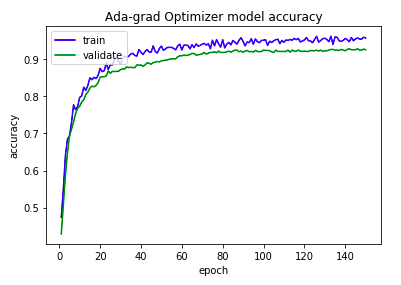
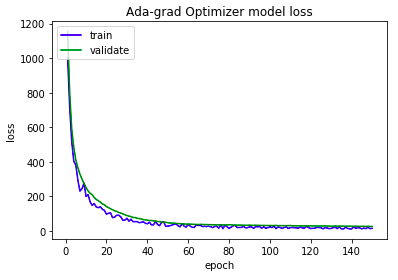
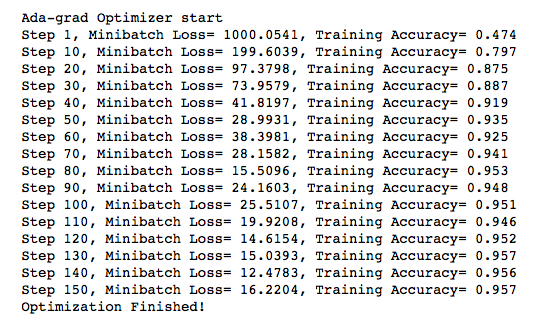
(ii) record precision, recall and F1-score as well as the averages of those criteria

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 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。

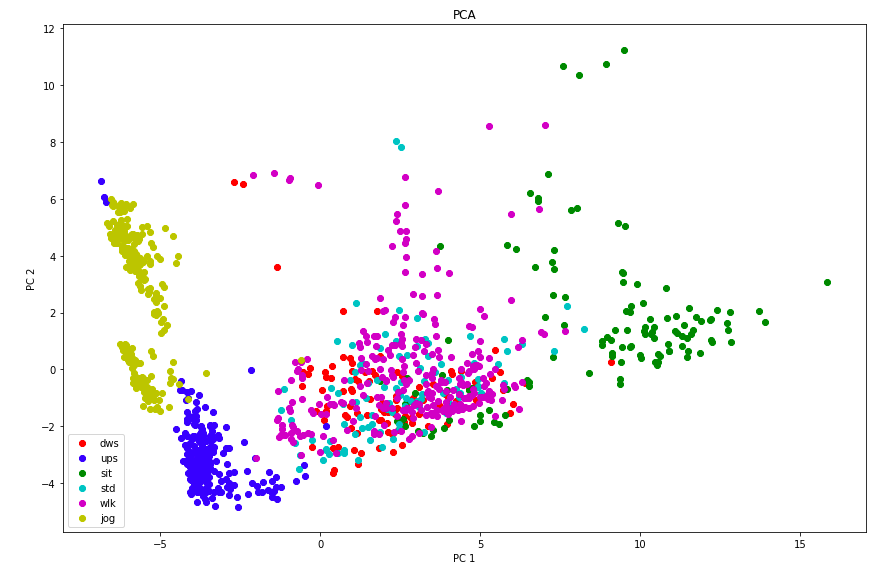
而Ada-grad在訓練時，training loss下降的速率為最慢，但其優點為其training loss和training accuracy為三者中最穩定，其需要較多的epoch才能將training loss下降到較低的值，其最終的training loss為0.943。

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

