Homework3 Image Sentiment Classification

Professor Pei-Yuan Wu
EE5184 - Machine Learning
姓名：柯上優
學號：b05902109

1. 請說明你實作的 CNN model，其模型架構、訓練過程和準確率為何？

- 我最終的模型是由7個CNN做ensemble，以下舉其中一個模型，是由VGG做為發想的CNN模型。

model 1
Con2D(16, (3,3), padding=same, BatchNorm, LeakyReLU)
Con2D(32, (3,3), padding=same, BatchNorm, LeakyReLU)
Con2D(64, (3,3), padding=same, BatchNorm, LeakyReLU)
MaxPooling2D((2, 2))
Dropout(0.1)
Con2D(128, (3,3), padding=same, BatchNorm, LeakyReLU)
MaxPooling2D((2, 2))
Dropout(0.2)
Con2D(256, (3,3), padding=same, BatchNorm, LeakyReLU)
MaxPooling2D((2, 2))
Dropout(0.2)
Con2D(512, (3,3), padding=same, BatchNorm, LeakyReLU)
MaxPooling2D((2, 2))
Dropout(0.2)
Dense(512, BatchNorm, LeakyReLU)
Dropout(0.5)
Dense(256, BatchNorm, LeakyReLU)
Dense(7, BatchNorm, LeakyReLU)

- 有做training data Affine，30%幅度晃動以及些微機率偏移等，以batch=128做epoch=300。

training	**public testing**	**private testing**
0.99	0.68654	0.69657

2. 承上題，請用與上述 CNN 接近的參數量，實做簡單的 DNN model，其模型架構、訓練過程和準確率為何？試與上題結果做比較，並說明你觀察到了什麼？

model 1
Dense(4096, input_shape(4848, 1), BatchNorm, LeakyReLU)
Dropout(0.5)
Dense(1024, BatchNorm, LeakyReLU)
Dropout(0.5)
Dense(7, BatchNorm, LeakyReLU)

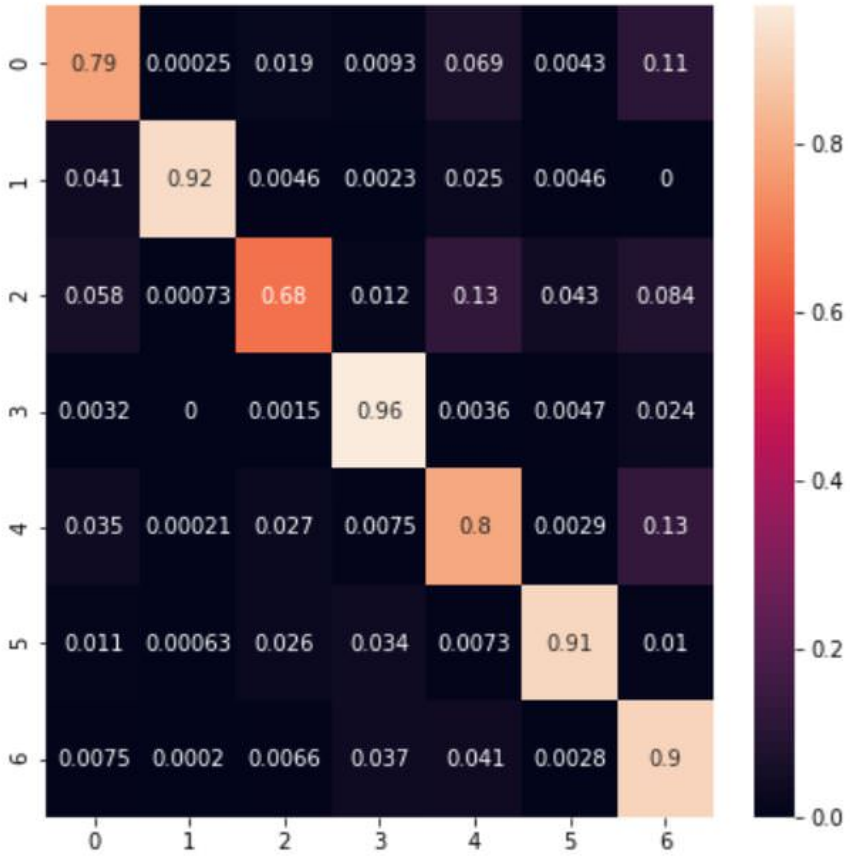
- CNN約四百萬個，DNN約四百二十萬個參數量
- 無法做如同DNN一樣地晃動資料，所以沒有做資料處理。一樣以batch=128做epoch=300。

training	public testing	private testing
0.99	0.41142	0.40714

- 觀察到DNN的結果非常失敗，嚴重的overfit，連dropout也救不了。可以理解到Convolution和maxPooling對於圖項特徵的萃取有多麼重要。

3. 觀察答錯的圖片中，哪些 class 彼此間容易用混？ 並說明你觀察到了什麼？ [繪出 confusion matrix 分析]

- 恐懼(2)容易被誤會成難過(4)，由作業投影片，我猜可能因為都有手的部分。
- 生氣(0)、恐懼(2)、難過(4)都有機會被誤會成中立(6)，中立可能代表著沒有特別的區分法，可以理解有些圖被誤會成中立。



4. CNN time/space complexity:

1. How many parameters are there in each layer(Hint: you may consider whether the number of parameter is related with)
   - Layer A: $2 \times 2 \times 6 \times 5$ filters, 6 bias, so total parameter is 126
   - Layer B: $2 \times 2 \times 4 \times 6$ filters, 4 bias, so total parameter is 100
2. How many multiplications/additions are needed for a forward pass(each layer).
   - Layer A:
     - multiplication $2 \times 2 \times 5$ per filter, $3 \times 3 \times 6$ filters, so total is 1080
     - additions: $2 \times 2 \times 5 - 1$ per filter, $3 * 3 * 6$ filters, so total is 1026
   - Layer B:
     - multiplication $2 \times 2 \times 6$ per filter, $1 \times 1 \times 4$ filters, so total is 96
     - additions: $2 \times 2 \times 6 - 1$ per filter, $1 \times 1 \times 4$ filters, so total is 92
3. What is the time complexity of convolutional neural networks?(note: you must use big-O upper bound, and there are l layer, you can use C_l, C_{l-1} as l th and $l-1$ th layer)
   - $O(\sum_{i=2}^{l}(\lceil \frac{n_{i-1}-k_i+2\times p_i}{s_i} + 1 \rceil)^2 \times (k_i)^2 \times C_i \times C_{i-1})$

5. PCA practice:Problem statement: Given 10 samples in 3D space.(1,2,3),(4,8,5),(3,12,9),(1,8,5),(5,14,2),(7,4,1),(9,8,9),(3,8,1),(11,5,6),(10,11,7)

1. What are the principal axes?
   - $0.3999 \times x_1 + 0.3376 \times x_2 - 0.8521 \times x_3 = 0$
   - $-0.6782 \times x_1 + 0.7344 \times x_2 - 0.0273 \times x_3 = 0$
   - $-0.6166 \times x_1 - 0.5888 \times x_2 - 0.5226 \times x_3 = 0$
2. Compute the principal components for each sample.
   $[-2.25104047, -1.37323947, 7.18658682]$
   $[-0.73022635, 0.94399334, 0.75871342]$
   $[-3.1883001, 4.45059025, -3.07034019]$
   $[-1.92979259, 2.97853006, 2.60849751]$
   $[4.25159619, 4.75401212, -1.82299166]$
   $[2.52755823, -3.91896138, 3.35457763]$
   $[-2.13952468, -2.55604371, -4.41464321]$
   $[2.27849363, 1.73131477, 3.46569126]$
   $[0.2038499, -6.03371503, -2.31359638]$
   $[0.97738622, -0.97648096, -5.75249521]$
3. Reconstruction error if reduced to 2D.(Calculate the L2-norm)
   - for each data, $[2.25, 0.73, 3.18, 1.92, 4.25, 2.52, 2.13, 2.27, 0.20, 0.97]$
   - total error: 7.39