學號：B04901019 系級： 電機三　姓名：梁書哲

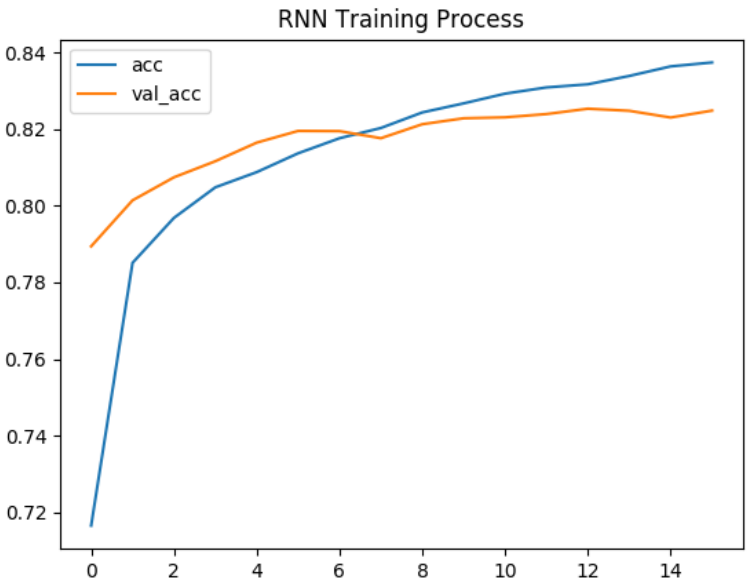
1. (1%) 請說明你實作的 RNN model, 其模型架構、訓練過程和準確率為何？(Collaborators: 無)

答：

Public: 0.82261 Private: 0.82031

100 epochs with early stopping, optimizer='adam', loss function='binarycrossentropy'

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 25, 256)	31874560
gru_1 (GRU)	(None, 25, 256)	393984
batch_normalization_1 (Batch	(None, 25, 256)	1024
gru_2 (GRU)	(None, 25, 256)	393984
batch_normalization_2 (Batch	(None, 25, 256)	1024
gru_3 (GRU)	(None, 256)	393984
batch_normalization_3 (Batch	(None, 256)	1024
dense_1 (Dense)	(None, 512)	131584
batch_normalization_4 (Batch	(None, 512)	2048
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 256)	131328
dropout_2 (Dropout)	(None, 256)	0
batch_normalization_5 (Batch	(None, 256)	1024
dense_3 (Dense)	(None, 1)	257



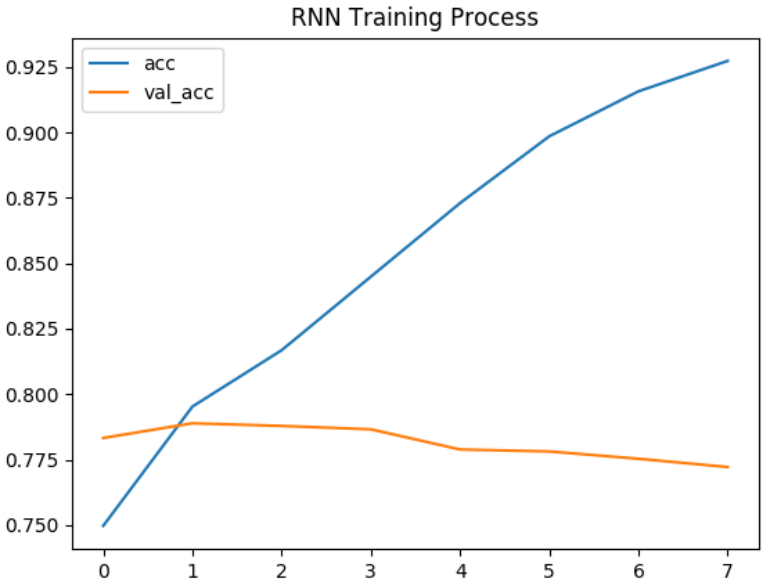
2. (1%) 請說明你實作的 BOW model, 其模型架構、訓練過程和準確率為何？(Collaborators: 無)

答：

Public: 0.78622 Private: 0.78631

100 epochs with early stopping, optimizer='adam', loss function='binarycrossentropy'

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 512)	1280512
batch_normalization_1 (Batch	(None, 512)	2048
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 512)	262656
dropout_2 (Dropout)	(None, 512)	0
batch_normalization_2 (Batch	(None, 512)	2048
dense_3 (Dense)	(None, 256)	131328
dropout_3 (Dropout)	(None, 256)	0
batch_normalization_3 (Batch	(None, 256)	1024
dense_4 (Dense)	(None, 1)	257



3. (1%) 請比較 bag of word 與 RNN 兩種不同 model 對於"today is a good day, but it is hot"與"today is hot, but it is a good day"這兩句的情緒分數，並討論造成差異的原因。
(Collaborators: 無)
答：

RNN:[0.3829788, 0.95893127]
BOW:[0.39831534, 0.39831534]

BOW 沒有考慮到文字的順序，因此同樣的文字組成的句子的情緒分數相同，而 RNN 模型會因為輸入順序而有所不同，第二句判斷為明顯的正面情緒。

4. (1%) 請比較"有無"包含標點符號兩種不同 tokenize 的方式，並討論兩者對準確率的影響。
(Collaborators: 無)
答：

有標點符號，Public: 0.82679 Private: 0.82688
無標點符號，Public: 0.82261 Private: 0.82031

有標點符號的準確率較好。推測為標點符號會造成句子語氣的轉變，而 Tokenize 有標點符號的狀況下，可以將這項資訊傳入 RNN 中，增加判斷的準確率。

5. (1%) 請描述在你的 semi-supervised 方法是如何標記 label, 並比較有無 semi-surpervised training 對準確率的影響。
(Collaborators: 無)
答：

有 semi-supervised，Public: 0.81948 Private: 0.81937
無 semi-supervised，Public: 0.82261 Private: 0.82031

Threshold 設為 0.03，iterate 五次，Training Process 如下

Iteration	Non-labeled Data	Labeled Data	val_acc
0	1178614	0	nan
1	840492	338122	0.8254
2	627995	550619	0.8208
3	539749	638865	0.8220
4	435025	743589	0.8215
5	409932	768682	0.8197