

# ML2017FALL-HW4-Report

學號：B03801039 系級：電機四 姓名：楊福恩

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: 廖宜倫 B03901001)

答：

準確率：

Training accuracy:82.18%  
Validation accuracy:82.23%  
Public accuracy:82.079%  
Private accuracy:82.048%

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 60, 100)	25417100
gru_1 (GRU)	(None, 60, 256)	274176
gru_2 (GRU)	(None, 128)	147840
dense_1 (Dense)	(None, 128)	16512
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 1)	129
Total params: 25,872,269		
Trainable params: 455,169		
Non-trainable params: 25,417,100		

訓練過程與模型細節：

首先利用 gensim 做 word to vector，取 size=100。在經過了無數次的實驗及與同學交流討論後，決定了左圖的模型架構，其中除了輸出層外每一層均加 dropout，參數設為 0.35，此參數為測試過 0.2, 0.3, 0.35, 0.4 後所決定的，此外 GRU 層中的 activation function 均為 tanh，Dense 層中的 activation function 則為 ReLU。在 training 的部分，因為有使用 early stopping，故 epoch 取 100；optimizer 取 Adam，實驗過程中嘗試了 RMSprop, adamax, nadam 等 optimizer，但效果均不如 Adam；因為輸出層的 activation function 選擇 sigmoid，故 loss function 選擇 binary crossentropy。

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: )

答：

準確率：

Training accuracy:92.32%  
Validation accuracy:75.37%  
Public accuracy:75.15%  
Private accuracy:75.114%

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 512)	16100864
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 2)	1026
Total params: 16,101,890		
Trainable params: 16,101,890		
Non-trainable params: 0		

Epoch:5  
Optimizer:Adam  
Dense 層 activation function:ReLU  
輸出層 activation function:sigmoid  
Loss function:binary crossentropy

#### 訓練過程與模型細節：

因為 bag of word 非常地佔記憶體，故 training 時僅用了 20% 的 training data 來 train，否則資料量太多便會出現 memory error，而所使用的 data 量少於 RNN model 及 bag of word 不考慮文字排列順序的特性，推測應為使得 bag of word model overfitting 較嚴重且 testing 上準確率低於 RNN model 的原因。

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: )

答：

以下兩個 model 輸出層的 activation function 皆為 sigmoid

Bag of word model:

第一句話預測分數為:0.7890，第二句話預測分數為:0.7890。

RNN model:

第一句話預測分數為:0.0501，第二句話預測分數為:0.9537。

觀察可知，題目中的第一句話應偏向負面情緒，故判斷後的分數應接近 0；而第二句話則偏向正面情緒，故分數應接近 1。由以上判斷後結果可看出，Bag of word model 因為其不管文字排列的特性，故兩句話判斷分數相同，並且可能因為兩句話中皆有「good」這個明顯為正面的字，故皆判斷為較接近 1 的值。而 RNN model 不僅考慮了文字內容也考慮了文字的排列順序，所以能較準確地判斷出兩句話的情緒，讓第一句話的判斷分數為接近 0 的值，第二句話判斷分數為接近 1 的值。

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: )

答：

有包含標點符號情況之準確率：

Training accuracy:80.78%

Validation accuracy:80.64%

Public accuracy:80.651%

Private accuracy:80.586%

結果和第一題不包含標點符號的 tokenize 方式相比，準確率下降了約 2%，推測原因應為要判斷句子所要表達的情緒主要看的是句子包含哪些文字及這些文字的排列順序，當把各種標點符號都考慮進去時，可能造成訓練後的 model 因為標點符號而誤判，所以造成準確率下降。

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

(Collaborators: 廖宜倫 B03901001 )

答：

有做 semi-supervised training 準確率：

Training accuracy:87.27%

Validation accuracy:82.20%

Public accuracy:81.393%

Private accuracy:81.313%

無 semi-supervised training 準確率：

Training accuracy:82.18%

Validation accuracy:82.23%

Public accuracy:82.079%

Private accuracy:82.048%

標記 label 方式：

用第 1 題 train 好的 model 來標記 unlabeled data 的標籤，threshold 設定為 0.95，即通過 sigmoid 後的分數若大於 0.95 則標記標籤為 1，若分數小於 0.05，則標記標籤為 0。由以上的準確率可看出，做了 semi-supervised training 後準確率些微下降，推測應為用 train 好的 model 預測後的結果來當作標籤，本身就存在有誤差，所以再加入這些可能有錯誤標籤的 data 後，可能造成原本 model 再經過 training 後 bias 變大，使得 public 及 private 準確率下降。