

學號：R06921005 系級：電機碩一 姓名：陳昱文

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:)

答：大體而言是 Embedding→RNN→Dense

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 50, 150)	3000150
spatial_dropout1d_1 (Spatial	(None, 50, 150)	0
bidirectional_1 (Bidirection	(None, 256)	285696
dense_1 (Dense)	(None, 64)	16448
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 32)	2080
dropout_2 (Dropout)	(None, 32)	0
dense_3 (Dense)	(None, 1)	33
Total params: 3,304,407		
Trainable params: 3,304,407		
Non-trainable params: 0		

上圖為使用一層 bidirectional 的架構 為了使用 ensemble method 亦建構不同的架構如下

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 50, 150)	3000150
dropout_1 (Dropout)	(None, 50, 150)	0
gru_1 (GRU)	(None, 50, 128)	107136
dropout_2 (Dropout)	(None, 50, 128)	0
bidirectional_1 (Bidirection	(None, 50, 256)	197376
dropout_3 (Dropout)	(None, 50, 256)	0
bidirectional_2 (Bidirection	(None, 256)	295680
dropout_4 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 256)	65792
dropout_5 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 64)	16448
dropout_6 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 1)	65
Total params: 3,682,647		
Trainable params: 3,682,647		
Non-trainable params: 0		

此為三層 GRU 模型，兩者接為先經過一 Embedding 層再進入不等數量的 recurrent 層

Accuracy	validation
Former RNN	0.807
Latter RNN	0.8089

為了使 ensemble 有更好的效益，尚有其他不同的 RNN 以及 CNN 等架構，不一一敘述

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：大體而言是 Embedding→Dense

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 256)	5120256
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 64)	16448
dropout_2 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 1)	65
Total params: 5,136,769		
Trainable params: 5,136,769		
Non-trainable params: 0		

利用 Keras 提供之 texts_to_matrix，再輸入如上圖所示之若干層 Dense Layers

Accuracy	validation
BOW	0.80

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：

得分	前句	後句
RNN	0.27670151	0.54705328
BOW	0.65368831	0.65368831

很明顯的，RNN 在這兩句話的表現上差異比 BOW 大，由於 BOW 的方式是不在意語序順序，因此預測出相同的結果是可預見的，而 RNN 因為訓練方式是有考慮語句順序的，鼓隊兩者預測的值有差異

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：

Accuracy	Private	Public
有	0.81052	0.81050
無	0.80779	0.80814

有加入標點符號的情況下，其準確率較高，估計是標點符號亦帶有部分資訊，比如說驚嘆號就表示前面該詞是重點之一，但部分不帶有情緒或過於稀少的標點圖形則可去除以避免佔據字典

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:)

答：

這次所使用 semi-supervised 的方法是 self-training，然而這個對我的 model 的預測率並沒有顯著的提升(僅約 0.2%的提升)，估計因為此種 training 的方法為採用自己的 model 預測後再拿回來使用，但由於我們所選取具有可信度的資料其實已經近乎完美地從該模型預測出，故使用它來校正模型的意義不大，比較好的方式或許為利用其他的模型估測出預測值再拿回來訓練此模型，這樣的校正效果應該會比較顯著