

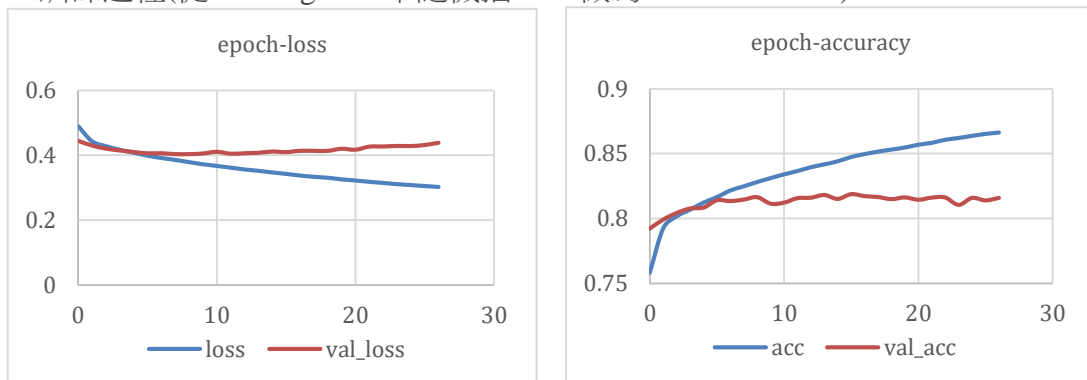
學號：B03901145 系級：電機四 姓名：郭恆成

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: )

模型架構:

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 41, 200)	49780800
masking_1 (Masking)	(None, 41, 200)	0
gru_1 (GRU)	(None, 256)	350976
dense_1 (Dense)	(None, 64)	16448
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 1)	65
Total params: 50,148,289		
Trainable params: 367,489		
Non-trainable params: 49,780,800		

訓練過程(從 training data 中隨機抽 10% 做為 validation data):



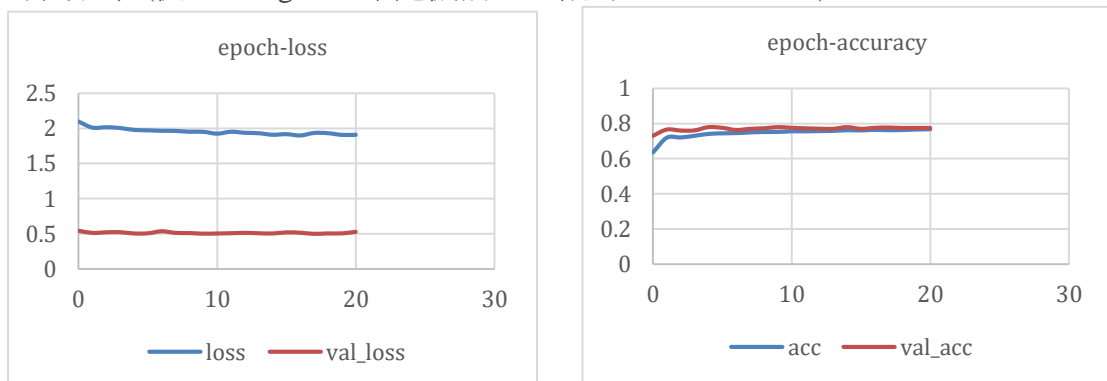
準確率: Kaggle 上 public score 0.82207, private score 0.82008

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: )

模型架構(取頻率最高的前 20000 個字做 bow)

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 41, 20000)	400020000
lambda_1 (Lambda)	(None, 20000)	0
dense_1 (Dense)	(None, 512)	10240512
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 256)	131328
dropout_2 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 1)	257
dropout_3 (Dropout)	(None, 1)	0
Total params: 410,392,097		
Trainable params: 10,372,097		
Non-trainable params: 400,020,000		

訓練過程(從 training data 中隨機抽 10% 做為 validation data):



準確率: Kaggle 上 public score 0.78316, private score 0.78494

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: )

	Sentence 1	Sentence 2
BOW	0.54712844	0.54712844
RNN	0.3246138	0.99374253

BOW 的 input 是所有字 one hot 的和，所以句子 1 跟句子 2 預測出來的分數完全相同，而 RNN 會學到 but 後才是句子的主要意思，而 good 會造成高分，hot 會導致低分。

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: )

不加標點符號 Kaggle 上 public score 0.82207, private score 0.82008

而加標點符號 Kaggle 上 public score 0.82832, private score 0.82688

可以發現標點符號對正負評論的準確率有正向的幫助。

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: )

先用 `training data` 訓練出一個 `model`，再用這個 `model` 去預測 `unlabeled data` 的分數，把分數大於 0.95 的標記成 1、小於 0.05 的標記成 0，丟進 `training data`，繼續去訓練 `model`。

而 `semi-supervised` 在 Kaggle 上 public score 0.81960, private score 0.81993，略低於未加入 `unlabeled data` 的成績，推測是 0.95 跟 0.05 得到的 `data` 已經很接近用 `training data` 認為的正確答案，加入後並不能使 `model` 更好的去預測 `testing data`，甚是有 `overfit` 的傾向。