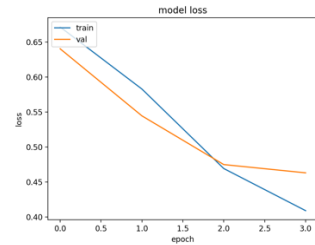
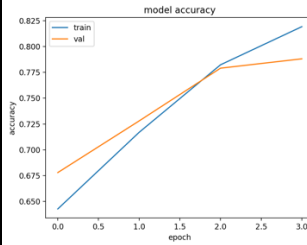


學號：R06921066 系級：電機碩一 姓名：劉宇閔

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：

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 50, 50)	5000000
gru_1 (GRU)	(None, 10)	1830
dense_1 (Dense)	(None, 1)	11
Total params: 5,001,841		
Trainable params: 5,001,841		
Non-trainable params: 0		

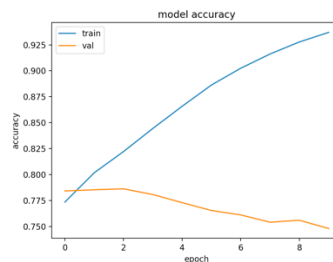


將單字轉為 50 維的向量來表示，每個 cell 內得 unit 設為 10，lr 使用 default 的 1e-3 並使用 adam 演算法來修正，最後將得到的 10 維 state 給 fully connected layer 來做分類，得到的值再透過 sigmoid 限制在 0~1 之間。此用此種訓練方法可以得到 public 約 0.8 的 accuracy。

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：

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 256)	1280256
dense_2 (Dense)	(None, 50)	12850
dense_3 (Dense)	(None, 1)	51
Total params: 1,293,157		
Trainable params: 1,293,157		
Non-trainable params: 0		



使用 bag of word，因 bow 並不具有時序性上的關係所以在此搭配 DNN 使用，可以發現使用這種做法雖然可以達到約 78% 的 validation accuracy 但達到峰值後就會逐漸下降，因此推論雖然 model 逐漸 fit 到 training data 上，但 bog 並無法有效地使用在 validation 上。

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：透過實驗可以知道使用 BOW 後得到的 pred 值都是 0.94 也正確驗證了 BOW 不具有時續性的關聯，只照著統計出現的次數來當作預測基準，而使用 RNN train 出來的 model 可以得到的分數分別為 0.65 和 0.96，雖然無法明確地表明出兩個的語義是不同，但第一句話對人來說也是不容易判讀的，所以分數也藉再模糊之間。

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: R06922096 洪子翔)

答：實測後 accuracy 並無明顯上升，故可知標點符號對於整體並沒有影響

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: R06922093 陳禹齊，R06922096 洪子翔)

答：使用的 semi-supervised 為將 no label 的檔案丟入 pre-train 好的 model 內，得到他們 predict 的值並將其當作 label 存起來，再將這個 label 的值重新和 pred 的值去計算 cross entropy 來取得 loss fine tune 整個 model，讓 model 可以 fit 到這整個答案上，實作出來可以讓 model 較不會 overfitting 在前面的 training data