

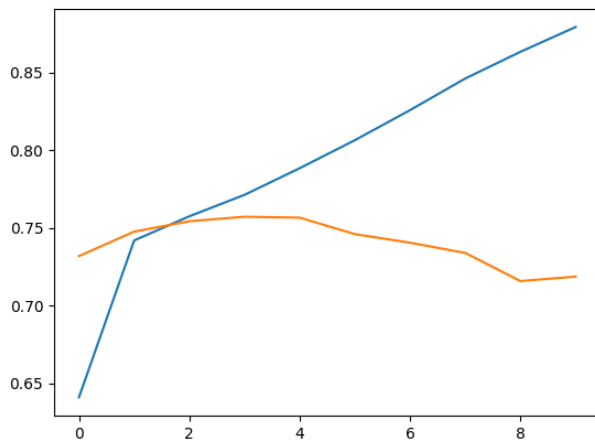
Machine Learning HW6 Report

學號：B05705053 系級：資管三 姓名：蔡涵如

1. (1%) 請說明你實作之 RNN 模型架構及使用的 word embedding 方法，回報模型的正確率並繪出訓練曲線

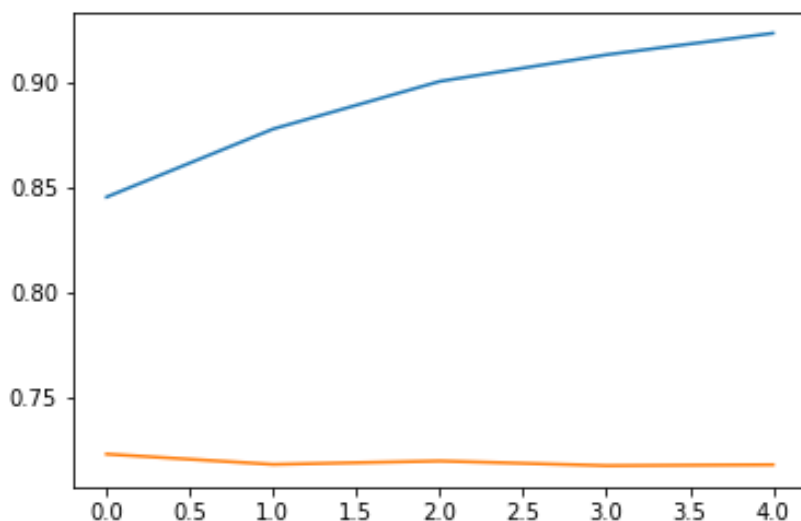
我使用的是word2vector embedding dim 100，使用4層的Bidirectional LSTM，每層使用 sequence 長度 50的方式。

Layer (type)	Output Shape	Param #
embedding_22 (Embedding)	(None, 50, 100)	4183400
conv1d_21 (Conv1D)	(None, 50, 50)	15050
max_pooling1d_10 (MaxPooling)	(None, 16, 50)	0
bidirectional_53 (Bidirectio	(None, 16, 400)	401600
batch_normalization_53 (Batc	(None, 16, 400)	1600
bidirectional_54 (Bidirectio	(None, 16, 400)	961600
batch_normalization_54 (Batc	(None, 16, 400)	1600
bidirectional_55 (Bidirectio	(None, 16, 400)	961600
batch_normalization_55 (Batc	(None, 16, 400)	1600
bidirectional_56 (Bidirectio	(None, 400)	721200
batch_normalization_56 (Batc	(None, 400)	1600
dense_22 (Dense)	(None, 1)	401
activation_22 (Activation)	(None, 1)	0



2. (1%) 請實作 BOW+DNN 模型，敘述你的模型架構，回報模型的正確率並繪出訓練曲線。模型架構為六層的dense，每一層512個units，使用超過五次以上的字，

總字元4622個，下去做one-hot-encoding，正確率大概0.722708左右，以下是training history的圖，可以看到從第一個epoch就val_acc下不去了。

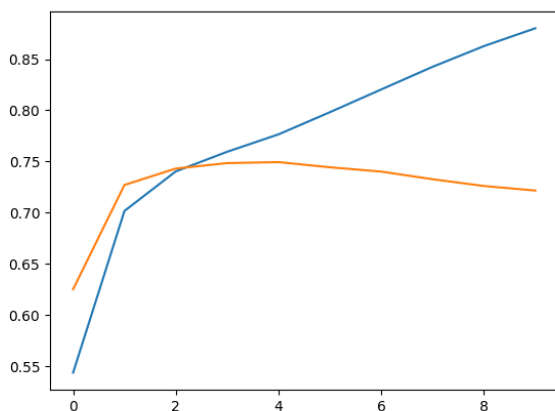


3. (1%) 請敘述你如何 improve performance (preprocess, embedding, 架構等)，並解釋為何這些做法可以使模型進步。

而最後用K-fold cross validation，將所有的model predict的機率平均，private : 0.75740 public: 0.75970，會想要用K-fold的原因是因為我的local validation大概都有0.763左右，但傳上去大概都只有0.75，所以implement了K-fold，發現就算用同一個model的架構，還是會有validation accuracy差距很大得現象，可以推測的難易程度差異很大。

4. (1%) 請比較不做斷詞 (e.g., 以字為單位) 與有做斷詞，兩種方法實作出來的效果差異，並解釋為何有此差別。

這是不斷詞得training history，最高可以到0.74940，雖然比我實做的用jieba斷詞



的(0.75740)還要遜色一點，但是也有可能是因為沒有調參數的關係。

5. (1%) 請比較 RNN 與 BOW 兩種不同 model 對於 "在說別人白痴之前，先想想自己"與"在說別人之前先想想自己，白痴" 這兩句話的分數 (model output)，並討論造成差異的原因。

- a. 使用BOW 預測出來的confidence皆為0.5570325，原因是因為這兩句話字的組成一模一樣，而BOW不會考慮詞跟詞中間的關聯性。

- b. 使用RNN 預測出來得confidence分別為0.71505669與0.71710439，後者的分數比較高，因為RNN會考慮字跟字間的前後關係。