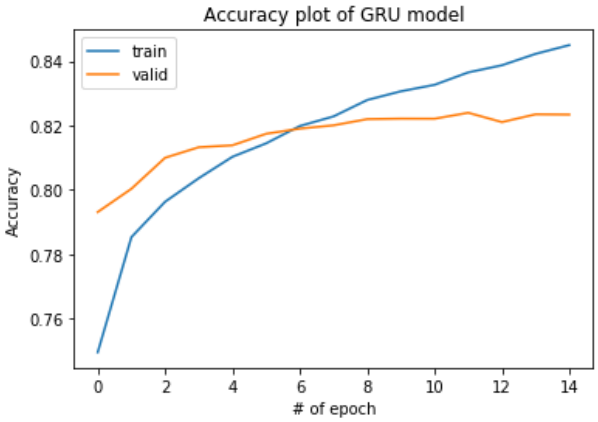
學號：R06922006 系級： 資工碩一 姓名：劉宏國

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: 自己)

答：模型架構&準確率：Accuracy: 88.03%、 val accuracy: 0.82398。

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 30, 300)	24000000
dropout_1 (Dropout)	(None, 30, 300)	0
gru_1 (GRU)	(None, 30, 300)	540900
gru_2 (GRU)	(None, 128)	164736
dense_1 (Dense)	(None, 64)	8256
dropout_2 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 1)	65

Total params: 24,713,957
Trainable params: 713,957
Non-trainable params: 24,000,000



訓練過程：(此 model 為未加 nolabel 的 model)

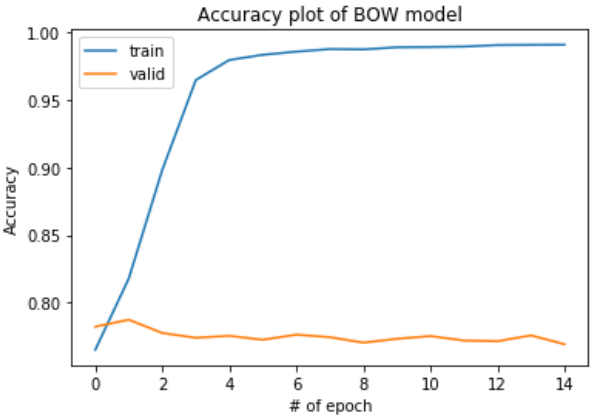
a. max_features = 80000、max_length = 30、embedding_size = 300、gru_output_size = 128、batch_size = 1000、epochs = 15。
b. Loss function：binary_crossentropy。
c. Optimizer：adam。

2.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: 自己)

答：模型架構&準確率：Accuracy: 99.19%、val accuracy: 0.78740。

Layer (type)	Output Shape	Param #
dense_44 (Dense)	(None, 4096)	12292096
dense_45 (Dense)	(None, 2048)	8390656
dropout_26 (Dropout)	(None, 2048)	0
dense_46 (Dense)	(None, 2048)	4196352
dropout_27 (Dropout)	(None, 2048)	0
dense_47 (Dense)	(None, 2048)	4196352
dropout_28 (Dropout)	(None, 2048)	0
dense_48 (Dense)	(None, 1)	2049

Total params: 29,077,505
Trainable params: 29,077,505
Non-trainable params: 0



訓練過程：(此 model 為未加 nolabel 的 model)

a. max_features = 3000、batch_size = 100、epochs = 15。
b. Loss function：binary_crossentropy。
c. Optimizer：adam。

比較：

為了將 BOW 的 params 調成和 GRU 相似，因此 BOW 的結果 overfitting 了，且從準確率(val accuracy)來看，GRU 的效果會比 BOW 好很多。

3.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: 自己)

答：(units=1)，機率靠近 1 的，預測結果為 1。

a. GRU：

today is a good day, but it is hot ：0.54。

today is hot, but it is a good day ：0.99。

b. BOW：

today is a good day, but it is hot ：0.75。

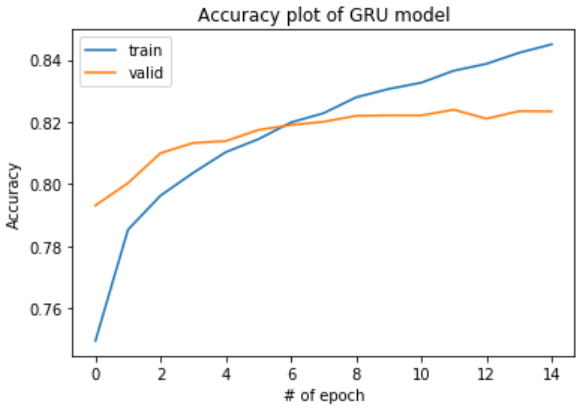
today is hot, but it is a good day ：0.75。

由上述結果來看，BOW 因為只在意 term 的出現次數，不考慮詞出現的順序，因此兩句話都出現相同的 term，只是順序不同，所以 BOW 的預測機率會相同而 GRU 則不會。

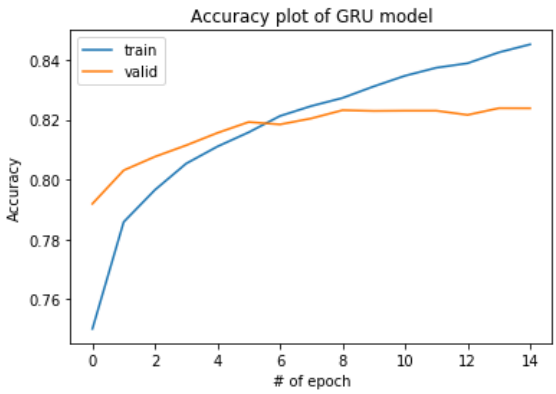
4.請比較"有無"包含標點符號兩種不同 tokenize 的方式，並討論兩者對準確率的影響。 (Collaborators: 自己)

答：

1.有 filter(去除標點符號)：Accuracy： 88.03%、val accuracy：0.82398。



2.無 filter(保留標點符號)：Accuracy：88.01%、val accuracy：0.82388。



上述的兩種 model 都在相同的訓練過程之下且都未包含 nolabel 的 data。而由上兩張圖所示，不管有無標點符號，對準確率來說，並無明顯的差異。

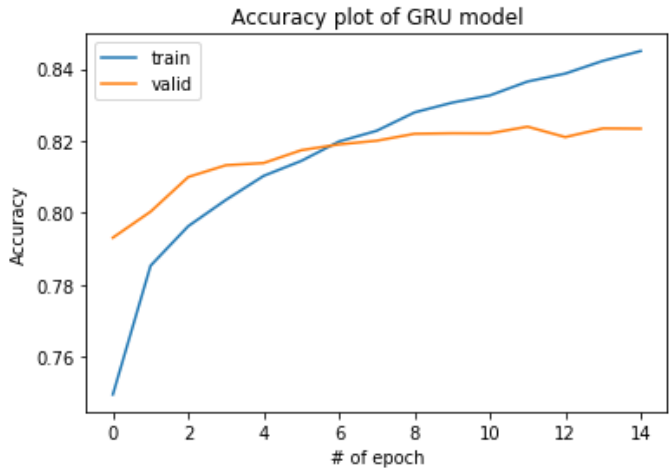
5.請描述在你的 semi-supervised 方法是如何標記 label，並比較有無 semi-surpervised training 對準確率的影響。(Collaborators: 自己)

答：

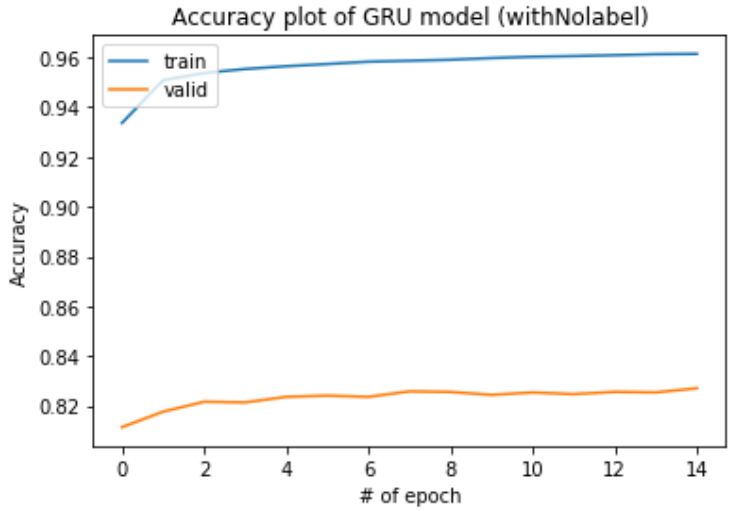
我是先用 training data 先 train 一個 model，再將此 model 對 nolabel 的 data 進行預測，加入預測的 label，threshold 值我設 0.95，大於等於 0.95 的 label=1，反之小於等於 0.05 的 label=0，不在此範圍的機率值會被過濾掉。

最後再將原本的 training data 再加入預測後的 nolabel data 重新 train 一次 model，進而得到最終的 model。

準確率(未加 nolabel data)：Accuracy：88.03%、val accuracy：0.82398。



準確率(加 nolabel data)：Accuracy：87.82%、val accuracy：0.82700。



由上圖所示，再加入 nolabel 的 data 後，val accuracy 從原本的 0.82398 提升到 0.82700。