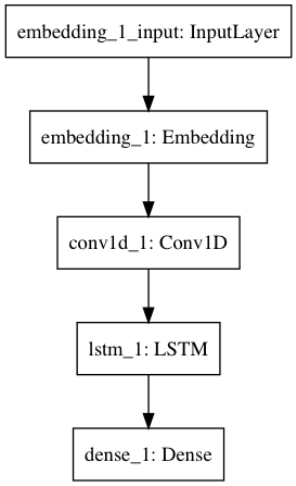
學號：B03901144 系級： 電機四 姓名：郭彥滕

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？



epoch：20

optimizer：adadelta

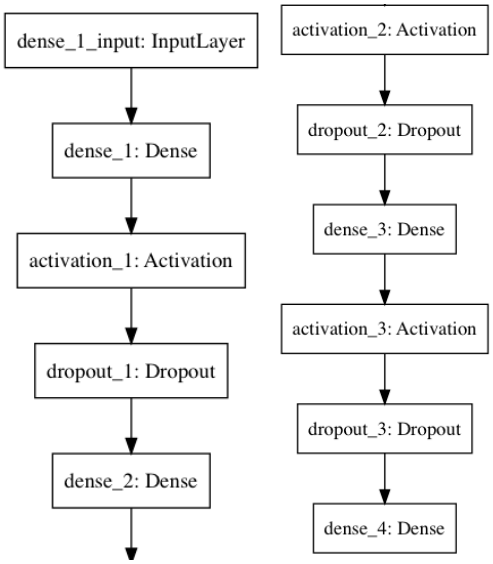
loss function：binary_crossentropy

accuracy：0.81812

Adding Conv1D seems to improve accuracy.

Using adadelta have good acc because using adam might drop too fast.

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？



epoch: 5

optimizer: adadelta

loss function: binary_crossentropy

accuracy: 0.7753

BOW have poorer performance.

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。

	today is a good day, but it is hot	today is hot, but it is a good day
BOW	0.837022	0.837022
RNN	0.767334	0.947018

For BOW, two sentences have totally same vectors, so the prediction would be the same.

For RNN, they have different vectors, so the prediction would be much more accurate.

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。

With punctuations and numbers: 0.79400

Without some punctuations(+$,’) and all numbers(0123456789): 0.818

We can find that punctuations have significant meaning in sentences.

Perhaps ? ! could give some information about emotion.

5. (1%) 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。

>0.9: let it be 1

<0.1: let it be 0

others: skip them

before semi-supervised: 0.81812

after semi-supervised: 0.81891

The method has little improvement on accuracy. Having a better semi-supervised algorithm might lead to good result.