

Machine Learning HW6 Report

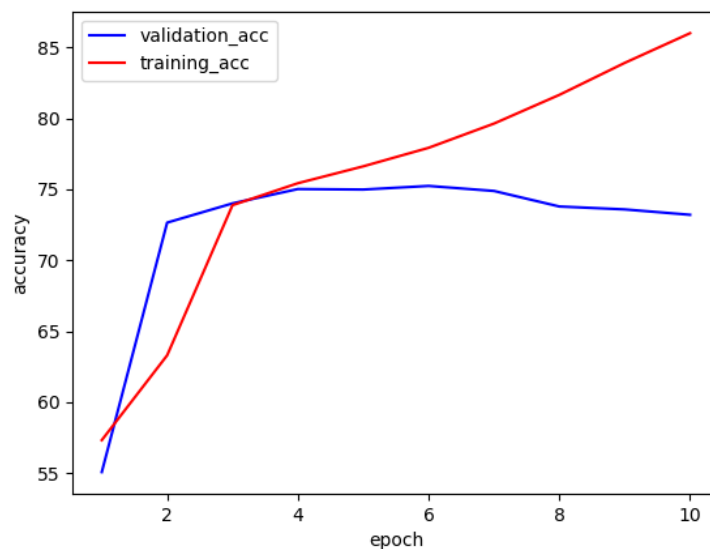
學號 b05902125 系級：資工三 姓名：葛洵為

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*

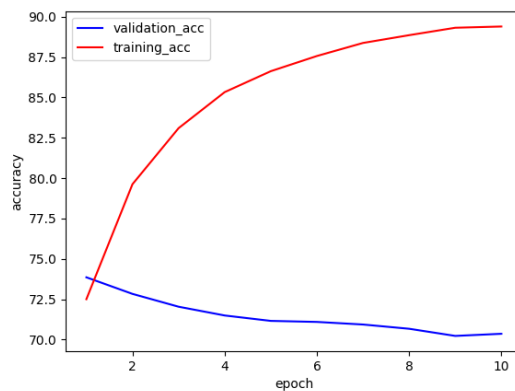
Ans :

| |
|---|
| nn.LSTM(embedding_dim, hidden_dim, num_layers=num_layers, batch_first=True, bidirectional=True, dropout=0.5) |
| nn.Dropout(dropout),nn.Linear(hidden_dim * 2, 1024), |
| nn.Dropout(dropout),nn.Linear(1024, 512), |
| nn.Dropout(dropout),nn.Linear(512, 1), |
| nn.Sigmoid() |

Embedding dim = 200, hidden_dim=200, num_layers=5, dropout=0.5
embedding 的 weight 來自事先 train 好的 word2vec model 。



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*。



```
nn.Linear(hidden_dim * 2, 1024)
```

```
nn.Linear(hidden_dim * 2, 1024)
```

```
nn.Linear(hidden_dim * 2, 1024)
```

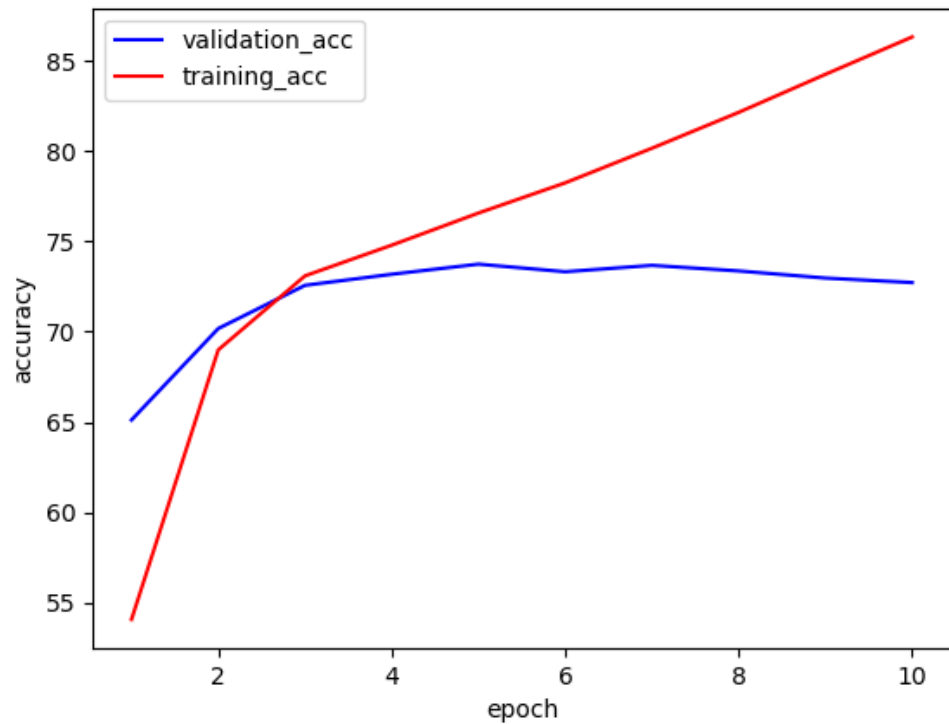
沒有任何的 dropout，可以看到 validation 不停地下降，但是還是可以 train 得起來。

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。

Ans：在 word2vec model 上加上 test data 的資料再去 train 一次使得資料更加準確，以及對每一維 lstm 的 output 取 max 以及 mean 然後與原本的最後一層全部 concat 起來，再丟進 activate function，這麼做原因是因為這樣對於 oov 的處理後的資料數值更加平均。

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。

embedding：



可以看到雖然 training accuracy 穩定上升，但是 validation accuracy 無法達到 75% 的水準以上。

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。