

Homework5 Report – Malicious Comments Identification

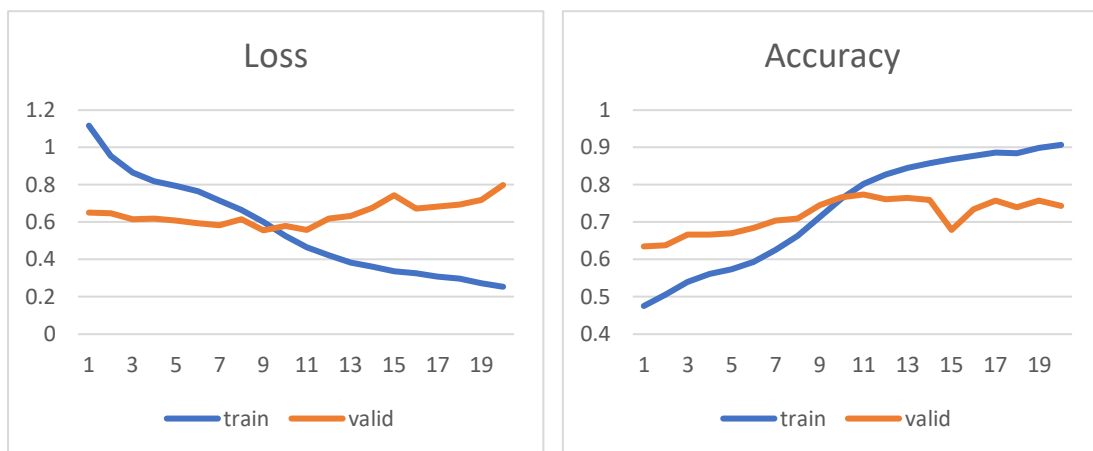
B04703117 財金五 謝昊辰

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。

Kaggle Private Score: 0.84651, Kaggle Public Score: 0.79069

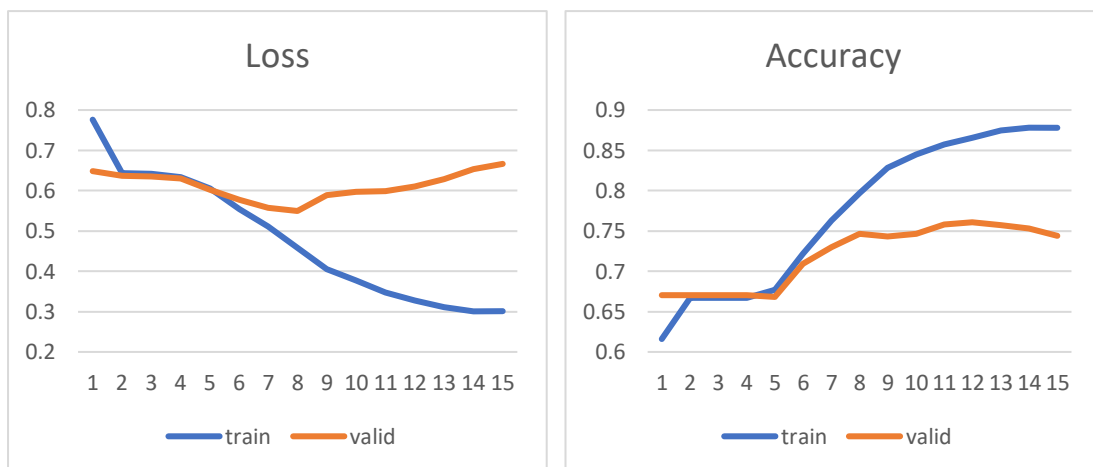
Embedding：先將利用 gensim 訓練出來的 word2vec 傳進模型當作 initial 參數，並傳入 hidden size 為 300 的 embedding

RNN：將 embedding 完的 output 傳入 hidden size 為 200 維的雙向 gru，最後再將第二層 gru 的兩個 output 合併傳入兩層 fully connected 的 layer



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。

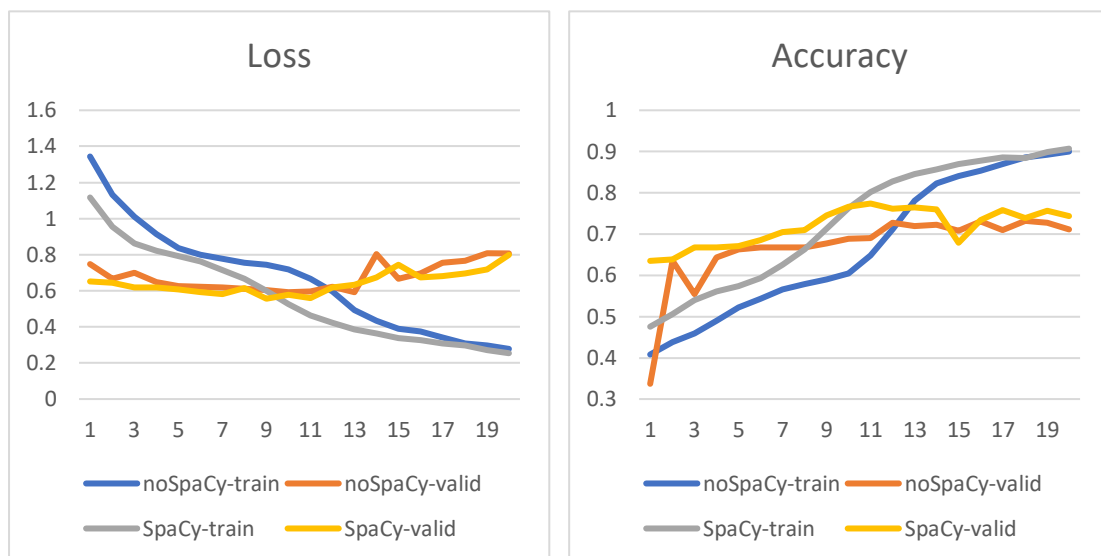
將長度為 6412 預訓練的 word2vec 匯入 model，並將單字藉由 word2index 轉換成以 index 表示的(13240, 6412)維的 array，接著傳入三層 fully connected layers，optimizer 使用 adam，activation function 使用 sigmoid，loss function 使用 cross entropy



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。
- (1) Preprocess：將句子中的標點符號剔除、字母都轉換成小寫，並且將重複超過三次的字母轉換成一次 (Ex: Heyyyyyyyy → Hey)，對於機器在理解句子有很大的幫助。
 - (2) Embedding：在 word2index 中增加 unknown 和 padding 的欄位，並將每一句的 sequence length 設為 40，若 sequence length 小於 40 則用 padding 補滿。
 - (3) 架構：使用兩層雙向的 gru 增加機器理解句子的能力，也有在 fully connected layer 中增加 dropout 避免 overfitting，並且對 gru 參數做 orthogonal initialization。
4. (1%) 請比較不使用斷詞套件 (e.g., 用空白分開) 與有使用斷詞套件，兩種方法實作出來的效果差異，並解釋為何有此差別。

	Private Score	Public Score
使用 SpaCy	0.84651	0.79069
不使用 SpaCy	0.79069	0.75581

可以看到沒有用 SpaCy 斷詞的效果不管 Private 還是 Public 都會比有用 SpaCy 來得低上許多，下圖也可以看到明顯有用 SpaCy 的 Accuracy 會略勝沒有使用 SpaCy。判斷是因為斷詞套件能夠比較有邏輯的去處理原始的 natural language，相較於只用空白來斷詞，SpaCy 顯然對於一些縮寫或是特殊文字，能夠有較好的處理。



5. (1%) 請比較 RNN 與 BOW 兩種不同 model 對於 "Today is hot, but I am happy" 與 "I am happy, but today is hot" 這兩句話的分數 (model output) , 並討論造成差異的原因。

	RNN	BOW
Today is hot, but I am happy	0.0120	0.5542
I am happy, but today is hot	0.1868	0.5542

因為兩個句子用到的字都一樣，因此 BOW model 出來的分數會是相同的，而 RNN 因為有參考前後文的語意以及先後順序，因此可以看到兩個分數的差別。

6.

$$\begin{aligned}
 w &= [0,0,0,1], & b &= 0 \\
 w_i &= [100,100,0,0] & b_i &= -10 \\
 w_f &= [-100,-100,0,0] & b_i &= 110 \\
 w_o &= [0,0,100,0] & b_i &= -10
 \end{aligned}$$

t	1	2	3	4	5	6	7	8
x^t	0	1	1	0	0	0	1	1
	1	0	1	1	1	0	1	0
	0	1	1	1	0	1	1	1
	3	-2	4	0	2	-4	1	2

$$f(z) = \frac{1}{1 + e^{-z}} \quad g(z) = z \quad h(z) = z$$

t	z	z _i	z _f	z _o	c' ₁	y
1	3	90	10	-10	3	0.000136
2	-2	90	10	90	1	1
3	4	190	90	90	-3	-3
4	0	90	10	90	-3	-3
5	2	90	10	-10	-1	-4.54
6	-4	-10	110	90	-1	-1
7	1	190	-90	90	-1	-1
8	2	90	10	90	-3	-3

7. Let V_{w_j} be the jth column of matrix W' , so $u_j = V_{w_j}^T \cdot h$

$$\begin{aligned}
 L &= -\log\left(\prod_{c=1}^C e^{u_c}\right) + \log\left(\sum_{i=1}^V e^{u_i}\right)^C = -\sum_{c=1}^C u_c + C \log\left(\sum_{i=1}^V e^{u_i}\right) \\
 \frac{\partial L}{\partial W_{ij}^{T'}} &= \frac{\partial L}{\partial u_j} \cdot \frac{\partial u_j}{\partial W_{ij}^{T'}}
 \end{aligned}$$

$$\rightarrow \frac{\partial L}{\partial u_j} = - \sum_{c=1}^c u_c + C \cdot \frac{1}{\sum_{i=1}^v e^{u_i}} \cdot \frac{\partial}{\partial u_j} \sum_{j=1}^v e^{u_j} = - \sum_{c=1}^c 1 + \sum_{j=1}^v y_j = y_j - t_j = e_j$$

$$\rightarrow \frac{\partial L}{\partial W_{ij}^{T'}} = e_j \cdot \frac{\partial u_j}{\partial W_{ij}^{T'}} = e_j \cdot \frac{\partial W_{ij}^{T'} * h_i}{\partial W_{ij}^{T'}} = e_j \cdot h_i$$

$$\frac{\partial L}{\partial W_{ij}^T} = \frac{\partial L}{\partial u_j} \cdot \frac{\partial u_j}{\partial W_{ij}^T} = \frac{\partial L}{\partial u_j} \cdot \frac{\partial u_j}{\partial h_i} \cdot \frac{\partial h_i}{\partial W_{ij}^T} = e_j \cdot W_{ij}^{T'} \cdot \frac{\partial W_{ij}^T * x_i}{\partial W_{ij}^T} = e_j \cdot W_{ij}^{T'} \cdot x_i$$