

Machine Learning HW5 Report

學號：r07323035 系級：經濟系碩二 姓名：王嚴

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*

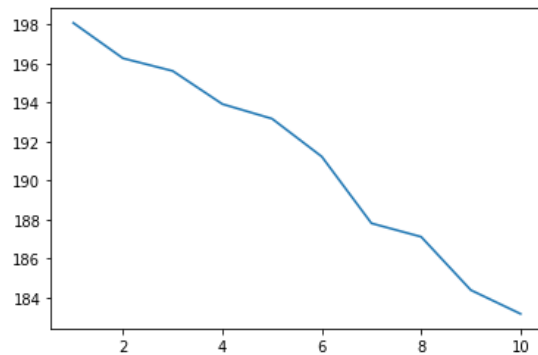
使用 LSTM

(embedding_dim, hidden_dim, n_layer, dropout) = (400, 256, 2, 0.5)

word embedding

將字詞轉為 400 維的向量來表示

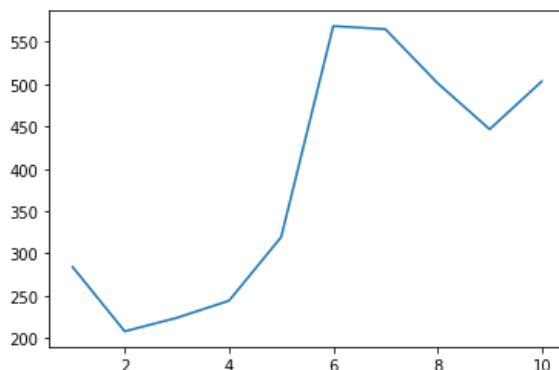
F1_score = 0.78139



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*。

使用三層 linear 配 dropout 最後再用 sigmoid 轉換

F1_score = 0.6308



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) 將句子中屬於 **stop word** 的詞給拿掉，這樣能夠去除多餘的無用資訊，使資料更精簡
- (2) 使用較高維度的 **hidden_layer**，相當於用更多的特徵來描述一個詞和句子

4. (1%) 請比較不做斷詞 (e.g.,用空白分開) 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。

用空白斷詞的效果會較差

因為單純用空白來做斷詞會遇到一些問題, ex : **#Trump** 和 **Trump** 在空白斷詞時會是兩個不同詞彙, 但是用 **spacy** 來斷詞可以正確的區分為**#**和 **Trump**

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。

使用 **RNN** 在這兩句話的分數會不一樣

因為 **RNN** 有考量句子的順序關係, 所以兩句話對 **RNN** 來說是不一樣的而 **BOW** 是單純考量這段句子具有得詞彙, 所以對 **BOW** 來說這兩句話是一樣的

LSTM Cell

$$\begin{aligned}
 t = 1, \quad x^1 &= (0, 1, 0, 3), \quad c = 0 \\
 z = 3 &\Rightarrow g(z) = 3 \\
 z^i = 90 &\Rightarrow f(z^i) = 1 \\
 z^f = 10 &\Rightarrow f(z^f) \approx 1 \Rightarrow c' = 3 * 1 + 0 * 1 = 3 \\
 z^o = -10 &\Rightarrow f(z^o) \approx 0 \Rightarrow y = 0 * c' = 0
 \end{aligned}$$

$$\begin{aligned}
 t = 2, \quad x^2 &= (1, 0, 1, -2), \quad c = 3 \\
 z = -2 &\Rightarrow g(z) = -2 \\
 z^i = 90 &\Rightarrow f(z^i) = 1 \\
 z^f = 10 &\Rightarrow f(z^f) \approx 1 \Rightarrow c' \approx 1 \\
 z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' \approx 1
 \end{aligned}$$

$$\begin{aligned}
 t = 3, \quad x^2 &= (1, 1, 1, 4), \quad c = 1 \\
 z = 4 &\Rightarrow g(z) = 4 \\
 z^i = 190 &\Rightarrow f(z^i) = 1 \\
 z^f = -90 &\Rightarrow f(z^f) \approx 0 \Rightarrow c' = 4 \\
 z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' = 4
 \end{aligned}$$

$$\begin{aligned}
 t = 4, \quad x^2 &= (0, 1, 1, 0), \quad c = 4 \\
 z = 0 &\Rightarrow g(z) = 0 \\
 z^i = 90 &\Rightarrow f(z^i) = 1 \\
 z^f = 10 &\Rightarrow f(z^f) \approx 1 \Rightarrow c' \approx 4 \\
 z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' \approx 4
 \end{aligned}$$

$$\begin{aligned}
 t = 5, \quad x^2 &= (0, 1, 0, 2), \quad c = 4 \\
 z = 2 &\Rightarrow g(z) = 2 \\
 z^i = 90 &\Rightarrow f(z^i) = 1 \\
 z^f = 10 &\Rightarrow f(z^f) \approx 1 \Rightarrow c' \approx 6 \\
 z^o = -10 &\Rightarrow f(z^o) \approx 0.000045 \Rightarrow y = 0.000045 * c' \approx 0.00027
 \end{aligned}$$

$$\begin{aligned}
 t = 6, \quad x^2 &= (1, 0, 1, -2), \quad c = 6 \\
 z = -4 &\Rightarrow g(z) = -4 \\
 z^i = -10 &\Rightarrow f(z^i) \approx 0.000045 \\
 z^f = 110 &\Rightarrow f(z^f) = 1 \Rightarrow c' \approx 6 \\
 z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' \approx 6
 \end{aligned}$$

$$t = 7, \quad x^2 = (1, 1, 1, 1), \quad c = 6$$

$$\begin{aligned}
z = 1 &\Rightarrow g(z) = 1 \\
z^i = 190 &\Rightarrow f(z^i) = 1 \\
z^f = -90 &\Rightarrow f(z^f) \approx 0 \Rightarrow c' = 1 \\
z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' = 1
\end{aligned}$$

$$\begin{aligned}
t = 8, \ x^2 &= (1, 0, 1, 2), \ c = 1 \\
z = 2 &\Rightarrow g(z) = 2 \\
z^i = 90 &\Rightarrow f(z^i) = 1 \\
z^f = 10 &\Rightarrow f(z^f) \approx 1 \Rightarrow c' \approx 3 \\
z^o = 90 &\Rightarrow f(z^o) = 1 \Rightarrow y = 1 * c' \approx 3
\end{aligned}$$

$$y = (0, 1, 4, 4, 0.00027, 6, 1, 3)$$