

Homework4 Report

Professor Pei-Yuan Wu
EE5184 - Machine Learning

姓名：顏宏宇

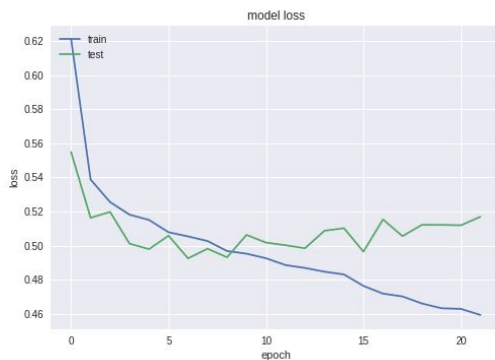
學號：r07942086

Problem 1. (0.5%) 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。(0.5%) 請實作 BOW+DNN 模型,敘述你的模型架構,回報正確率並繪出訓練曲線。

RNN: public score為0.76165

單層LSTM後接Dense layer最後接softmax

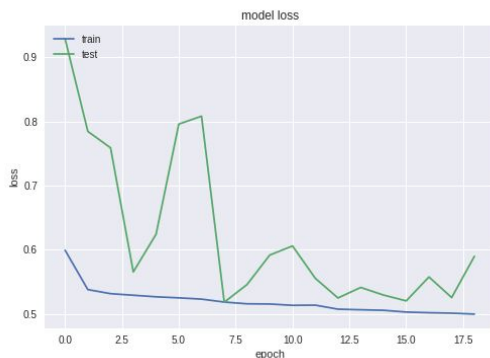
word embedding使用jieba斷字、用word2vec轉換成vector、取前120個word轉成vector當成input



BOW+DNN: public score為0.73827

純Dense layer最後接softmax

word embedding使用jieba斷字、用word2vec轉換成vector、把前120個word vector全部相加當成input, 概念上應該跟BOW一樣



Problem 2. (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。

在斷字之前我會先把全行字都換成半形、把句號頓號逗點都換成空白符號。

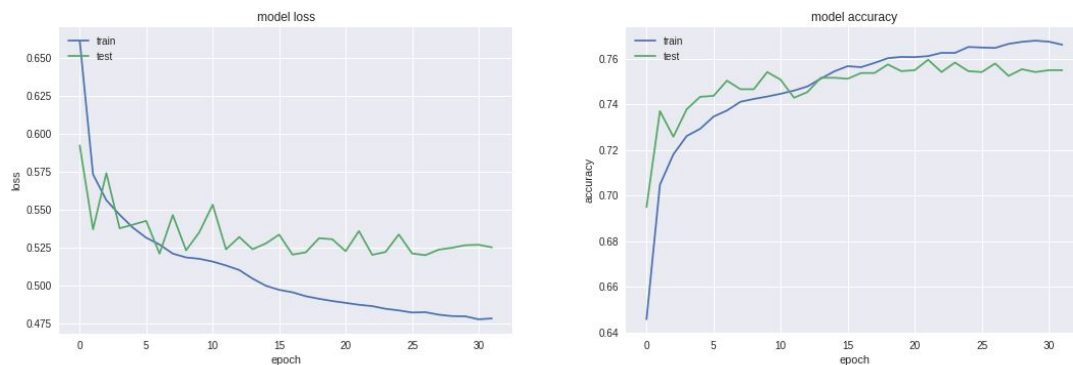
在斷字之後有把重複的字詞過濾掉, 例如連續相同的emoji只會抓第一個。

接著是把常常出現的「b01」或「B01」這種沒有特別意義的詞用python re統一換成「Dcard用戶」減少雜訊。

另外我還有試過各種不同的word2vec參數、input的長度等超參數, 想辦法找到最好的。

另外也多加幾層Dense以及Dropout, 再來設定LSTM的Dropout, 效果都會變得更好。

Problem 3. (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。



public score為0.75082, 比最好的RNN少了0.01左右, 原因可能是因為中文的字跟詞是不太一樣的單位, 若是只看一個個字沒辦法精確的考慮到差異而被誤導, 例如「白癡」跟「白色」都有白字, 但兩者的意義相差甚遠。

Problem 4. (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。

RNN: 前一句0.45447743、後一句0.5177116

BOW: 前一句0.53776234、後一句0.5377624

BOW只考慮詞出現的頻率, 而沒有考慮到順序關係, 因此兩句話的分數會非常接近。而RNN則會考慮到上下文關係, 所以可以看到RNN前一句分數低於0.5判斷不是惡意留言而後一句則高於0.5判斷為惡意留言。

Problem 4,5 collaborator : d07946003 王嘉澤

Problem 5.

x 0 1 2 3 4 5 6 7 8 9
 y + - + + + - - + - -

$t=1$
 f_1
 u_1 1 1 1 1 1 1 1 1 1 1
 $f_1(x)$ + + + + + - - - - -
 $z_1 = 10$
 $\varepsilon_1 = \frac{z_1}{10} = \frac{1}{5}$, $d_1 = \sqrt{\frac{4}{1}} = 2$
 $\alpha_1 \doteq 0.693$

$t=2$
 f_2
 u_2 $\frac{1}{2}$ 2 $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$
 $f_2(x)$ - - + + + + + + + +
 $z_2 = 8$
 $\varepsilon_2 = \frac{z_2}{8} = \frac{5}{16}$, $d_2 = \sqrt{\frac{11}{5}} \doteq 1.483$
 $\alpha_2 \doteq 0.394$

$t=3$
 f_3
 u_3 0.7415 1.348 0.337 0.337 0.337 0.7415 0.7415 1.348 0.7415 0.7415
 $f_3(x)$ + + + + + - - - - -
 $z_3 = 7.4145$
 $\varepsilon_3 = \frac{z_3}{7.4145} \doteq 0.364$
 $d_3 = \sqrt{\frac{0.636}{0.364}} \doteq 1.322$
 $\alpha_3 \doteq 0.279$

$\Rightarrow H(x) = \text{Sign}\left(\sum_{t=1}^3 \alpha_t f_t(x)\right)$

x 0 1 2 3 4 5 6 7 8 9
 $H(x)$ + + + + + - - - - - #

Problem 6.

$z = w_0 x + b$, $z_1 = w_1 x + b_1$, $z_2 = w_2 x + b_2$, $z_3 = w_3 x + b_3$

$y = f(z_0)h(c')$ $c' = f(z_1)g(z_2) + c f(z_3)$

$W = [0, 0, 0, 1]$ $b = 0$
 $w_1 = [100, 100, 0, 0]$ $b_1 = -10$
 $w_2 = [100, -100, 0, 0]$ $b_2 = 110$
 $w_3 = [0, 0, 100, 0]$ $b_3 = -10$

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

$t =$ 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

$t =$ 1 2 3 4 5 6 7 8
 z 3 -2 4 0 2 -4 1 2
 z_1 90 90 190 90 90 -10 190 90
 z_2 10 10 -90 10 10 110 -90 10
 z_3 -10 90 90 90 -10 90 90 90

$g(z)$ 3 -2 4 0 2 -4 1 2
 $f(z_1)$ 1 1 1 1 1 5×10^{-5} 1 1
 $f(z_2)$ 1 1 0 1 1 1 0 1
 $f(z_3)$ 5×10^{-5} 1 1 5×10^{-5} 1 1 1 1
 c 0 3 1 4 4 6 ≈ 6 1
 c' 3 1 4 4 6 ≈ 6 1 3
 $h(c')$ 3 1 4 4 6 6 1 3
 y 1.5×10^4 1 4 4 3×10^4 6 1 3 #