

HW4 Report

電機三 B05901043 莊鎧爾

Problem 1

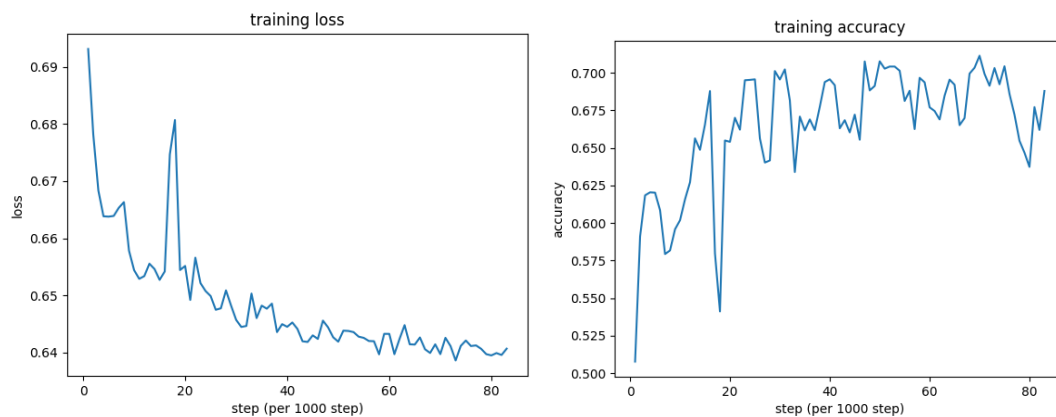
RNN 架構：

Layer type	Input (dimension)	Output (dimension)	Activation
Word Segmentation	sentence	List of word	
Word embedding	word	250	
LSTM	250	1000	
Linear	1000	500	Relu
Linear	500	250	Relu
Linear	250	2	Softmax

Word embedding 的方法：

先用 jieba 把中文句子剪斷成詞，之後放入 Word2Vec 實作 word embedding

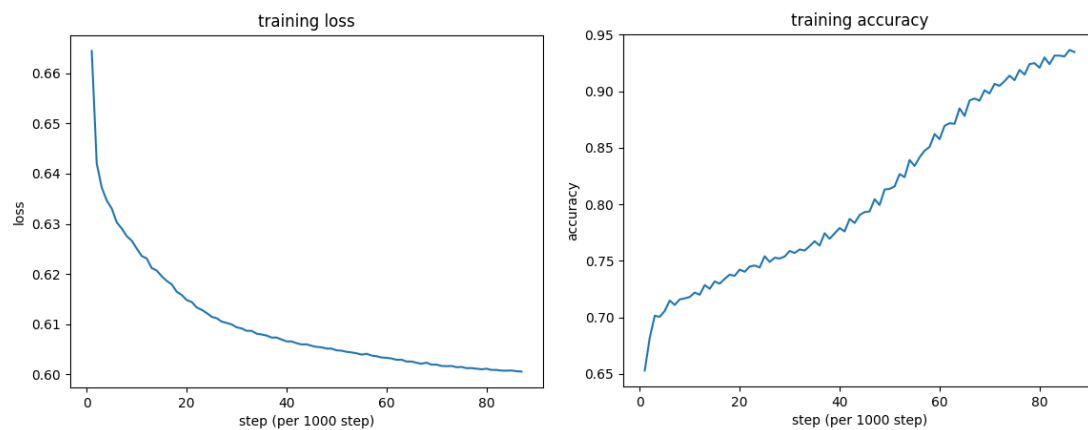
Training 正確率曲線：



BOW+DNN 模型：

Layer type	Input (dimension)	Output (dimension)	Activation
Word Segmentation	sentence	List of word	
Bag of Word	word	3087	
Linear	3087	1000	Relu
Linear	1000	500	Relu
Linear	1000	500	Relu
Linear	500	250	Relu
Linear	250	2	Softmax

Training 正確率曲線:



兩者架構比較：

	RNN	BOW+DNN
Testing set accuracy	0.72277	0.70127

雖然 training set DNN+BOW 可以表現的比較好，但是在 testing set 上，卻是 RNN 表現的較好

Problem 2

在 LSTM 層後，我增加 DNN 層數與每層的輸出數量，以此使模型 testing set 上表現進步。

Problem 3

	有斷詞	沒斷詞
Testing set accuracy	0.72277	0.69862

可以知道有做斷詞比沒做斷詞的結果好很多，因為有做斷詞可以較容易把相近意思的詞訓練出來，沒有斷詞可能需要較大 model 才可以達到一樣的效果。

Problem 4

	RNN	DNN+BOW
在說別人白痴之前，先想想自己	不是惡意	惡意
在說別人之前先想想自己，白痴	惡意	不是惡意

可以看出 RNN 可以得到較正確的答案，此外，直觀上 BOW 中兩個句子應該要得到一樣的結果，但是實際實驗得到不一樣的結果，推測是因為在做 word segmentation 會有不一樣的結果。

Problem 5

t	u_t	Classifier	error rate ϵ_t	α_t
1	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	$f_1 = \begin{cases} +1, & \text{if } x \leq 4 \\ -1, & \text{if } x > 4 \end{cases}$	0.2	0.6931
2	(0.5, 2, 0.5, 0.5, 0.5, 0.5, 0.5, 2, 0.5, 0.5)	$f_2 = \begin{cases} +1, & \text{if } x \geq 2 \\ -1, & \text{if } x < 2 \end{cases}$	0.3125	0.3942
3	(0.742, 1.348, 0.337, 0.337, 0.337, 0.742, 0.742, 1.348, 0.742, 0.742)	$f_3 = \begin{cases} +1, & \text{if } x \leq 0 \\ -1, & \text{if } x > 0 \end{cases}$	0.318	0.3815

Final classifier:

$$f = \text{sign}\left(\sum_{t=1}^3 \alpha_t \times f_t\right) = \begin{cases} -, & \text{if } x \geq 5 \text{ or } x = 1 \\ +, & \text{if } 2 \leq x < 5 \text{ or } x = 0 \end{cases}$$

Problem 6

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
z^t	3	-2	4	0	2	-4	1	2
z_i^t	90	90	190	90	90	-10	190	90
z_f^t	10	10	-90	10	10	110	-90	10
z_o^t	-10	90	90	90	-10	90	90	90
c^t	0	3.0	1.0	4.0	4.0	6.0	6.0	1.0
c'^t	3.0	1.0	4.0	4.0	6.0	6.0	1.0	3.0
y^t	3.0	1.0	4.0	4.0	0.000272	6.0	1.0	3.0