

1. (1%)請問 softmax 適不適合作為本次作業的 output layer? 寫出你最後選擇的 output layer 並說明理由。

我覺得 softmax 不適合作為這次作業的 output layer, 因為 softmax 會將讓最大值有最高的機率而壓縮其他的機率, 適合用在很多個分類選一個的情況, 但是對這題而言其實每個 tag 是二選一的狀況, 因此我覺得 Sigmoid 的表現會更好, 因此我選擇 Sigmoid 當最後的 output layer。

2. (1%)請設計實驗驗證上述推論。

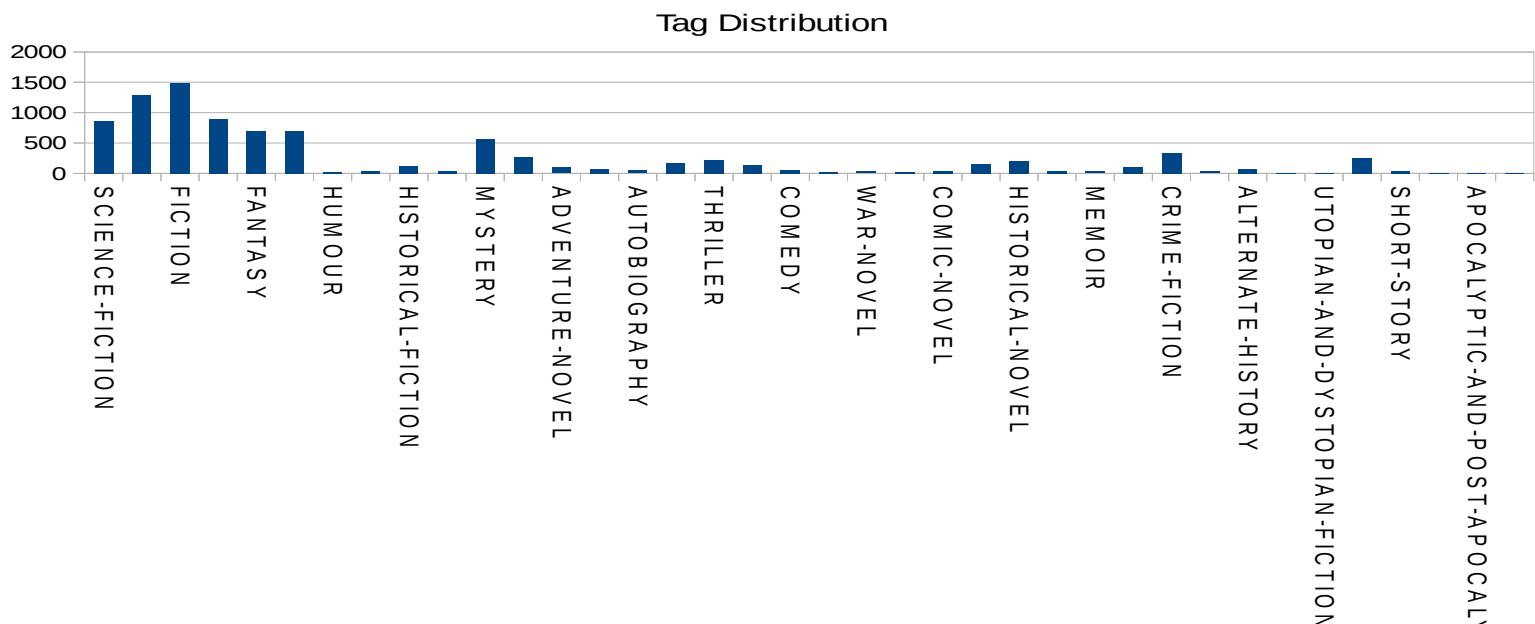
Kaggle Public Score:

Sigmoid: 0.50426

Softmax: 0.47646

Sigmoid 表現較好

3. (1%)請試著分析 tags 的分布情況(數量)。



最多的是 Fiction 和 Speculative fiction

4. (1%)本次作業中使用何種方式得到 word embedding?請簡單描述做法。

我使用的是 Glove 所訓練好的 word embedding, Glove 本身是處於 Count Base 的方法, Glove model 主要是利用大量的統計資料在一個字與字之間的 co-occurrence matrix 上訓練, 將每個文字都轉換成一個多維向量。兩個文字向量的內積愈高則代表兩個字之間愈相似, 而兩個字向量之間

的差也可以代表兩者之間的關係（相似關係的字組會有相似的差）。

5. (1%)試比較 bag of word 和 RNN 何者在本次作業中效果較好。
bag of words 表現較好。