學號：R05945039 系級： 生醫電資所碩一 姓名：張宇軒

* 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。

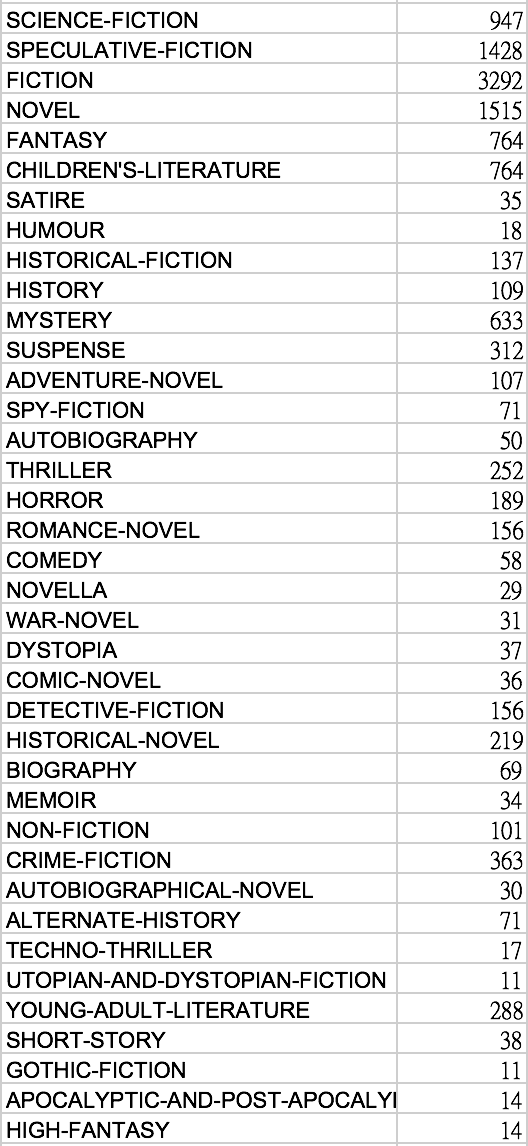
Softmax會把向量中的每一個元素的範圍處理在(0,1)，並且所有元素的合為1，結果會變成一個機率分佈，最後選用sigmoid當作output activation

* 1. (1%)請設計實驗驗證上述推論。

實際用softmax當作最後一層就會發現出來的結果被轉成機率分佈，threshold會很難設

* 1. (1%)請試著分析tags的分布情況(數量)。

從train data來看tag主要都是 FICTION，再來是NOVEL和SPECULATIVE-FICTION，其他的tag佔的比例都不多，從rnn輸出的結果來看也是這三種占的比例最高，其他的tag看起來很難被分類出來。



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

用glove 作word embedding，glove是global vectors for word representation的縮寫，glove的context vector跟word2vec不同，context裡只有一個word，glove認為兩個word的同現機率的比值可以作為向量差來當作編碼訊息

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

這次的simple baseline用rnn在助教的code上再加一層就會過，用bag of word則一直無法過simple baseline，但rnn的隨機性讓結果差異蠻大的，用bag of word作每一次的結果都可以輕鬆再現。