

RNN

我的模型用了兩層 Bidirectional LSTM 然後在用一個 time distributed softmax layer 輸出一個 sequence. 每一個 sequence 用了 500 字的 padding, batch size 用了 16, optimizer 是 adam optimizer. 第一次實驗, 每一個 blstm layer 用了 64 個單位, 最後的 test levenstein distance score 是 24%左右. 第二次 blstm layer 改用了 128 各單位, 大幅提高了最後的 test score, 到了 13% . 多加了幾個單位的原因是想讓神經網路, fit 得更好, 然後爲了防止 overfitting, 我兩次都利用了 dropout, dropout rate 用了 0.25, recurrent dropout rate 用了 0.5.

跟普通 unidirectional LSTM 相比, 我發現 bidirectional 在學習 sequence 的能力更強. 在實驗的過程中發現平均的 validation accuracy 比普通的 lstm 高 10%.

其他用的 lib 包括 pandas 來讀寫 csv/data, numpy data manipulation, sklearn 用來做 one hot encoding, 還有學習架構是用 keras 2.0, backend 是 tensorflow-gpu.