MLDS\_HW3 Report

Group name :

城隍府ＹＥＥ

Student ID\_Name :

B02901144\_\_林呈翰

B02901074\_\_陳奕禎

B02901030\_\_陳尚甫

B02901071\_\_黃淞楓

1. **組員貢獻**

林呈翰：transprob

林呈翰，黃淞楓：修改readLabel

陳奕禎，陳尚甫：viterbi

黃淞楓：打報告

1. **程式架構**
2. Data structure & algorithm design
3. hmmFunc.py

functions: transprob，用於統計transmission probability。

(二) hmmMain.py

functions: viterbi

我們先將之前dnn測出來的48個column的output存入observationProb，以及transprob()跑出來的二維機率矩陣存入transMatrix，然後用forward()一步一步做viterbi並把機率和路徑分別記錄在value和trace裡面，再把最大機率的路徑存取並回傳

(三) IO.py

modifying functions: readTrainLabel, deleteSil

將readTrainLabel的存取方式從dictionary改為二維list

deleteSil則是把sentence裡的sil拿掉

1. Comparison with DNN and RNN

DNN只要有夠深夠多的neuron和夠多的時間就能得到好的效果，但RNN則很難train出來，有時候還會比DNN還差，HMM則是僅僅將DNN的結果進行修正，搭配字句之間的關聯性讓結果更接近人會說出的句子，因此跟RNN相比是能以較少的時間得到較好的效果。

1. **實驗設計及結果**
2. 實驗結果：Kaggle上的score = 8.52027
3. 與RNN比較：我們的RNN並沒有很好的結果，調過很多參數(learning rate, neuron數, NAG的momentum大小)計算出的cost都大於三
4. 與DNN比較：經過smoothing, Kaggle score = 13.92632，HMM有較好的結果
5. 與RNN+HMM比較：