

數位語音處理概論 hw2

姓名：鄭澄遠

系級：機械四

學號：b03502059

Part I – baseline

```
Terminal Terminal File Edit View Search Terminal Help
frank@frank-VirtualBox: ~/Windows_data/SpeechSignalProcessing/dsp_hw2-1/dsp_hw2_baseline/result
===== HTK Results Analysis =====
Date: Mon May 14 00:18:54 2018
Ref : labels/answer.mlf
Rec : result/result.mlf
===== Overall Results =====
SENT: %Correct=38.54 [H=185, S=295, N=480]
WORD: %Corr=96.61, Acc=74.34 [H=1679, D=13, S=46, I=387, N=1738]
=====

"accuracy" 8L, 397C
```

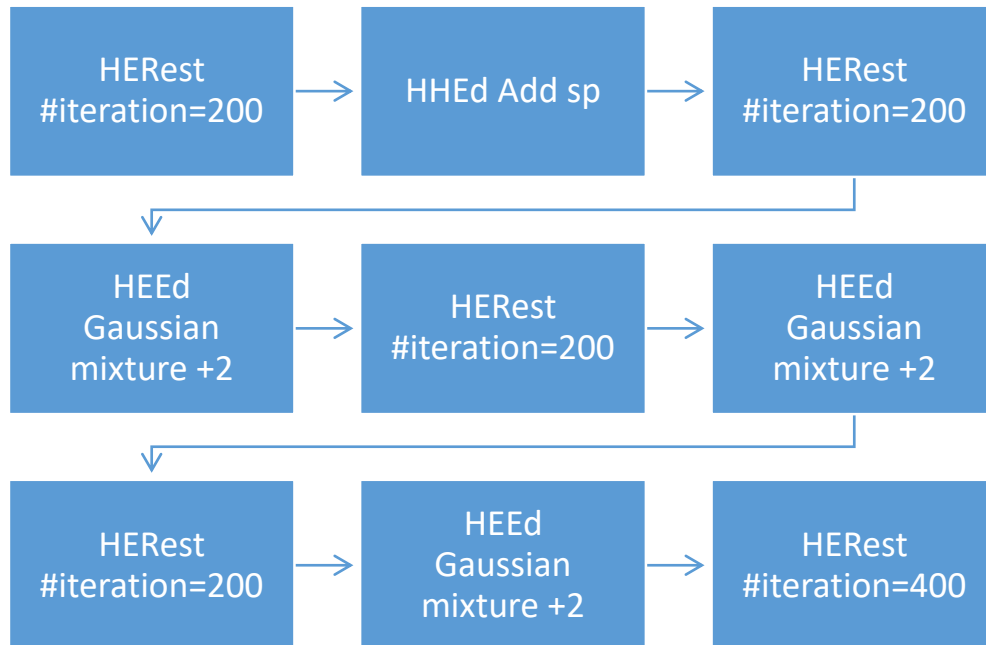
Part II – Improve Accuracy

```
Terminal Terminal File Edit View Search Terminal Help
frank@frank-VirtualBox: ~/Windows_data/SpeechSignalProcessing/dsp_hw2-1/dsp_hw2/result
===== HTK Results Analysis =====
Date: Mon May 14 09:09:55 2018
Ref : labels/answer.mlf
Rec : result/result.mlf
===== Overall Results =====
SENT: %Correct=85.21 [H=409, S=71, N=480]
WORD: %Corr=98.27, Acc=95.68 [H=1708, D=20, S=10, I=45, N=1738]
=====

"accuracy" 8L, 395C
```

Part III – Report

一開始的 HMM 只有 5 個 State，且 training 的 Re-estimate 數也都只有 3~6 次、所以我一開始的想法是把 State 慢慢增加，一開始嘗試加到 10 個 State，而且把訓練的 Re-estimate 數調整到(50, 50, 100)，Accuracy 達到了 91.25%，有大幅的進步。後來我發現 Gaussian mixture 的數量只有兩個好像太少了，就改變 train.sh 的架構如下：



State 的數量我調整到 14 個 State

這個訓練方式的 Accuracy 達到了 95.68%，但 State 的數量我覺得有點太多，這導致訓練的 Iteration 數必須很大才能收斂到比較好的結果，也增加了 Model 的複雜性可能會導致 Overfitting，這是必須注意的地方。