HW2

Install:

將CTC Speech Recognition的github檔案

git clone URL

進入檔案位置:

cd ctc-speech

把一些會使用的函式庫使用pip install 下載

準備data:

wget URL

解壓縮檔案

輸入指令

python generate\_audio\_cache.py –audio\_dir vctk-p225

開始訓練:

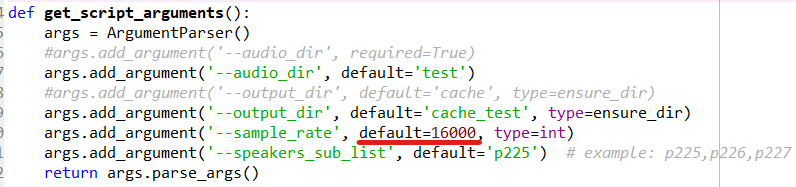
python ctc\_tensorflow\_example.py

我是在windows作業系統訓練資料

使用anaconda 的spyder來訓練

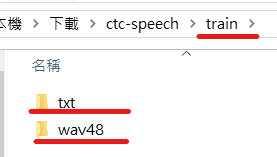
---------------------------------------------------------------------

將generate\_audio\_cache.py 畫紅線部分更改為16000

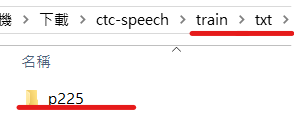


訓練音檔:

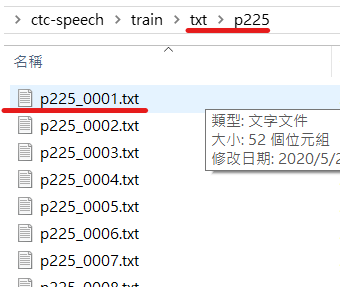
train資料夾下放這兩個檔案



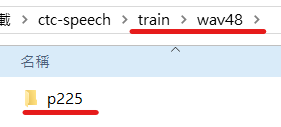
txt檔案下有p225



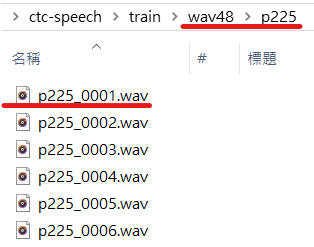
p225檔案底下將名字改為p225\_xxxx.txt



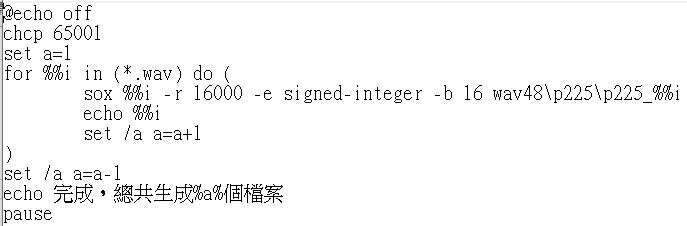
wav48檔案下有p225



p225檔案底下將名字改為p225\_xxxx.wav



音檔轉16k

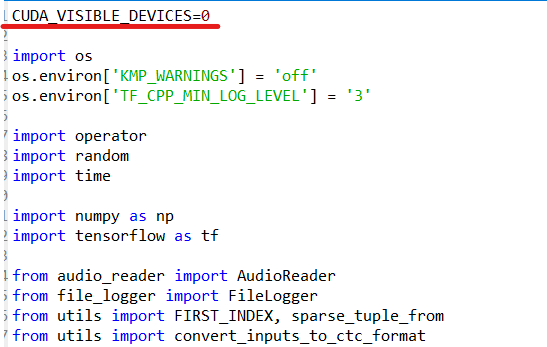


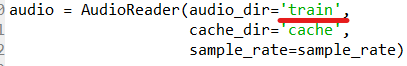
將這個檔案存為.bat檔

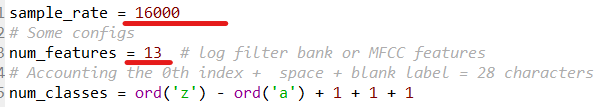
---------------------------------------------------------------------

train的部分:

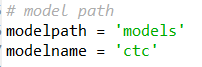
按照助教的ppt

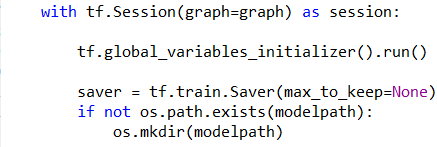


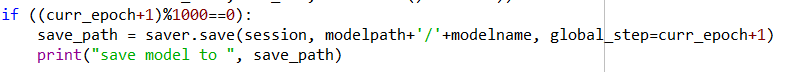




儲存模型:







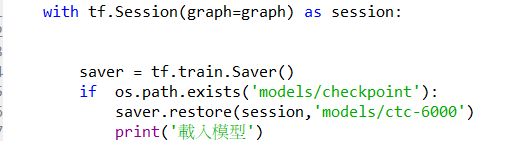
---------------------------------------------------------------------

test的部分

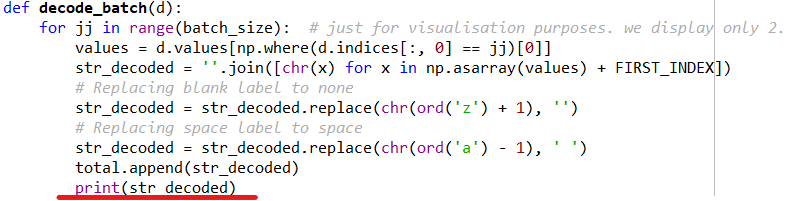
generate\_audio\_cache.py



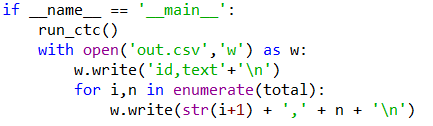
test.py



印出346行



讓輸出到csv檔案



---------------------------------------------------------------------

結果:

在kaggle上的Private Leaderboard的成績是14名，分數是5.49794，Public Leaderboard的成績是12名，分數是4.15533。