HW1

安裝kaldi

網路下載srilm-1.7.3.tar.gz

利用指令重新命名(mv srilm-1.7.3.tar.gz srilm.tgz)放到 ~/kaldi/tools/ 下

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

在 kaldi/egs/ 下創建新的資料夾(taiwanese)

kaldi/egs/taiwanese/ 下創建新資料夾(s5)

s5資料夾中有以下:

從formosa複製這些檔案(conf,local,cmd.sh,path.sh,run.sh)

從wsj複製這些檔案(steps,utils)

創建資料夾(data)

data資料夾裡創建(train,test),這兩個資料夾裡面有text.txt

創建資料夾(lexicon)

lexicon資料夾裡有lexicon.txt

創建資料夾(train,test),這兩個資料夾裡面有wav資料夾,wav資料夾裡面有(\*.wav,text.txt)

將csv音檔轉成16 kHz sampling, signed-integer, 16 bits指令

find . -name '\*.wav' -exec sox {} -r 16000 -e signed-integer -b 16 ~/kaldi/egs/taiwanese/s5/data/train/wav/{} \;

text.txt是從train-tonless.csv轉檔過去

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

cmd.sh :

留下這兩行

export train\_cmd=run.pl

export decode\_cmd=run.pl

其他註解

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

run.sh :

找到這兩個(train\_dir,eval\_dir),將第三行eval\_key\_dir = ...刪除

train\_dir = train/wav

eval\_dir = test/wav

#Data Preparation

echo "$0: Data Preparation"

local/prepare\_data.sh --train-dir $train\_dir --eval-dir $eval\_dir || exit 1; (這邊要更改)

# mfcc

if [ $stage -le -1 ]; then

echo "$0: making mfccs"

for x in train test ; do (這邊要更改)

steps/make\_mfcc\_pitch.sh --cmd "$train\_cmd" --nj $num\_jobs data/$x exp/make\_mfcc/$x $mfccdir || exit 1;

steps/compute\_cmvn\_stats.sh data/$x exp/make\_mfcc/$x $mfccdir || exit 1;

utils/fix\_data\_dir.sh data/$x || exit 1;

done

fi

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

data/train/text.txt :

總共有3120行包含(id,text)

data/test/text.txt :

總共有347行包含(id,text)

local/prepare\_data.sh :

更改這兩行

train\_dir=train/wav

eval\_dir=test/wav

# have to remove previous files to avoid filtering speakers according to cmvn.scp and feats.scp

rm -rf data/local/train (這邊要更改)

mkdir -p data/local/train (這邊要更改)

# make utt2spk, wav.scp and text (以下三行要更改)

find -L $train\_dir -name \*.wav -exec sh -c 'x={}; y=$(basename -s .wav $x); printf "%s %s\n" $y $y' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/train/utt2spk

find -L $train\_dir -name \*.wav -exec sh -c 'x={}; y=$(basename -s .wav $x); printf "%s %s\n" $y $x' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/train/wav.scp

find -L $train\_dir -name \*.txt -exec sh -c 'x={}; y=$(basename -s .txt $x); printf "%s " $y; cat $x' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/train/text

utils/fix\_data\_dir.sh data/train (這邊要更改)

(中間有一段可刪除)

# for LM training

echo "cp data/train/text data/local/train/text for language model training"

cat data/train/text | awk '{$1=""}1;' | awk '{$1=$1}1;' > data/local/train/text

# preparing EVAL set. (以下三行要更改)

find -L $eval\_dir -name \*.wav -exec sh -c 'x={}; y=$(basename -s .wav $x); printf "%s %s\n" $y $y' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/test/utt2spk

find -L $eval\_dir -name \*.wav -exec sh -c 'x={}; y=$(basename -s .wav $x); printf "%s %s\n" $y $x' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/test/wav.scp

find -L $eval\_dir -name \*.txt -exec sh -c 'x={}; y=$(basename -s .txt $x); printf "%s " $y; cat $x' \; | sed 's/\xe3\x80\x80\|\xc2\xa0//g' | dos2unix > data/test/text

utils/fix\_data\_dir.sh data/test (這邊要更改)

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

local/prepare\_dict.sh :

source\_dir=lexicon

dict\_dir=data/local/dict

rm -rf $dict\_dir

mkdir -p $dict\_dir

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

local/nnet3/run\_ivector\_common.sh :

stage=0

train\_set=train

test\_sets="test" (這邊要更改)

gmm=tri5a

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

local/chain/run\_tdnn.sh :

# training options

num\_epochs=6

initial\_effective\_lrate=0.00025

final\_effective\_lrate=0.000025

max\_param\_change=2.0

final\_layer\_normalize\_target=0.5

num\_jobs\_initial=3

num\_jobs\_final=8 (這邊要更改)

minibatch\_size=64

frames\_per\_eg=150,110,90

remove\_egs=false

common\_egs\_dir=

xent\_regularize=0.1

dropout\_schedule='0,0@0.20,0.5@0.50,0'

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

開始訓練:

進入資料夾

cd kaldi/egs/taiwanese/s5

./run.sh

結果:

在kaggle上的Private Leaderboard的成績是26名，分數是4.57613，Public Leaderboard的成績是25名，分數是5.01941。