Assignment 3 Word boundary Detection
ELL888 Advanced Machine Learning

-Krishna Prasanth 2016EET2630
**-Manoj Rathor 2016EEY7520**
-Priyank Soni 2016EET2639

Dataset:

In the given data, we have wav files and wrd files for different speakers. Wav file contains the audio signal of a speaker. Wrd file contains the words uttered by the speaker with word boundaries. The audio signal contains amplitude of sound at each time instance(given in number of samples). Sampling rate is 16000 Hz.

Specification : Given data is time series data so we have to apply those techniques which uses time series dependency. LSTM is the most common technique for these type of data.
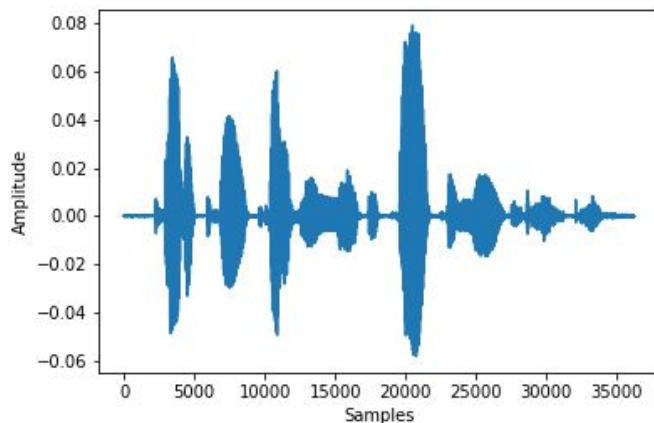
Problem : How to give labels to data as there can be many methods for giving labels to data.Our Approach : we gave each sample a class. These classes will depends on the location of that sample in the data.

We have tried 4 different ideas to implement this approach.

Idea 1: We have labeled the data such that each sample will get either of 4 numbers. If the sample is starting of the word then we assign 0 as the label. If it is ending of the word we assign 1 as the label. If it is middle of the word we assign 2 as the label. If it is a space then we assign 3 as the label. We solved the above approach using both regression and classification. In regression we gave 0,1,2,3 as our output for each sample and in classification we used 0,1,2,3 as the class of each sample.

Problems : In this approach data becomes unbalanced dataset because we were giving 2 for middle of the word and maximum samples are middle of the word.

Solution : instead of giving a four classes to samples we gave each sample a class according to their word position in the dictionary(index of the dictionary corresponding to that word)

1. As a Regression problem:

Filename:
traindata_new.py
LSTM.py

Here accuracy is defined as detecting the correct label corresponding to each sample.

Method	samples	Values	epoch	Validation Accuracy(%)	Testing Accuracy(%)
LSTM	250	Word indexes in vocabulary	6	20.2	30.87
LSTM	200	Word indexes in vocabulary	10	20.65	38.96

We are taking a threshold while scanning each window, if its less than threshold inside window than 1 word else more than one word.

2. As a classification problem:

Foldername: classification using 4 classes

Here accuracy is defined as detecting the correct label corresponding to each sample.

Method	samples	class	epoch	Validation Accuracy(%)	Testing Accuracy(%)
LSTM	400	4	5	56.58	56.10
LSTM	400	4	20	40.28	59.80

Idea 2: First we have created a dictionary of words using training and test data. Now we have index for each word of training and test data. Then we have labeled the data such that each sample of a word will assigned the corresponding word index. After labelling the data, we used two different approaches.

1. As a Regression problem:

Filename:
traindata_dict.py
LSTM_final.py
testdata_final.py

Here accuracy is defined as detecting the correct label corresponding to each sample.

Method	samples	Values	epoch	Validation Accuracy(%)	Testing Accuracy(%)
LSTM	300	Word indexes in vocabulary	5	22.65	30.18
LSTM	250	Word indexes in vocabulary	15	19.44	35.10

2. As a classification problem:
traindata_dict_one_hot.py
LSTM.py

Here accuracy is defined as detecting the correct label corresponding to each sample.

Method	samples	class	epoch	Validation Accuracy(%)	Testing Accuracy(%)
LSTM	400	Equal to size of vocabulary	5		56.10