Assignment 5 - Audio Classification

Create Neural Network model for the problem of “Wake-up word detection” 

For example, you have a personal assistant device Alexa which listens for a wakeup word from you like “Alexa” or a custom word like “Ping” or “Siri”. Sometimes you want to confuse Ms. Alexa by saying “Pong” instead of “Ping” (or “free” instead of “Siri”). The motivation is that you have to train an audio classification model with positive and negative samples as “Ping vs Pong” or “Alexa vs Telaxa” to make it more robust to the similarly sounding words.

Note: This is an interesting research problem as audio features are similar! 😊

Steps

1. Create a dataset of short audio files (like 1 sec to 3 sec), 10 for “Ping” and 10 for “Pong”. You can use your mobile to record the audio samples and further upload to your PC/Laptop. The format of the audio can be anything such as .wav, .mp3.
2. Read the audio using any library (sklearn has wav audio file reader). Extract the STFT features from the audio. STFT (Short Term Fourier Transform) is a frequency feature representation for audio. Since audio samples differ in length, STFT will have differing lengths. Hence extra entries of matrix can be removed from the STFT features.
3. Split the dataset (75% training and 25% testing) into training and testing sets with STFT audio features as input, audio class as target label.
4. Train a Neural Network model of your choice to do the classification.
5. Run different Neural Network model with different number of nodes such as 8, 16, 32, 64, 128 etc. and different number of layers such as 2, 3, 4 etc. For all the models plot the number of parameters learned, training accuracy, testing accuracy and running time for testing in bar chart. Analyse the results and discuss what you discovered!

Suggested Libraries:

Python packages such as Numpy, Sklearn, Keras

Submission:

Submit your files in **Single ipython Notebook** in LMS before **Saturday 17th Feb, 11.59 pm.**

Marking:

Marking is based on both **performance during the lab hours** as well as **complete submission in LMS**.