Analyzing Voice Samples to Predict Age

Milestone Report

Raj Singh

Problem Statement

The project is for a company called Neurolex Labs. Neurolex analyzes voice samples to help diagnose diseases such as schizophrenia, Parkinson's disease, Alzheimer's disease, and depression. In this case, we are analyzing human voice samples to determine the age of the speaker. We may want to know the age range of the individual so we can get specialized care for people in that age range. Moreover, knowing the age of the individual can give us a better context for the nature of their disease, and how likely they are to have a particular disease.

The Data

The name of the dataset is called the Common Voice Dataset by Mozilla. It contains about 200,000 voice samples. It contains the feature we are interested in, which is the age of the speaker. In order to featurize this audio data, we will be using two Python libraries: pyAudioAnalysis and Librosa.

Data Cleaning

The data can be downloaded as a .csv file from the internet, so we can load in the data as a Pandas Dataframe. Not all entries of the dataset were labeled with age categories, so those entries must be deleted. In order to featurize the data for our multilayer perceptron model, we use pyAudioAnalysis to featurize the data. An example of this can be found here. We also use Librosa to convert the audio files into log-mel-spectrograms, which is essentially image data with two channels. An example of this can be found here. Generally speaking, there were not many data cleaning steps except for those that involved featurizing the audio data.