Assignment #2 - Audio Sensing
EE382V Activity Sensing and Recognition - Fall 2020

*Due **Oct 5th 2020***

Turn-in your assignment as a zip file including your code and data. Label which files correspond to which answers very carefully. Make it intelligible. If this is not clear, you will not receive credit for your solution.

Q1. You have just been hired at Audio Intelligence Co. as a machine learning researcher. Your first job is to design an auditory scene analysis system that leverages ambient audio to identify whether a person is in one of four possible environments: *at a coffee shop*, *in the kitchen*, *at a soccer game, or at a party*. You are asked to build this system using Python, scikit-learn, and librosa (https://librosa.github.io).

Training Data

To complete this task, four files have been made available to you: *coffee.wav*, *kitchen.wav, soccer.wav, and party.mp3*. The files are located on Canvas (Files > Assignments > A2).

Problem

a) Build a multi-class classifier that can recognize the environment a person is in given these 4 possible choices and their respective audio files. Evaluate the classifier using the split train/test metric (80% train, 20% test) with 3 learning algorithms of your choice.

For this problem, you will need to extract audio features from each of the audio files. There are a number of audio feature extraction tools available in Python, but you must use librosa to facilitate our grading process. Audio extraction and classification parameters such as frame size, whether you perform feature selection or not, are up to you. *Bonus: include a document (pdf format) explaining the rationale for the design choices you made.*

Grading

Your grade will be based not only on whether your results are reasonable but also on the design choices that you make regarding features, pre-processing, frame extraction, classifier, etc. *There is not one right answer, but Dr. Thomaz reserves the right to award Tiff's Treats to the designer of the best performing classifier.*

Note that one of the ways we will test your system/classifier is with audio samples that were not made available to you. This will be a true test to check how well your audio classifier generalizes to data it has never seen.