**ECE 6254**

**Final Project – Ideas / Discussion / Process**

This file is for us to exchange ideas, discuss progress and will help making the final report a lot easier as we’ll be documenting things throughout.

File rules:

Color coded fonts so that there’s no confusion as to who is saying/doing what

Ananya – Orange/Yellow (this color)

Varun – Blue

Tillson – Black

Use red color for:

* TODO – Self explanatory – just to list what needs to be done and by whom
* Questions – questions by one person to the rest of the group

11/21/2022 Updates

* TG: First off, I added the data files to the GitHub repo and wrote a script to pre-process them into numpy. This should help in testing multiple models so we have a consistent dataset.
  + Because the files are large, I used Github LFS (<https://git-lfs.github.com/>) to store them. If you have issues setting this up, you’ll probably be ok just re-generating the data using `create\_dataset\_v02\_AB.m` followed by `python preprocess\_data.py [output\_of\_previous\_script]`.
* Also added a simple neural network and plotted loss/accuracy over epochs. Definitely more we can do here with hyperparameters and architecture, but the framework is now there so we can play with it.

Day 1 Updates (11/14/2022)

I created a dataset – based on 8 different signals (classes), then expanded that to 24 – by randomly distributing where in the time vector the signals appear.  
Each input/signal is 10000 samples in length, or d = 10000.

The I duplicated the data by adding white gaussian noise at various amplitudes (signal to noise ratios) and a strong reflector also distributed randomly in time (again with varying amplitudes) to yield a dataset of 1200 samples.

You need to run the matlab file to create the dataset. I have also put that on GitHub for the first go around.

Question 1 – Is the dataset size okay? Just based on VC dimension analysis? Since our dataset is synthetic (for now, I do have real data as well, but to start off I was thinking we can work with just this), I can add as much data as we need.

Question 2 – What classification techniques do we want to test and compare the performance of?

TODO – Signal Processing Approach Implementation as a comparison of the traditional method – for Ananya