**CSC 417 Unit 3 Day 4 Outline**

1. Natural Language Processing
   1. Voice Recognition
      1. Introduction
         1. Spoken languages presents many challenges for detection/recognition
            1. Inconsistent sounds
            2. Accents
            3. Speed
         2. Language is not uniform across all human speakers
      2. Process
         1. Preparing data
            1. Necessary to encode audio waves in digital form

Mel Spectrogram – “picture” of sound

* + - 1. Data Augmentation
         1. Speech data is time consuming to record – how can we vary existing data for a more robust training set?

Artificially adjust parameters (pitch, speed, reverb, etc.)

Spectrogram augmentation

Randomly “cut out” parts of spectrogram

Provides many variations of the same speech for training

* + - 1. Deep Learning Model
         1. CNNs and RNNs often employed
         2. Softmax method assigns decimal probabilities to each class in a multi-class problem (must add up to 1)