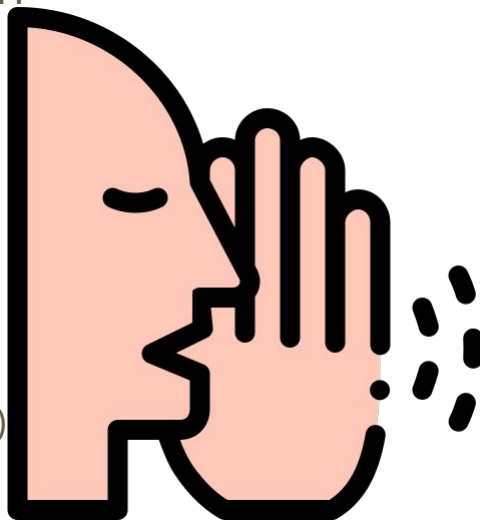
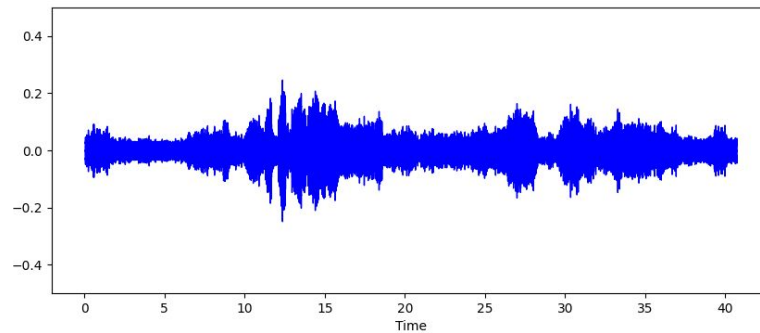
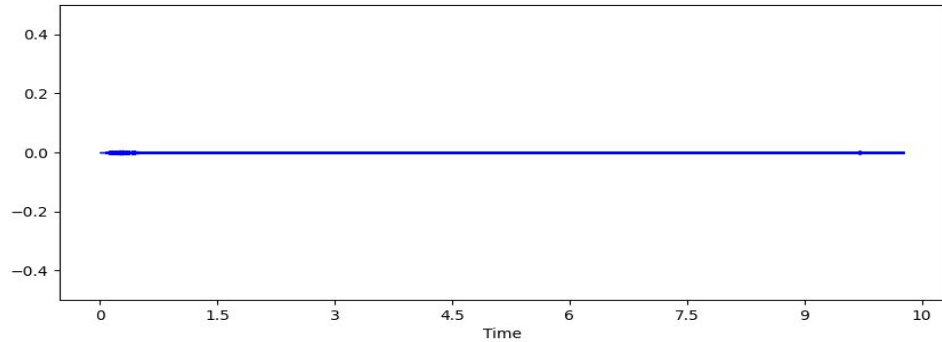

Classifying Speech

Lauren Cockey, Sarah Hecker,
Ashanti Roeung

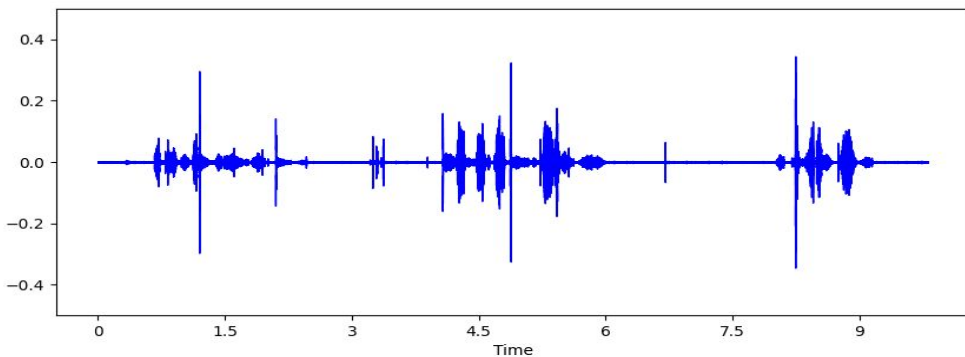
Our Project

- Our goal is to create a program to classify the type of speech present in a voice file
- Possible outcomes of the program
 - Talking
 - Not talking
 - Whispering
- Each group member provided multiple different audio files
 - Multiple 10s samples of each audio(talking, not talking, and whispering)

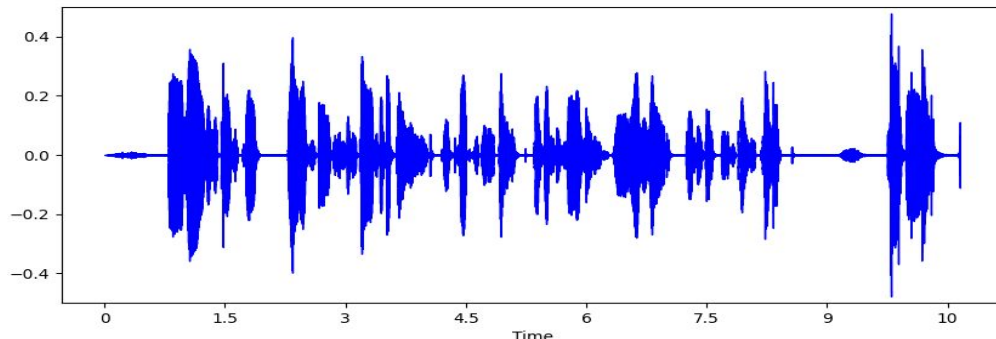




Not Talking (the 2 above, left is silence, right is city noise/traffic)



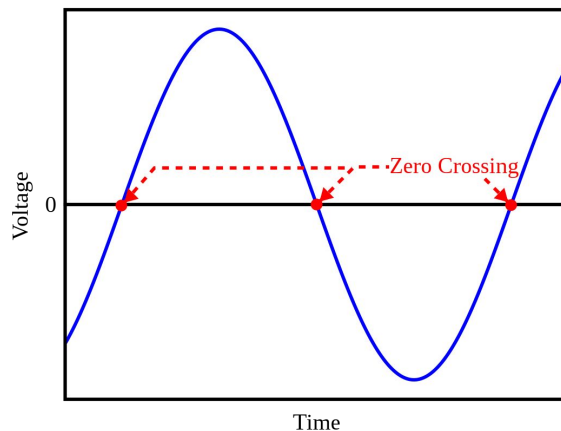
Whispering

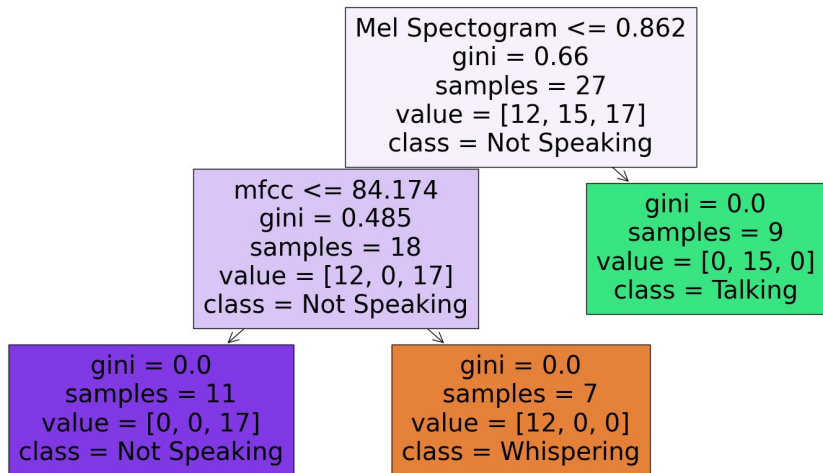


Talking

Methods

- Utilized zero crossings, MFCC, amplitude, RMS, spectral contrast
- Based off of the assignment 4 code, which detected features.
- Used the audio files we recorded to test the code
- Used the random forest from assignment 4

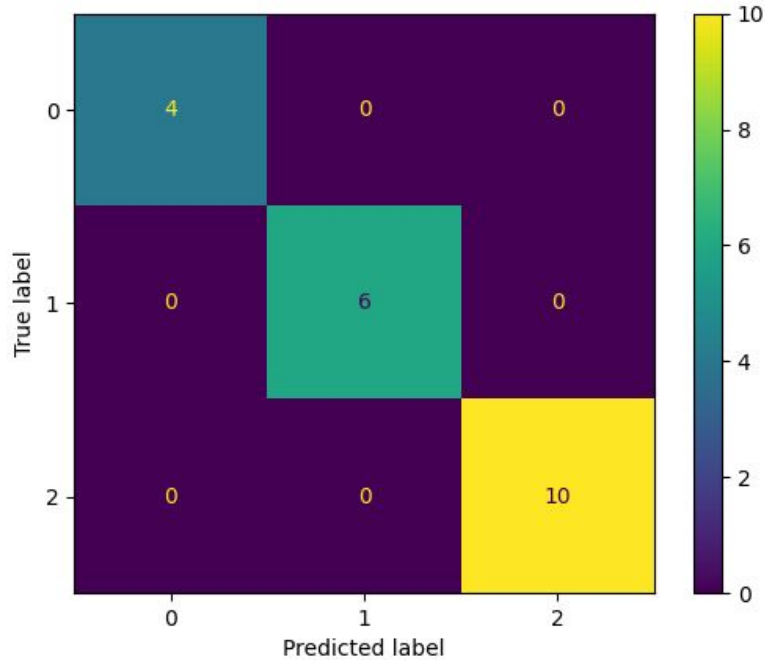




Decision Tree

- Because a random forest was used, this is the visual representation for one of the decision trees

* Updated with city noise
under not talking



Confusion Matrix

Results



- Could be utilized for auto captioning and audio description in television or videos
- What we learned:
 - The features are easily distinguishable between the three types of speech
- Could be used by smart devices and voice assistants
 - Alexa has a whisper mode where it detects if the user is whispering or not to it and will respond in the same manner
- We could detect more types of speech, did only three because of time constraints
- **UPDATE:** Added code, most important features were the Mel Spectrogram and MFCC
- **UPDATE:** Added traffic noise/other ambiance