

Thomas Powell
Bill Pan
Ryan Casey

Project Overview

The Spoken Emotion Distinguisher, or SED, aims to assist people with understanding and producing emotion in their professional or day-to-day speech. The project is interesting because it combines machine learning models with user-friendly interactives to provide helpful information about emotion in speech and tips on how a user can reach a desired feeling while talking. The project aims to create a centralized hub for analyzing and improving emotion in speech. SED will benefit people with communication disorders or disabilities that affect their speech and improve their emotion recognition and projection. SED will also be a great tool for people who do public speaking or voice acting, allowing them to grow their abilities to apply emotion to their voice. Our project is going to be free, made for personal use, and can accept input audio to be analyzed for emotional content.

Major Components

Software

- Flask (Python)
 - Our website's back-end
 - Handles user input files
 - The preprocessing necessary for our neural network
 - Neural network model
 - Routing between different web pages
- HTML
 - Our web pages front-end
 - Form for user input
 - Text, Images and Videos
 - Site Navigation
- CSS
 - Our web pages styling
- Python
 - Built our neural network
 - This will be used to manage, manipulate and extract data
- Visual Studio Code
 - This will be used to write code for Flask, Python, HTML and CSS
 - This will be used to test the website and machine learning algorithm
- Github

- This will be used as the repository for the code to be shared between group members

Hardware

- Computer with an Internet connection
 - This will be used to access our code on the repository and test our website

Network

- Website Hosting
 - Allows our website to be accessed from any device with Flask and the directory housing web application
- PHP
 - Hosts our website on the local machine

Design Decisions

Website Frontend

- Displays home web page
 - Prompts the user to input a wav audio file
 - Displays predictive models emotional analysis of inputted audio file
- Displays web page of tips on how to reach desired emotion in speech
- Displays web page of an about me of the project
- Designed using HTML and styled using CSS

Website Backend

- Data preprocessing for user input
- Neural network model prediction
 - Model
 - Encoder
 - Scalar
- Web page routing
- Designed using Flask

Machine Learning

- RAVDESS database
- Categorical
- Trained based on audio features and eight different emotions
- Neural network model
 - Sequential
 - Four Pooling Layers
 - Four 1D Convolution Layers
 - A Dropout Layer
 - Trained Scaler
 - Trained Encoder

Project Milestones

- Complete a working machine-learning model for our dataset
 - A high success rate in predicting emotion from speech
 - Works with English
 - Works with both sexes
 - Works with adolescents and elderly people
- Complete the front end for our website
 - Allows user input of audio files
 - Displays the result of our machine-learning model
 - Gives tips to the speaker on how to achieve a desired emotion
- Complete the back end of our website
 - Sends user input to a machine-learning model
 - Gets results from machine-learning model
 - Routing between web pages