Project Proposal

Members: Alan Cai, Ali Saleh, Manraj Singh, Mekdes Teklehaimanot

Project Title: Speech Emotion Recognition: An Analysis of Performance and Applicability

Project Objectives: Develop ML models to recognize emotions from speech audio. This project aims to create an accurate and efficient machine learning model that can classify emotions (e.g., happiness, sadness, anger, surprise, etc.) in spoken language, enabling applications such as sentiment analysis, virtual assistants, and customer service automation. This project will also explore dimensionality reduction techniques to reduce the complexity of the data and improve the performance of the different models.

Data Source: The data for this project can be obtained from [Kaggle.](#) This portion of the RAVDESS contains 1440 files: 60 trials per actor x 24 actors = 1440. The RAVDESS contains 24 professional actors (12 female & 12 male), vocalizing two similar linguistic statements in a neutral North American accent. Speech emotions include calm, happy, sad, angry, fearful, surprise, and disgust expressions. Each expression is produced at two levels of emotional intensity (normal & strong), with an additional neutral expression.

Methods/Packages: Feature extraction techniques such as Mel-frequency cepstral coefficients (MFCCs), spectral features, and prosodic features will be employed to transform raw audio data into a suitable format for machine learning models. Data preprocessing and feature engineering will be performed to extract relevant information from the raw audio data, using techniques such as PCA, UMAP, etc. Python libraries such as LibROSA (library for audio/music analysis), TensorFlow, and Keras will be used for implementing the MLP and deep neural network models and evaluating their performance.

Expected Outcome: Demonstrate the effectiveness of MLP and neural network-based models for speech emotion recognition, providing insights into their performance and applicability for various audio processing tasks. Success in this project will be measured through appropriate accuracy metrics. The project will also offer valuable experience in working with audio data and feature extraction techniques, and their approaches in real-life application.