Emotion detection from speech data is an emerging field in machine learning that focuses on recognizing emotions through voice patterns. This project aims to build a model for detecting emotions using a dataset composed of four widely-used databases: CREMA, RAVDESS, SAVEE, and TESS. These datasets provide a diverse range of emotional speech recordings in terms of age, gender, and accents, contributing to a more generalized emotion recognition model.

The process involves preprocessing the audio files by extracting relevant features like Mel-frequency cepstral coefficients (MFCCs), which capture the essence of speech signals. These features are fed into machine learning models like support vector machines (SVM) or deep learning architectures such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs).

The project focuses on detecting key emotions like happiness, anger, sadness, fear, and surprise, with the goal of achieving high accuracy in real-time applications such as virtual assistants, customer service, and mental health monitoring. By leveraging these datasets, this project aims to contribute to advancements in emotion recognition systems, improving human-computer interaction.