# Gender Classification from Voice Recordings

## Project Overview

This project aims to develop a gender classification system based on voice recordings using deep learning techniques. The system utilizes a pre-trained Wav2Vec2 model for audio processing and gender prediction. The workflow encompasses several stages, including voice activity detection, noise reduction, audio normalization, and gender classification.

## Components

### 1. Voice Activity Detection

Voice activity detection (VAD) is implemented using the Silero VAD model, which identifies segments of audio that contain speech. This step is crucial for filtering out silence and noise, ensuring that only relevant audio segments are processed for gender classification.

### 2. Noise Reduction

To enhance the quality of the audio recordings, a noise reduction process is applied. This step minimizes background noise, improving the clarity of the speech and the performance of the subsequent classification model.

### 3. Audio Normalization

Audio normalization ensures that the volume levels of the audio files are consistent, facilitating better feature extraction during model inference. This process involves adjusting the amplitude of the audio signals to a standard level.

### 4. Gender Classification

The core of the project is the gender classification model, which is built on the Wav2Vec2 architecture. This model is trained to differentiate between male and female voices based on the features extracted from the audio recordings. The model leverages deep learning techniques to achieve high accuracy in predictions.

## Workflow

The workflow of the project is structured as follows: - Load the audio files for analysis. - Apply voice activity detection to filter out non-speech segments. - Execute noise reduction and normalization on the detected speech segments. - Use the Wav2Vec2 model to predict the gender of the speaker based on the processed audio.

## Model Information

The finetuned Wav2Vec2 model utilized in this project is stored in a zip file, which contains all necessary components for inference. The model has been fine-tuned on a gender classification dataset to ensure optimal performance.

## Installation and Requirements

To run this project, ensure you have the following libraries installed: - torch - torchaudio - transformers - librosa - numpy

Use pip or conda to install these dependencies.

## Conclusion

This project demonstrates the effectiveness of using deep learning models for audio classification tasks, specifically gender prediction. The combination of voice activity detection, noise reduction, and normalization enhances the quality of the audio data, leading to improved classification results.

## Model Download Link

You can download the finetuned Wav2Vec2 model from the following link:

[Download Wav2Vec2 Model](https://drive.google.com/file/d/1ievcRB8JpGoOULBHjdKRLabMS3R5uiSB/view?usp=sharing)