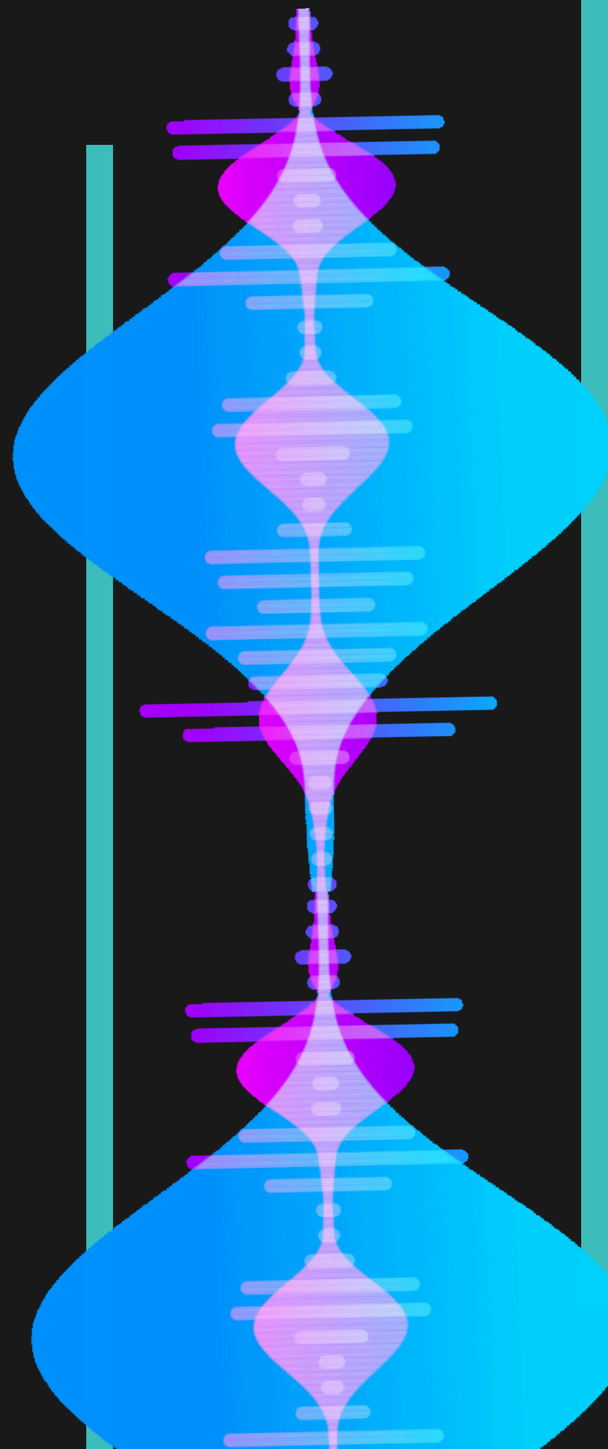
TALKDOC

A.I speech therapy assistant



PROBLEM STATEMENT

We are here to design an AI speech therapy tool that aids people with speech difficulties, in particular stuttering and stammering. The tool analyses the user's speech and based on inherent data suggests courses and feedback to improve the pronunciation and communication skills of the individual. The data is a collection of speech samples from individuals affected by stuttering.The tool also helps individuals learn at their pace, provides personalised therapy plans, interactive sessions, and progress tracking, is accessible at all times, and is cost-efficient. We think a clear speech helps us express ideas with confidence and that we are never too old to learn something the correct way.

PRIMARY FOCUS

5-Step Process

1

Dataset collection and storage

2

Use of python libraries to analyse and provide speech feedback

3

Selecting threshold to match input features with reference features

4

Train appliation to detect 80 commonly mispronounced Words and Phrases in English with meaning

5

Develop improvement and feedback tracker

IMPLEMENTATION

1

1. Collect Stuttering and Non-Stuttering Speech Data: Gather a dataset of speech samples that include both stuttering and non-stuttering speech. You may need to collaborate with speech therapists or organizations to obtain appropriate data.

2.Data Preprocessing: Audio Preprocessing: Preprocess the audio data to ensure it's in a suitable format for analysis. This may include resampling, noise reduction, and audio feature extraction.

2

Extract Audio Features: Common libraries in python:

⁎Vosk:Vosk is a speech recognition library that is particularly useful for developing automatic speech recognition (ASR) systems.

⁎soundfile:It can be used to handle audio data and store audio recordings.

⁎speechbrain:It provides a wide range of pre-trained models, as well as tools for training your own speech models.

3

Extract Audio Features: Extract relevant audio features from the speech data. Common features used in speech analysis include:

- Mel-frequency cepstral coefficients (MFCCs)
- Pitch and fundamental frequency (F0)
- Spectral features
- Prosodic features
- Duration and pause information

4

1. **Machine Learning Model:** Train a machine learning model to classify speech as stuttering or non-stuttering. Common machine learning algorithms used for audio classification include:
   - Support Vector Machine (SVM)
   - Random Forest
   - Neural Networks (e.g., with libraries like TensorFlow or PyTorch)
2. **Model Optimization:** Fine-tune the model parameters and perform feature selection to improve classification accuracy.

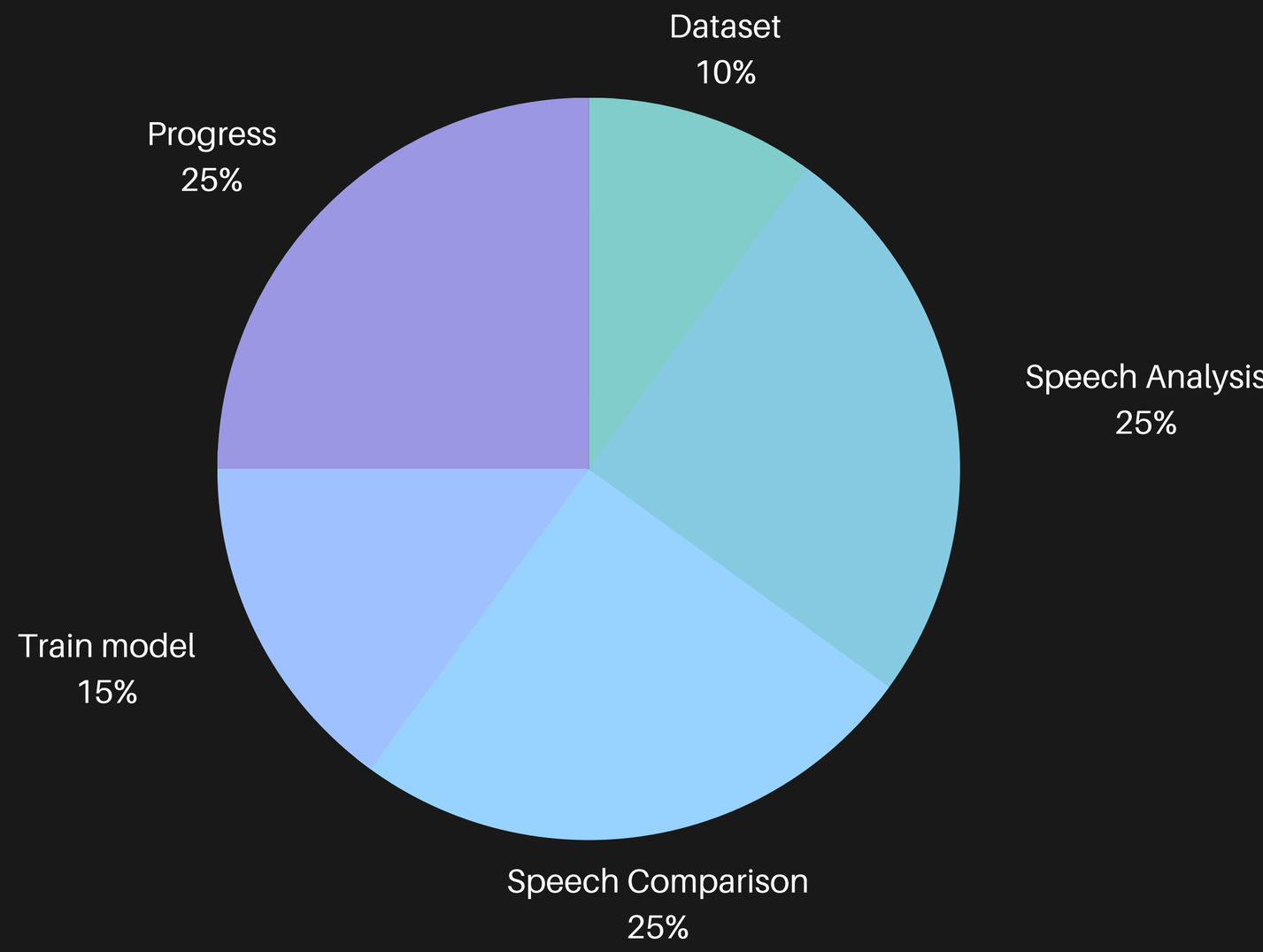
5

11. Create a User Interface: Develop a user-friendly interface for users to interact with the stuttering and stammering detector.

Real-time Analysis: 12. Real-time Analysis (Optional): If the application is meant for real-time detection, design it to analyze speech in real-time using microphones or audio input devices.

Continuous Improvement: 13. Continuous Improvement: Continuously update and improve the model based on user feedback and evolving speech analysis techniques.

TIME PLAN


Dataset
10%
Progress
25%
Speech Analysis
25%
Train model
15%
Speech Comparison
25%

RESOURCES

Dataset

1

The UCLASS archive of stuttered speech - PMC

Python libraries

2

Vosk
librosa
soundfile
speechbrain

Similarity metrices

3

Cosine similarity
Euclidean distance
Dynamic Time Warping (DTW)
Cross-correlation
Mel-frequency cepstral coefficients (MFCC) distance

Common errors

4

https://www.yourdictionary.com/articles/mispronounced-words

Tracker

5

Based on stored data

THANK YOU