UE21CS320A – Capstone Project Approval

Project Title : Generating Clear speech from Speech impaired Audio
Project ID : PW24_RS_03
Project Guide : Dr. Ramamoorthy Srinath
Project Team : Roseline Jerry A, T P Shriambhikesh, Tanmay Praveen Udupa, Vandana S

Outline

- Problem Statement
- Scope and Feasibility study
- Applications/Use cases
- Expected Deliverables
- Capstone (Phase-I & Phase-II) Project Timeline
- Any other information

Problem Statement

The project addresses the challenge of impaired speech communication, where various speech impairments hinder clarity and understanding. Individuals with speech difficulties, whether due to medical conditions or other factors, encounter barriers in effective communication. This problem statement emphasizes the need for a solution that classifies and enhances different types of speech impairments, ensuring clearer and more understandable communication.

Scope and Feasibility study

The project aims to develop an application designed to classify distinct speech impairments by transforming audio input into text while identifying attributes influencing clarity and ease of understanding. A Machine Learning model is used to enhance and refine the speech, producing a clear and easily understandable output. The system will benefit from user feedback, ensuring continuous refinement and higher accuracy of speech.

Key components of the scope:

- Speech impairment classification
- Attribute based text generation
- Machine learning model for speech enhancement
- User feedback integration
- Positive/ Negative response analysing

Scope and Feasibility study

Feasibility Study

- **Data collection:** The feasibility of the project depends on collection of diverse dataset which includes conversations and their correct translation.
- **NLP Techniques:** these techniques can be used to preprocess, analyse, extract features from the dataset which is to be used in the model.
- **Sentiment Analysis:** Determining positive or negative response, requires sentiment analysis. Which is a well-established field in NLP.
- **Model Training:** Training a model to accurately predict the destination user's response based on the source user's message is feasible using supervised learning techniques.

Challenges

- **Data Limitations:** Limited data can lead to underfitting or overfitting, which would affect the model's accuracy.
- **Ethical Considerations:** It should be ensured that privacy and consent in data collection is crucial. Ethical guidelines must be followed and the users must be informed about the purpose of their data.
- **Real world noise:** Preprocessing techniques such as noise reduction, can be used to improve the model's performance.
- **Subjective nature of clarity:** Clarity of speech often varies on user's perspective.

Applications/Use cases

- **Communication Aid:** People with speech impairments like stammering or old people who have difficulty in articulating speech can use this application to communicate with other people more effectively.
- **Education:** This application can be used to support students with speech impairments, so that they can participate actively in classroom discussions.
- **Employment Setting:** Provides support to help the speech impaired candidates effectively convey their thoughts to the interviewer.
- **Courtrooms:** People who are facing speech challenges can fully participate in legal proceedings.
- **Voice enabled technologies:** This application makes it easier for speech impaired people to use the voice enabled feature in applications.

Expected Deliverables

- Capstone-I deliverables

- Project Plan
- Project Literature Survey
- Project Requirements Specification
- High Level Design Document of the Project
- Timeline and workflow outline
- Speech Enhancement Algorithm
- Cleaned and Pre-processed Dataset
- Preliminary Proof-of-Concept

Expected Deliverables

- Capstone-II deliverables

- Low-level Design Document
- Generative Model for Speech Enhancement
- Multi-agent Speech Correction Model
- Real-time Speech Processing Application
- Application with a User-Friendly Interface

Any other information

- Speech disability accounted for 7.6% of the total disabilities in India.
- According to National Institute on Deafness and Communication (NIDCD), in 2016 approximately 7.5 million people have language impairments.
- There are more than 250,000,000 people with non standard speech needing accessible speech technologies.
- Models are usually trained on specific disorders, not generic.

Thank
You