CSCE 5290: Natural Language Processing
Project Proposal

TITLE: Speech-to-Text Conversion with Automated Response Grading.

Instructor: Dr. Zeenat Tariq (zeenat.tariq@unt.edu)

Group Members : Sai Gottumukkala, Sai Charan Merugu, Karishma Bollineni, Swetha Guntupalli

Motivation:

The main issue that this project seeks to address is the need for efficient evaluation of spoken responses, particularly in educational settings or customer service scenarios where verbal communication plays a crucial role. Traditional methods of assessing responses often rely on manual grading, which can be time-consuming and subjective. By automating the process through speech-to-text conversion and automated response grading, we aim to streamline evaluation procedures and improve efficiency. This project's justification lies in the increasing demand for scalable and reliable evaluation methods in various domains, including education, training, and customer support.

Significance:

This project addresses a critical need in multiple fields where verbal communication assessment is essential. By automating response grading, we can carry out the evaluation processes without an actual person. The project's outcomes have the potential to revolutionize educational assessment methods, improve training programs, and enhance customer service interactions.

Achieving the project's goals will facilitate the adoption of speech recognition technologies in diverse applications.

Objectives:

The objectives of this project include:

1. The first objective is to develop a robust speech-to-text conversion system capable of accurately transcribing spoken responses into text format.
2. Implement algorithms for automated response grading, considering factors such as content relevance, coherence, and language proficiency.
3. Evaluate the performance of the speech-to-text conversion and response grading systems using benchmark datasets and metrics such as accuracy, precision, and recall.
4. Integrate the developed systems for easy deployment and usage in educational or customer service environments.
5. Validate the effectiveness and reliability of the automated grading system by testing the professed system.

Features:

The project's technical characteristics include:

1. Development of speech recognition algorithms leveraging deep learning architectures such as suitable neural networks.
2. Implementation of natural language processing techniques for semantic analysis to enhance response grading accuracy.
3. Milestones include algorithm development, dataset acquisition and preprocessing, system integration, testing, and evaluation phases.

Deliverables:

Deliverables include a fully functional speech-to-text conversion, response grading system and integration as three phases along with documentation and tutorials for system usage and deployment.

Dataset:

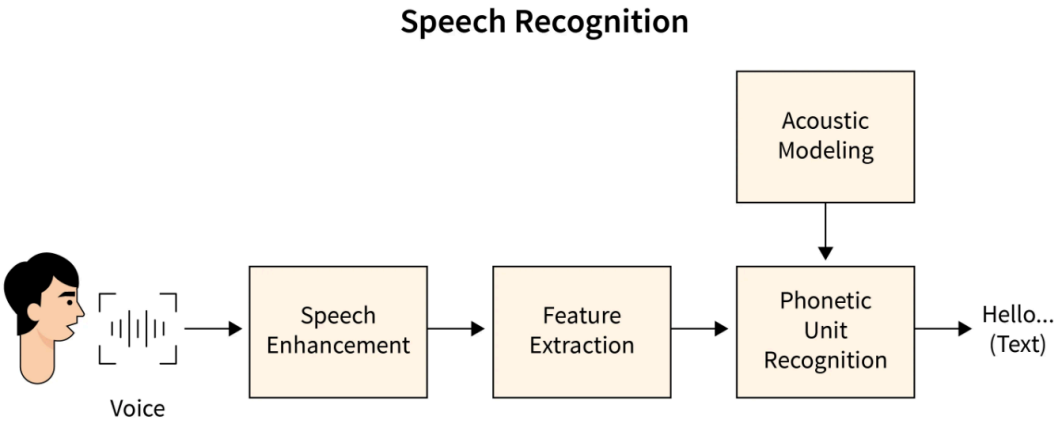
The dataset intended for this project comprises spoken responses collected from various datasets. The dataset will include a diverse range of speech samples covering different accents, languages, and communication styles to ensure robustness and generalization of the speech recognition and grading models. Preprocessing steps will involve audio normalization, transcription, and annotation of response labels for training and evaluation purposes.

For the text essay evaluation component of the project, we will utilize a dataset consisting of essays or written responses across various domains, such as education, standardized testing, or professional assessments. The dataset will contain a diverse range of already graded essays covering different topics, genres, and writing styles to ensure the robustness and generalization of the response grading model. Sources for the dataset may include educational repositories, standardized testing platforms, or crowdsourced collections of writing samples.

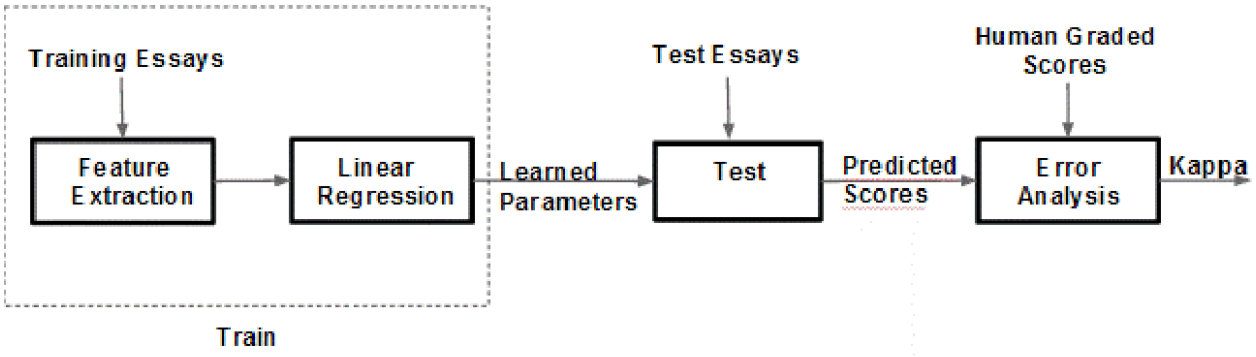
Visualization:

Below are the two modules of the system:

1. Speech Recognition to text:
   Explanation: First we train a model that inputs audio files and converts the audio data into text data using a dataset and use this module to achieve the first objective that is to convert speech to text conversion.



2. Text response grading:
   Explanation: On the other hand, we also train another model that is capable of grading an essay on the basis of different metrics.



Finally, we integrate these two models to develop a system that can grade the audio response from the user. Here, the output of the first system shall be the input to the second system.

Github Link:
https://github.com/saig06/NLP_Project/