PROJECT ABSTRACT

**Title:**

*AI-Driven Speech Feedback System with Enhanced Evaluation using NLP*

**Abstract:**

Effective communication, particularly spoken language, is a vital skill in both professional and personal contexts. However, providing consistent, real-time feedback on speech delivery and content remains a challenge. This project presents an *AI-Driven Speech Feedback System* that integrates Natural Language Processing (NLP) with automated speech recognition to offer enhanced evaluations of spoken communication. The system is designed to provide real-time, detailed feedback on various aspects of speech, including fluency, tone modulation, grammatical accuracy, and coherence.

The system's front-end flow involves forking and cloning the code from the owner’s repository, making updates, and pushing changes to the user’s repository. It includes a feature where the user can interact with the speech feedback system by submitting audio files for evaluation. Once the audio is uploaded, the back-end processes the input by performing noise reduction, normalization, and segmentation. After pre-processing, the audio is converted into text using an in-house speech recognition engine (e.g., CMU Sphinx or DeepSpeech). The system then analyses the text output using NLP techniques, providing feedback on clarity, grammar, tone, and structure.

The AI-driven back-end flow processes the audio file in several stages: from receiving the file, converting it to text, post-processing the transcription, and validating it for accuracy. The final output, including detailed analysis and scores, is stored securely in a database, allowing users to access and download the feedback for further review.

The findings demonstrate that this system significantly improves the efficiency and depth of speech evaluation, offering valuable, real-time insights that help users enhance their speaking skills. This system is applicable in public speaking training, language learning, and interview preparation, reducing the need for time-consuming manual feedback.

**Accepted Abstract:**

In the realm of spoken communication, whether in public speaking, language learning, or professional presentations, providing consistent and actionable feedback has always been a challenge. This research introduces an *AI-Driven Speech Feedback System* that utilizes Natural Language Processing (NLP) and advanced speech recognition algorithms to provide real-time, comprehensive evaluations of speech. The system allows users to upload their audio, which is processed through multiple stages of analysis to offer detailed feedback on pronunciation, grammar, fluency, tone, and pacing.

The project workflow starts with the front-end interface, where the user forks and clones the system’s repository, updates the necessary code, and pushes their updates to their local repository. Once the user uploads an audio file for feedback, the back-end processes it by applying pre-processing techniques like noise reduction and normalization. Following this, the system converts the audio to text using an in-house speech recognition engine and applies NLP-based algorithms to analyse the text’s linguistic accuracy, tone, and structure. Post-processing steps ensure that unnecessary elements such as filler words are removed and that the final text is checked for accuracy before being stored in the system’s database for further access.

The system’s ability to generate detailed, real-time feedback fills the gap between subjective human-based assessments and automated systems that only provide limited feedback. Its real-time capabilities make it ideal for public speaking training, language learning, and interview preparation. The research demonstrates the effectiveness of AI-driven evaluation methods in improving communication skills across a wide range of applications.

**Base Paper:**

**IEEE Access**

**Waiting Experience: Optimization of Feedback Mechanism of Voice User Interfaces Based on Time Perception**

**Problem Statements:**

1. **Lack of Real-Time, Comprehensive Feedback in Speech Evaluation:** Traditional methods of evaluating speech rely heavily on subjective feedback from humans or limited software capabilities, which can be inconsistent or lacking in depth. This creates a need for an automated, objective system capable of providing real-time, multi-faceted feedback on speech performance.
2. **Manual Workflows for Speech Analysis are Time-Consuming:** Manually analyzing speech quality for elements such as tone, grammar, and fluency requires significant time and expertise, which is impractical for large-scale or real-time feedback. An AI-based solution can automate this process, allowing for immediate and consistent evaluation.
3. **Challenges in Processing Noisy and Unstructured Audio:** Audio inputs are often noisy or unstructured, which impacts the accuracy of speech-to-text conversion and subsequent analysis. The system must incorporate robust pre-processing techniques to ensure high-quality transcription and analysis.
4. **Need for Effective Evaluation Tools for Non-Native Speakers:** Non-native speakers learning a language often struggle with pronunciation, grammar, and fluency. Traditional feedback systems are not equipped to provide the detailed level of analysis required to help non-native speakers make meaningful improvements.
5. **Inefficiency in Public Speaking and Communication Training:** In public speaking or communication training, instructors may not always have the time or resources to give personalized feedback. An AI-driven system that provides real-time feedback can address this issue by allowing users to self-evaluate and improve over time.
6. **Limited Scope of Current Speech Feedback Systems:** Most existing systems provide feedback on basic metrics like word count or grammatical correctness, but few offer a holistic evaluation that includes tone modulation, emotional analysis, and pacing—all of which are crucial for effective communication.

**Related Topics:**

1. **Natural Language Processing (NLP) in Speech Evaluation:** NLP is the backbone of this project, particularly for analysing linguistic content, grammar, and tone. Related research includes the use of NLP in grammar correction, sentiment analysis, and language modeling for speech processing.
2. **Automatic Speech Recognition (ASR) Technologies:** ASR technologies such as CMU Sphinx and Mozilla DeepSpeech are critical to converting speech to text. Improvements in ASR accuracy, noise handling, and speed are directly applicable to this project.
3. **Real-Time Speech Feedback Systems:** Real-time speech coaching platforms that utilize AI are becoming more prevalent. These systems offer immediate feedback on spoken communication and are often used in applications like public speaking, customer service training, and language learning.
4. **AI in Public Speaking and Presentation Training:** AI-based tools for public speaking training focus on evaluating speech delivery, tone, and engagement. These tools are related to the speech feedback system as they aim to improve both content and delivery in real-time.
5. **Audio Pre-Processing for Speech Recognition:** Noise reduction, normalization, and segmentation are critical steps in processing raw audio for speech recognition. Research on audio pre-processing techniques improves the quality of the transcription, which in turn enhances feedback accuracy.
6. **Speech Delivery and Emotional Tone Analysis:** Analyzing the emotional tone of speech using NLP and AI is an emerging area of research. This includes detecting stress, tone modulation, and sentiment to provide a more nuanced evaluation of spoken communication.
7. **Language Learning Applications Using AI:** AI-powered language learning tools like Duolingo leverage speech recognition and NLP to provide real-time feedback to learners. These systems are closely related to this project as they share similar goals of improving language fluency and accuracy.
8. **Data Storage and Management for Speech Feedback Systems:** Storing large amounts of processed speech data securely while maintaining quick access is a challenge for such systems. This involves database management, data compression, and security protocols for sensitive information.
9. **Speech-to-Text Conversion and Its Challenges:** Accurate speech-to-text conversion is crucial for evaluating spoken content. Challenges such as noisy environments, accents, and speech pacing make this an important area of research in developing robust speech feedback systems.
10. **User Experience in AI-Driven Feedback Systems:** Designing intuitive user interfaces for interacting with AI-driven systems is key to user adoption. Research in human-computer interaction (HCI) and user experience design focuses on making complex AI systems more accessible to non-technical users.