SPEECH SUMMERIZATION

- Abhinav Shandilya – 1PE15CS002
- Rajdeep Chatterjee – 1PE15CS118
- Syed Najid Salman – 1PE15CS172
- Yashasvi M – 1PE15CS188

Abstract

The motive behind developing this project of simplifying complex phrases and figures of speech using Machine Learning to ease the process of understanding for people with hearing impairments or learning disabilities. The project aims to recognize and classify various sound patterns and words in the speech, simplify the same to produce a concise and intelligible output to the target audience.

The input is taken as a voice and converted to text and a paragraph is formed, which is then summarized using machine learning and a output is given in the form of text/speech.

The Text to Speech model is trained with a high volume of training audio files (typically .wav) consisting of common sound patterns, complex terms/idioms. The algorithm employs *keyword spotting* to aid in capturing the overall intent of the input speech. The model is expected to classify various terms in the input and produce a simplified and easily understandable gist of it.

The Summarization model is trained rigorously with a high volume of text articles (eg : DUC dataset, CNN dataset). It employs weighted important keyword identification to make the summary better.

Further, the accuracy of the model can be made better by training it on noisy dataset or by employing noise elimination. Also, the model can be trained to summarize in multiple languages.