School of Computer Science Engineering and Technology

Course- B. Tech   Type- Elective
Course Code- CSET346   Course Name: Natural language
                                                        processing
Year-   2022   Semester- odd
Date: 07-11-2022   Batch- ALL

Lab Assignment 10 – Speech Summarizer

CO-Mapping

Exp. No.	Name	CO1	CO2	CO3
10	Speech Summarizer	✓	✓	✓

Objective: The motivation for this lab is generate an optimal and unambiguous summary from text speech using the concept of NLP.

Introduction:

Because everyone going to words shortcut-way except research scientist, so required a summary of every product. For, example if a patient is admitted to hospital and the health insurance company wants a summary about a patient for the claim and that process is time-consuming so we are going to do Text Summarization Using NLP which helps us to summarize.

In Natural Language Processing, or NLP, Text Summarization refers to the process of using Deep Learning and Machine Learning models to synthesize large bodies of texts into their most important parts. Text Summarization can be applied to static, pre-existing texts, like research papers or news stories, or to audio or video streams, like a podcast or YouTube video.

Before lets shortly discussed the types of summarizations. Broadly classified into two groups-

- Extractive Summarization
- Abstractive Summarization

Extractive Summarization
These methods rely on extracting several parts, such as phrases and sentences, from a piece of text and stack them together to create a summary. Therefore, identifying the right sentences for summarization is of utmost importance in an extractive method. This method based on the weight of the most important words in that given sentence. Various algorithms are used to define the specific weights for the word and further rank them based on important.

Abstractive Summarization
These methods use advanced NLP techniques to generate an entirely new summary. This method based on semantic means selects the specific words using the semantic method. Sometime words did not appear in documents. The main aim is to produce important material in a new.

School of Computer Science Engineering and Technology

Lets us solve the problem to generate the speech summary.

1. Use the Abstractive summarization method, and it required a cosine similarity to find out the distance between two similar words. Cosine similarity is a measure of similarity between two non-zero vectors. Since we will be representing our sentences as a bunch of vectors, we can use it to find the similarity among sentences. It measures the cosine of the angle between vectors. The angle will be 0 if sentences are similar.

2. New Take a audio speech of any news or sports commentary, compute the similar frames and generate the summary.