University of North Texas

Fall-2022 Semester

CSCE5290- Natural Language Processing

Instructor: Zeenat, Tariq

Project Title:

Text Summarization from the Real time Speech Recognition

Team Members:

Achala Samudrala –11519259 (Team Lead)

Pavani Mangugari –11542749

Tharun Puri – 11558533

Venkata Sai Preetham Bonthala – 11550407

GOALS AND OBJECTIVES:

1. **Motivation**:
   In this new era, where everything going on the Internet, it is most important to provide the improved mechanism to extract the information quickly and most efficiently. It is very difficult for human beings to memorize things and manually extract the summary of a large documents of text or video. There are plenty of text and video material available on the Internet. So, it will take long time in search of similar documents from the available documents and concluding similar text from it. It is crucial to use automatic text summarizing to address the two issues. But what about videos? Speech Recognizer to rescue. In speech recognition means converting the speech to text in the real time scenario i.e., exactly converting the words with means of pronunciation. Text summarization is defined as the process of finding the meanings of difficult words and gets summarised to the shorter version where we can find them easily. Also, we can enhance this project by adding Q&A module so that we can see more accurate text from summarized data, we can implement Q&A model along with SR and Text Summarization. Question answering module is a crucial NLP problem. QA module allows the end user to derive a question in NLP and gets quick and short explanation.

2. **Significance**:

Before going to the Text summarization, Speech Recognition first we, must know that what a summary is and what is Speech Recognition is. A summary is a text that is produced from one or more texts, that conveys important information in the original text, and it is of a shorter form. The goal of automatic text summarization is presenting the source text into a shorter version with semantics. The most important advantage of using a summary is, it reduces the reading time. Text Summarization methods can be classified into extractive and abstractive summarization. An extractive summarization method consists of selecting important sentences, paragraphs etc. from the original document and concatenating them into shorter form. An Abstractive summarization is an understanding of the main concepts in a document and then express those concepts in clear natural language. The length of informative summary is 20 to 30 percent of the main text. The speech is primary mode of communication among human being and the most natural and efficient form of exchanging information among human in speech. By classifying the speech with voiced, unvoiced and silence (VAS/S) an elementary acoustic segmentation of speech which is essential for speech can be considered. Speech processing is the study of speech signals, and the various methods which are used to process them. In this process various applications such as speech coding, speech synthesis, speech recognition and speaker recognition technologies; speech processing is employed. In order to extract and determine the linguistic information conveyed by a speech wave we have to employ computers or electronic circuits. This process is performed for several applications such as security device, household appliances, cellular phones ATM machines and computers.

3. **Objectives**:

We are going to implement this project in 3 steps.

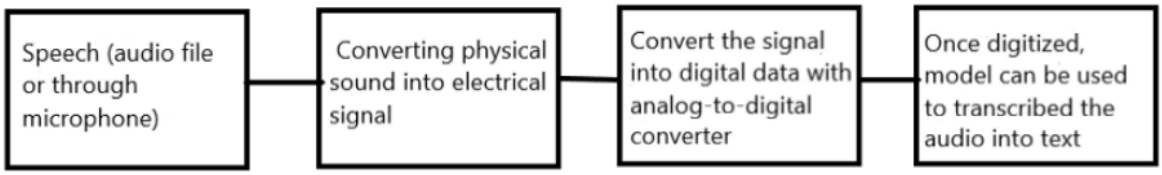
- Speech to Text
- Text Summarization
- Q&A module

- **Speech to Text:** he ability of a machine or program to recognize words and phrases in spoken language and translate them into machine-readable format is known as speech recognition. The following criteria can be used to categorize speech recognition systems.
- **Speaker**: Every speaker has an own voice. As a result, the models are either created for a particular speaker or an independent speaker.

- **Vocal Sound**: Speaking recognition also takes into account the speaker's speech pattern. Some models have the ability to distinguish between single

utterances and different utterances interspersed with silence.

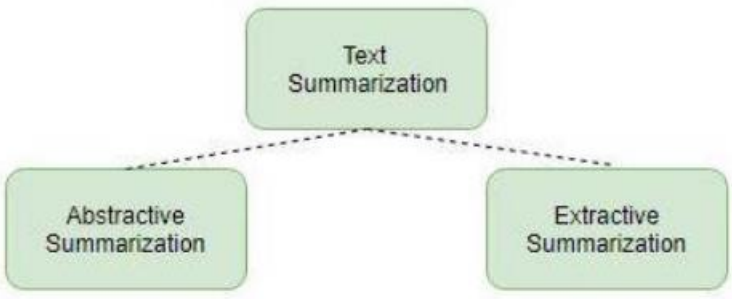
How does Speech recognition work?



We'll decode analog signals using the Python module base64, then turn speech into text using speech recognition.

4.Text Summarization: Tools like NLTK, SPACY, Genism, and Hugging Face can be used to accomplish this. According to studies from several blogs, HUGGINGFACE's abstractive technique can be outperformed by GENSIM and SPACY.

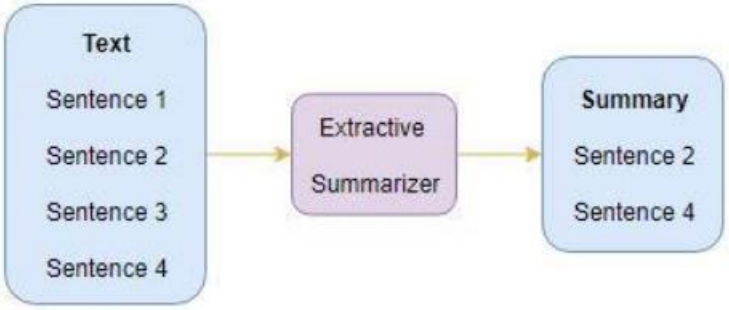
Types of Summarizations:



How to perform text summarization?

a) Convert the paragraph into sentences
b) Text processing
c) Tokenization

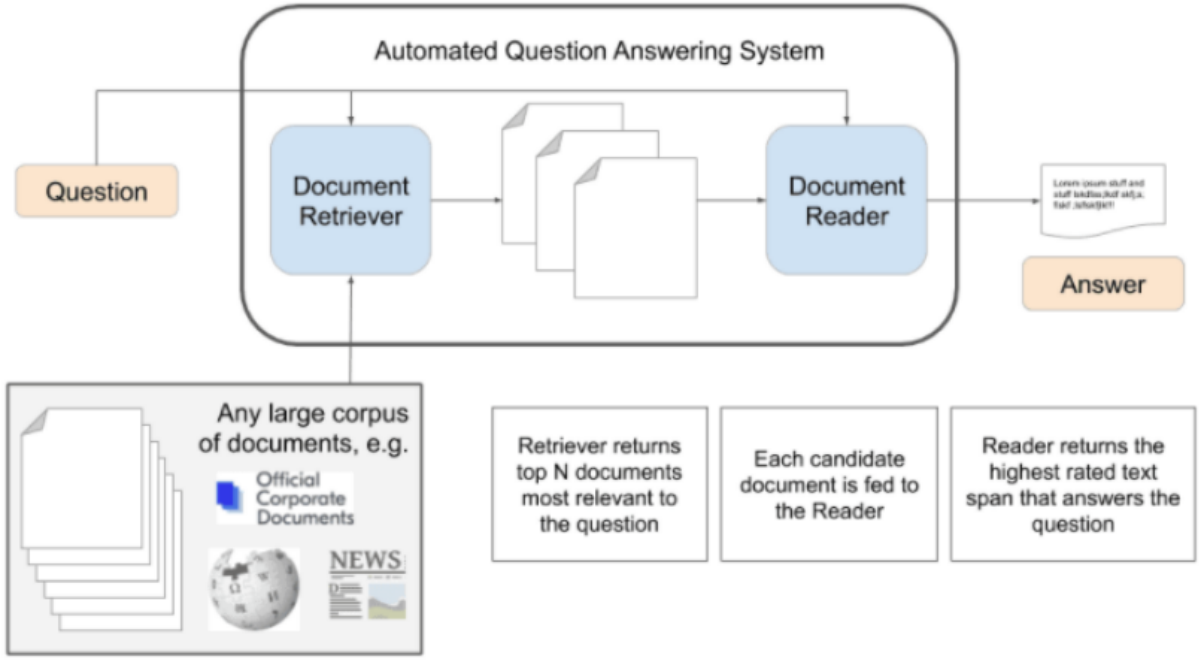
d) Evaluate the weighted occurrence frequency of the words
e) Substitute words with their weighted frequencies



Q & A Model Implementation:

QA systems are now present in phone conversational interfaces and search engines, and they can respond to little informational snippets. However, for more challenging queries, these typically only go as far as giving a list of snippets that the users must then read through to discover the solution to their problem.

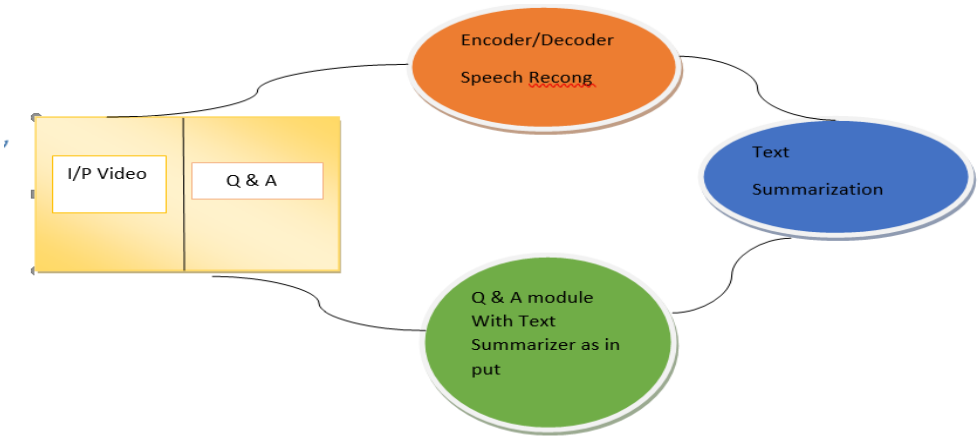
To train our Q&A model, we'll use the Stanford Question Answering Dataset (SQuAD). Although the steps are fairly lengthy, the following illustration can help us grasp the Q&A module.



5. Features:

1. Video Input will be given
2. Video will be decoded as it will be iv .wav file

3. After decoding, file will be transferred to Speech Recogniser
4. Now we will be having data, which will again input for Text Summarization
5. After summarization, Q&A will accept it as input and get trained on it.
6. At end, after asking Q, we will get relatable answer.



CONCLUSION:

A beneficial technology that will soon be prevalent is speech-to-text conversion. To reach our goal, we are combining several NLP techniques, including Text Summarization and Q&A. It is simple to construct applications with this tool using Python, one of the most well-liked programming languages worldwide. As we advance in this field, we are establishing the foundation for a time when spoken commands as well as fingertips can be used to access digital information.

S.no	Person	Task Description	Contribution
1.	Achala Samudrala	Worked on Motivation, and basic project idea	25%
2.	Pavani Mangugari	Given the basic idea of the project(Significance)	25%
3.	Tharun Puri	Provided required Objectives for the project	25%
4.	Venkata Sai Preetham Bonthala	Given the required Features of the project.	25%

References:

https://ieeexplore.ieee.org/document/9622156

https://paperswithcode.com/paper/predicting-subjective-features-from-questions

https://paperswithcode.com/paper/iit-uhh-at-semeval-2017-task-3-exploring

https://paperswithcode.com/paper/qdee-question-difficulty-and-expertise