Business Meeting Summarization Using Natural Language Processing(NLP)

Prof. Swapnil Waghmare
Department of computer engineering
Pillai Hoc college of engineering and technology,
Rasayani

Mr. Chaitanya Pathak
Department of computer engineering
Pillai HOC college of engineering and technology,
Rasayani

Mr. Raj Kshirsagar
Department of computer engineering
Pillai HOC college of engineering and technology,
Rasayani

Mr. Suyog Malkar
Department of computer engineering
Pillai HOC college of engineering and technology,
Rasayani

Abstract— Text summarization aims to condense a source text into a shorter version. Automatic data summarization is part of data mining. In order to build a corpus for this task, it is necessary to obtain the transcription of each meeting, and then to segment and align it with the corresponding manual report to produce training examples suitable for training. In this work, an MMS method combining the techniques of natural language processing (NLP), speech processing, computer vision and advanced encryption standard (AES) encryption is used to explore the rich information contained in multi-modal data, to improve the quality and security as well the key idea is to bridge and lessen the semantic gaps between multi-modal data. For audio, speech transcriptions are used. Finally, all the multi-modal aspects are considered to generate a textual summary by maximizing the salience, readability, non- redundancy. The summary so generated by text, audio is encrypted using AES encryption method (to make it secure) and sent to members of meeting on mail, from there the user can retrieve it whenever required by providing the decryption key.

Keywords—**Abstractive Summarization, AES algorithm, Natural Language Processing (NLP), Multimedia.**

I. INTRODUCTION

Today's world is centralized on computers and data. Data are our intangible thoughts and imagination. We are producers and consumers of data at the same time. Every little thing in our mundane lives are either a source or receiver of data. For, example when we drive there's data involved, the speed of the car, mileage, distance traveled, etc. Since 20th century, data has been a significant part of our lives, but these days we infer more from data. We store and access them through electronic and wireless systems.

Automatic meeting summarization is valuable, especially if it takes advantage of multi-modal sensing of the meeting environment, such as microphones to capture speech each participant's voice. It is the task of writing the report corresponding to a meeting. We focus on so-called exhaustive reports, which capture, all the information of a meeting, keeping chronological order and speakers' interventions. Such reports are typically written by professionals based on their notes and the recording of the meeting. We use abstractive text summarization method for summary generation, In Abstractive Summarization they do not select sentences from the originally given text passage to create the summary. Instead, they produce a paraphrasing of the main contents of the given text, using a vocabulary set different from the original document. We built such a corpus using Speech Recognition to generate transcription, then aligning segments manually and automatically.

When NLP processing has been used as a technique to summarize text documents, we see that python libraries such as nltk, spacy, and OpenNMT, Tkinter, Pyaudio, Speech Recognition has been used. Text summarization has grown into a crucial and appropriate engine for supporting and illustrate text content in the latest speedy emergent information age. It's far very complex for humans to physically summarize oversized documents of text. We use AES Encryption and decryption algorithm to provide security to generated summary. The Advanced Encryption Standard (AES) is a symmetric block cipher, is implemented in software and

hardware throughout the world to encrypt sensitive data. AES is very fast and secure, and it is the de facto standard for symmetric encryption.

II. LITERATURE SURVEY

Various Text summarizations were developed with different methods and implemented in different fields. Different Text summarization techniques are sentence compression, template and graph-based methods, sentence framing, etc.

Rahul, Surabhi Adhikari have used a sentence framing method to generate EXT text summaries. In this method, the documents are examined for their applicability and scored accordingly. Similarly, sentences were clustered collectively to find the most meaningful sentences, and Documents were selected based on sentence scores. The summaries extracted using these approaches may not always accurate. In some conditions, it's also inappropriate to the native documents.[1]

Soe Soe Lwin, Khin Thandar Nwet proposed Extractive Summarization for Myanmar Language system, which was the Myanmar text Summarizer developed using latent semantic analysis model. There are various projects developed with text summarization for English and European languages and also in other languages such as Chinese and Japanese. But, there are very few projects are implemented in Myanmar text summarization. This project compares the sentence selection methods of latent semantic analysis but mainly focused on a single document.[2]

Ravali Boorugu, Dr. G. Ramesh Presented "A Survey on NLP based Text Summarization for Summarizing Product Reviews". This paper recounted the most important techniques of summarizations. Basically, the project compares all the single document summarization, multi-document summarization, and domain-specific summarization. Different hybrid classifiers such as Naive Bayes and SVM are used to develop summaries.[3]

Madhuri, Ganesh Kumar. R proposed a system for text summarization which is implemented n by using a statistical novel method that is based on sentence ranking. The highly-rated sentences are taken out from the original documents to generate a high- quality test summary. The summaries of documents are extracted and converted into audio form. The summaries of documents are extracted and converted into audio form. This method gives more accuracy as compared to a traditional approach.[4]

Kavya Kishore, Greeshma N Gopal, and Neethu P H Implemented Document Summarization mainly focused on Malayalam sentences and grammar. This system is developed with a sentence framing approach and the whole framework depends on the linguistic treatment of the original document. The system provides benefited for both abstractive and extractive approaches and also summarizes multiple documents of the same topic.[5]

III. IMPLEMENTED SYSTEM

We will be discussing various roles of the system in overview. *Text summarization* means the technique of making short the long pieces of text. In the System, the meeting conversation will record by using python libraries such as speech recognition and pyaudio.
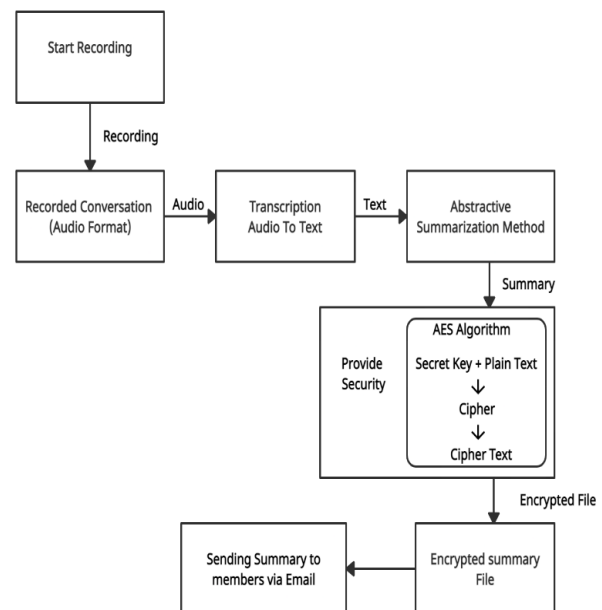

Fig 1: Block Diagram

Speech Recognition is a library for performing speech recognition, with support for several engines and APIs, online and offline. If you want to use microphone as input the speech recognition is required.

After the recording the audio file is generated. Then we will transcript the audio file (Audio to text conversion) using pyaudio library.

The summary will generate from transcription by using abstractive summarization method and for that we use Spacy library. To provide security to the generated summary we used AES algorithm. For encryption and decryption of the summarized file the Pycrypto library is required. We made one model for decrypt the encrypted file. This model and the encrypted file will send to the meeting members via Email. We have used smtplib library for sending mail.

We created GUI of this system using Tkinter library, Tkinter has several strengths.   Tkinter is most

commonly used. The layered approach used in designing Tkinter gives Tkinter all of the advantages of the TK library. Therefore, at the time of creation, Tkinter inherited from the benefits of a GUI toolkit that had been given time to mature.

Abstractive text summarization is the technique of generating a summary of a text from its main ideas or main words. The *AES algorithm* is a symmetrical block *cipher algorithm* that takes plain text in blocks of 128 bits and converts them to cipher text using keys of 128, 192, and 256 bits.

After generating the summary, system will take Email id of members of Meeting and send over an email automatically and also create the text file (.txt) of summary. This system will reduce the human efforts.

IV. RESULT

This window is the home window of the model, in that two buttons are there i.e., Record and Browse. When we click Record button, it will go to Recording window and the browse button is for selecting the recorded audio file.
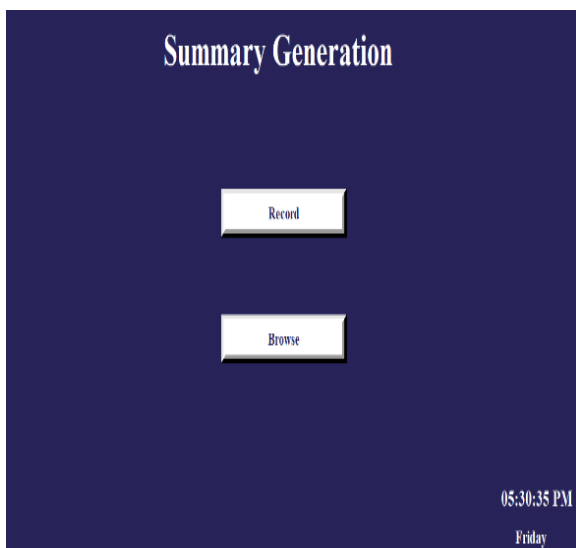

Fig. 2.1 : Home Window

After clicking Record Button of home window this Recording window will open, this window has two buttons are Start and Stop. The start button is use to start recording and stop button is use to stop the recording.
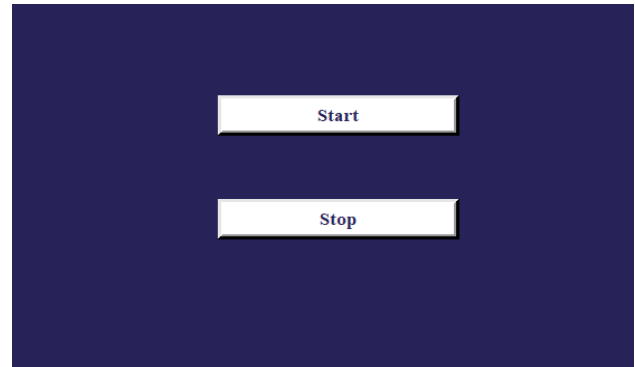

Fig. 2.2: Recording Window

The browse window appears when we click browse button. The Browse window is for select a recorded audio file. after selecting recorded audio file, the summary generates and sends to meeting members via mail.
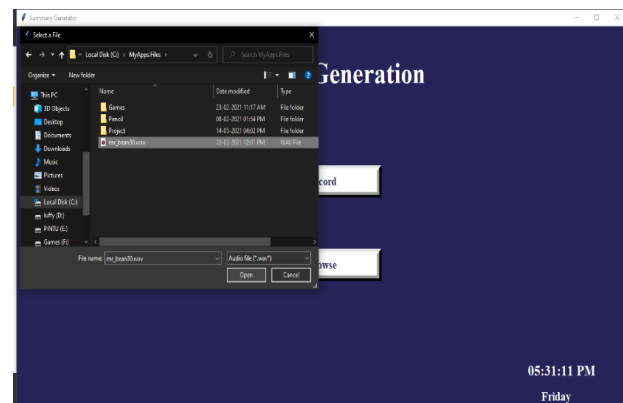

Fig. 2.3: Browse file

The message box opens after sending mail to meeting members successfully and this window shows "The Summary has been sent!" message.
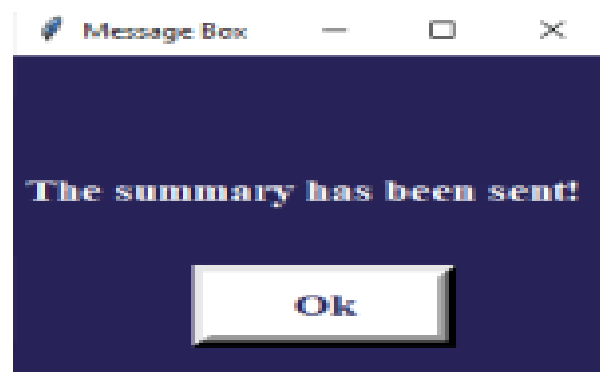

Fig. 2.4: Message Box

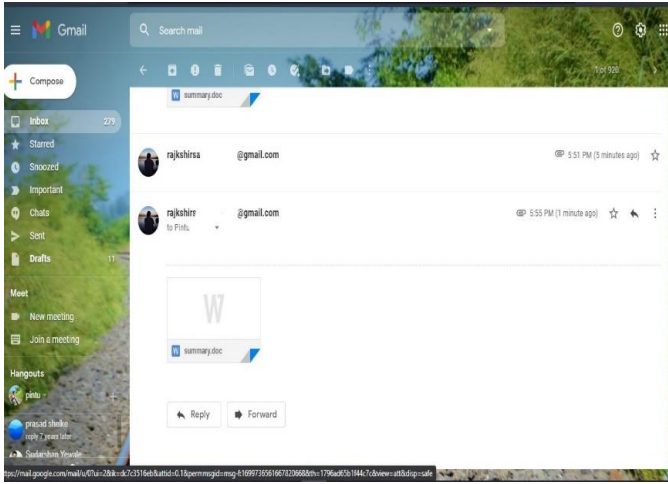In the below screenshot (fig2.5), the mail has been successfully sent to the intended participants.

Fig. 2.5: Mail

V. COMPARISON TABLE

Parameter	Existing System	Implemented System
Text Summarization Method	Extractive summarization Method	Abstractive Summarization Method
Security	No	Yes (AES Algorithm)
Mail	Manually mail send	Automatically mail send

Table 1: Comparison Table

VI. REFERENCES

1. Rahul, Surabhi Adhikari, Monika" NLP based Machine Learning Approaches for Text Summarization", International Conference on Computing Methodologies and Communication 2020.

2. Soe Soe Lwin, Khin Thandar Nwet" Extractive Summarization for Myanmar Language", IEEE 2018

3. Ravali Boorugu, Dr. G. Ramesh" A Survey on NLP based Text Summarization for Summarizing Product Reviews", International Conference on Inventive Research in Computing Applications, IEEE 2020

4. Madhuri, Ganesh Kumar. R" Extractive Text Summarization Using Sentence Ranking", IEEE 2019

5. Kavya Kishore, Greeshma N Gopal, and Neethu P H" Document Summarization in Malayalam With Sentence Framing", IEEE 2016