Voice Meeting Using Text Summarization

Vallepu Nithish, Dodla Nithin Reddy,
Peddireddy Harshith Reddy and Thummoju Shashank

Voice Meeting Using Text Summarization

Vallepu Nithish
Department of Computer science
and Engineering
Artificial Intelligence
Parul University
Vadodara, India
vallepunithish@gmail.com

Dodla Nithin Reddy
Department of Computer science
and Engineering
Artificial Intelligence
Parul University
Vadodara, India
nithinreddy0409@gmail.com

Peddireddy Harshith Reddy
Department of Computer science
and Engineering
Artificial Intelligence
Parul University
Vadodara, India
harshithpeddireddy06@gmail.com

Thummoju Shashank
Department of Computer science
and Engineering
Artificial Intelligence
Parul University
Vadodara, India
Shashankshiva76@gmail.com

.

Abstract —**The proposed project is aimed to automate the complete process which records the discussions of the meeting and summarizes them to draft minutes of the meeting and disseminate the information to the concerned. The process involves automatic recognition of the participants voices and recording the opinions of all the participants converting them into text and uploading it into website and summarize collected data Users can access the minutes by logging in with their respective credentials. In this voice meeting, we will leverage text summarization technology to enhance our communication and collaboration. Text summarization allows us to condense complex information into concise summaries, facilitating a more efficient and focused discussion. Participants are encouraged to actively engage in the conversation, while the summarization tool provides real-time summaries to ensure everyone stays informed. This innovative approach aims to streamline communication, making our meetings more productive and accessible. Feel free to inquire about the summarization process for a clearer understanding. Let's embrace this tool to optimize our discussions and decision-making. The proposed project is aimed to automate the complete process which records the discussions of the meeting and summarizes them to draft minutes of the meeting and disseminate the information to the concerned.**

Keywords—Voice Meeting, Collaboration Tools Text Summarization Natural Language Processing(NLP),Speech-to-Text(STT), Automatic Speech, Recognition(ASR), Summarization Algorithms, Keyword Extraction, Meeting Transcription.

I. INTRODUCTION

Text summarization has reached a relatively mature stage; there are well established methods for summarization of a single document and many researchers are working on techniques for summarizing a set of related documents.

We present approaches used in text summarization, showing how they can be adapted for speech summarization and where they fall short. Given errors resulting from speech recognition and the fact that spoken language is often less formal than written language, the most widely used method for single document text summarization, sentence extraction, cannot be directly applied to speech summarization

However, if systems exploit the additional information that can be derived from the speech signal and from dialog structure, extractive methods can be extended for spoken language and augmented by new methods that focus on extracting particular kinds of information and reformulating it appropriately. We present ongoing work at Columbia on summarization for two different types of spoken sources, broadcast news and meetings.

utilizing text summarization techniques to efficiently capture and communicate key information. Text summarization is a powerful tool that allows us to distill large volumes of text into concise summaries, enabling us to save time and focus on the most important Throughout the meeting, we encourage active participation and engagement from all participants. The text summarization tool will serve as a valuable aid, providing us with real-time summaries to facilitate more effective communication and decision-making.

As we engage in discussions and share information, the text summarization feature will help us generate quick and accurate summaries, making it easier for everyone to stay informed and actively participate in the conversation. This technology enhances our ability to streamline communication and ensure that we extract the essential details from the discussions.
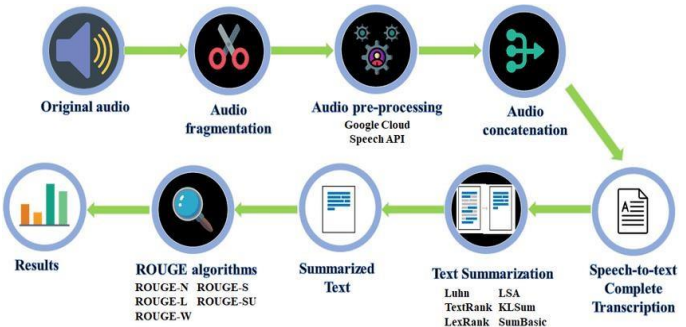

Fig1-Speech to text through extractive text techniques

COMPARATIVE STUDY

Papers	Merits	Demerits
M. S. Badhan. [1]	Trends and advancements in Automatic Speech Recognition.	Fail to provide a holistic overview of the subject.
A. J. Joshi, S. Kadam.[2]	It could offer a comprehensive solution for extracting valuable information from meeting.	Speech recognition component suffers from significant errors.
K. F. Wong, K. Hui, and D. Lee.[3]	1. System successfully integrates information from audio, video, and text modalities.	1. Handling multi-modal input can introduce additional complexity to the system 2. Video data raises privacy concerns and
Y. H. Lee.[4]	Speaker role identification helps in understanding the dynamics and interactions between different participants	Identifying speaker roles might be challenging in situations where the roles are not clearly defined
M. Pan and S. Renals.[5]	Language models can provide a more holistic representation of the meeting content,	The Processing video and audio features can be computationally intensive
Y. Liu, L. Wang, and T. S. Chua.[6]	Allowing the summarization system to extract relevant content from the most influential participants in the meeting.	Speaker identification, poor audio quality could impact the accuracy of the identification process
S. Li, L. Wan, and J. Wei.[7]	The vanishing Utilizing Graph Convolutional Networks allows the system to model the relationships and interactions between different elements in the meeting	The availability and quality of training data. Limited or biased datasets may affect the model's generalization.
M. Lui, K. Sagae, and M. Johnson.[8]	Approach that captures variations and differencs in meeting content,	Potentially leading to increased development complexity.
M. Diab and M. Kearns.[9]	Combines multiple decision trees, providing a robust and accurate summarization model.	Can be a limitation in certain applications where interpretability is crucial.
Y. H. Lee.[4]		

Z. Luo and X. Liu.[10]	Can allow the system to focus on different aspects in each stage,	Can allow the system to focus on different aspects in each stage,
S. Kim, J. Choi, and C. Kim.[11]	Summarization can provide a more comprehensive understanding of the meeting	Potentially requiring sophisticated algorithms for effective fusion.
A. Dowd et al.[12]	In comparison to cloud computing, edge computing is quicker, more effective, and more secure.	1.Misclassificati on of emotions Realtime detection is not done
P.Chouhan, R. Gupta, [13]	This can enhance the summarization process by addressing various aspects of the content.	This may pose challenges in terms of development and maintenance.
J. Deng, R. T. Li, and X. Yang,[14]	The hierarchical and sequential structure of the content.	This complexity may impact system development, training, and maintenance.
H. Wu, C. Hsu, and C. Hsu,[15]	Such as intonation and pitch variation, which may provide additional context for summarization	Requiring sophisticated algorithms and potentially impacting ease of implementation.

II. PROBLEM STATEMENT

Regular meetings are a must for the growth of a company or an organization. Usually meetings will be held with a fixed agenda. The number of persons attending the meeting is also decided before the meeting. Recording the minutes of meeting and following up the decisions taken in a meeting is a tedious and meticulous job. Implementation of the minutes of meeting is very important for the growth of an organization. Automation of the whole process will not only save the manpower but it will also speeds up dissemination process without any errors.

Furthermore, Meetings generate a substantial amount of spoken content, making it challenging for participants to recall and retain important details.The data input consists of audio recordings, transcriptions, or real-time speech-to-text conversions of the spoken content in meetings.

The goal is to develop a text summarization system capable of extracting essential information from the meeting discussions.Summarization should consider factors such as relevance, importance, and context to produce concise and coherent summaries.

successful voice meeting summarization system should not only extract key information but also preserve the context, tone, and nuances of the conversation to ensure that the summary accurately reflects the meeting dynamic.
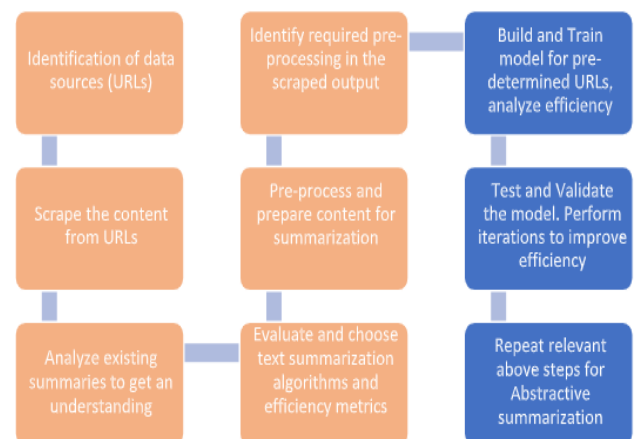

Fig2-Text summarization using deep learning techniques

In some cases, there might be additional input modalities, such as video or screen-sharing, which need to be considered for a comprehensive summary. The solution requires advanced natural language processing techniques to understand the nuances of spoken language, identify key topics, and extract meaningful information.

The system should be able to identify different speakers during the meeting to attribute statements accurately. whether the summarization is performed in real-time during the meeting or as a post-meeting process, influencing the system's responsiveness and resource requirements.

III. METHODOLOGY

To address the voice meeting text summarization Problem, a proposed method would involve a combination of speech processing, natural language processing (NLP), and summarization techniques. Here's a general outline of a potential approach

Speech-to-Text Conversion:
Utilize automatic speech recognition (ASR) systems to convert spoken content in the meeting into text. This step is crucial for obtaining a textual representation of the audio.
Speaker Identification:

Preprocessing:
Perform text preprocessing to clean and enhance the quality of the transcribed text. This may involve removing irrelevant filler words, correcting errors from ASR, and ensuring uniformity in the representation.

Topic Extraction:
Apply topic modeling techniques to identify key themes and topics discussed during the meeting. This step helps in understanding the context and structure of the conversation.

Named Entity Recognition (NER):
Use NER to identify entities such as names, locations, dates, and other relevant information. This helps in recognizing important details like action items, deadlines, and specific topics.

Sentiment Analysis (Optional):
Optionally, perform sentiment analysis to gauge the emotional tone of the meeting, which can provide additional context to the summarized content.

Abstractive Summarization:
Apply abstractive summarization techniques to generate concise and coherent summaries. This involves understanding the context of the conversation and generating human-like summaries that capture the essence of the discussion.

Extractive Summarization (Optional):
Consider incorporating extractive summarization methods, which involve selecting and presenting important sentences or phrases directly from the meeting transcript.

Contextual Understanding:
Develop mechanisms to understand the contextual dependencies between statements, ensuring that the summary maintains coherence and accurately represents the flow of the meeting.

Real-time or Post-meeting Implementation:
Depending on the requirements, choose between real-time summarization during the meeting or post-meeting summarization. Real-time summarization may involve incremental updates to the summary as the meeting progresses.

User Customization and Feedback:
Allow users to customize the summarization process based on their preferences and requirements. Implement feedback mechanisms to improve the system over time.

Privacy and Security Measures:
Implement robust privacy and security measures, especially if the summarization process involves handling sensitive information. Consider on-premises solutions or secure communication protocols.

Evaluation Metrics:
Define and use appropriate evaluation metrics to assess the quality of the summaries, such as ROUGE scores, BLEU scores, or domain-specific metrics tailored to the organization's needs.

User Interface Integration:
Design an intuitive and user-friendly interface for presenting the summarized content, allowing users to easily navigate and understand the key points discussed in the meeting.

Scalability and Adaptability:
Ensure that the system is scalable and adaptable to different meeting formats, languages, and industries.

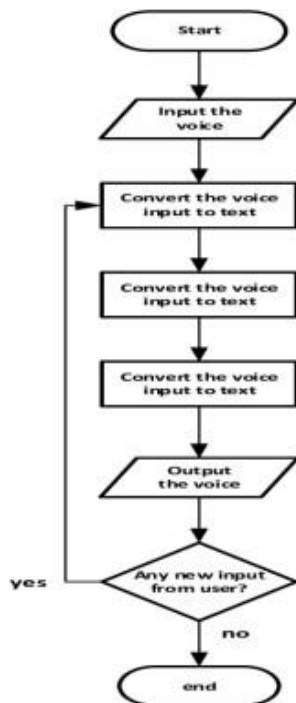analysis to gain insights into consumer preferences and behaviours.

Fig3- Flowchart describing the voice chatbot process

IV. CONCLUSION

In conclusion, The integration of speech-to-text technology enables the accurate conversion of spoken words into text, facilitating efficient information retrieval and searchability. Speaker identification techniques, such as diarization, contribute to a richer understanding of the meeting context by attributing spoken content to specific participants.

The summarization process aids in distilling the essence of discussions, promoting better collaboration among participants by focusing on key points and action items.

By automating the summarization process, the system saves time and resources that would otherwise be spent manually reviewing lengthy meeting transcripts. User customization features, such as adjusting summary length or prioritizing certain topics, enhance the user experience and cater to individual preferences.

The availability of summarized meeting content increases accessibility, allowing participants to quickly catch up on discussions, especially in scenarios where they may have missed parts of the meeting.

The system's adaptability to diverse meeting scenarios, languages, and accents ensures its utility across a range of organization

As technology continues to advance, the application of text summarization in voice meetings is likely to play a significant role in transforming how information is processed and shared during collaborative. streamlined decision-making processes, and improved accessibility to meeting discussions. The effectiveness of such a system relies on the seamless integration of natural language processing techniques, ensuring that the summarization process accurately captures the essence of spoken conversations

REFERENCE

1. M. S. Badhan, V. Kumar, and N. Kumar, "Automatic Speech Recognition and Text Summarization: A Review," in 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), 2019.

2. A. J. Joshi, S. Kadam, and R. R. Manthalkar, "Speech Recognition and Text Summarization System for Smart Meeting Environment," in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN),2019.

3. K. F. Wong, K. Hui, and D. Lee, "Intelligent Meeting Summarization with Multi-Modal Audio, Video and Text," in 2019 International Conference on Asian Language Processing (IALP), 2019

4. M. Pan and S. Renals, "Meeting Summarization Using Video and Audio Features and Language Models, " in 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011.

5. Y. Liu, L. Wang, and T. S. Chua, "Meeting Summarization via Dominant Speaker Identification and Key Sentence Extraction," in 2014 ACM International Conference on Multimedia Retrieval, 2014.

6. S.Li, L. Wan, and J. Wei, "Meeting Summarization via Graph Convolutional Networks," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020.

7. M. Lui, K. Sagae, and M. Johnson, "Automatic Meeting Summarization by Modeling Divergence," in Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2012.

8. M. Diab and M. Kearns, "Automatic Meeting Summarization Using Random Forests," in Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 2009.

9. Y. H. Lee, "Summarizing Multi-Party Spoken Conversations using Speaker Role Identification and Multi-Document Summarization," in Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2017.

10. M. Lui, K. Sagae, and M. Johnson, "Automatic Meeting Summarization by Modeling Divergence," in Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2012

11. Z. Luo and X. Liu, "A Two-Stage Approach to Automatic Meeting Summarization," in Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2019B. Li et al., "Meeting Summarization by Extracting Reliable Utterances," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012.

12. S. Kim, J. Choi, and C. Kim, "Speaker and Summarization System Using Text and Speech," in 2019 IEEE International Conference on Consumer Electronics (ICCE), 2019.

13. P. Chouhan, R. Gupta, and T. Sinha, "Automatic Text Summarization of Speech Based Lecture using Improved Hybrid Approach," in 2021 7th International Conference on Advanced Computing Communication Systems (ICACCS), 2021.

14. J. Deng, R. T. Li, and X. Yang, "Multi-Document Meeting Summarization based on Structural and Functional Analysis," in 2019 International Conference on Asian Language Processing (IALP), 2019.

15. H. Wu, C. Hsu, and C. Hsu, "Improving Meeting Summarization Performance by Incorporating Prosodic and Lexical Features," in 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2019