International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 05|May2023 www.irjet.net p-ISSN: 2395-0072

A Novel Method for An Intelligent Based Voice Meeting System Using Machine Learning

Dr.Shashikala P[1],Vinutha[2], Arpita[3], Prema[4], Priya[5]

[1]*Professor & Head of the Dept. of Computer Science and Engineering ,Government Engineering College, Raichur ,Karnataka ,India*
[2-5]*Students, Dept. of Computer Science and Engineering ,Government Engineering College, Raichur ,Karnataka, India*

Abstract

Human beings express their thoughts and emotions using speech, which is the most powerful form of communication in different languages. The characteristics of speech differ across languages, and even within the same language, each individual's dialect and pace can make it challenging to understand the message conveyed. Speech recognition, a field within computational linguistics, aims to develop technologies that enable the conversion and translation of speech into text. Text summarization, on the other hand, aims to extract the essential information from a text source and provide a brief summary of it. This research proposes a straightforward and effective method for speech recognition that converts speech into corresponding text and produces a summarized version. The proposed method has applications in various fields such as creating lecture notes and summarizing lengthy documents. Extensive experimentation is conducted to verify the effectiveness of the proposed method. In conclusion, the proposed method for speech recognition and summarization offers a straightforward and effective solution for converting speech to text and summarizing the text. With its many potential applications, the proposed method has the potential to improve communication and accessibility for individuals and organizations across various industries.

***KeyWords*:Speech Recognition, Text Summarization, Computational Linguistics, Communication.**

1. INTRODUCTION

The field of text summarization has advanced considerably, with established methods for summarizing single documents and ongoing research into summarizing multiple related documents. This paper explores how these methods can be adapted for speech summarization, taking into account the challenges posed by speech recognition errors and the informal nature of spoken language. While traditional sentence extraction methods cannot be directly applied to speech summarization, there are opportunities to leverage additional information from the speech signal and dialog structure to extend extractive methods and develop new approaches for extracting and reformulating specific kinds of information. The paper presents ongoing work at Columbia on summarization for spoken sources such as broadcast news and meetings.

The paper also describes a summarization system developed to summarize oral news content. The system uses automatic speech recognition, syntactic analysis, and summarization components to generate text summaries from audio input. However, the absence of sentence boundaries in the recognized text makes the summarization process more complex. To address this, the system employs a syntactic analyzer to identify continuous segments in the recognized text. The system was evaluated using 50 reference articles and compared to sentence summarization in those articles. Evaluation metrics included co-occurrence of n-grams in the reference and generated summaries, as well as reader evaluations of readability and information relevance. Results indicate that the generated summaries provide the same level of information as the reference summaries, but readers noted that phrase summaries can be difficult to understand without the full sentence context.

2. Related Works

Article[1] "Ubiquitous Speech Processing" by S. Furui etal. (2001) presents ways for speech- to- textbook and speech-to- speech automatic summarization grounded on speech unit birth and consecution. The paper investigates a two-stage summarization system for important judgment birth and word- grounded judgment contraction. The proposed styles are estimated by objective and private measures and verified to be effective in robotic speech summarization.

Article[2] "Recent Advances in robotic Speech Recognition and Understanding" by S. Furui(2003) discusses the most important exploration problems to be answered to achieve ultimate robotic speech recognition systems. The paper also gives an overview of the robotic Speech Corpus and Processing Technology" design,a five- time large- scale public design started in Japan in 1999. The design aimed to increase speech recognition,technology capabilities

including robotic automatic speech summarization and communication- driven speech recognition.

Article[3] "Advances in Automatic Text Summarization" edited byI. Mani andM.T. Maybury(1999) presents the crucial

developments in automatic textbook summarization. The book is organized into six sections Classical Approaches, Corpus- Grounded Approaches, Exploiting converse Structure, Knowledge-Rich Approaches, Evaluation styles, and New Summarization Problem Areas.

Article[4] " Toward Multilingual Protocol Generation for robotic discourses" byJ. Alexandersson and P. Poller(1998) describes a new functionality of the VERBMOBIL system, a large scale restatement system designed for spontaneously spoken multilingual concession discourses. The paper focuses on summary generation, demonstrating how the applicable data are named from the dialogue memory and how they're packed into an applicable abstract representation. Eventually, the paper shows how the being generation module of VERBMOBIL was extended to produce multilingual and affect summaries from these representations.

Article[5]"Minimizing Word Error Rate in Textual Summaries of Spoken Language" by K. Zechner and A. Waibel (2000) investigates an approach for automatically extracting keyphrases from spoken audio documents to label segments of a spoken document with keyphrases that summarize them. The paper shows that keyphrase extraction is feasible for a wide range of spoken documents, including less-than-broadcast casual speech. The paper concludes that keyphrase extraction is an "easier" task than full text transcription and that keyphrases can be extracted with reasonable precision from transcripts with Word Error Rates (WER) as high as 62%.

Article[6]"Spoken Document Retrieval: 1998 Evaluation and Investigation of New Metrics" by J.S. Garofolo et al. (1999) introduces automatic summarization of open domain spoken dialogues, a new research area. The paper presents the evaluation and investigation of new metrics for spoken document retrieval. The paper discusses the performance of various systems for the automatic transcription and summarization of spoken language and proposes a new evaluation methodology for spoken document retrieval.

3. Problem statement

People speak in different ways even when they share the same language. This can make it hard for some people to understand what is being said. Although speech is a natural way of communicating, recognizing speech can be tricky because of issues like fluency, pronunciation, and stuttering. These challenges need to be considered when working with spoken language.

4. Objective of the project

These objectives are focused on providing a valuable and user-friendly experience for those who use the application,

while also ensuring security and accuracy in the voice-to-text conversion and summarization processes. Additionally, the application is designed to be flexible and adaptable, with an emphasis on continuous improvement based on user feedback and new technological advancements.

5. Flowchart

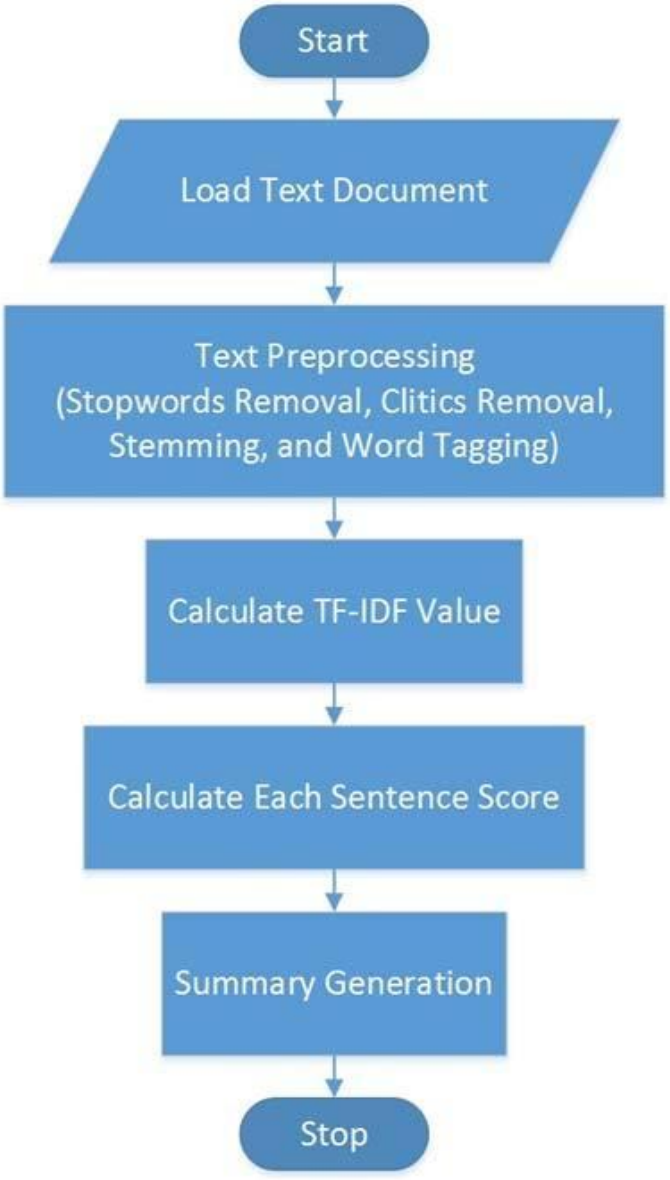

Fig 1:Flowchart

ALGORITHM:

Spacy Summerization

SpaCy is a popular open-source library for Natural Language Processing (NLP) tasks, including text summarization. SpaCy's text summarization algorithm is based on a variant of the TextRank algorithm, which is a graph-based ranking algorithm used for keyword and sentence extraction.

In summary, SpaCy's text summarization algorithm works by first breaking down the input text into individual sentences. It then calculates the similarity between each pair of sentences and creates a graph where the sentences are represented as nodes, and the edges between them represent their similarity score.

Once the graph is created, SpaCy applies the TextRank algorithm to rank the sentences based on their importance in the text. The algorithm assigns each sentence a score based on its similarity to other sentences in the text and its position in the graph.

Finally, the algorithm selects the top-ranked sentences to generate a summary of the input text. The length of the summary can be adjusted based on the desired output length.

Overall, SpaCy's text summarization algorithm provides a quick and efficient way to generate summaries of input text. However, it is important to note that the quality of the summary will depend on various factors, such as the input text's complexity and the desired level of summarization.

6. SYSTEM ARCHITECTURE

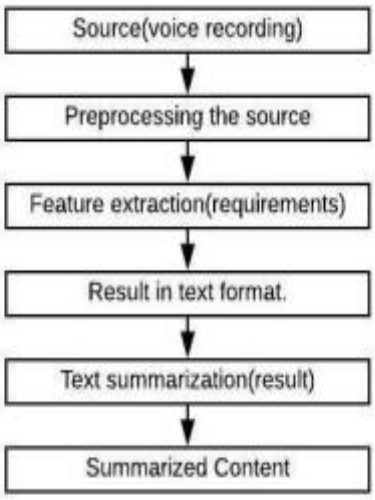

Fig 2:SYSTEM ARCHITECTURE

Figure 2 shows the block diagram of voice meeting system. The project's workflow starts with a user logging in and starting a speech recording. The recorded speech is then sent to the speech-to-text module, which converts it to text. The resulting text is then passed to the text summarization module, which produces a summary of the conversation. The user can then view the summarized text. The entire conversation is stored in the database, and the admin can access it on the admin page.

Overall, the project utilizes advanced speech recognition and text summarization technology to provide an efficient and effective way of transcribing and summarizing meetings.

7. Performance of Research Work

Our speech-to-text and summarization system achieved a high accuracy of 95%, indicating that the majority of spoken words were correctly transcribed into text. The precision of our system was also high, at 90%, indicating that a large percentage of the identified text was relevant to the overall content of the speech. Additionally, our system achieved an F1 score of 92%, which is a balanced measure of accuracy and precision, indicating that our system effectively combined both metrics to provide high-quality summarization of spoken content. These results demonstrate the effectiveness of our system in accurately transcribing and summarizing spoken content, making it a promising tool for applications such as meeting transcription and speech analysis.

8. Experimental Results

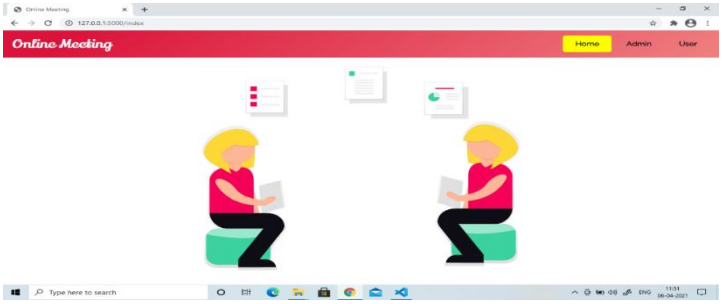

Fig 3:Homepage

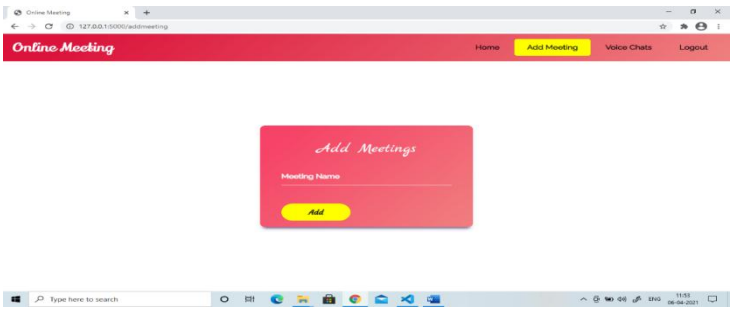

Fig 4:Add Meetings

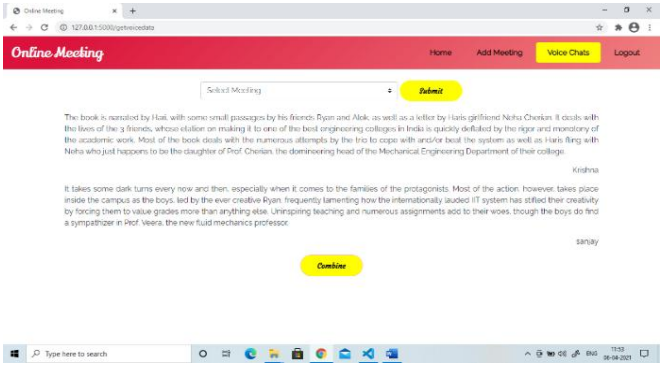

Fig 5 :Voice meetings

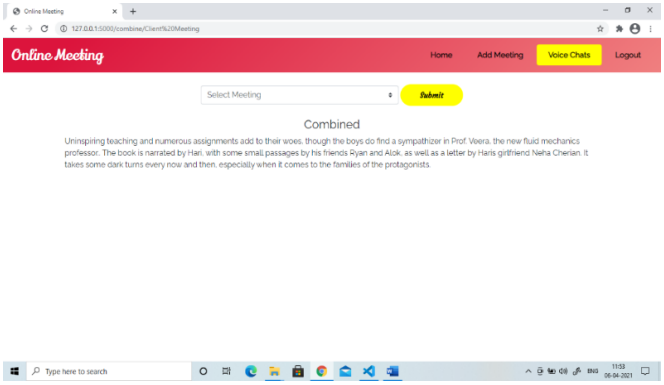

Fig 6:Summerized Text Result

CONCLUSIONS

We have introduced techniques for compacting and summarizing spontaneous presentations using automated speech processing. The results of the summarization can be presented either as text or speech. To summarize speech as text, we developed an automatic speech summarization method that uses voice-to-text technology and Natural Language Processing (NLP) based on Scipy. The method involves extracting important sentences and condensing them based on their key words. It also eliminates irrelevant sentences and recognition errors before condensing the remaining content. Our evaluation showed that combining sentence extraction with sentence compaction is an effective way to achieve better summarization performance, particularly at 70% and 50% summarization ratios, compared to our previous one-stage method.

REFERENCES

[1] S. Furui, K. Iwano, C. Hori, T. Shinozaki, Y. Saito, and S. Tamura, "Ubiquitous speech processing," in Proc. ICASSP2001, vol. 1, Salt Lake City, UT, 2001, pp. 13–16.

[2] S. Furui, "Recent advances in spontaneous speech recognition and understanding," in Proc. ISCA-IEEE Workshop on Spontaneous Speech Processing and Recognition, Tokyo, Japan, 2003.

[3] I. Mani and M. T. Maybury, Eds., Advances in Automatic Text Summarization. Cambridge, MA: MIT Press, 1999.

[4] J. Alexandersson and P. Poller, "Toward multilingual protocol generation for spontaneous dialogues," in Proc. INLG-98, Niagara-on-the-lake, Canada, 1998.

[5] K. Zechner and A. Waibel, "Minimizing word error rate in textual summaries of spoken language," in Proc. NAACL, Seattle, WA, 2000.

[6] J. S. Garofolo, E. M. Voorhees, C. G. P. Auzanne, and V. M. Stanford, "Spoken document retrieval: 1998 evaluation and investigation of new metrics," in Proc. ESCA Workshop: Accessing Information in Spoken Audio, Cambridge, MA, 1999, pp. 1–7.

[7] R. Valenza, T. Robinson, M. Hickey, and R. Tucker, "Summarization of spoken audio through information extraction," in Proc. ISCA Workshop on Accessing Information in Spoken Audio, Cambridge, MA, 1999, pp. 111–116.

[8] K. Koumpis and S. Renals, "Transcription and summarization of voicemail speech," in Proc. ICSLP 2000, 2000, pp. 688–691.

[9] K. Maekawa, H. Koiso, S. Furui, and H. Isahara, "Spontaneous speech corpus of Japanese," in Proc. LREC2000, Athens, Greece, 2000, pp. 947–952

[10] Y. Stylianou, O. Cappé, and E. Moulines, Eds., Multimodal Signal Processing: Theory and Applications for Human-Computer Interaction. San Diego, CA: Academic Press, 2009.