2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)

Automatic meeting minutes generation using Natural Language processing

1st Satish Muppidi
Dept. of CSE
GMR Institute of Technology
Rajam, Vizianagaram, India
satish.m@gmrit.edu.in

2nd Jayanthi Kandi
Dept. of CSE
GMR Institute of Technology
Rajam, Vizianagaram, India
20341A0583@gmrit.edu.in

3rd Bhavani Sankar Kondaka
Dept. of CSE.
GMR Institute of Technology
Rajam, Vizianagaram, India
20341A0599@gmrit.edu.in

4th Chakradhar Kethireddy
Dept. of CSE
GMR Institute of Technology
Rajam, Vizianagaram, India
20341A0592@gmrit.edu.in

5th Sai Eswar Kandregula
Dept. of CSE.
GMR Institute of Technology
Rajam, Vizianagaram, India
20341A0584@gmrit.edu.in

2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT) | 979-8-3503-1341-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/EASCT59475.2023.10393102

***Abstract*—In recent years, online meetings have become a significant component of communication in a variety of organizations. To optimize the generation of meeting minutes, we propose a novel approach based on NLP in this work. A model that converts speech to text, summarizes meetings has been proposed in the existing work. The goal of this proposed work is to automate speech recognition, extract key points, summarize the meeting, and extract action items. The approach begins with extracting text from a meeting that is in transcript or audio format. The proposed work utilizes summarization models like BART (Bidirectional Auto-Regressive Transformers), T5 (Text-to-Text Transformer Model), and a Summarization Pipeline. Summarization pipeline model performed well for generating meeting summary. Finally, the proposed work generates concise meeting minutes for the meeting which can be in the form of transcript or Audio. ROUGE Score and Human evaluation will be used to evaluate these models.**

***Keywords*—Transcript, speech Recognition, summarization, Key points Extraction, Action items.**

I. Introduction

In recent times online meetings are playing a key role in many of the organizations where the people are virtually collaborating, discussing, decision-making. They are typically conducted using video conferencing software, such as Zoom, Google Meet, or Microsoft Teams. These online meetings are bridged between the people in the far areas and decrease the time, efforts, travelling to the place of meeting. Important key information would be discussed in the meeting, but it is not that much easy to sit straight away for hours to listen to the discussion in the meeting. Meetings will consist of some of the both important and unimportant information and sometimes the participants will be unable to attend the meeting due to technical issues or any reasons or a lot of discussion on a particular topic will be takes place so that the meeting participants will get confused or feel ambiguous about the decisions and the important information that is delivered in the meeting. They will lose the information that is being discussed now and at this time meeting minutes will play a major role in preparing efficient meetings. Meeting minutes are the record of meaningful discussions, decisions, and key takeaways from a meeting. There are two types of meeting minutes formal and informal meetings. These meeting minutes are used for summarizing the meeting for those who couldn't attend the meetings, important discussions and key points that are discussed in the meeting, consists of participants who attended the meeting and the discussion which are used for the future meetings to prepare agenda for the next meeting. Automatic meeting minutes generation systems have the potential to save time and effort, improve the accuracy of meeting minutes, and make meeting minutes more accessible. Meeting minutes are generated from the meeting recordings or the transcripts that are available in the meeting calendar or history. The system is developed so that the meeting is manually recorded using the microphone or the mobile recorder if both are not available in the meeting history.

II. Existing work

Zhang in 2023[1] The Hybrid-View Class Minutes Automatic Generation Model (HVCMM) for Teaching, which is the suggested work, seeks to produce short class minutes by extracting the essential information from the texts of the class records using text summarizing technology.

Long texts are encoded using the HVCMM model's multilevel encoding technique to prevent memory overflow. To address the problem of confusion brought on by the excessive number of students in the class, it also integrates coreference resolution and role vectors. Machine learning algorithms are used to evaluate the subject and body of texts as well as gather structural data. Testing on the improved multiparty interaction (AMI) dataset and the Chinese class minutes dataset (CCM) reveals that the HVCMM model performs better on the ROUGE metric than other baseline models. This proposed model performs well in several areas, such as the ability to generalize in non-teaching scenarios and to compare several abstract models in a teaching context. Furthermore, it exhibits stability in situations involving several conversations.

Devika in 2021[2] The article offers a survey of the literature on key word extraction, concentrating on papers from various fields as well as current events, news, and online content. In this overview, supervised and unsupervised key word extraction techniques are covered, including graph-based models and machine learning models. The use of deep learning algorithms for Twitter data, which have showed promise in enhancing the efficiency of numerous NLP jobs, is also mentioned in the article. The literature review also briefly examines the shortcomings of previous key word extraction studies, especially in scientific material due to insufficient human-annotated corpora.

Biswas in 2022[3] The paper provides a literature review of related research in the domain of automatic text summarization. The review broadly categorizes the research into two categories: extractive summarization and abstractive summarization. The paper focuses on extractive summarization, which involves selecting and condensing the most important information from the original text. This literature review covers a wide range of research studies, including early work from the 1950s to more recent studies that use deep learning architectures such as BERT and recurrent neural networks. The limits of current summarizing techniques are also discussed in the study, notably in the context of call transcripts, which present special difficulties that are not sufficiently handled by most open-source automatic text summarizers.

Koay in 2021[4] This essay will outline a sliding window method for creating meeting minutes on demand. Using a sliding window and a neural abstractive summarizer, the method searches through the transcripts for relevant information. It summarizes some of the current research in this field and gives some background information on the topics of text summarization and speech synthesis. By automatically combining notes and producing meeting minutes, the suggested method might help meeting attendees save time and effort. It sets the proposed strategy in perspective and emphasizes the need for more study in this area.

Manuel in 2021[5] The authors propose an automated system called AMBOC, the Automated Minute Book Creation system, leverages machine learning to automate the creation of meeting transcripts and minutes. It employs Deep Neural Networks (DNN) to convert audio into plain text, identifies speakers using Mel Frequency Cepstral Coefficients (MFCC), and utilizes Transformers for summarizing meeting transcripts into concise minutes. The system segments meeting recordings based on pauses and converts them to text. To condense meeting minutes, it uses T5 transformers for abstractive text summarization, eliminating irrelevant stop words and content.

FM in 2022[6] The primary goal of the article is a pretrained big Transformer-based encoder-decoder model named PEGASUS, which would boost similarity over manual minutes of meetings and simplify summarization by breaking up text data into subtext pieces. In comparison to extractive summary approaches, the employment of abstractive summarization techniques can result in more meaningful summaries and more stable language usage.

Arianto in 2023[7] In this study, the Latent Dirichlet Allocation (LDA) method is employed to identify topics. The research involves calculating coherence and similarity scores using the LDA approach to assess topic quality. The proposed method can assist notetakers in concluding online meetings and automatically generate meeting minutes, which can save time and resources. In the conducted research, the average values for the coherence score and similarity score stand at 64.56% and 57.91%, respectively.

Jung in 2023[8] By allowing for the use of information other than input statements, the suggested technique seeks to improve summarization performance in comparison to conventional methods. An embedding-based metric has been developed to assess the soft overlap between contextual bidirectional encoder representations from transformer (BERT) embeddings of tokens between the reference and the produced summaries. To partially satisfy CML (Constraint Markup Language) restrictions, several summaries are generated.

Liu in 2020[9] Using the recorded audio from the meeting, draft minutes. Future work plan extraction, meeting voice text transcription, meeting minutes extraction, and speech detection during meetings are the four sections that make up the system. The system may extract crucial key phrases and abstract sentences from the complete meeting transcript, remove descriptions of negative emotions, and offer the meeting's next task agenda.

Jha in 2022[10] The goal of the article is to provide a mechanism for quickly and fluently summarizing business meetings while retaining important information and the meeting's overall purpose. The study also focuses on enhancing the voice-to-text conversion of recorded audio files utilizing the Rev-AI Speech-to-Text API. The next focus of this work will be to investigate how the suggested system might be used to summarize various kinds of audio information, like speeches, interviews, and lectures.

Kawai in 2021[11] The purpose of the study paper is to make recommendations for a Question-Answering System that would allow citizens to receive responses to their questions regarding the local government without having to read the meeting minutes. The benefits of the suggested Question-Answering System are not stated in the study clearly.

Cho in 2021[12] This work demonstrated a potent framework (VQ- VAE) for deep neural network model learning. The goal is to use the String Hover summarizer and BERT model to extract summaries from big datasets or Livestream transcripts. It might concentrate on enhancing summarizing tools to give viewers a focused perspective of streamed content.

Agrawal in 2021[13] The goal of the paper is to use NLP methods to summarize the Google Meet content. Based on the ROUGE ratings, it is a comparative study of both extractive and abstractive summarization. The article offers a Chrome addon that simplifies the transcription process and produces clear, brief text. The text analyzed by the application is additionally translated into the user's preferred language using Google Translation.

Ohsugi in 2022[14] A machine learning model with a copy mechanism for abstractive extraction of agenda automatically from assembly minutes. Extracted agendas can be utilized as information retrieval indexes, which can help you rapidly understand the meeting's topics of discussion. extending the suggested method to extract more sorts of information from assembly minutes, such as action items or meeting decisions.

Fu in 2023[15] The goal of the study is to create a complete model that can produce better meeting minutes by including structural information. It also aims to jointly optimize summarization and segmentation tasks and to provide a benchmark dataset for Chinese meeting minutes generation research. Future The goal of this research is to investigate how other structural elements, such speaker information and dialogue acts, might be used to enhance meeting summaries' accuracy.

Ghosal in 2022[16] The automatic creation of meeting minutes from multi-party meeting transcripts is the topic of this study. The purpose of the article is to examine the many issues surrounding the assignment and their potential fixes from the standpoint of a multi-year joint community project. creating meeting-specific summarising or minuting models that are better equipped to tackle the difficulties of creating minutes from transcripts of multi-party meetings.

Savelieva in 2020[17] To produce abstractive summaries of narrated instructional videos is the fundamental goal of the BERT Sum model. This report makes various recommendations for future research, including developing benchmark models to supplement the human valuation framework with human-curated summaries and examining the use of these summarization models in conversations between humans and chatbots.

Bharti in 2021[18] In this research, the author described a method for producing audio, transcripts, and summaries from YouTube video clips using an HMM-based technique. This method is utilized to produce meeting minutes from online meetings, lecture notes for online instruction, and narrative/subtitle production from entertainment videos. Existing online meeting platforms like Google Meet and Microsoft Teams can be enhanced in terms of functionality and use by integrating the suggested method as an add-on.

Yamaguchi in 2021[19] The automated minuting method used by the system involves segmenting a transcript into topic-based blocks, which are then summarized using a pre-trained BART model that has been tweaked using a corpus of chat discussion. The resulting minutes are then carefully organized and coherent utilizing an argument mining technique. Future work on the study will focus on the summary module in order to increase the suggested automatic minuting system's grammatical accuracy and fluency.

Ajallouda in 2022[20] The purpose of the paper is to present KP-USE, an unsupervised method based on the Universal Sentence Encoder (USE) embedding methodology for extracting key phrases from lengthy documents. KP-USE is an unsupervised method that can be used in a variety of areas and languages because it doesn't require labelled data. The document makes recommendations for further research into key phrase prediction using KP-USE.

Mahajan in 2022[21] This essay aims to create meeting minutes from a supplied transcript of the proceedings. The authors generated minutes by utilizing cutting-edge text summarizing models and a transfer learning strategy. The goal of this work's future iterations is to integrate these systems into a GAN-like architecture to provide summaries and boost system efficiency.

Yao in 2021[22] The system includes a text tendency analysis algorithm based on a genetic algorithm language model, a Doc2vec text feature representation algorithm integrating the LDA model, and a parallelization model of the algorithm using the Spark big data platform. The system includes a Doc2vec text feature representation algorithm integrating the LDA model, which improves the recognition degree of the fused text representation model.

Bhatnagar in 2022[23] The paper's goal is to introduce DeepCon, an end-to-end multilingual toolbox for automatic meeting minute-taking of multi-party interactions. The study makes numerous recommendations for future development on the DeepCon toolkit, including expanding the toolkit to support additional languages like Telugu, Spanish, and Romanian.

III. Proposed Work

In this work, we focus on meeting minutes generation through four major tasks.

1. we collected the audio files from a meeting and transcribed the audio into text.
2. Extracted the key points in the meeting.
3. Generating summary of the meeting using both extractive and abstractive summarization.
4. Action items extraction from the meeting.

This whole work involves several stages from data collection, text preprocessing, applying different techniques link part of speech (POS) tagging. Named Entity recognition, Latent Dirichlet allocation for topic modelling. Various Transformers models like BART, T5 are applied for extracting the key information from the meeting and generating the summary of the meeting.

The proposed workflow is given in fig 1.

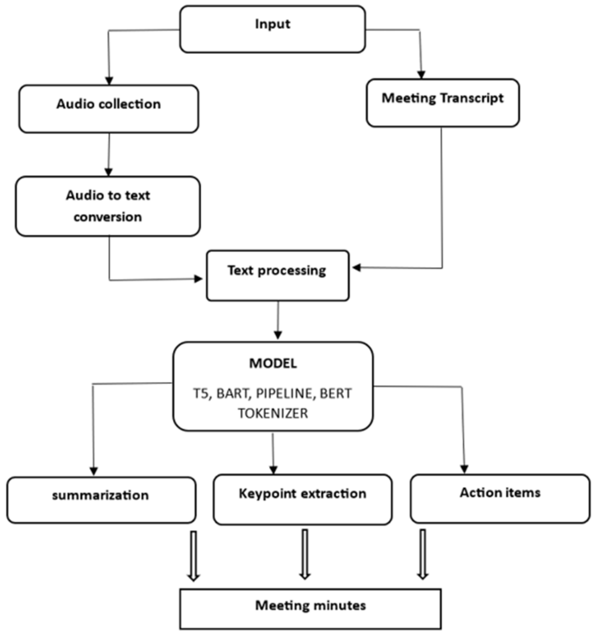

Fig. 1. Flow chart of proposed work

IV. Methodology

A. *Speech Recognition (Audio to text)*

In this task various types of meeting audio files are collected and transcribed into text using PyDub and Speech Recognition using audio segmentation. PyDub is a powerful library for audio processing. To segment an audio file, you can use it to split the audio into smaller, manageable chunks. This code will split the audio into segments based on silence, which is useful for separating different spoken words or phrases. With Speech Recognition that you have segmented audio can use the Speech Recognition library to transcribe each segment into text.

B. *Keypoints Extraction*

1) Using NLTK: In order to extract key information, natural language processing techniques like part-of-speech tagging are used in large part. This indicates the word's part of speech in the text. Keywords, which are often nouns, verbs, and adjectives, can be found using this data. This recognizes named entities in the text, such as persons, locations, and organizations, using NER (named entity recognition). These things may be thought of as keywords as well.TF-IDF: Term Frequency-Inverse Document Frequency is the acronym for this. It is a statistical indicator of a word's importance in a manuscript. More significance is attached to words that are used often in one text but not in many others. Using Text rank, this is an algorithm that ranks the words in a document based on their importance. The algorithm considers the frequency of the words, as well as the proximity of the words to each other.

2) Using Spacy: We employed the 'en_core_web_sm' model, which is a small-sized English model offered by spacy for different Natural language processing applications, in this key points extraction. The collected text is first loaded into the spacy model. The entities and keywords are then extracted from the loaded text. The weight of each important point is then determined by combining the keywords and entities and considering both its frequency and where it appears in the text. The relative relevance of each key point is calculated by dividing the frequency of the key point by the total number of tokens in the provided text. Third, assign a score to each phrase in the input text based on the importance of each one's main ideas. The content is divided into sentences depending on the score as sentence boundaries which results in extracting of important points.

C. Meeting Summarization

We have done this summarization task in two ways extractive and abstractive summarization where in extractive summarization summary is generated using the word present in the text only, no new text will be produced as a summary. A summarization pipeline is applied for generating the summary where the meeting is long a large text is generated so to ease the process of summarization a pipelined model is used which divides the large text into small chucks each respective chuck will be summarized. For abstractive summarization, pre-trained transformers models like BART, T5, transformers pipeline are applied to generate summary of the meeting as a new text.

1) T5 Model: The procedure starts with the text you wish to summarize as input. This text might be a document, an essay, a news report, or any other written work that you wish to summarize. [i]Text-to-Text Approach: T5 uses a special "text-to-text" strategy in which the input and output are both regarded as text. When text summarization is used, the input text serves as the source document, and the objective is to provide a succinct summary of this material. [ii]Tokenization: The input text is tokenized using a pre-trained tokenizer into smaller units, usually words. Tokenization changes the text into a model-compatible format where each token represents a distinct word or sub-word in the input. [iii] Encoding: The T5 encoder is then used to encode the tokenized input. The encoder is made up of several layers of feedforward neural networks and self-attention processes. It records the connections between various text segments while it analyses the incoming text in parallel. [iv]Decoding: After encoding the input, T5 generates the summary using the model's decoder component. Additionally, the decoder has numerous layers of feedforward and self-attentional networks. It considers both the encoded input and previously created tokens as it creates the summary tokens step by step. [v]Text Generation: The decoder predicts the following token in the summary during text generation using the supplied text and context. The model is capable of producing summary tokens up until a predetermined stopping condition, such as a maximum length or an end-of-sequence token. Fig-2 shows the T5 model architecture.

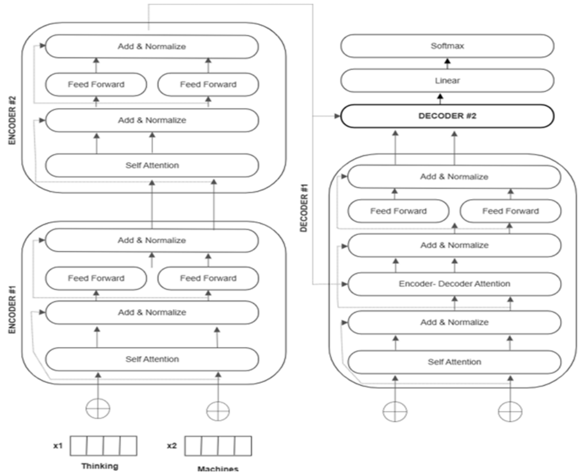

Fig. 2. T5 Model Architecture

2) BART MODEL: Preprocessing entails deleting superfluous material from the text and dividing it into smaller parts like sentences or paragraphs. The text data has been tokenized, which means it has been divided up into individual words or sub-word units. Sub-word tokenization is a technique used by BART, frequently with tokenizers like Sentence-Piece or Byte-Pair Encoding (BPE). The data are prepared for entry into the model in this stage. This model is trained on a large volume of text data during pre-training in order to learn language representations. Because the model's design consists of both an encoder and a decoder, it may be used to comprehend and produce text. The dataset of source (original) and target (corresponding summary) texts is used to fine-tune the pre-trained BART model for summarization. The model mentioned in fig-3 learns to provide succinct and coherent summaries depending on the input text during fine-tuning.

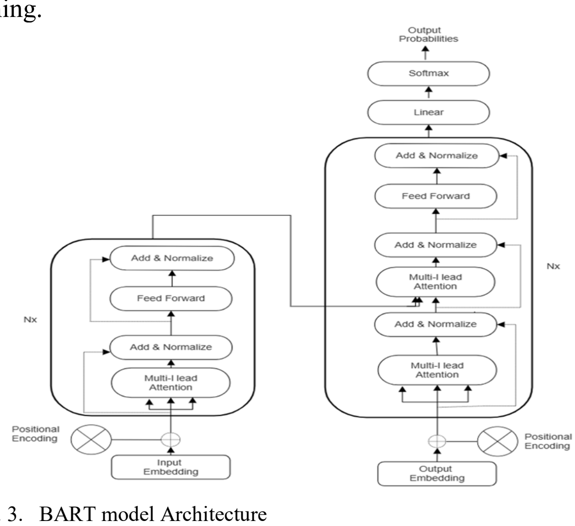

Fig. 3. BART model Architecture

3) Transformer Summarization Pipelining: Summarization using transformers through pipelining is an NLP technique that leverages pretrained models to generate concise summary of the long text provided as input. This task selects a suitable pretrained transformer which tokenizes the text into smaller units and encodes it into a format that the model can understand. Pipeline that encapsulates the entire summarization process. Decoding of model's output which is typically sequence of token into readable text. The length of the summary can be controlled by specifying the maximum length or other parameters. Fig-4 shows the workflow of the whole summarization.

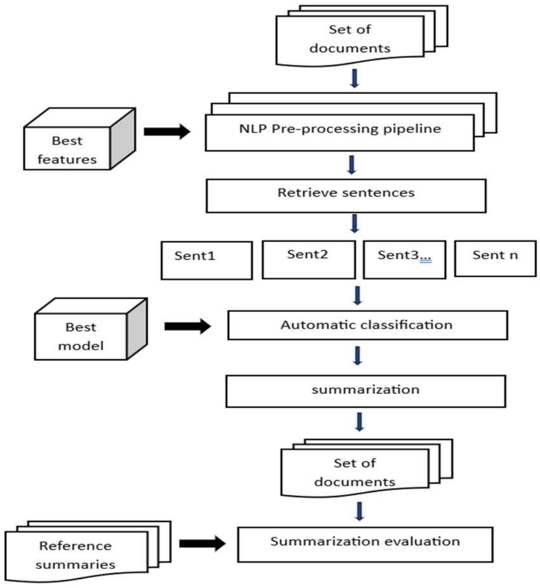

Fig. 4. Summarization pipeline

D. *Action Items Extraction*

In this work, extracting the action items is the main task. Action items will acknowledge the deliverables, work assignments, conclusion of the meeting. For these action items extraction, a BERT based pretrained that is Bert-based-uncased model, is trained to classify the whole meeting into the action items and non-action items using which is a binary classification. using this pretrained model is fine-tuned by the labels of the dataset. Based on the meeting input action items are extracted from the meeting text.

V. Implementation

A. *Speech recognition.*

Using Pydub and audio segmentation the audio file, any conversation recorded and converted into text. How long may be the meeting file size this process will produce the text from the speech by dividing the large audio into several chunks. Some of the chucks of data is shown in fig-5.

```
audio-chunks/chunk46.wav : On june 12th 2023 in the early morning hours.
audio-chunks/chunk47.wav : Officer craig watson an officer rob cuneta were
dispatched to an aed call.
audio-chunks/chunk48.wav : The call notes indicated to patient a 73 year
old female was unconscious and not breathing.
audio-chunks/chunk49.wav : Officer watson and officer cuneta arrived on
scene within 2 minutes.
audio-chunks/chunk50.wav : When they entered the residents they met the
reporting party and they found the patient lying on the floor inside the
home.

audio-chunks/chunk51.wav : Officer watson and officer conetta immediately
started cpr.
```

Fig. 5. Audio to text chunks

B. *Datasets*

For text summarization datasets from hugging faces are used by using the pretrained models.

Dataset used for the summarization are CNN/DM and YouTube Transcripts.

For action item extraction an action dataset is created where it consists of two labels action item or non-action item as binary values (0,1) and the other label is the dialogue related to the action item. The dataset used for the action items extraction in Fig-6.

1	Let's work together today to get this done Phillip
1	you have my approval
0	Lucy, We can discuss your email later.
0	How is progress on creating the spreadsheets.
0	You will probably need to close the file before you attach to an email.
0	It is 2:00.
0	I really want to make some progress on these two files.

Fig. 6. Action items dataset

VI. Results And Discussions

In this proposed work the generated summary is evaluated using Rouge Scores and Human evaluation. Human evaluation by the experts to identify the similarity, Clarity and Topics that are missed out or not. This system will generate meeting minutes for both Audio or Transcript of the meeting. We have employed both Extractive and Abstractive techniques for the generation of minutes. In conversion from speech to text noisy data is lost this is one of the limitations of this work which can be cleared by employing a separate model for reducing the noise and finding out the missed dialogues from the meeting.

A. *Evaluation Metrics*

The text which is generated from all the procedures is evaluated based on ROUGE-1, ROUGE-2, ROUGE-L scores. Table-1,2,3 shows different ROUGE Scores for the summarization. These ROUGE scores are calculated based on the similarity and overlapping n-grams. ROUGE-1 refers to overlap of unigram (each word), ROUGE-2 refers to overlap of Bi-gram (two words) and ROUGE-L refers to overlap of the longest sequence of the generated summary with reference summary.

TABLE I. ROUGE-1 SCORE

Model	ROUGE-1		
	precision	***Recall***	***F1-score***
Summarizer-pipeline	0.82	0.43	0.57
BART	0.80	0.42	0.55
T5	0.5	0.177	0.26

TABLE II. ROUGE-2 SCORE

Model	ROUGE-2		
	precision	***Recall***	***F1-score***
Summarizer-pipeline	0.75	0.34	0.47
BART	0.66	0.29	0.40
T5	0.15	0.048	0.072

TABLE III. ROUGE-L SCORE

Model	ROUGE-L		
	precision	***Recall***	***F1-score***
Summarizer-pipeline	0.47	0.43	0.57

BART	0.40	0.40	0.53
T5	0.072	0.13	0.19

B. Comparitive Analysis

Comparative analysis of the proposed work with the existing work based on ROUGE scores are shown in Table-4.

TABLE IV. COMPARATIVE ANALYSIS BETWEEN THE PROPOSED AND EXISTING MODELS

Model	Pretraining Data	Test Set	Rouge-1	Rouge-L	Content-F1
Transformer [14]	Q/A of TMA	Q/A of TMA	22.8	22.8	35.1
Seq2Seq [11]	Q/A of TMA	Q/A of TMA	32.3	33.2	-
SW (Human Trans) / BART [4]	CNN/DM	CNN/DM	26.1	27.4	33.1
ESAMM [15]	UNGA	UNGA	44.60	43.05	71.9
Distilbert- Base- Uncased	SAMSum	SAMSum	47.4	39.8	-
SVM and BERT [19]	CNNDailyMail	SAMSum	28.2	15.9	32.5
Pipeline	CNN/DM and YouTube Transcripts	CNN/DM and Youtube Transcripts	50.5	48.2	57.1
BART	CNN/DM and YouTube Transcripts	CNN/DM and Youtube Transcripts	49.6	46.5	55.2
T5	CNN/DM and YouTube Transcripts	CNN/DM and Youtube Transcripts	28.1	15.4	20.2

C. System Requirements and Software Requirements

System Requirements are 8GB/16GB RAM and Windows, macOS or Linux operating system. For large dataset 16GB is required. For software requirements Programming language is Python. Text-processing tools such as NLTK and spaCy. Hugging face pre-trained models, transformer libraries tokenizers, TensorFlow etc., It would be good experience by using GPU which can speed up the process.

Fig-7 shows the Graphical representation of different Rouge Scores.

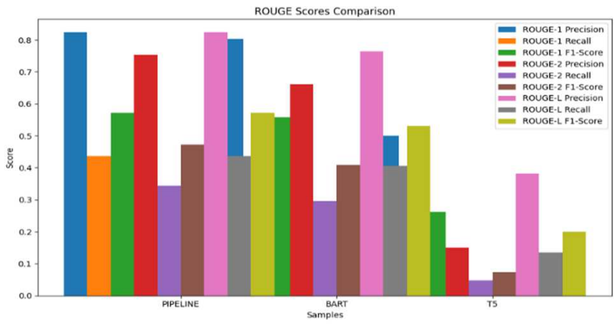

Fig. 7. Bar graph representing ROUGE_SCORES of each model

D. Example of meeting minutes generated

The below fig-8 is an example of meeting minutes generated form a small meeting conversation.

Meeting text = """
Sarah: Good morning, John. Thank you for joining us today. We're excited to discuss the progress of your project. Michael and I have prepared some updates for you.
John: Good morning, Sarah and Michael. I'm looking forward to hearing how things are going.
Michael: Let's start with the project progress. As of now, we've completed 75% of the development work, and our team is on track to meet the project's milestones.
Sarah: That's right, John. We've encountered a few minor issues, but our team has been proactive in addressing them. Quality control measures are in place to ensure the final deliverable meets your expectations.
John: I appreciate the transparency. How are we doing on the budget and timeline?
Michael: Regarding the budget, we are currently within the allocated budget. However, there might be some additional costs if we decide to implement the requested feature changes, we discussed last week. We'll provide you with a detailed estimate for those changes.
Sarah: As for the timeline, we're still on track to complete the project by the agreed-upon deadline. Barring any unforeseen issues, we should meet that target.
John: Thank you for the update. I'd like to discuss those feature changes and get an estimate as soon as possible. Also, I have a few questions about the user interface design. Can we address those?
"""

Key Points Extraction:
1.John: Good morning, Sarah and Michael.
2.However, there might be some additional costs if we decide to implement the requested feature changes we discussed last week.
3.I'd like to discuss those feature changes and get an estimate as soon as possible.
4.Sarah: Good morning, John.
5.Michael and I have prepared some updates for you.

Summarization:
The project has completed 75% of the development work, and is on track to meet the project's milestones. John: "We've encountered a few minor issues, but our team has been proactive in addressing them" John asks: "How are we doing on the budget and timeline?"

Action-items:
1: As of now, we've completed 75% of the development work, and our team is on track to meet the project's milestones.
2: We'll provide you with a detailed estimate for those changes.

Fig. 8. Generated meeting minutes for a small conversation

VII. CONCLUSION

The proposed work generated precise meeting minutes includes Speech to text conversion, extraction of key points, Summarization and extraction of action items from the meeting. In the task of summarization, transformers using summarization pipeline model performed well. In this work extracting action items from the meeting is a new method that can help to plan for preparing next meeting's Agenda. The scope of the work can be improved by reducing the noise when converting the speech to text. Speaker identification using advanced techniques can give maximum scope for generating meeting minutes. Employing different model for converting the meeting minutes into multi-languages can be a great future scope for the existing work.

REFERENCES

[1] Zhang, H., Liu, S., Li, Y., Ren, W., Tong, H., Cao, J., & Zhou, Q. (2023). HVCMM: A Hybrid-View Abstractive Model of Automatic Summary Generation for Teaching. IEEE Transactions on Neural Networks and Learning Systems.

[2] Devika, R., Vairavasundaram, S., Mahenthar, C. S. J., Varadarajan, V., & Kotecha, K. (2021). A deep learning model based on BERT and sentence transformer for semantic keyphrase extraction on big social data. IEEE Access, 9, 165252-165261.

[3] Biswas, P. K., & Iakubovich, A. (2022). Extractive summarization of call transcripts. IEEE Access, 10, 119826-119840.

[4] Koay, J. J., Roustai, A., Dai, X., & Liu, F. (2021). A sliding-window approach to automatic creation of meeting minutes. arXiv preprint arXiv:2104.12324.

[5] Manuel, M., Menon, A. S., Kallivayalil, A., Isaac, S., & KS, D. L. (2021). Automated generation of meeting minutes using deep learning techniques. International Journal of Computing and Digital System, 109-120.

[6] FM, M. F. A., Pawankumar, S., Guruprasath, M., & Jayaprakash, J. (2022, March). Automation of Minutes of Meeting (MoM) using Natural Language Processing (NLP). In 2022 International Conference on Communicatio n, Computing, and Internet of Things (IC3IoT) (pp. 1-6). IEEE

[7] Arianto, R., Abdullah, A., Nurhasan, U., & Wakhidah, R. (2023). Drafting of Online Meeting Minutes Based on Video Recording Using Topic Modelling. SISFORMA, 10(1)

[8] Jung, J., Seo, H., Jung, S., Chung, R., Ryu, H., & Chang, D. S. (2023, March). Interactive User Interface for Dialogue Summarization . In Proceedings of the 28th International Conference on Intelligent User Interfaces (pp. 934-957).

[9] Liu, H., Liu, H., Wang, X., Shao, W., Wang, X., Du, J., ... & Salim, F. D. (2020). Smart meeting: a novel mobile voice meeting minutes generation and analysis system. Mobile Networks and Applications, 25, 521-536.

[10] Jha, A., Temkar, S., Hegde, P., & Singhaniya, N. (2022). Business Meeting Summary Generation Using NLP. In ITM Web of Conferences (Vol. 44, p. 03063). EDP Sciences.

[11] Kawai, T., Akiba, T., & Masuyama, S. (2021, September). Development of Political QA Systems aimed at Assembly Minutes based on Abstractive Summarization . In 2021 8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA) (pp. 1-5). IEEE.

[12] Cho, S., Dernoncourt, F., Ganter, T., Bui, T., Lipka, N., Chang, W., ... & Liu, F. (2021). StreamHover: Livestream transcript summarization and annotation. arXiv preprint arXiv:2109.05 160

[13] Agrawal, Y., Thakre, A., Tapas, T., Kedia, A., Telkhade, Y., & Rathod, V. (2021). Comparative analysis of NLP models for Google Meet Transcript summarization. EasyChair Preprint,(5404)

[14] Ohsugi, R., Akiba, T., & Masuyama, S. (2022, September). Abstractive Approach for Agenda Extraction from Assembly Minutes. In 2022 9th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA) (pp. 1-5). IEEE

[15] Fu, M., & Zhang, J. (2023, April). End-to-end Structure-Aware Model for Meeting Minutes Generation. In 2023 8th International Conference on Computer and Communication Systems (ICCCS) (pp. 1147-1152). IEEE.

[16] Ghosal, T., Hledíková, M., Singh, M., Nedoluzhko, A., & Bojar, O. (2022, July). The second automatic minuting (automin) challenge: Generating and evaluating minutes from multi-party meetings. In Proceedings of the 15th International Conference on Natural Language Generation: Generation Challenges (pp. 1-11)

[17] Savelieva, A., Au-Yeung, B., & Ramani, V. (2020). Abstractive summarization of spoken and written instructions with BERT. arXiv preprint arXiv:2008.09676.

[18] Bharti, N., Hashmi, S. N., & Manikandan, V. M. (2021, September). An Approach for Audio/Text Summary Generation from Webinars/Online Meetings. In 2021 13th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 6-10). IEEE.

[19] Yamaguchi, A., Morio, G., Ozaki, H., Yokote, K. I., & Nagamatsu, K. (2021). Team hitachi@ automin 2021: Reference-free automatic minuting pipeline with argument structure construction over topic-based summarization. arXiv preprint arXiv:2112.02741

[20] Ajallouda, L., Fagroud, F. Z., Zellou, A., & Lahmar, E. B. (2022). Kp-use: an unsupervised approach for key-phrases extraction from documents. International Journal of Advanced Computer Science and Applications, 13(4).

[21] Mahajan, P., Singh, M., & Singh, H. (2021). Team autominuters@ automin 2021: Leveraging state-of-the-art text summarization model to generate minutes using transfer learning. Proceedings of the First Shared Task on Automatic Minuting at Interspeech, 2021, 1-7.

[22] Yao, J., & Xu, J. (2022). English Text Analysis System Based on Genetic Algorithm. Mobile Information Systems, 2022.

[23] Bhatnagar, A., Bhavsar, N., & Singh, M. (2022, September). DeepCon: An End-to-End Multilingual Toolkit for Automatic Minuting of Multi-Party Dialogues. In Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue (pp. 396-402)