The Meeting Wizard

Unlock the Power of In-person Meeting Using LLM + LAM

Abdullah Alwehaibi, Abdullah Alhabib, Abdulrahman Alghamdi, Fahad Alotaibi

Abstract

In the aftermath of the COVID 19 outbreak virtual meetings have become cornerstone of professional communication. However, Managing the outcomes of these meetings poses a challenge which could result in inefficiencies and missed opportunities. To tackle this issue, we created The Meeting Wizard, an AI driven system that summarizes meetings, extracts tasks from meetings, and automate them. By leveraging Large Language Models (LLM) for text summarization and task detection The Meeting Wizard sifts through. Emphasizes vital information, decisions and tasks while also integrating APIs to streamline simple tasks. The Meeting Wizard filter and highlight information, decisions and tasks and also integrate APIs where simple tasks can be automated. This tool does not only boost efficiency but also reduce post-meeting activities which minimize misunderstandings and enhance overall productivity. Our findings show promising accuracy, in transcription and summarization with applications across sectors like business, education and healthcare.

Table of Contents

1. Introduction 3

2. Related Work 3

3. Methodology (Design) 4

3.1 Data Preparing and Transcribing 4

3.2 Text Summarization 4

3.3 Tasks Extraction 4

3.4 API Integration 5

4. Results 5

4.1 Speech-to-text Accuracy 5

4.2 Summarization Quality 5

5. Conclusion and Future Work 6

6. References 7

1. Introduction

The shift in workplace dynamics post the global COVID 19 crisis has significantly increased the need for virtual meetings as a means of communication. While virtual meetings facilitate real time collaboration opportunities, they often leave participants grappling with documenting discussions pinpointing action items and ensuring communication and execution of decisions. This procedure takes up a lot of time and is also prone to mistakes, which can cause inefficiencies that may impact the performance of an organization.

The Meeting Wizard initiative was created to tackle these issues by utilizing artificial intelligence specifically Large Language Models (LLMs) to automate the transcription, summarization, and extraction of tasks from virtual meetings. Our goal is to boost efficiency by saving time for important tasks enhancing the accuracy of meeting summaries and ensuring clear identification and completion of action items.

Our system captures meetings using quality equipment to ensure clear recordings even in varied environments. The recorded audio is transcribed using OpenAI's Whisper model, known for its accuracy in recognizing speakers. The transcribed text is then summarized with LLMs that focus on extracting information. This method does not only highlight points and decisions but also categorizes tasks as simple or complex

tasks enabling the automation of simple tasks through API integrations.

This AI powered solution is especially beneficial, for organizations holding meetings where manual documentation and task management can become burdensome. By automating these tasks The Meeting Wizard helps reducing the strain on participants and ensures that no important details are missed and facilitates decision making and task completion.

The potential uses of this system are wide ranging. spanning across industries such, as business, education, healthcare and any other field where effective communication and task coordination is crucial. As we fine tune our system we aim to broaden its functionalities to include support for languages and expand the scope of automated tasks thereby boosting its usefulness, in global and diverse organizational environments.

2. Related Work

thanks to Natural Language Processing (NLP) methods, Meeting summarization has made a large progress. Different strategies have been created to tackle the issues that come with speech during meetings like fragmented sentences, background noise and repetition.

A known method is the ClusterRank [1] algorithm, which is based on the TextRank algorithm and tailored for summarizing meeting transcripts. ClusterRank tackles the challenges of noise and repetition by organizing transcripts into groups of linked statements before creating a graph-based model, for summarization. This technique has shown enhancements, in effectiveness compared to the TextRank especially when applied to meeting information from the AMI corpus.

Another method that has sparked interest is the Seq2Seq models [2] for extractive

summarization. These models prove to be efficient in creating summaries by encoding the context of the meeting in vectors, which are subsequently transformed into summaries. The Seq2Seq framework is especially suited at managing huge datasets and generating summaries that is not only concise but also preserve the crucial details from the meetings.

Lastly, recent research conducted on the QMSum benchmark [3] presented an assessment framework, for evaluating summarization models in meetings. The benchmark emphasizes generating summaries based on user queries of summarizing the meeting, which proves to be beneficial when users require specific information, rather than an overall summary.

3. Methodology (Design)

Our project, The Meeting Wizard, focuses on creating an AI powered system that can transcribe, summarize, and extract tasks in meetings automatically and automate their execution. To accomplish this goal we have divided our approach into multiple phases which are: data preparing and transcribing, Text Summarization, Tasks extraction, and API Integration for task automation.

3.1 Data Preparing and Transcribing

In the first phase of our approach, we gathered high-quality audio recordings of meetings. We utilized recording tools to capture the audio clearly for precise transcription. Maintaining quality was crucial due to the varied settings where meetings take place.

After gathering the audio, we used OpenAI's Whisper for converting speech-to-text. Whisper is an Automatic Speech Recognition (ASR) system that can accurately transcribe language even when there are multiple speakers. This transcription serves as the foundation for the summarization and the task extraction in the next stages.

3.2 Text Summarization

To summarize the transcript obtained from the first phase, we used a Large Language Model (LLM) in an attempt to filter out the most important content from the discussions of the meeting. The chosen LLM was fine-tuned by us to highlight only the key points, decisions, and the tasks that were considered during the meetings. The quality of the generated summarized text was assessed by means of the standard evaluation tools such as ROUGE, which measures the overlap between the system-generated summaries and reference summaries. Moreover, the usage of human reviews was made in order to check whether the summaries meet the particular business needs of the users. The process of generating the summaries was conducted in Python with help of summarizer.py module. This module links to the LLM with the transcriptions data, analyze the text and produces an abstract, which convey the most important information about the meeting.

3.3 Tasks Extraction

In this phase we aimed to extract tasks from the transcript. We tested out both Large Language Models (LLMs) and Large Action Models (LAMs) to recognize and categorize tasks. As LAMs are still in the stages of development, we decided to

utilize an LLM model for this task. The model was trained to differentiate between simple and complex tasks which will allow for better task management.

The task extraction process took place within the extractor.py module, where the model analyzes the text, identify tasks, classify them based on their complexity. The identified tasks are then saved for handling and automation.

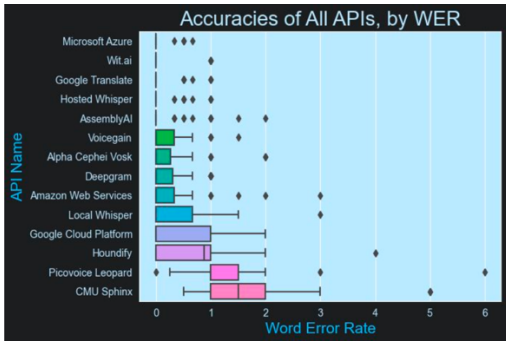
3.4 API Integration

In order to boost efficiency within the organization we incorporated APIs to automate the performance of simple tasks as identified by the system. The integration process was carried out through the show_tasks.py module, which establish communication, with services to streamline activities such as sending emails arranging meetings and updating task management system

4. Results

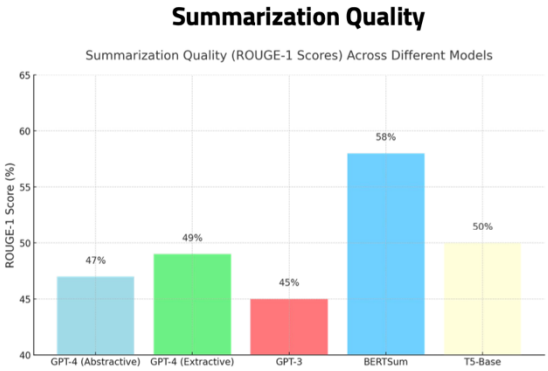
The results of our study demonstrate the effectiveness of the tools and models we selected for converting speech to text and summarizing the resulting text. We prioritized accuracy and reliability in these processes to ensure that the subsequent stages of task extraction and automation were based on dependable data.

4.1 Speech-to-text Accuracy



As depicted in the first graph, we evaluated various APIs for speech-to-text conversion, focusing on their Word Error Rates (WER). Among the options, OpenAI's Whisper model, both in its hosted and local versions, demonstrated the lowest WER, indicating superior accuracy in transcribing speech from meetings. This accuracy is crucial as it ensures that the transcriptions, which serve as the foundation for further processing, are precise and reliable. The Whisper model's performance exceeded that of other popular APIs such as Google Cloud Platform, Microsoft Azure, and Amazon Web Services, making it the optimal choice for our system.

4.2 Summarization Quality



Summarization Quality For the text summarization and task extraction processes, we compared several Large Language Models (LLMs), including GPT-3, GPT-4, BERTSum, and T5-Base. The graph above presents the ROUGE-1 scores, a standard metric for evaluating summarization quality. Although BERTSum achieved the highest score at 58%, we chose to implement GPT-4 (both its extractive and abstractive versions) in our system due to its balanced performance. GPT-4 scored 47% for the

abstractive model and 49% for the extractive model, offering a good compromise between summarization effectiveness and advanced natural language understanding. This balance is essential for ensuring that the summaries not only capture the key points and decisions but also maintain a coherent and contextually appropriate narrative.

In conclusion, our selection of Whisper for transcription and GPT-4 for summarization allowed us to build a robust and reliable AI-driven system. These tools were instrumental in producing accurate transcriptions and high-quality summaries, which in turn facilitated the efficient extraction and automation of tasks discussed during meetings. These results underscore the system's potential to enhance productivity and streamline post-meeting activities across various sectors.

5. Conclusion and Future Work

In the context of our research, we have developed The Meeting Wizard which is the AI-based system designed to facilitate meetings by providing services such as transcription, summarization of the meeting and identification of the tasks which were discussed during the meeting and automation them.

By combining the speech to text tools such as OpenAIs Whisper for transcribing speeches presented in meetings, and by using the most advanced LLMs which can summaries and extract information, manages to deal with the difficulties of documenting and organizing meetings' results.

Using our tool, we observed the high potential for the recognition of language during meetings even if it is of varying and complicated nature. The summaries produced were brief but comprehensive in a way that captured points, decisions, and tasks with accuracy that was evaluated using ROUGE scores as well as the human evaluation feedback. However, the integration of the API based task automation carried out the simple tasks of meeting further enhancing participant's workload. Thus, observing the results achieved by our system, there are features where its performance needs to be further refined and its effectiveness has to be extended. One of the areas of improvement may be to build up the task extraction model that is more suitable to capture complicated tasks by integrating the theoretical changes in Large Action Models (LAMs) over time. In addition to this, to increase the capability of the systems in other areas such as transcription, summarization that would increase the usefulness of the system in places of work across the globe. Furthermore, the future research could explore real-time summarization as well as the identification of the tasks that the system might provide feedback during the meetings. Another useful line of development would be to apply sentiment analysis to get a sense of the mood and likely emotions in meetings that could help in decisions and follow up actions. In general it can be stated that The Meeting Wizard can be regarded as a improvement of the conventional approaches to meeting management with its uses in various fields. Probability of this system evolving could also enhance its position as a utility in increasing interaction and output, in the current workplace.

6. References

[1] Garg, N., Favre, B., Reidhammer, K., & Hakkani-Tuer, D. (2009). ClusterRank: a graph based method for meeting summarization. *ClusterRank: A Graph Based Method for Meeting Summarization*. https://doi.org/10.21437/interspeech.2009-456

[2] Fm, M. F. A., S, P., M, G., & J, J. (2022). Automation of Minutes of Meeting (MoM) using Natural Language Processing (NLP). *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*. https://doi.org/10.1109/ic3iot53935.2022.9767933

[3] Marklynn, V., Sebastian, A., Tan, Y. L., Bae, W. D., Alkobaisi, S., & Narayanappa, S. (2024). A framework for abstractive summarization of conversational meetings. *A Framework for Abstractive Summarization of Conversational Meetings*. https://doi.org/10.1109/ccwc60891.2024.10427755