



North South University

CSE499A: Senior Design Project I

Project Design

“Autonote: Transformative meeting summarization and highlighting points based on NLP “

Group Members:

Serial	Name	NSU ID	Section
1	Md. Saiyem Raiyan	2012468042	10
2	Samia Sultana	2014048042	29
3	Sheikh Mohammed Wali Ullah	2021186042	10
4	Zobaer Ahammod Zamil	2021796042	10



Overview

In today's fast-paced business environment, effective communication and documentation of meetings are essential for collaboration and decision-making. However, the process of manually summarizing meetings can be time-consuming and prone to errors. To address this challenge, our project, Autonote, aims to leverage Natural Language Processing (NLP) techniques to automate the summarization of meetings. The goal of NLP meeting summarization is to make it easier for people to understand and use the information that is generated in meetings. This can lead to a number of benefits, such as improved communication, collaboration, decision-making, and productivity.

What is Summarization?

Summarization is the process of condensing a longer piece of text, such as an article, book, or document, into a shorter version while retaining its essential information, main points, and key ideas. This is a useful skill for students and professionals, allowing for quick overviews, key points identification, comparisons, summary creation, and report generation. It saves time, enhances learning, and improves communication.

In the context of our project, "**Autonote**" summarization refers to the automated process of condensing the content of meetings into shorter, more digestible versions while retaining the essential information, key discussion points, action items, and decisions. This feature of Autonote streamlines the review and reference of meeting content, making it easier for participants to grasp the main ideas without the need to sift through the entire meeting transcript. Summarization in Autonote enhances efficiency, promotes effective collaboration, and facilitates better decision-making by providing concise and coherent summaries of meetings.

What is Natural Language Processing (NLP)?

Natural language processing (NLP) refers to the branch of computer science and more specifically, the branch of artificial intelligence or AI concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. NLP combines computational linguistics rule based modeling of human language with statistical, machine learning, and deep learning models. Together, these technologies enable computers to process human language in the form of text or voice data and to 'understand' its full meaning, complete with the speaker or writer's intent and sentiment.

Several NLP tasks break down human text and voice data in ways that help the computer make sense of what it's ingesting. Some of these tasks include the following:

- **Speech recognition:** It is also called speech-to-text, is the task of reliably converting voice data into text data. Speech recognition is required for any application that follows voice commands or answers spoken questions.
- **Part of speech tagging:** It is also called grammatical tagging, is the process of determining the part of speech of a particular word or piece of text based on its use and context. Part of speech identifies 'make' as a verb in 'I can make a paper plane,' and as a noun in 'What make of car do you own?'



- **Word sense disambiguation:** It is the selection of the meaning of a word with multiple meanings through a process of semantic analysis that determines the word that makes the most sense in the given context. For example, word sense disambiguation helps distinguish the meaning of the verb 'make' in 'make the grade' (achieve) vs. 'make a bet' (place).
- **Named entity recognition (NEM):** Identifies words or phrases as useful entities. NEM identifies 'Kentucky' as a location or 'Fred' as a man's name.
- **Coreference resolution:** The task of identifying if and when two words refer to the same entity. The most common example is determining the person or object to which a certain pronoun refers (e.g., 'she' = 'Mary'), but it can also involve identifying a metaphor or an idiom in the text (e.g., an instance in which 'bear' isn't an animal but a large hairy person).
- **Sentiment analysis:** It attempts to extract subjective qualities, attitudes, emotions, sarcasm, confusion, suspicion from text.
- **Natural language generation:** It is sometimes described as the opposite of speech recognition or speech-to-text; it's the task of putting structured information into human language.

How does NLP work?

Computational linguistics involves understanding and constructing human language models using computers and software tools. Machine learning trains computers to understand human language features like sarcasm, metaphors, and grammar. Deep learning, a specific field of machine learning, teaches computers to learn and think like humans. NLP implementation involves gathering and preparing unstructured text or speech data, pre-processing using techniques like tokenization, stemming, lemmatization, and stop word removal, training models with large data samples, and deploying and inferring the model's output. This process allows for the application to be run on live data and obtain the required output.

Approaches to natural language processing

Natural language processing (NLP) approaches include supervised, unsupervised, natural language understanding, and natural language generation. Supervised NLP trains software with labeled input and output, while unsupervised NLP uses a statistical model to predict patterns in non-labeled input. Natural language understanding analyzes sentence meanings, while natural language generation produces conversational text based on specific keywords or topics. These approaches help software categorize documents, predict sentence patterns, and produce conversational text.



Problem Clarification

In modern workplaces, meetings are essential for collaboration and decision-making. Automatic meeting summarization, powered by Natural Language Processing (NLP), aims to address this challenge by distilling key meeting insights into concise summaries. Sometimes we miss meetings that are happening online due to some unavoidable circumstances. Again, sometimes we miss out on some important points of the meeting because of various reasons. However, before our important work or study time is very crucial so we need to do it in a faster or optimal way. That's why we need to summarize a long meeting so that we can be able to understand easily and a faster way within our due time. Therefore, a gist of an online meeting is often very important to those who missed it or those who try to recall it. There can be a system which can make a summary of a meeting while it is running online.

I. What is NLP doing?

Basically, this NLP **summerizing** develops a system that can automatically generate accurate and informative summaries of meetings.

II. Generated Summary

Introducing a revolutionary concept: a dynamic system that crafts comprehensive summaries of online meetings as they unfold. The generated summary should capture the essence of the meeting while filtering out redundant or less relevant information.

III. System Ability:

The Autonote project seeks to automate the summarization of meetings to:

1. Provide participants with concise and coherent summaries that capture the main discussion points, key decisions, and action items.
2. Reduce the time and effort required for manual note-taking during meetings.
3. Enhance the tracking and management of action items, improving accountability and task follow-through.
4. Facilitate knowledge sharing and decision-making by making meeting insights easily accessible.

By addressing these problem statements, Autonote aims to revolutionize the way organizations conduct and benefit from meetings, ultimately leading to increased efficiency, better collaboration, and more informed decision-making.



Summarization with NLP

Summarization with Natural Language Processing (NLP) refers to the use of NLP techniques and algorithms to automatically generate concise and coherent summaries of longer texts, documents, or spoken content. NLP-based summarization methods can be broadly categorized into two main approaches:

- **Extractive Summarization:** In extractive summarization, the algorithm selects and extracts sentences or phrases directly from the original source text. It identifies the most important sentences or passages by assessing factors like sentence importance, relevance, and redundancy. Extractive summarization methods do not create new sentences or rephrase the content but assemble the summary from the existing material.
- **Abstractive Summarization:** Abstractive summarization takes a more creative approach by generating summaries that may not use the exact words or phrases from the source text. Instead, it rephrases and restructures the content to create a more human-like and coherent summary. Abstractive summarization requires a deeper understanding of the source text and often involves natural language generation techniques.

NLP-based summarization has a wide range of applications, for example:

- **News Summarization:** Automatically generating brief summaries of news articles to provide readers with the main points.
- **Document Summarization:** Condensing lengthy reports, research papers, or legal documents for quick reference.
- **Meeting Summarization:** As in your Autonote project, summarizing the content of meetings to capture key discussions and action items.
- **Content Curation:** Creating concise descriptions for content aggregation, such as blog posts or product descriptions.
- **Search Engine Snippets:** Generating brief descriptions for search engine results to help users quickly assess the relevance of web pages.

Approaches: It employs two primary methods—abstractive summarization, which generates summaries in its own words, and extractive summarization, which selects and combines sentences from the source text.

Applications: NLP-based summarization has diverse applications, including news summarization, document condensation, content recommendation, and enhancing search engine results.

Technologies: Advanced NLP models like BERT and GPT have revolutionized abstractive summarization, while extractive summarization relies on statistical methods, machine learning, and graph-based algorithms.

Benefits: It streamlines information retrieval, aids in decision-making, and saves time by providing clear and concise summaries of complex textual data, making it a valuable tool in various domains.

NLP-based summarization continues to evolve with advancements in machine learning and deep learning, making it an invaluable tool for processing and extracting valuable insights from vast amounts of textual data.



Risk and Constraints

Natural Language Processing (NLP)-based meeting summarization can offer many benefits, but it also comes with various risks and constraints that need to be considered. Here are some of the key risks and constraints associated with NLP meeting summarization:

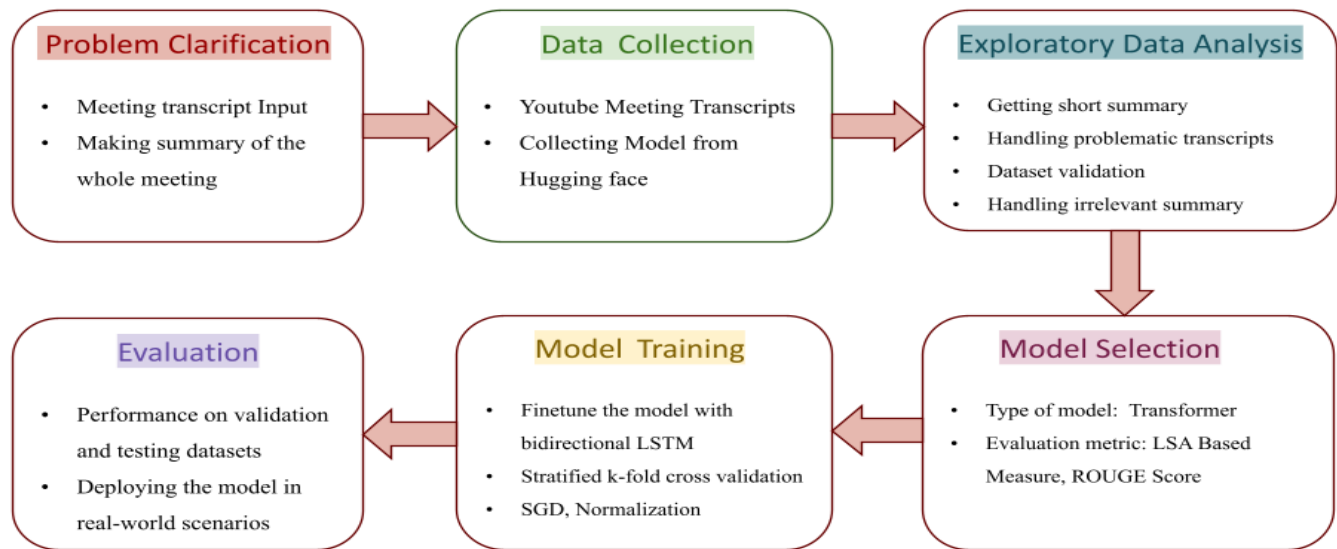
NLP can be used to summarize meetings, but it has several drawbacks. Data privacy and security are crucial, as it may include sensitive information. Misinterpretation of context can lead to inaccurate summaries, and biases in training data can affect the summaries. Incomplete or omitted information can occur, and human judgment is often lacking. Technical constraints, such as computational resources, can affect the timeliness and accuracy of summaries. Multilingual and multimodal challenges, such as speaker identification, can also impact the summaries. Customization and domain specificity are also important, as meeting diverse user expectations can be challenging. Real-time feedback and correction may not be possible, potentially leading to misunderstandings or misrepresentations.

Estimate Budget Plan after implementing

The approximate estimated budget for the meeting summarization project can vary significantly based on factors such as the project's complexity, the scale of implementation, and the organization's specific requirements. Cost considerations include development costs, data acquisition and annotation, NLP tools and libraries, model training and fine-tuning, hardware and infrastructure costs, maintenance and updates, testing and evaluation, user interface development, legal and compliance expenses, training and support, scale and infrastructure growth, security and data protection, human reviewers, and miscellaneous expenses. NLP projects can range from tens of thousands to millions of dollars, depending on the scope and complexity. It is advisable to work, project managers, and financial professionals to create a detailed budget plan tailored to your specific project goals and constraints. Additionally, consider whether to build the system in-house or collaborate with external vendors, as this can impact the budget.



Flowchart



Plan Per Week

Week	Task/Goal
Week 1	Project Idea
Week 2	Project Idea
Week 3	Project Idea
Week 4	Reserarch Paper Analysis
Week 5	Analysis + Related Works
Week 6	Analysis + Related Works
Week 7	Analysis + Related Works
Week 8	Project Design
Week 9	Project Design
Week 10	Project Design
Week 11	Fine Tuning
Week 12	Ethical and professional responsibility
Week 13	Social, Environmental effects
Week 14	Final demo
Week 15	Final presentation
Week 16	Final report

Conclusion

NLP meeting summarization automates the process of condensing discussions into concise summaries, saving time and making content more accessible. It improves decision-making, knowledge retention, searchability, and customization. However, challenges like data privacy, accuracy, and bias need to be addressed. Budgeting and resource allocation are crucial for the system's development and maintenance. Autonote is an NLP project that aims to summarize meetings and provide participants with concise and coherent summaries that capture the main discussion points, key decisions, and action items. It addresses the problems of manual note-taking, incomplete or inaccurate meeting summaries, forgotten action items, and difficult-to-access meeting insights. *Autonote* uses a variety of NLP and machine learning techniques to summarize meetings accurately and efficiently. It is a valuable tool for teams of all sizes, helping them to save time, improve productivity, and make better decisions. Despite these challenges, incorporating NLP meeting summarization into an organization's workflow can significantly enhance collaboration, decision-making, and knowledge management.

Reference

- I. *What is Natural Language Processing?* | IBM. (n.d.). <https://www.ibm.com/topics/natural-language-processing>
- II. Foster, K. (2023). Text Summarization for NLP: 5 Best APIs, AI Models, and AI Summarizers in 2023. *News, Tutorials, AI Research*. <https://www.assemblyai.com/blog/text-summarization-nlp-5-best-apis/>
- III. *What is Natural Language Processing? - NLP Explained - AWS*. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/what-is/nlp/>
- IV. Wikipedia contributors. (2023). Natural language processing. *Wikipedia*. https://en.wikipedia.org/wiki/Natural_language_processing
- V. Collison, J. (2023, May 2). *Maximizing Productivity with Text Summarization: How Chat GPT Can Help You Save Time*. The Average Guy Network. <https://theaverageguy.tv/2023/05/01/maximizing-productivity-with-text-summarization-how-chat-gpt-can-help-you-save-time/>
- VI. Singh, A. (2022, August 28). *Natural language processing (NLP) - DataSciences.io*. DataSciences.io -. <https://datasciences.io/course/natural-language-processingnlp/>

