

Judicial Speech Summarization Tool

Geetha N¹, S Sanjushree², Priya Dharshini M³ and Sneha R⁴

¹⁻⁴Atria Institute of Technology, CSE, Bengaluru, Karnataka, India

Email: geetha.n@atria .edu, {sanjushree2907, magdamag0327, sneharavi0206}@gmail.com

Abstract—Audio-to-text summary entails turning spoken knowledge into brief written summaries to help in effective information retrieval and understanding. Recapitulating an oral argument in court In audio, spoken legal discourse is condensed into a brief written form yet important details and ideas are maintained for effective understanding and reference. Court oral argument records must be summarized in order to conduct case analysis, legal research, and transparency all properly. Long speeches are condensed so that legal professionals can swiftly grasp and recollect important information. This research investigates the critical steps involved in precisely and concisely transcription of a complicated legal argument from audio. For both academics and legal professionals, accurate transcribing of legal discourse is crucial. This research investigates these methodologies, from state-of-the-art Natural Language Processing (NLP) software to conventional transcription procedures. It emphasizes how crucial this technology is to ensuring courtroom transparency, enhancing case analysis, and speeding up legal research. This research contributes to a revolution in the understanding and accessibility of legal knowledge at a time when legal technology is undergoing rapid change. Our project focuses on producing the transcript and summary of the court oral argument audio.

Index Terms— Audio Summarization, Natural Language Processing (NLP), Extractive Summarization, Rouge score.

I. INTRODUCTION

Being able to summarize well is crucial for retaining the core idea while condensing complex details. It comprises choosing the most important information and salient features from a larger body of oral, written, or visual sources. This process is crucial in numerous domains as it enhances comprehension and enables efficient retrieval of knowledge. In scholarly research, professional reports, news articles, and casual chats alike, concise summaries are crucial for effective communication and information dissemination. In an era where information is abundant, it is critical to swiftly record spoken information and put it in writing. Audio-to-text summarization, which makes it possible to efficiently extract key ideas and details from audio recordings, is a crucial step in this process.

This ground-breaking method has a plethora of applications, from expediting material analysis to making the hard of hearing more approachable to professionals and scholars. The foundation for exploring the various applications and methods of audio summarization is laid forth in this introduction. The precise and efficient transcription of oral arguments into written text is critical in the field of legal discourse. This process, known as audio-to-text summarization, is an essential tool for academics, attorneys, and members of the public. Rapid access to essential information is facilitated, speeding up case analysis and legal research. Additionally, it

encourages transparency by providing a clear record of court proceedings. The framework for a study of the methods and applications of audio summarization in the legal industry is established by this introduction. Technology development offers a promising chance to enhance the efficiency and accessibility of legal processes, leading to a more open and inclusive legal system that is advantageous to all parties. ROUGE measures are used to combine multiple standards for evaluating summary texts. The ROUGE (Recall-Oriented Understudy for Gisting Evaluation) ratings evaluate the quality of summaries to reference summaries. They are: ROUGE-N (N-gram overlap), ROUGE-L (longest common subsequence), and ROUGE-W (weighted LCS). Higher scores imply greater summary quality, which is often used in natural language processing to assess text summarizing systems.

A. Summarization Techniques

Summarizing techniques are methods for condensing and removing the most important information, ideas, and details from a lengthy audio, video, or text. The two primary types of summarization techniques are abstraction summarization, which creates new terms for the summary, and extractive summarizing, which synthesizes the phrases extracted from the original into a summary.

Extractive Summarization: Extraction summarization is the process of selecting and reducing phrases or words from the original text. It involves figuring out which sentences best convey the main points and important facts. This method ensures an accurate summary that matches the original material by preserving the original language and structure. It is particularly useful in circumstances where maintaining correctness and integrity to the original context is crucial, such as in legal documents or scientific investigations. TF-IDF, graph-based methods, and machine learning models are just a few of the algorithms used in extraction summarization to find and extract the most important data for a concise and understandable summary.

Abstractive Summarization: Abstractive summarizing is a technique that generates brief, cogent summaries by reinterpreting and rephrasing the original material. Unlike extractive techniques, it doesn't rely solely on repeating sentences, which gives greater freedom and creativity to the summarizing process. This approach uses complex Natural Language Processing (NLP) and deep learning models, such as transformers and sequence-to-sequence architectures, to comprehend the context and generate summaries that resemble those of a human. Since abstractive summarization can produce summaries that may or may not be present in the original text, it is incredibly useful for tasks like creating attention-grabbing headlines, simplifying complex information, and creating creative content that effectively conveys the spirit in comparison to the original source material.

Audio Summarization: The process of "audio summarization" condenses lengthy audio recordings into smaller, more digestible segments while keeping all the relevant information. To extract important content and generate a summary, techniques for audio processing are used with natural language processing. Among other things, this technology can be used to record and summarize podcasts, interviews, seminars, and meetings. It uses methods like speaker diarization, content extraction, and speech-to-text conversion to identify significant sections. Sophisticated algorithms prioritize important data, identify important words, and analyze voice patterns. It is also possible to teach audio summarizing algorithms to distinguish between different voices and identify speaker identities. By automating the summary process, this method enhances accessibility and expedites the retrieval of information from audio data, benefiting sectors including journalism, content creation, and education.

B. AssemblyAI and Streamlit

Our project looks at how cutting-edge methods can be used to courtroom oral sessions, with a focus on turning audio recordings into succinct text summaries. With the help of AssemblyAI, a cutting-edge Automatic Speech Recognition (ASR) system, we hope to streamline and enhance the oral transcription process used in courts. Additionally, a fluid and interactive interface where users can obtain condensed content is made possible by the integration of Streamlit, an approachable web application framework.

In addition to expediting the transcribing and summary process, the integration of AssemblyAI's potent ASR technology with Streamlit's UI design offers legal professionals and stakeholders an intuitive platform for quickly obtaining and comprehending crucial litigation data. This study looks at the techniques and how they work together to transform audio into readable written summaries, which will ultimately improve the processing and retrieval of legal information.

The process of turning courtroom audio recordings into text summaries includes a number of methodical processes, the first of which is uploading the audio recordings to a platform that has sophisticated speech recognition software. AssemblyAI, a prominent supplier of automated speech recognition technology, facilitates this crucial beginning. AssemblyAI translates spoken words into text when audio is uploaded by using sophisticated algorithms, which simplifies a difficult process into a format that is easier to handle. After

transcription, text summarizing techniques are applied to alter this content. These algorithms are essential for gleaning crucial information from lengthy transcripts and creating clear, enlightening text summaries. This thorough procedure not only makes oral proceedings more accessible, but it also gives attorneys a useful tool for finding information fast. Examining this methodology's specifics reveals that the employment of AssemblyAI in conjunction with sophisticated speech recognition and text summarizing approaches provides a potent remedy for the problems caused by big audio data in legal contexts.

II. LITERATURE REVIEW

We searched through the surveys that are currently accessible on text summarizing and audio to text summary and highlighted a few of the articles. Literary works employ a range of techniques to generate summaries for a range of situations and document types. Text summarizing techniques that are extractive and abstractive were studied by Ujjwal Rani et al. [1]. There are several workable solutions or approaches for summarizing content in different dialects. To be able to outline the records, the abstractive summary requires standard language preparation tools. Heuristic strategies are required in the extraction summary in addition to being factual and etymological in order to put the phrases. Three categories—Extractive, Abstractive, and Hybrid text summarization—as well as particular techniques within each were examined by Suad Alhojely et al. [6]. This survey has examined several strategies and techniques. In comparison to employing just one strategy, it was found that combining two approaches or procedures is likely to yield positive outcomes and raise the standard of the summaries. A technique for automatically summarizing audio data utilized in contact centers is covered by Jatin Pardhi et al. [3]. It describes a procedure that uses the AssemblyAI API to convert audio data to text, which is then summarized using graph theory. The best appropriate voice recognition API, according to the authors' comparison of several, is AssemblyAI. Additionally, they discuss text summarizing strategies, emphasizing extraction-based approaches. Future research to extend this method to additional languages is suggested by the authors.

A technique for creating informative audio summaries using probability distribution divergence (DJS) and audio features is presented by Carlos-Emiliano González-Gallardo et al. [4]. Mel-frequency Cepstral Coefficients (MFCC), among other 277 audio characteristics, are used in the process of mapping informativeness. Further research into the links between summary quality and segment quality, as well as multilingual extension, may be part of future study. A study by Paheli Bhattacharya et al.[7] investigated and contrasted different algorithms for summarizing court rulings. ..It is likely that a variety of summarizing techniques were employed in this study, examined, and contrasted to ascertain how well they reduced the length of drawn-out legal texts into manageable summaries. This kind of research is helpful in enhancing the retrieval of legal information and making it simpler to retrieve the substance of court cases. Burman et al., Aman The goal of the project was to apply several NLP algorithms to improve the accuracy of legal document summary. In order to produce more accurate summaries, the researchers adjusted the T5 model on the gathering of legal documents. In the legal sector, legal documents are very important, so automating the summarizing process is essential. When they compared how many NLP models were implemented, BERT emerged as the most successful model. In the end, they decided to greatly enhance the T5 algorithm's performance by fine-tuning it to summarize legal papers. The final model may be accessed via an intuitive interface, making it a helpful tool for lawyers and other professionals.

III. PROPOSED SYSTEM AND ARCHITECTURE

The suggested method looks over the audio input, extracts the essential terms, and summarizes them. The suggested approach would support real-time audio files. For all pertinent audio inputs, the accuracy and efficiency of the suggested solution are good. We build a web application with Streamlit that requests the user's audio file in order to provide a transcription of the audio content. Using Streamlit, a free open-source framework that lets you write Python apps with just a few lines of code, the objective is to create an application.

A. Uploading an audio file

Using AssemblyAI's web interface or API, uploading an audio file is a straightforward process. We are going to upload an audio file to the website we made using their API. It is shown in fig. 1 below.

API access:

Authentication: Obtain an API key from AssemblyAI so that we may authenticate our requests using the API.

Uploading: The endpoint / upload is what we'll utilize to upload our audio files. We include audio files and our API keys in the request. MP3, WAV, and FLAC are just a few of the audio formats that AssemblyAI supports.

Response: Following a successful upload, AssemblyAI provides a response with the uploaded file's unique ID,

which we utilize in future API calls.

B. Transcript and summary of audio file

There has to be transcribing for this section. We now send POST queries rather than GET requests to the AssemblyAI transcript endpoint. AssemblyAI requires the post method to be used in order to load sound URLs. This step is required in order to obtain the transcript and summary, which are produced by default if auto is set to True. We will receive the transcript only if we set autocount to False.

Transcription: AssemblyAI converts an audio file to text using cutting-edge voice recognition technology. To identify the words uttered, the audio is segmented into smaller segments and subjected to analysis. To ensure an accurate transcription, the transcribing process considers a number of variables, including speech patterns, speaker accents, and background noise.

Text processing: Following the audio transcription, AssemblyAI modifies the text to enhance its correctness and readability. This could entail formatting the content, fixing typos, and eliminating pointless filler words.

Summary: Following the transcription, AssemblyAI use natural language processing (NLP) methods to provide an overview of the text as a summary for the users. You can use abstract or extractive approaches to perform a summary. In contrast to abstract summarizing, which bases the summary on the text's content, extractive summarizing involves selecting key words or sentences from the text. We have used extractive approach in our project. The method of summarizing aims to produce a concise and educational synopsis of the audio material.

C. Web Application

Streamlit will be utilized to develop an online application that asks for the user's audio file in order to provide a transcription of the audio recording. Using Streamlit, a free open-source framework that lets you write Python apps with just a few lines of code, the objective is to create an application.

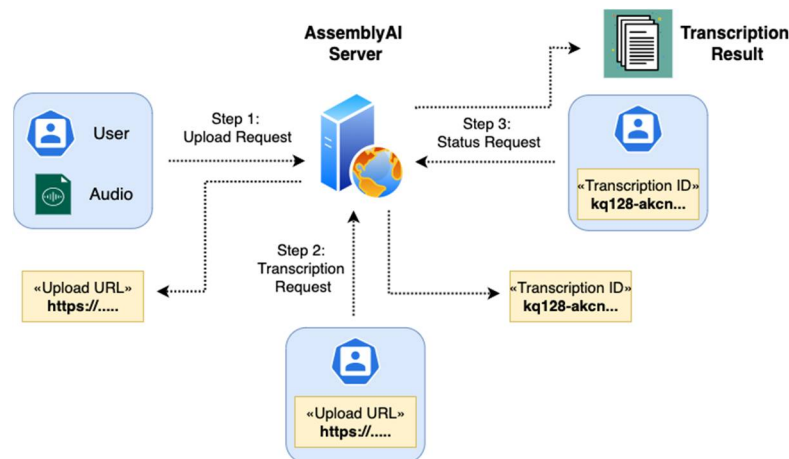


Figure 1. Process of getting transcript using Assembly AI

IV. RESULT

The primary input for processing the summary given in this article is the audio file. The court's oral speech, which has already been recorded in the court, is uploaded as the input. Once the audio file has been uploaded, the application will take all the required actions to compile and transcribe the audio of the oral court session once the button is pushed. After a little while, the results will be visible on the website.

The result shows the transcript text file that can be downloaded, summary notes of the audio uploaded, and a complete short summary of the uploaded audio, as well as the rouge measures at the end. The summary notes include the chapters that has the heading and the summary at the given time chunks. The below figures shows the website and the result page.

A summarization system's performance is shown by its ROUGE scores shown in below table. Merely 4% of the unigrams produced by the system were included in the reference summary, as shown by the ROUGE-1 Precision of 0.04. The system recalled every unigram in the reference summary if the recall value was 1.00. ROUGE-1's F-Measure, however, is low at 0.08, indicating potential for improvement. With a Precision of 0.04, Recall of 0.93, and F-Measure of 0.07, ROUGE-2 exhibits comparable patterns and suggests that bigram matching has to be improved.

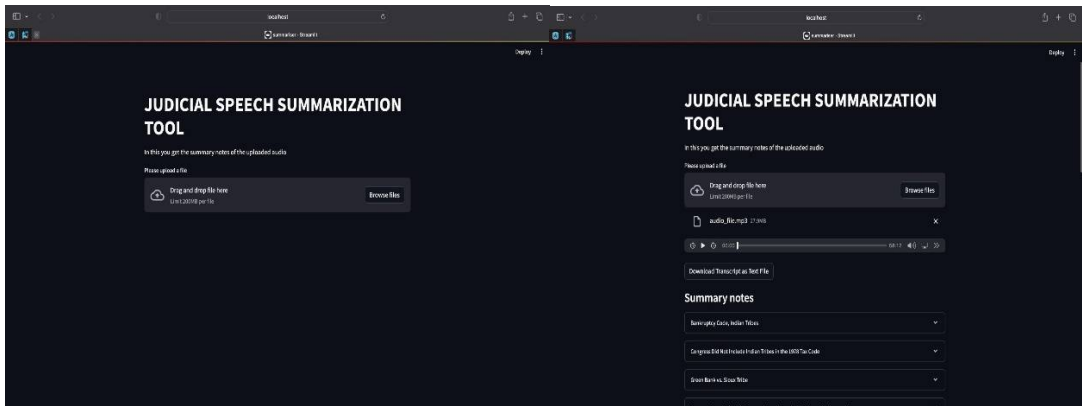


Figure 2. Website page

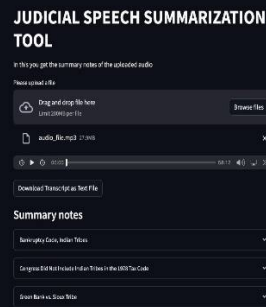


Figure 3. Result page

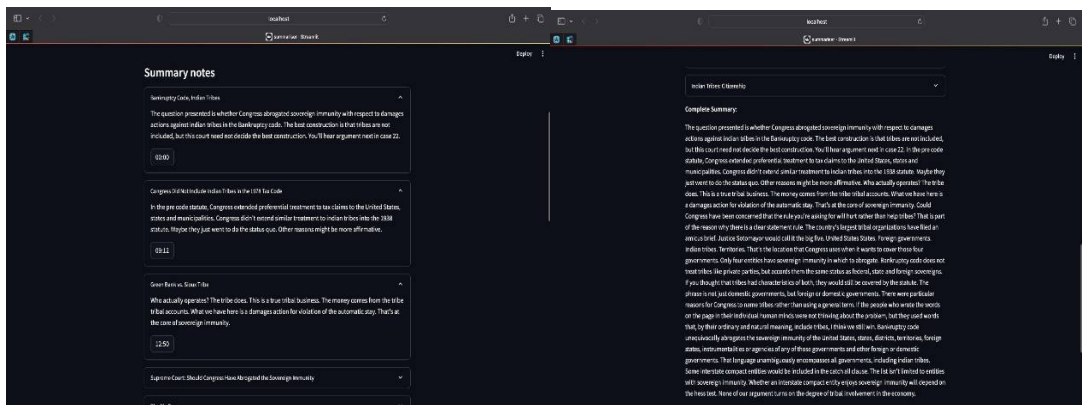


Figure 4. Summary notes shown

Figure 5. Complete summary shown

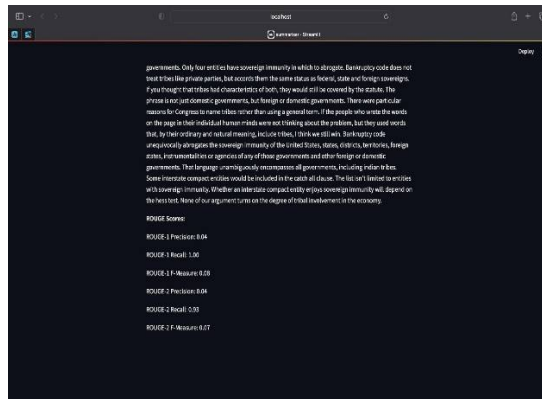


Figure 6. Shows Rouge scores at the end of the page

TABLE I. ROUGE SCORES

Metric	ROUGE - 1	ROUGE - 2
Precision	0.04	0.04
Recall	1.00	0.93
F- Measure	0.08	0.07

V. CONCLUSION

To sum up, the goal of our project is to create succinct written summaries from judicial oral statements that are recorded in court. After audio data is acquired, it goes through pre-processing to transform it into a format that

can be analyzed. The audio content is then converted into text using the speech, and the text is then further pre-processed to make it more polished. Text summary: Using an algorithm, whether it be an extract or an abstract, is essential to producing insightful and useful summaries. This method will be helpful in successfully resolving a complex legal dispute by enabling legal professionals to obtain pertinent information fast. These techniques' smooth integration will contribute to the development of a voice-to-text court summary system that will streamline legal procedures, provide access to important legal information, and offer a useful way to enhance court proceedings and documentation. We can enhance text data gathering by utilizing sophisticated audio pre-processing, reliable voice recognition systems, and strong text summarizing techniques. Legal experts may now efficiently comprehend and manage the large amount of audio data by using a succinct summary that is provided by this technology, in addition to making it easier to create accurate and searchable transcripts. The method improves the accessibility and effectiveness of the legal process in addition to saving time and money.

ACKNOWLEDGMENT

We express our sincere gratitude to **Dr. Rajesha S**, Principal, Atria Institute of Technology, for providing congenial environment to work in. We also express our gratitude to **Dr. Devi Kannan**, Head of the Department, Computer Science and Engineering, Atria Institute of Technology, for her continuous support and encouragement. We express our gratitude to **Mrs. Geetha N**, Assistant Professor, Department of Computer Science and Engineering, Atria Institute of Technology, our guide for her continued support, advice and valuable inputs during the course of this project work.

REFERENCES

- [1] Ujjwal Rani and Karambir Bidhan, "Review Paper on Automatic Text Summarization," International Research Journal of Engineering and Technology (IRJET), Vol. 7, No. 4, 2020
- [2] Balaji N, N Megha, Deepa Kumari, Sunil P, Bhavatarini N, A Shikah, "Text Summarization using NLP Technique," in The 2022 International Conference on Robotics, Electrical Circuits, VLSI, and Distributed Computing (DISCOVER), IEEE, 2022
- [3] Jatin Pardhi and Tanmay Kharik, "The automated method of Summarization of Audio Data." By Creative Commons License 4.0 International, Springer, 2021
- [4] Carlos-Emiliano González-Gallardo, Romain Deveaud, Eric Sanjuan, Juan-Manuel Torres-Moreno, "Audio Summarization with Audio Features and Probability Distribution Divergence," 20th Computational Linguistics and Intelligent Text Processing International Conference, 2019, hal-02171387
- [5] Athanasia Zlatintsi, Elias Iosif, Petros Maragos and Alexandros Potamianos, "Audio salient event detection and summarization using audio and text modalities," in The 23rd European Conference on Signal Processing (EUSIPCO), 2015
- [6] Suad Alhojely and Jugal Kalita, "Recent Progress on Text Summarization," Computational Science and Computational Intelligence (CSCI) International Conference, 2020
- [7] Paheli Bhattacharya, Kaustubh Hiware, Subham Rajgaria, Nilay Pochhi, Kripabandhu Ghosh, Saptarshi Ghosh, "A Comparative Study of Summarization Algorithms Applied to Legal Case Judgments," The Advances in Information Retrieval, Springer, 2019, pp. 413-428
- [8] Aman Burman and Eric Bradford, "Building an Optimized algorithm that provides summaries of legal documents," ResearchGate, 2022
- [9] P. R. Dedhia, H. P. Pachgade, A. P. Malani, N. Raul and M. Naik, "Study on Abstractive Text Summarization Techniques," in The Conference of International on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-8, doi: 10.1109/ic-ETITE47903.2020.087
- [10] Zlatintsi, Nancy & Maragos, Petros & Potamianos, Alexandros & Evangelopoulos, Georgios, "A saliency-based approach to audio event detection and summarization," in Conference on European Signal Processing, 2012
- [11] N. Nazari and M. A. Mahdavi, "A survey on Automatic Text Summarization," Journal of AI and Data Mining, Vol 7, No 1, 2019, pp. 121-135
- [12] Aghdam, Mehdi Hosseinzadeh. "Automatic Extractive and Generic Document Summarization Based on NMF" Journal of The Artificial Intelligence and Soft Computing Research, vol.13, no.1, 2023, pp.37-49. <https://doi.org/10.2478/jaiscr-2023-0003>
- [13] Gambhir, Mahak & Gupta, Vishal, "Recent automatic text summarization techniques: a survey," Review of Artificial Intelligence. 47, 2017
- [14] Sirohi, Neeraj & Rajshree, Mamta & Rajan, Siddhi, "Text Summarization Approaches Using Machine Learning & LSTM," Revista Gestão Inovação e Tecnologias. 11, 2021
- [15] Jatin Pardhi, Tanmay Kharik. "The automated method of Summarization of Audio Data" 22 April 2022, PREPRINT (Version 1) available at ResearchSquare [<https://doi.org/10.21203/rs.3.rs-1565580/v1>]

- [16] Liya B.S, Seenuvasan V, Sathya Prakash K, Siva B, "Audio summarization in real time for podcast, speeches and audio books," *Journal of Survey in Fisheries Sciences (SFS)*, Vol. 10 No 4S, (2023)
- [17] Seth Polsley, Pooja Jhunjhunwala and Ruihong Huang, "CaseSummarizer: A System for Automated Summarization of Legal Texts," *International Conference on Computational Linguistics* 2016
- [18] Pravin Khandare, Sanket Gaikwad, Aditya Kukade, Rohit Panicker and Swaraj Thamke, "AUDIO DATA SUMMARIZATION SYSTEM USING NATURAL LANGUAGE PROCESSING," in *International Research Journal of Engineering and Technology (IRJET)*, Vol. 6, No. 9, 2019
- [19] Samrat Babar, M Tech-Cse, Rit, "Text Summarization:An Overview," *ResearchGate*, (2013)
- [20] Neto Joel, Freitas Alex and Kaestner Celso, "Automatic Text Summarization Using a Machine Learning Approach," *ResearchGate*, 2002
- [21] Shaz Furniturewala, Racchit Jain, Vijay Kumari and Yashvardhan Sharma, "Legal Text Classification and Summarization using Transformers and Joint Text Features," *CEURWS*, Vol. 3159, 2021
- [22] Aries, Abdelkrime, and Walid Khaled Hidouci. "Automatic text summarization: What has been done and what has to be done." *arXiv preprint arXiv:1904.00688* (2019)
- [23] Dalal, Vipul, and Latesh Malik. "Semantic graph based automatic text summarization for hindi documents using particle swarm optimization." In *The Conference of International on Intelligent Systems using Information and Communication Technology*, Springer, Cham, 2017
- [24] Moratanch, N., and S. Chitrakala. "A survey on abstractive text summarization." in *2016 Conference International on Circuit, power and computing technologies (ICCPCT)*. IEEE, 2016
- [25] Grusky, Max, Mor Naaman, and Yoav Artzi. "Newsroom: A dataset of 1.3 million summaries with diverse extractive strategies." *arXiv preprint arXiv:1804.11283* (2018)
- [26] GoyalTanya, Jiacheng Xu1, Junyi Jessy Li and Greg Durrett, "Training dynamics for text summarization models." *arXiv preprint arXiv:2110.08370* (2021)
- [27] Parida S, Motlicek P (2019), "Idiap abstract text summarization system for German text summarization task," *SwissTex*, pp 1–5
- [28] K. S. Umadevi, R. Chopra, N. Singh, L. Aruru and R. J. Kannan, "Text Summarization of Spanish Documents," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp.1793-1797, doi:10.1109/ICACCI.2018.8554839
- [29] Pochhi, Nilay, Kripabandhu Ghosh, and Saptarshi Ghosh, "A Comparative Study of Summarization AlgorithmsApplied to Legal Case Judgments," 2019 .
- [30] Alexander R. Fabbri, Irene Li, Tianwei She, Suyi Li, Dragomir R. Radev, "Multi-news: A large-scale multi-document summarization dataset and abstractive hierarchical model." *arXiv preprint arXiv:1906.01749* (2019).