Summarizing Lecture Videos Using LLMs

Author
Department of Computer Science & Engineering
Sahyadri College of Engineering & Management
Author's E-mail address

ABSTRACT

This paper introduces a novel system for summarizing long-form lecture videos and educational content using transformer-based Large Language Models (LLMs). The system was made through a three-stage methodology, comprising model selection, transcript retrieval, and summarization. We first discuss the importance of selecting an appropriate LLM for the task, considering factors such as model architecture, pre-training data, and computational resources. Next, we detail the process of extracting transcripts from lecture videos or PDFs, emphasizing techniques for accurate and efficient transcription. Finally, we describe the summarization algorithm, which leverages the power of LLMs to distill key information from the transcripts into concise summaries. We evaluate the system's performance through quantitative metrics and qualitative analysis, demonstrating its effectiveness in condensing lengthy educational content into digestible formats. Our approach offers a valuable solution for enhancing learning efficiency and accessibility in online education platforms.

1. INTRODUCTION

In today's digital age, the landscape of education has undergone a profound transformation, propelled by the widespread adoption of digital learning platforms and the rise of remote education. As learners increasingly turn to online resources for their educational needs, the accessibility and digestibility of educational content have emerged as critical factors in facilitating effective learning experiences.

Open online courses represent invaluable sources of knowledge across a myriad of subjects, offering learners a wealth of information at their fingertips. However, the sheer length and complexity of these resources often present significant challenges for learners striving for efficient comprehension. Navigating through hours of video content or dense textual materials can be daunting, leading to information overload and cognitive fatigue. Techniques like using summaries of the content can lead to better comprehension of the material, and more efficient learning as well. [11]

In response to this pressing need, we introduce a novel system designed to streamline the process of summarizing lengthy educational materials, leveraging the transformative capabilities of Large Language Models (LLMs) [10]. These advanced models hold the promise of revolutionizing content summarization by automating the extraction of key insights and distilling them into concise summaries.

Our system represents a departure from traditional methods of content summarization, which often rely on manual techniques or simplistic algorithms. By harnessing the power of state-of-the-art LLMs, we aim to transcend the limitations of conventional approaches and unlock new possibilities for enhancing learning outcomes.

One of the primary sources of educational content in our digital age is YouTube, a platform that hosts a vast array of free courses covering a diverse range of subjects. YouTube offers learners access to high-quality educational content from leading experts and institutions worldwide. However, navigating through lengthy video lectures can be time-consuming and challenging, underscoring the need for effective summarization solutions to aid in content digestion and comprehension. [21]

In addition to the abundance of educational resources available online, the importance of note-taking and revising cannot be overstated in the learning process. Taking effective notes and revisiting them systematically play a pivotal role in reinforcing learning, aiding memory retention, and promoting deeper understanding of the material. By providing learners with concise summaries of educational content, our solution complements these essential learning strategies, enabling students to optimize their study sessions and accelerate their learning progress. [12]

Through empirical evaluation and qualitative analysis, we demonstrate the efficacy and utility of our approach in condensing long-form educational materials into digestible summaries. By offering a scalable and adaptable solution, our system holds the potential to empower learners in their quest for knowledge, revolutionizing the way they engage with and assimilate educational content across diverse subjects and contexts.

2. RELATED WORK

The task of text summarization has been extensively studied in natural language processing (NLP) research, with a focus on various domains such as news articles, scientific papers, and social media posts. Early approaches to automatic text summarization primarily focused on extractive methods, where sentences or phrases from the original text are selected and concatenated to form a summary. Classic techniques such as TF-IDF (Term Frequency-Inverse Document Frequency) have been widely used in extractive summarization [13]. TF-IDF calculates the importance of each word in a document based on its frequency in the document and across a collection of documents, thus identifying key terms for inclusion in the summary.

In the domain of educational content summarization, several approaches have emerged to address the unique challenges posed by lecture videos and academic documents. Traditional methods often rely on heuristic-based algorithms or keyword extraction techniques, which may lack the semantic understanding required for generating coherent and informative summaries.

More recently, with the advent of deep learning and large language models (LLMs), abstractive summarization techniques have gained prominence [22]. Abstractive methods generate summaries by paraphrasing and rephrasing the original text, allowing for more flexibility and creativity in summarization. Models such as the Transformer architecture [14] and its variants, including BERT [15] and GPT [10], have demonstrated remarkable capabilities in abstractive summarization tasks.

Various models have been developed to improve abstractive summarization, each with its strengths and limitations. The Fine-Tuned PEGASUS model, based on the Transformer architecture, has shown proficiency in generating abstractive summaries but is constrained by input size limitations, affecting its processing speed [1]. In a similar vein, Sequence Level Contrastive Learning for Text Summarization shows BART, a model pretrained with a denoising autoencoder, which exhibits commendable performance in abstractive summarization tasks [2].

Exploring alternative approaches, such as pre-trained seq2seq models, offers avenues for abstractive summarization, although challenges exist in capturing extensive contextual dependencies [3]. Abstractive Review Summarization introduces Pointer-Generator Networks, combining extractive and abstractive methods to excel in directing attention to specific words while generating novel content [4].

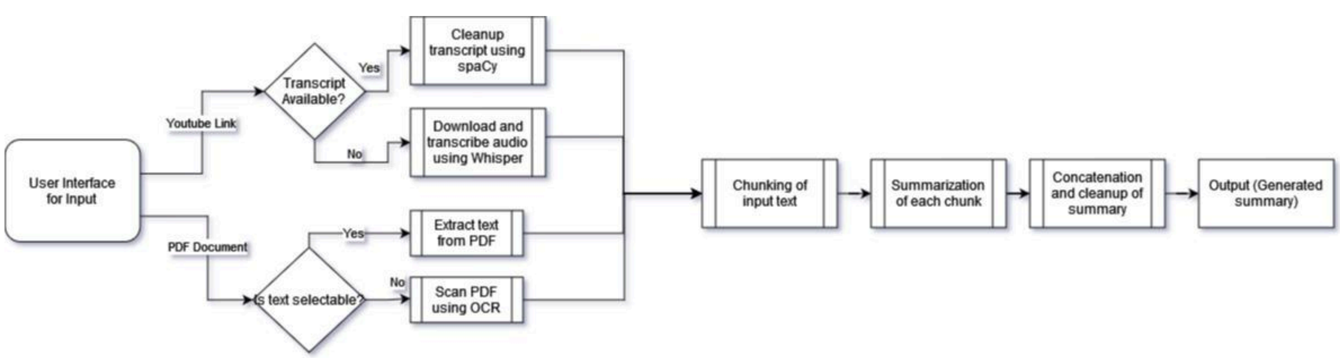
Traditional methods like TF-IDF have been effective in extracting key sentences based on word frequency but are limited in their capacity to generate new or original content [5]. Enhanced TextRank builds upon the TextRank algorithm, aiming to better capture essential information from the input text by employing graph-based techniques [6].

Longformers introduce a linearly scaling attention mechanism for handling long sequences, setting new benchmarks in processing extensive documents [7]. Similarly, Pretrained Encoders utilize a BERT-based document-level encoder for both extractive and abstractive summarization tasks, achieving state-of-the-art results [8]. A Python-based service employing BERT for extractive text summarization, acknowledges areas for improvement, particularly in lecture summarization [9], and further state-of-the-art models were used to improve on abstractive summarization tasks. [23]

Building upon this foundation, our work extends the current state-of-the-art by proposing a comprehensive methodology tailored specifically for summarizing long-form lecture videos and educational content. By integrating model selection, transcript retrieval, and summarization techniques, our approach offers a holistic solution to the challenges of educational content summarization, paving the way for enhanced learning outcomes in online education environments.

3. PROJECT DESCRIPTION

To create our summarization system we followed the following steps. Figure 3.a shows a broad overview of the design of our summarization system.



3.1 Model Selection

In the context of text summarization, the choice of model is pivotal in determining the effectiveness and performance of the summarization system, particularly when dealing with long-form educational content. Recognizing the importance of abstractive summarization over extractive methods, our goal was to develop a system capable of generating concise and informative summaries by capturing the essence of the content, especially prevalent in lecture videos. It was also important to ensure that the model was not overfitted to any particular domain, and general enough for use with any type of content.

For our solution, we opted for a fine-tuned variant of the Longformer Encoder-Decoder (LED) model. The LED model is specifically designed to handle long sequences of text efficiently, making it well-suited for summarizing the content of lengthy lecture videos and PDF documents commonly encountered in educational settings. By incorporating mechanisms such as global attention and sparse attention patterns, LED can process extensive textual inputs while mitigating computational constraints, thus enabling more effective summarization [7]. Given the nuanced nature of educational content and the necessity for abstractive summarization, fine-tuning the LED model was deemed necessary. Pre-trained models like BERT and T5 were initially tested, but their outputs did not meet our expectations in terms of quality and coherence and neither were they efficient enough. Hence, the choice of a LLM capable of being fine-tuned became imperative to ensure the production of satisfactory summaries.

The variant of the LED model, specifically LED-Large, we selected was fine-tuned on a dataset called BookSum [18] which is a corpus of book summaries. This fine-tuning strategy aimed to capitalize on LED's inherent capabilities while tailoring its parameters to be general enough while providing concise summaries. The choice of book summaries as the fine-tuning data aligns with the structural and content similarities between books and educational materials, thereby facilitating the model's ability to extract salient information from educational content.

Furthermore, the availability of the LED model through the HuggingFace library provided ease of integration and accessibility for developers and researchers. The open-source nature of the LED model fostered collaboration and experimentation, enabling continuous improvement and innovation in educational content summarization. By selecting a fine-tuned variant of the LED model, our aim was to harness its capabilities in efficiently summarizing long-form educational materials, thereby enhancing the accessibility and digestibility of educational content for learners and educators alike. Through empirical evaluation and validation, we sought to demonstrate the effectiveness of this model in generating high-quality summaries that meet the diverse needs of educational stakeholders.

3.2 Transcript Retrieval

Transcripts serve as the foundational input for our summarization system, enabling the extraction of textual content from lecture videos and PDF documents. Our approach to transcript retrieval employs a two-step methodology tailored to the diverse formats and sources of educational materials.

1. YouTube Video Transcripts:
    a. Given a YouTube video link through the interface, it first attempts to retrieve the transcript directly from the YouTube API, which can either be auto-generated on the YouTube backend, or added by users.
    b. If the transcript is available through the YouTube API, we proceed to clean it up using spaCy [16], a natural language processing library, to remove noise, correct errors, and enhance readability.
    c. This direct retrieval approach ensures efficiency and accuracy in obtaining transcripts from lecture videos hosted on YouTube, streamlining the subsequent summarization process.
2. Video-to-Text Conversion:
    a. In cases where no transcript is available through the YouTube API, our system initiates a secondary process of video-to-text conversion.

   b. The system downloads the video associated with the provided YouTube link and extracts the audio component.
   c. The audio is then processed using the Whisper model [17], a speech recognition model trained specifically for generating accurate transcriptions from spoken language.
   d. The resulting transcript undergoes further cleaning and normalization to ensure consistency and coherence, preparing it for subsequent summarization steps.
3. PDF Text Extraction:
   a. We also allow the user to upload PDF documents to be summarized. For these, our interface employs techniques for text extraction to retrieve the content embedded within the document.
   b. If the PDF contains searchable text, we perform necessary cleanup operations, such as removing extraneous characters, formatting inconsistencies, and irrelevant metadata.
   c. In cases where the PDF consists of scanned images or non-searchable text, we employ Optical Character Recognition (OCR) technology to extract text from the images, converting them into machine-readable format. For this we use the Python package pytesseract.
   d. The extracted text is then subjected to preprocessing steps to enhance its quality and coherence, facilitating accurate summarization.

By adopting a two-step approach to transcript retrieval, our system accommodates the diverse formats and sources of educational materials, ensuring robustness and reliability in obtaining the requisite textual content for summarization. This comprehensive methodology lays the groundwork for generating informative and concise summaries that cater to the needs of learners and educators across various educational contexts.

3.3 Summarization Methodology

Our summarization methodology leverages the fine-tuned Longformer Encoder-Decoder (LED) model, tailored to distill key insights from lengthy educational content into concise summaries. The methodology comprises several key steps aimed at accommodating varying input sizes, generating informative summaries.

In order to summarize large documents that exceed the limits of most LLMs that can be run locally, we tried different techniques, including chunking the input, adding summaries of the previous chunk into the next chunk to be summarized, and using extractive summarization techniques before summarizing using the model. Finally, we decided to use an adaptive chunking technique to keep the summaries consistent in length despite the size of the input.

1. LED Model Selection and Fine-tuning:
    a. We selected the Longformer Encoder-Decoder (LED) model for its ability to handle long sequences of text efficiently, making it well-suited for summarizing lengthy educational materials.
    b. The LED model was fine-tuned on a corpus of book summaries to adapt its parameters to summarizing general content without it being too overfitted to any one domain.
2. Adaptive Chunking for Input Handling:
    a. To address the challenge of varying input sizes, we implemented a dynamic chunking mechanism.
    b. Depending on the size of the input, the content is segmented into chunks of appropriate lengths, ensuring that the resulting output summaries remain consistent in size and

quality. The chunks also overlap each other by some margin to ensure not too much context is lost.

    c. This adaptive chunking strategy optimizes the summarization process for inputs of different lengths, maintaining coherence and readability across diverse educational materials.

3. Summarization Process:
    a. The chunked input is fed into the LED model, which generates summaries for each segment independently.
    b. Each summary encapsulates the salient information extracted from the corresponding chunk, capturing the essence of the content while preserving its context and relevance.
    c. The summaries from all chunks are then concatenated to produce the final summarized output, providing a comprehensive overview of the entire input document.

Our summarization methodology offers a robust and comprehensive approach to condensing long-form educational content into digestible summaries. We also implemented a cleanup function after the generation of the summary as in some instances there may be an unfinished sentence at the end, or artifacts in the summary. These are removed before the summary is output to the user.

3.4 Summarization Interface

Following the development of our comprehensive summarization methodology, the next crucial step was to build a user-friendly frontend interface and a robust backend system to operationalize the system. To achieve this, we utilized modern web development technologies to create a seamless user experience while ensuring flexibility in deployment options.

For the frontend, we opted for React, a popular JavaScript library for building interactive user interfaces.

React's component-based architecture enabled us to modularize the interface and maintain a clean and organized codebase. Leveraging React's rich ecosystem of libraries and components, we designed an intuitive and responsive user interface that allows users to interact with the summarization system effortlessly. Users can input the link or even upload a PDF document and get the summary of the content.

On the backend, we implemented a RESTful API using Flask, a lightweight and extensible web framework for Python. Flask's simplicity and flexibility made it an ideal choice for building the backend infrastructure of our summarization system. We designed the API endpoints to handle incoming requests for summarization tasks, orchestrate the processing pipeline, and deliver the generated summaries back to the user.

In summary, by building the frontend using React and backend using Flask, we have created a versatile summarization solution that can even be run locally on a laptop or deployed onto cloud infrastructure and run there as well. This flexible deployment architecture ensures accessibility and usability for users across diverse environments, empowering them to leverage the summarization system according to their specific needs and preferences.

4. EVALUATION

In this section, we present the evaluation of our chosen summarization model based on some standardized datasets for evaluating summarization performance [19].

We compute Recall-Oriented Understudy for Gisting Evaluation (ROUGE) scores to measure the overlap between the generated summaries and reference summaries. The dataset used was the CNN/DailyMail news summary dataset [20].

ROUGE-1: 32.877

ROUGE-2: 13.371

ROUGE-L: 20.436

We then performed qualitative analysis on the outputs generated by the model. We assess the comprehensiveness of the generated summaries by comparing them to the original educational content and reference summaries. Further, we evaluated the coherence and readability of the summaries through manual inspection, considering factors such as sentence fluency, logical flow, and grammatical correctness.

By evaluating our summarization tool using a combination of quantitative metrics, qualitative analysis, and comparison with baseline models, we aim to assess its performance comprehensively and demonstrate its efficacy in summarizing long-form educational content effectively.

5. RESULTS AND DISCUSSION

While evaluating summarization techniques using scores like ROUGE can give us some idea of the effectiveness of those techniques, ultimately we have to resort to human evaluation to determine whether the produced summary is fluent and understandable. Since our goal was to provide summaries that helped in comprehending and recalling the important points in a given lecture video, we fine-tuned the output to be conducive to that.

The output was seen to be satisfactory for a large variety of topics, and with enough length and clarity to facilitate our goals as stated above. However, there were instances where the output would get truncated before the completion of a sentence, and instances where new or rarely used words would be misspelled in the final summary. While these problems are to be expected from an LLM running locally that has such a general scope, these are some areas where the model can be improved.

5.1 Performance

We evaluated the performance of our summarization solution on a sample set of 10 different 30-45 minute long lecture videos to assess its inference time across different hardware configurations. Specifically, we measured the inference time on a laptop CPU, a laptop GPU, and a cloud provider with Nvidia T4 GPU. The impact of the Whisper transcribing model should also be taken into account for the rare instances where the input link doesn't have any transcripts available. In those cases, the running time is usually doubled.

On average, the inference time for summarizing a single lecture video on a laptop CPU was on average 493.4 seconds. This duration reflects the computational overhead associated with processing extensive textual inputs and running inference on a CPU-bound environment.

In contrast, utilizing a laptop GPU yielded a significant reduction in inference time, with an average processing time of approximately 103.3 seconds per lecture video. The parallel processing capabilities of GPU hardware accelerated the summarization process, resulting in faster turnaround times compared to CPU-based execution.

Furthermore, leveraging cloud infrastructure with Nvidia T4 GPU provided a substantial improvement in performance, with an average inference time of approximately 16.3 seconds on average. The powerful compute capabilities of Nvidia T4 GPU, coupled with optimized cloud infrastructure, enabled rapid and efficient summarization of educational content at scale.

Hardware configuration	Average Inference Time (in seconds)
Laptop (CPU)	493.4
Laptop (GPU)	103.3
Cloud (Nvidia T4 Runtime)	16.3

Table 5.a

These results in Table 5.a highlight the impact of hardware configuration on the performance of our summarization system, with GPU acceleration and cloud-based deployment offering significant advantages in terms of speed and scalability. By harnessing the computational resources of modern hardware and cloud platforms, our system can efficiently process lengthy lecture videos and enhance the accessibility and digestibility of educational content for learners and educators alike. While it may be too slow to deploy on a laptop in CPU mode (without a GPU), it is reasonably fast to be run on a GPU-enabled laptop (with Nvidia CUDA) and even faster when deployed on cloud infrastructure accelerated with a GPU like the Nvidia T4.

5.2 Future Development

Incorporating multimodal capabilities into our summarization system represents an exciting avenue for future development. By integrating audio and visual cues alongside textual content, we can enrich the summarization process and provide more comprehensive insights to learners. For example, leveraging video analysis techniques to identify key visual elements such as slides, diagrams, and gestures can enhance the context-awareness of the summarization process. Similarly, incorporating audio processing capabilities to extract speech patterns, intonations, and emphasis can offer additional context for generating more informative summaries. By embracing multimodal approaches, our solution can cater to diverse learning preferences and provide a more immersive and engaging learning experience.

Implementing our summarization solution in a real-time system, where it listens to a lecture as it happens and provides the summary immediately afterward, holds immense potential for enhancing learning experiences. By leveraging speech recognition technology and live transcription capabilities, it can dynamically process spoken content in real-time and generate summaries on the fly. This real-time summarization feature can facilitate active engagement during lectures, allowing learners to quickly grasp key concepts and reinforce their understanding as the lecture unfolds. Additionally, integrating feedback mechanisms to adjust summarization parameters based on user preferences and comprehension levels can further personalize the learning experience and optimize knowledge acquisition in real-time.

Augmenting it with a Question/Answer (Q&A) generation model can significantly enhance learning outcomes by promoting active engagement and knowledge retention. By generating structured queries based on the summarized content, learners can deepen their understanding through self-assessment and inquiry-based learning. Additionally, providing contextually relevant answers to generated questions enables learners to clarify concepts, reinforce learning, and bridge knowledge gaps effectively. Furthermore, incorporating adaptive learning mechanisms that tailor the Q&A generation process to individual learning styles and proficiency levels can further optimize the learning experience and facilitate personalized knowledge acquisition.

In summary, embracing multimodal approaches, real-time summarization, and Q&A generation capabilities represents promising avenues for further enhancing the functionality and effectiveness of our summarization system. Adding support for more languages other than English is also something to be looked into. By continuously innovating and exploring new methodologies, we can advance the field of educational content summarization and empower learners to achieve their learning goals more efficiently and effectively.

6. CONCLUSION

In conclusion, the landscape of text summarization has seen significant advancements driven by innovative models and approaches using the new advancements like LLMs. It is possible to generate summaries of large texts accurately and abstractly even on a simple modern laptop by leveraging these technologies. The summaries are concise enough to be used for educational purposes.

Traditional methods like TF-IDF have their place but are limited in their ability to generate original content. However, as demonstrated by our use of the Longformer based LED model, fine-tuned on book summaries, to take in lecture transcripts as input and produce concise and fluent summaries, abstractive summarization tasks can yield remarkable results. Yet, there is room for improvement, as evidenced by the acknowledgment of areas for enhancement in the output of the model.

Moving forward, the future of text summarization lies in embracing multimodal approaches, real-time processing capabilities, and integrating Question/Answer generation models to foster deeper engagement and knowledge retention. By continuously innovating and refining methodologies, we can unlock new possibilities for summarizing diverse types of content and empower learners and researchers to navigate the vast sea of information more efficiently and effectively.

7. REFERENCES

[1] Alsuhaibani, M. 2023. ”Fine-Tuned PEGASUS: Exploring the Performance of the Transformer-Based Model on a Diverse Text Summarization Dataset.” In Proceedings of the 9th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS’23), Brunel University, London, United Kingdom (pp. 5870–5875). Paper No. CIST 117. DOI: 10.11159/cist23.117

[2] Xu, S., Zhang, X., Wu, Y., Wei, F. 2022. ”Sequence Level Contrastive Learning for Text Summarization.” In The Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22). IIIS, Tsinghua University, Beijing, China; Microsoft Research Asia, Beijing, China; Shanghai Qi Zhi Institute, Shanghai, China.

[3] Xu, L., Zhang, Y., Hong, L., Cai, Y., Sung, S. 2021. ”ChicHealth @ MEDIQA 2021: Exploring the

limits of pre-trained seq2seq models for medical summarization." In Proceedings of the BioNLP 2021 workshop (pp. 263–267). June 11, 2021. Association for Computational Linguistics.

[4] Shobana, J., Murali, M. 2021. "Abstractive Review Summarization based on Improved Attention Mechanism with Pointer Generator Network Model." Webology, 18(1), WEB18028. ISSN: 1735-188X. DOI: 10.14704/WEB/V18I1/WEB18028

[5] Gabriela Nathania H., Ryan Siautamaa., Amadea Claire I. A., Derwin Suhartonoa. 2020. "Extractive Hotel Review Summarization based on TF/IDF and Adjective-Noun Pairing by Considering Annual Sentiment Trends." In Proceedings of the 5th International Conference on Computer Science and Computational Intelligence 2020. Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480.

[6] Yulianti, E., Pangestu, N., Jiwanggi, M. A. 2023. "Enhanced TextRank using weighted word embedding for text summarization." International Journal of Electrical and Computer Engineering (IJECE), 13(5), 5472–5482. ISSN: 2088-8708. DOI: 10.11591/ijece.v13i5.pp5472-5482

[7] Beltagy, Iz, Matthew E. Peters, and Arman Cohan. 2020. "Longformer: The Long-Document Transformer." Allen Institute for Artificial Intelligence, Seattle, WA, USA, beltagy,matthewp,armanc@allenai.org.

[8] Liu, Yang, and Mirella Lapata. 2019. "Text Summarization with Pretrained Encoders." Institute for Language, Cognition and Computation, School of Informatics, University of Edinburgh, yang.liu2@ed.ac.uk, mlap@inf.ed.ac.uk.

[9] Miller, Derek. 2019. "Leveraging BERT for Extractive Text Summarization on Lectures." Georgia Institute of Technology, Atlanta, Georgia, dmiller303@gatech.edu.

[10] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I., 2019. Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), p.9.

[11] McNamara, D.S., Kintsch, E., Songer, N.B. and Kintsch, W., 1996. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and instruction*, *14*(1), pp.1-43.

[12] Donoghue, G.M. and Hattie, J.A., 2021, March. A meta-analysis of ten learning techniques. In *Frontiers in Education* (Vol. 6, p. 581216). Frontiers.

[13] Salton, G. and Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Information processing & management*, *24*(5), pp.513-523.

[14] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, *30*.

[15] Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

[16] Honnibal, M. & Montani, I., 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

[17] Radford, A., Kim, J.W., Xu, T., Brockman, G., McLeavey, C. and Sutskever, I., 2023, July. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning* (pp. 28492-28518). PMLR.

[18] Hugging Face. 2021. BookSum dataset. Retrieved from https://huggingface.co/datasets/kmfoda/booksum

[19] Hugging Face. 2023. pszemraj/led-large-book-summary. Retrieved from https://huggingface.co/pszemraj/led-large-book-summary

[20] Hugging Face. 2021. CNN_DailyMail dataset. Retrieved from https://huggingface.co/datasets/cnn_dailymail

[21] McNamara, D. S. and Kintsch, W. (1996) ‘Learning from texts: Effects of prior knowledge and text coherence’, *Discourse Processes*, 22(3), pp. 247–288. doi: 10.1080/01638539609544975.

[22] Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., Affandy, A., & Setiadi, D. R. I. M. (2022). Review of automatic text summarization techniques & methods. Journal of King Saud University - Computer and Information Sciences, 34(4), 1029-1046. https://doi.org/10.1016/j.jksuci.2020.05.006.

[23] Liu, Y., Liu, P., Radev, D. and Neubig, G., 2022. BRIO: Bringing order to abstractive summarization. *arXiv preprint arXiv:2203.16804*.