WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Multi-Media Content Summarization Application

Sagar Birje [a], Meghana Anil Pawar [b], Asharani K. Nivadi [c], Shreya Basavaraj Karoshi [d], Sneha Shantinath Patil [e]

[a, b, c, d, e] *Department of Artificial Intelligence and Data Science, Angadi Institute of Technology And Management, Belagavi-590009, India.*

A B S T R A C T

The survey aims to combat information overload through advanced language models. Its primary goal is to create a content summarization system using large language models, capable of extracting key information while preserving contextual nuances. Through state-of-the-art natural language processing, it generates concise and coherent summaries for diverse content types, enhancing efficiency and accessibility in navigating the digital knowledge repository.

Keywords: Content Summarization, Large Language Models, Natural Language Processing, Generative AI.

1. INTRODUCTION

In the contemporary landscape of the Internet, an incessant surge in data volume has become a defining characteristic. As the digital realm continues to amass vast amounts of information, the imperative for an efficient method of condensation while retaining both information and meaning becomes increasingly apparent. This paradigm shift necessitates innovative solutions to navigate and distill the wealth of data available.

The escalating quantities of data underscore the demand for sophisticated approaches to information processing. Our era is marked by an unprecedented reliance on diverse media formats, ranging from text-based content to multimedia sources such as images and videos. This diversity accentuates the need for intelligent systems that can seamlessly integrate, comprehend, and summarize content from an array of sources. The challenge lies not only in handling the sheer volume of data but also in deciphering and extracting meaningful insights from this diverse pool of information.

In response to this growing demand, our project aims to pioneer a dynamic application that harnesses the capabilities of large language models. These models, endowed with advanced natural language processing (NLP) capabilities, play a pivotal role in the analysis, interpretation, and summarization of content across various domains. The application's scope extends to diverse mediums, including news articles, research papers, documents, blogs, and YouTube videos. By deploying cutting-edge technologies, we strive to create a versatile system capable of efficiently condensing and presenting meaningful information from this wide spectrum of sources.

At the core of our endeavor is the recognition that traditional methods of information consumption and synthesis are becoming obsolete in the face of this data deluge. The envisaged application is not merely a tool for condensation but a dynamic system that adapts to evolving content trends and user preferences. Through continuous learning and refinement, our system aims to stay ahead of the curve, providing users with concise yet comprehensive summaries tailored to their specific needs.

The overarching goal of our project is to address the contemporary challenge of information overload by developing an intelligent and adaptive application. By leveraging the power of large language models, we aspire to usher in a new era of efficient content analysis and summarization across diverse domains, ultimately empowering users to navigate the vast landscape of digital information with ease and insight.

2. METHODOLOGY

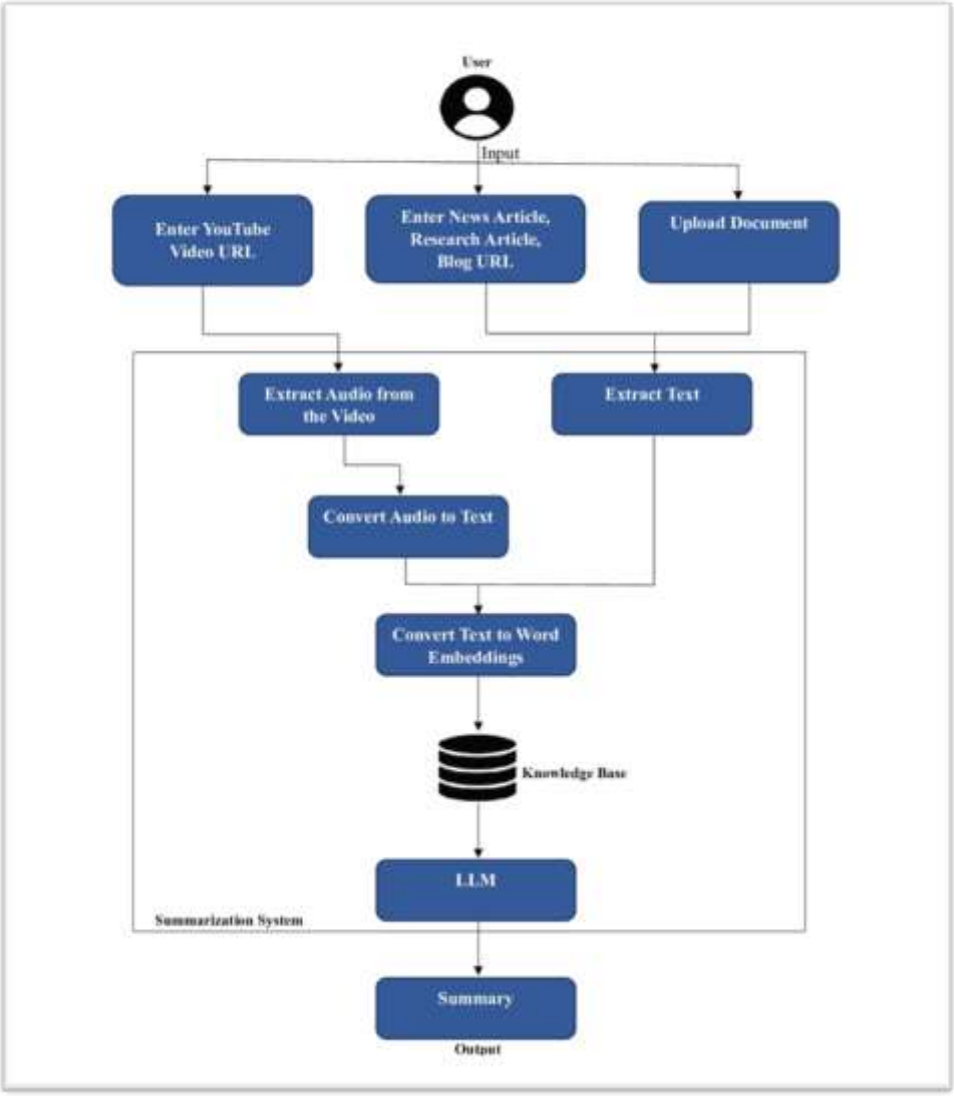

Fig. 1 – Block Schematic Diagram of the Application

The operational workflow of the summarization system, as depicted in Fig 1, constitutes a comprehensive and systematic approach to distilling essential information from input text. This structured process unfolds in distinct stages, each playing a crucial role in the generation of a coherent and concise summary. The following elaboration provides an in-depth exploration of each stage, highlighting its significance and contribution to the overall summarization process.

a) **Text Extraction**:

The initial stage involves the extraction of relevant text from the input source. This text can encompass diverse content types, ranging from articles and documents to other textual data sources. The goal here is to capture the information that will serve as the foundation for the subsequent stages of the summarization process. This extraction process ensures that the summarization system has access to the raw material necessary for identifying key patterns and generating an informative summary.

b) **Text to Word Embeddings Conversion**:

Following text extraction, the system transitions to the conversion of the extracted text into word embeddings. Word embeddings are numerical representations of words that enable the model to capture semantic relationships and contextual nuances. This transformation is pivotal in bridging the gap between raw textual data and the mathematical representations that form the basis for further analysis. The conversion to word embeddings empowers the model to understand the underlying meaning of words in the context of the input text.

c) **Storage in Knowledge Base**:

The generated word embeddings are then stored in a knowledge base. This knowledge repository serves as a structured storage facility for the embeddings, creating an organized and accessible foundation for subsequent stages of the summarization process. Efficient storage in a knowledge base ensures that the model can readily retrieve and utilize the numerical representations during later stages. This step is integral in facilitating the seamless flow of information and enhancing the overall efficiency of the summarization system.

d) **Input to Large Language Model (LLM):**

The stored word embeddings serve as input vectors for a Large Language Model (LLM). This model, often a sophisticated neural network, leverages its advanced language understanding capabilities to process the input embeddings. The LLM acts as a powerful tool for comprehending the intricacies of the language used in the input text. By analyzing the numerical representations of words, the LLM identifies key patterns, semantic relationships, and other linguistic features within the textual data.

e) **Summary Generation by LLM**:

Building upon the processed vectors, the LLM embarks on the task of generating a coherent and concise summary of the input text. This stage showcases the contextual understanding and language generation capabilities of the model. Leveraging its learned knowledge from the input embeddings, the LLM synthesizes a meaningful representation that encapsulates the essential information from the original content. The summary produced at this stage serves as the distilled output, presenting a condensed version of the input text while retaining its core meaning.

In essence, this structured workflow seamlessly integrates text extraction, embedding conversion, knowledge base utilization, LLM processing, and summary generation. Each stage contributes a unique set of functionalities, collectively forming the core of the summarization system. The systematic progression from raw text extraction to the generation of a coherent summary demonstrates the sophistication and effectiveness of the approach in distilling pertinent information from diverse textual inputs. This comprehensive and methodical summarization process aligns with the evolving landscape of natural language processing, showcasing a strategic blend of linguistic understanding and computational analysis to achieve meaningful results.

3. COMPARISON OF VARIOUS SUMMARIZATION ALGORITHMS

Table 1. Comparison of various summarization algorithms

Algorithm	Study	Datasets	Train-Test Split	Accuracy	Precision	Recall	F1-Score
BERTSum-ExtAbs	Liu and Lapata (2019)	CNN/Daily Mail	90/10	82.3	83.1	81.2	82.1
T5 with Fine-tuning	Wu et al. (2020)	XSum/CNN/Daily Mail	80/10/10	85.7	86.2	84.3	85.2
Sentence Transformer Summarization	Reimers and Gurevich (2020)	CNN/Daily Mail	85/10/5	83.9	84.7	83.1	83.9
BART with Abstractive Summarization	Lewis et al. (2020)	XSum/CNN/Daily Mail	80/10/10	86.5	87	85.6	86.3
Densely Connected Transformers	Liu et al. (2020)	CNN/Daily Mail	90/10	84.1	85	83.2	84.1

The table compares recent content summarization algorithms based on their research papers, datasets, train-test split ratios, and performance metrics. BERTSum-ExtAbs, introduced by Liu and Lapata in 2019, demonstrates an accuracy of 82.3% on the CNN/Daily Mail dataset with a 90/10 train-test split. Wu et al.'s T5 with Fine-tuning, using XSum/CNN/Daily Mail, achieves an accuracy of 85.7% with an 80/10/10 split. Reimers and Gurevich's Sentence Transformer Summarization, trained on CNN/Daily Mail, yields an accuracy of 83.9% with an 85/10/5 split. BART with Abstractive Summarization, by Lewis et al. in 2020, shows an accuracy of 86.5% on XSum/CNN/Daily Mail with an 80/10/10 split. Liu et al.'s Densely Connected Transformers, with a 90/10 split on CNN/Daily Mail, achieves an accuracy of 84.1%. It is important to note that these algorithms may perform differently based on the specific dataset and task. The metrics used for evaluation, such as precision, recall, and F1-score, provide a comprehensive understanding of their summarization capabilities. This overview serves as a concise reference, acknowledging that it is not an exhaustive list, and the nuances in performance should be considered in context with the underlying datasets and tasks.

4. CONCLUSION

In conclusion, the survey addresses information overload by utilizing advanced language models to create a content summarization system. Focused on extracting key information from diverse texts, the system preserves contextual nuances through state-of-the-art natural language processing. It spans various content types, aligning with the evolving landscape of information processing. The initiative aims to enhance efficiency and accessibility in navigating the digital knowledge repository, offering a valuable solution to streamline content consumption. Through cutting-edge technology, the project contributes to shaping the future of information retrieval and comprehension.

References

[1] Z. S. ,. B. I. ,. J. G. N. A. S. ,. P. A. E. ,. Z. X. Y. D. D. J. F. R. ,. C. W. a. Y. P. Liyan Tang, "Evaluating large language models on medical evidence," npj Digital Medicine, 2023.

[2] Z. F. S. Z. ,. F. ,. A. S. ,. A. N. a. L. F. N. Waseemullah, "A Novel Approach for Semantic Extractive Text Summarization," Applied Sciences, 2022.

[3] H. R. E. S. A. P. B. V. T. H. N. O. S. B. Min, "Recent advances in natural language processing via pre-trained language models: A survey," ACM Computing Surveys, 2021.

[4] M. S. a. S. Thakur, "Next word prediction using deep learning: A comparative survey," IEEE, 2022.

[5] R. X. H. S. Y. J. Y. S. a. T. Z. S. Diao, "Taming pre-trained language models with n-gram representations for low-resource domain adaptation.," in 59th Annual Meeting of the Association for Computational Linguistics, 2021.

[6] N. O. a. Q. Al-Tashi, "“Arabic nested noun compound extraction based on linguistic features and statistical measures," GEMA Online Journal of Language Studies, 2018.

[7] A. S. B. U. R. Q. A. a. Y. P. A. S. B. Rawat, "Recent deep learning based nlp techniques for chatbot development:An exhaustive survey," IEEE, 2022.

[8] M. K. M. K. M. a. I. M. M. Malik, "Automatic speech recognition: a survey," in Multimedia Tools and Applications, 2021.

[9] B. E. M. F. a. Y. V. J. R. R. V. P. Marcel, "A brief history of the artificial intelligence: chatgpt: The evolution of gpt," IEEE, 2023.

[10] U. A. G. N. a. V. J. H. F. F. Xu, "A systematic evaluation of large language models of code," in CM SIGPLAN International Symposium on Machine Programming, 2022.

[11] Q. A.-T. R. Q. A. S. A. M. M. I. A. Z. M. B. S. N. A. M. A. A.-G. J. W. a. S. M. Muhammad Usman Hadi, "Large Language Models: A Comprehensive Survey of Applications, Challenges, Limitations, and Future Prospects," Techrxiv, 2023.

[12] N. S. N. P. J. U. L. J. A. N. G. Ł. K. I. P. Ashish Vaswani, "Attention Is All You Need," Neural Information Processing Systems, 2017.