Vidyavaridhi's College of Engineering & Technology

Department of Information Technology Engineering

Certificate

This is to certify that the following students

SIDDHI VIJAY CHAVAN (07)
PARAG SANJAY NANDOSKAR (29)
VIPUL PREMSING PATIL (43)
HARSH VIKRAM SINGH (59)

have submitted mini project report entitled

TEXT SUMMARIZATION APPLICATION

as a part of their mini-project in partial fulfillment of ITM401: Mini Project – 1B of Semester IV of Bachelor of Engineering in Information Technology during academic year 2023-2024.

Supervisor : Mrs. Vaishali Shirsath ( )

Internal Examiner :_____ ( )

External Examiner :_____ ( )

TEXT SUMMARIZATION APPLICATION

Parag Sanjay Nandoskar
Department of Information Technology
Vidyavardhini's College of Engineering & Technology Vasai, India
parag.223424106@vcet.edu.in

Vipul Premsing Patil
Department of Information Technology
Vidyavardhini's College of Engineering & Technology Vasai, India
vipul.223564104@vcet.edu.in

Siddhi Vijay Chavan
Department of Information Technology
Vidyavardhini's College of Engineering & Technology Vasai, India
siddhi.223204201@vcet.edu.in

Harsh Vikram Singh
Department of Information Technology
Vidyavardhini's College of Engineering & Technology Vasai, India
harsh.223724101@vcet.edu.in

Abstract:
Text summarization holds immense potential in revolutionizing how we consume and utilize information. By condensing lengthy texts into concise summaries, it serves as a powerful tool for enhancing productivity and comprehension across various domains. One unique application lies in personalized education, where complex topics can be distilled into digestible snippets tailored to individual learning styles, fostering deeper understanding and retention.

Moreover, in the realm of research and development, text summarization accelerates knowledge dissemination by extracting key insights from voluminous scientific literature, enabling researchers to stay abreast of the latest advancements efficiently. Additionally, in the rapidly evolving landscape of news and media, summarization algorithms offer readers a quick overview of diverse perspectives on a given topic, promoting informed decision-making and critical thinking.

I. INTRODUCTION

A. This web application aims to provide users with a seamless and intuitive platform for generating summaries from inputted text or URLs. Leveraging advanced natural language processing (NLP) techniques, the application will offer both extractive and abstractive summarization methods to accommodate different user needs and text complexities.

B. The significance of this text summarization web application lies in its ability to empower users across various sectors, including academia, journalism, business, and research, to efficiently distil key insights from voluminous textual content. By enhancing content digestibility and promoting informed decision making, this application seeks to contribute to a more accessible and productive digital environment.

II. PROBLEM STATEMENT

The problem at hand is the need for an accessible and effective solution to summarize textual content, catering to diverse user needs and text complexities. Existing summarization tools often lack user-friendliness, customization options, or fail to produce coherent summaries that preserve essential context and meaning. Additionally, many tools are not readily available as web applications, limiting accessibility and usability across different devices and platforms. Therefore, the aim is to develop a text summarization web application that addresses these shortcomings and provides users with a seamless and intuitive platform for generating concise summaries from inputted text or URLs. This application should incorporate advanced natural language processing (NLP) techniques to offer both extractive and abstractive summarization methods, ensuring flexibility and accuracy in summarization results

III. DETAILED DESCRIPTION

A. Application: It is an application where we have used core concepts of Python and NLTK.

B. Requirement: Standard personal computer or laptop with sufficient processing power and memory for development and testing.

IV. FEATURES OF THE PROJECT

1. Enhancing accessibility: Develop a userfriendly web interface that allows users to input text or URLs and obtain summarized content with ease.
2. Improving summarization quality: Implement robust NLP algorithms to produce accurate and coherent summaries that preserve essential information and context.
3. Customization options: Provide users with the ability to customize summary length language preferences and summarization parameters to suit their specific needs.

4. Seamless integration: Ensure compatibility and seamless integration across different devices and platforms, enabling users to access the application from anywhere.

5. Security and privacy: Implement stringent security measures to safeguard user data and ensure confidentiality throughout the summarization process.

The amount of textual material on the web and other libraries is growing tremendously daily. Information utilization has become an expensive and time-consuming activity since data expands in a large quantity at a time and includes irrelevant content or noise. Text summarization is a method used to summarize the data. A manual text summarization process is undoubtedly an effective way to preserve the meaning of the text; however, this is a time-consuming activity. Another approach is to utilize the automatic text summarization. In ATS, different practical algorithms can be programmed into computers to produce summaries of information. Thus, text summarization creates a brief and accurate overview of a lengthy text document by concentrating on the essential parts that provide valuable details by maintaining the overall context. In natural language processing (NLP), automatic text summarization is a method of evaluating, comprehending, and extracting information from human language.

IV. LITERATURE SURVEY

Text summarization, a vital aspect of natural language processing, aims to condense large volumes of text into shorter, coherent representations while preserving essential information.

Scholars such as Nenkova and McKeown (2011) have extensively explored extractive and abstractive summarization techniques. Extractive methods select important sentences or phrases from the original text, while abstractive methods generate summaries by paraphrasing and rephrasing the content. These approaches often utilize machine learning algorithms, including neural networks, to enhance summarization quality.

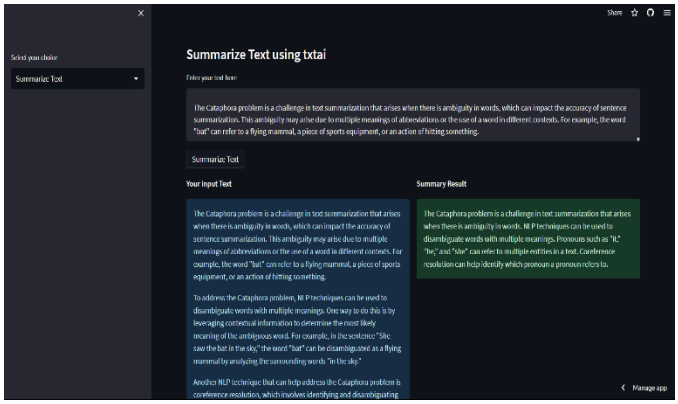
V. MOTIVE AND APPLICATION

This study aims to provide an overview of current research in NLPs and, precisely, ATS to accelerate knowledge about it. In addition, it allows the creation of new tools, methods, datasets, and resources that meet the needs of the research and industrial sectors. The advancement of NLPs made automatic text summarization usable for a regular text document summary and sentiment analysis. Moreover, ATS promotes a versatile approach to research various fields such as machine learning, natural language, cognitive science, and psychology. With multiple sources of information, ATS discusses cutting-edge work and future directions in this exciting area. These collective findings are the motivation behind this research. An essential part of research on ATS is the application, which is presented in the following section

VI. CONCLUSION

The text summarization web application represents a valuable resource for individuals and organizations seeking to streamline information consumption, enhance decision-making, and increase productivity in the digital landscape. As technology continues to evolve, we remain committed to further enhancing the application's capabilities and exploring new avenues for innovation in text summarization. With continued development and refinement, this application has the potential to make a significant impact on how we consume and utilize textual information in the future.

VII. RESULT



VIII. REFERENCES

- P. Shah and N. P. Desai, "A survey of automatic text summarization techniques for Indian and foreign languages," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 2016, pp. 4598- 4601, doi: 10.1109/ICEEOT.2016.7755587
- P. Raundale and H. Shekhar, "Analytical study of Text Summarization Techniques," 2021 Asian Conference on Innovation in Technology (ASIANCON), PUNE, India, 2021, pp. 1-4, doi: 10.1109/ASIANCON51346. 2021.9544804
- J. N. Madhuri and R. Ganesh Kumar, "Extractive Text Summarization Using Sentence Ranking," 2019 International Conference on Data Science and Communication (IconDSC), Bangalore, India, 2019, pp. 1-3, doi: 10.1109/IconDSC.2019.8817040
- K. Lanyo and A. Wausi, "A Comparative Study of Supervised and Unsupervised Classifiers Utilizing Extractive Text Summarization Techniques to Support Automated Customer Query QuestionAnswering," 2018 5th International Conference on Soft Computing & Machine Intelligence (ISCMI), Nairobi, Kenya, 2018, pp. 88-92, doi: 10.1109/ISCMI.2018.8703237.
- R. Barzilay and K. R. McKeown, “Sentence fusion for multidocument news summarization,” Comput. Linguist., vol. 31, no. 3, pp. 297–328, Sep. 2005. [Online]. Available: http://dx.doi.org/10.1162/089120105774321091
- S. H. Finley and S. M. Harabagiu, “Generating single and multidocument summaries with gistexter,” in In U. Hahn & D. Harman (Eds.), Proceedings of the workshop on automatic summarization, 2002, pp.30–38