

GenAI-Powered, Real Time News Aggregator Using NLP

Aryan Shukla, Harsh Kumar Singh, Anmol Yadav, Sarika Singh

Department of Computer Science and Engineering

Shri Ramswaroop Memorial College of Engineering and Management Lucknow, Uttar Pradesh, India

aryan4shukla2003@gmail.com, 4421hars@gmail.com,

yadavanmol81@gmail.com, sarikasingh2494@gmail.com

Abstract: *The expansion of digital communication technologies and online news platforms has resulted in a massive growth in the availability of information for readers. Every day, a large number of news articles are published across various categories such as politics, technology, business, sports, and entertainment. Although this widespread availability of information provides users with convenient access to current events and knowledge, it also leads to the issue of information overload. Readers often struggle to go through long articles and identify the most relevant details within a limited amount of time. Consequently, there is an increasing demand for intelligent systems capable of processing large volumes of textual information and presenting the most important content in a shorter and more understandable format. This research introduces a Generative AI-based news aggregation and summarization platform that utilizes Natural Language Processing (NLP) techniques to simplify the process of consuming digital news. In the proposed system, users can provide the URL of a news article through a web interface. The system retrieves the textual content from the webpage, performs preprocessing to clean the data, and then applies automatic summarization algorithms to generate a brief summary that reflects the key points of the original article. In addition to summarization, the system also performs sentiment analysis to determine the overall tone of the article and categorize it into positive, negative, or neutral sentiment.*

The platform further incorporates functionalities such as secure user authentication, storage of previously generated summaries, and an analytics dashboard that visually represents summary statistics and sentiment distribution. These features improve the usability of the system and allow users to re-view previously summarized articles conveniently. The proposed framework demonstrates how artificial intelligence and NLP techniques can enhance digital news consumption by enabling users to quickly understand essential information without reading complete articles. Future developments may include multilingual summarization capabilities, integration with real-time news APIs, and the adoption of advanced transformer-based models to further improve summarization accuracy.

Keywords: *digital communication*

I. INTRODUCTION

The internet has evolved into one of the primary sources of information in contemporary society. Digital news platforms publish a vast number of articles every day across multiple domains, including politics, economics, technology, sports, and entertainment. While this continuous flow of information allows users to stay updated on global events, it also introduces the challenge of information overload. Readers are often required to examine numerous lengthy articles in order to understand important developments, which can be time-consuming and inefficient. Therefore, there is a growing demand for intelligent systems capable of processing large amounts of textual data and presenting the most relevant information in a shorter and more understandable form.



Automatic text summarization has become an important research topic within the field of Natural Language Processing (NLP). The main objective of text summarization is to create a condensed version of a document while preserving its essential ideas and important details. Earlier summarization approaches mainly relied on extractive techniques, where the system selects key sentences directly from the original document based on statistical indicators such as word frequency, sentence position, and keyword relevance. Although these methods are computationally efficient, they often produce summaries that lack coherence and fail to capture the complete meaning of the source text.

Recent developments in artificial intelligence and deep learning have greatly enhanced the performance of text summarization systems. Advanced NLP models based on trans-former architectures are capable of producing abstractive summaries that closely resemble human-written content. These models analyze contextual relationships within text and generate meaningful summaries by understanding the semantic structure of sentences. The emergence of generative AI technologies has further improved the capability of automated summarization systems.

Along with summarization, sentiment analysis plays a crucial role in understanding the emotional tone expressed in textual data. Sentiment analysis enables computational systems to determine whether a piece of text reflects a positive, negative, or neutral perspective. When applied to news articles, this technique helps readers quickly grasp the overall tone or opinion conveyed by the article. Combining summarization with sentiment analysis allows users to receive both condensed information and insights into the emotional context of the news.

This research focuses on designing an AI-based news aggregation and summarization platform that integrates several Natural Language Processing techniques. In the proposed system, users can provide the URL of a news article through the application interface. The system automatically retrieves the textual content from the webpage, processes it using NLP algorithms, and generates a concise summary that highlights the most important information. Additionally, sentiment analysis is applied to categorize the article's tone.

The platform also incorporates several practical features that improve usability and interaction. These include user authentication, storage of previously generated summaries, and an analytics dashboard that presents visual insights related to summarized articles and sentiment distribution. By combining these functionalities, the system creates an integrated environment that supports efficient news consumption.

The primary aim of this research is to demonstrate how generative artificial intelligence and Natural Language Processing techniques can be used to develop an intelligent system capable of summarizing news content effectively. The proposed framework seeks to reduce the time required to read lengthy articles while maintaining the essential meaning of the original text. Moreover, the system highlights the potential of AI technologies in improving digital information processing and enhancing user experience in modern news platforms.

II. LITERATURE REVIEW

Text summarization has been widely studied within the field of Natural Language Processing (NLP) as researchers attempt to automatically condense large volumes of textual information. Over the years, numerous techniques have been proposed to generate summaries from long documents. Early approaches to text summarization primarily relied on extractive methods. In extractive summarization, the system identifies and selects the most relevant sentences from the original document and combines them to produce a shorter representation of the content. These approaches commonly use statistical indicators such as word frequency, sentence placement within the document, and keyword significance to determine which sentences should appear in the summary. Although extractive methods are relatively simple and computationally efficient, they often produce summaries that lack logical flow and deeper contextual understanding.

With the development of machine learning techniques, researchers began investigating more sophisticated methods for automatic summarization. Neural network-based models, including recurrent neural networks (RNNs) and sequence-to-sequence architectures, were introduced to improve the ability of systems to understand relationships between words and sentences. These models enabled systems to learn contextual patterns within documents and generate summaries that better reflect the meaning of the original text. However, these early neural approaches faced challenges when processing lengthy documents or capturing complex contextual dependencies.



The introduction of attention mechanisms marked an important milestone in the evolution of summarization models. Attention-based architectures allow models to focus on the most relevant parts of the input text while generating summaries. This innovation significantly improved the performance of neural summarization systems and led to the development of transformer-based architectures. Unlike traditional recurrent models, transformer architectures rely entirely on attention mechanisms to capture relationships between words across the entire document.

Modern transformer-based models such as BERT, BART, and T5 have demonstrated strong capabilities in natural language understanding and text generation tasks. These models enable abstractive summarization, where the system produces new sentences that represent the meaning of the original document rather than simply copying sentences from it. As a result, abstractive approaches often generate summaries that are more coherent and closer to human-written text compared to extractive techniques.

Apart from summarization, sentiment analysis has also become an important tool for analyzing textual information. Sentiment analysis focuses on identifying the emotional tone or opinion expressed in text and classifying it into categories such as positive, negative, or neutral. This technique is widely used in applications such as product review analysis, social media monitoring, and news content evaluation. When applied to news articles, sentiment analysis allows readers to quickly understand the tone and perspective conveyed by the article.

Despite the progress made in both summarization and sentiment analysis research, many existing solutions remain limited to experimental environments and lack practical user interfaces for real-world usage. There is a growing need for integrated platforms that combine multiple NLP capabilities and present results through accessible web-based systems. The system proposed in this research addresses this limitation by developing a web-based news aggregation platform that integrates automatic summarization, sentiment analysis, summary history tracking, and analytics visualization. This approach demonstrates how advanced NLP technologies can be applied to create practical systems that improve the efficiency of digital news consumption.

III. CHALLENGES AND RESEARCH GAPS

Although considerable progress has been achieved in the field of Natural Language Processing (NLP) and automatic text summarization, several challenges continue to exist in the development of efficient and dependable news summarization systems. One of the major challenges is generating summaries that are brief while still preserving the essential meaning, context, and key information of the original article. Traditional extractive summarization approaches typically select sentences directly from the source document based on statistical indicators such as word frequency, sentence position, and keyword relevance. While these techniques are computationally efficient and relatively simple to implement, they often produce summaries that lack logical flow and contextual coherence. As a result, the generated summaries may not fully capture the deeper meaning of the original content.

Another significant challenge arises from the diversity and complexity of news articles available on digital platforms. News content is published across a wide range of domains including politics, technology, business, sports, and entertainment. Each domain often follows a different writing style, vocabulary structure, and narrative format. This variation creates difficulties for summarization systems that must operate consistently across multiple topics and writing patterns. Additionally, many existing summarization models encounter difficulties when processing long and complex documents that contain intricate relationships between sentences and paragraphs. Capturing these relationships accurately remains a challenging task for many NLP models.

Scalability and real-time data processing also represent important obstacles for modern summarization systems. On-line news platforms continuously publish large volumes of information every minute, and automated summarization tools must be capable of processing this incoming data efficiently without compromising the quality of the generated summaries. Achieving a balance between processing speed and summarization accuracy is therefore a critical research challenge. Furthermore, maintaining grammatical correctness and semantic consistency in automatically generated summaries remains a complex issue for many machine learning and deep learning models.



Another important research gap exists in the integration of multiple NLP functionalities within a unified platform. Many previous studies concentrate mainly on improving summarization algorithms without incorporating complementary features such as sentiment analysis, user interaction mechanisms, and data visualization tools. As a result, most research prototypes remain limited to experimental environments and lack practical systems that can be deployed for real-world usage. In many cases, the absence of user-friendly interfaces also limits the accessibility of these systems for general users.

The system proposed in this research attempts to address these challenges and research gaps by integrating several components into a single web-based platform. The platform combines news article extraction, automatic text summarization, sentiment analysis, secure user authentication, summary history tracking, and analytics visualization. By integrating these functionalities, the system aims to improve both the usability and effectiveness of automated news summarization tools. This integrated approach demonstrates how advanced NLP technologies can be applied to create practical solutions that enhance the efficiency of digital news consumption and assist users in quickly understanding large volumes of information.

IV. PROPOSED SYSTEM

The proposed system is an artificial intelligence-driven news aggregation and automatic summarization platform designed to simplify the process of reading and understanding lengthy online news articles. With the rapid growth of digital media and online news websites, an enormous amount of information is published every day. Readers often face difficulty in processing such a large volume of content within a limited amount of time. As a result, there is a growing demand for intelligent systems capable of automatically analyzing news articles and presenting the most important information in a concise format. The proposed system addresses this problem by applying Natural Language Processing (NLP) techniques to automatically generate summaries of news articles.

The platform enables users to provide the URL of a news article through an interactive web-based interface. Once the URL is submitted, the system retrieves the article content from the webpage using web scraping and text extraction methods. These techniques allow the system to collect the primary textual content of the article while ignoring irrelevant elements such as advertisements, navigation menus, or embedded media components. The extracted article text is then passed through multiple processing stages that include text preprocessing, summarization, and sentiment analysis.

During the preprocessing stage, the system cleans and prepares the textual data for further analysis. This process involves removing HTML tags, punctuation symbols, special characters, and common stop words that do not contribute significant meaning to the text. Cleaning the text ensures that the data is structured and suitable for the summarization algorithm. After preprocessing, the system applies a text summarization technique to identify the most relevant information within the article and generate a shorter summary that retains the core ideas and essential details of the original content.

In addition to generating summaries, the system also performs sentiment analysis on the news article. Sentiment analysis examines the textual content and determines whether the overall tone of the article is positive, negative, or neutral. This functionality allows users to quickly understand the emotional context or perspective conveyed in the news article without reading the entire text.

To further enhance usability and system functionality, the proposed platform includes several additional features. A user authentication module is implemented to allow secure user registration and login. The system also maintains a summary history database where previously summarized articles are stored, enabling users to revisit their past summaries when needed. Furthermore, an analytics dashboard is integrated into the platform to present visual insights such as the total number of summarized articles, sentiment distribution across processed news content, and user activity patterns.

Overall, the proposed system integrates several important components including article extraction, NLP-based summarization, sentiment analysis, secure data storage, and analytics visualization within a single web-based environment. By



com-bining these features, the platform provides an efficient and user-friendly solution for managing digital news consumption and enables users to quickly access the most important information from lengthy news articles.

V. METHODOLOGY

The development of the proposed news summarization system follows a structured methodology consisting of multiple stages that collectively process online news articles and generate meaningful summaries. The overall methodology focuses on gathering news content from online sources, processing textual information using Natural Language Processing (NLP) techniques, generating concise summaries, and performing sentiment analysis to identify the tone of the article. By combining these processes, the system provides users with an efficient way to understand lengthy news articles without reading the entire content.

The first stage of the methodology begins with news article input. The platform provides a web-based interface through which users can submit the URL of a news article they want to analyze. After receiving the URL, the system retrieves the corresponding webpage and extracts the textual information from it. Web scraping techniques are applied to collect the primary content of the article while excluding irrelevant components such as advertisements, navigation menus, or embedded multimedia elements.

The second stage involves text extraction and preprocessing. The extracted article text usually contains unwanted elements such as HTML tags, punctuation marks, special characters, and formatting symbols. During preprocessing, the system cleans the text by removing these unnecessary components and standardizing the format of the data. The text is converted into lowercase form, and common stop words that do not carry meaningful information are eliminated. This preprocessing step improves the quality of the input data and ensures that the summarization process operates more accurately.

The third stage focuses on text summarization. In this phase, the processed textual content is analyzed using NLP-based summarization techniques to determine the most significant sentences within the article. Linguistic and statistical features are used to evaluate sentence relevance and importance. Based on this analysis, the system generates a shorter summary that retains the main ideas and critical information of the original article while significantly reducing its length.

The fourth stage involves sentiment analysis, which is used to identify the emotional tone expressed in the news article. The system analyzes the textual data and classifies the sentiment into predefined categories such as positive, negative, or neutral. This functionality helps users quickly understand the perspective or tone conveyed in the article without reading the entire content.

The fifth stage addresses data storage and summary history management. After generating the summary and sentiment analysis results, the system stores the information in a database. This allows users to access previously summarized articles through a history feature. Storing this information also enables the system to generate statistical insights related to the processed news content.

The final stage focuses on data visualization and user interaction. The system includes an analytics dashboard that visually presents summarized data using charts and graphical representations. The dashboard provides insights such as the number of summarized articles, sentiment distribution patterns, and user activity trends. These visualizations help users better understand patterns within the processed news content.

Overall, the methodology integrates several technologies including web scraping, text preprocessing, NLP-based summarization, sentiment classification, database management, and data visualization. By combining these components, the proposed system creates an intelligent platform that simplifies the process of reading, analyzing, and understanding large volumes of online news information.



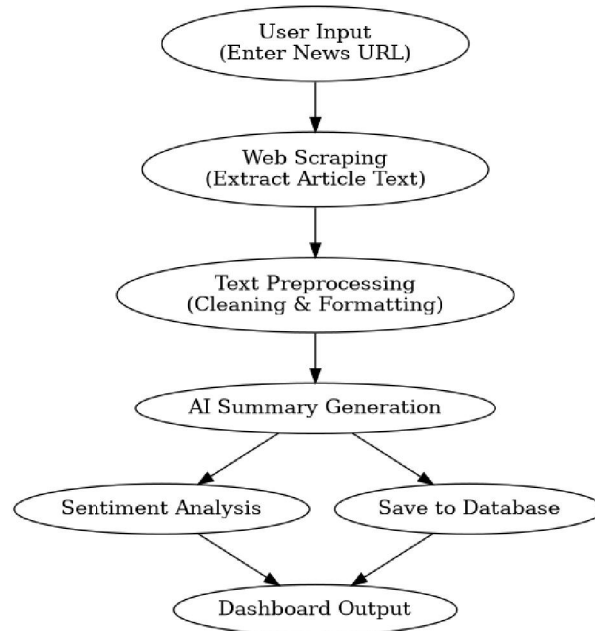


Fig. 1. Workflow Diagram of News Summarization System

VI. SYSTEM WORKFLOW

The workflow of the proposed AI-based news aggregation and summarization system is organized into a sequence of processing stages that work together to analyze news articles and produce informative summaries. The primary objective of the system is to automate the extraction, processing, and pre-sentation of news content using Natural Language Processing (NLP) techniques. By combining multiple processing steps, the system allows users to obtain important information from lengthy articles in a simplified and structured manner.

The workflow begins when a user provides the URL of a news article through the web-based interface of the platform. After the URL is submitted, the backend server accesses the webpage and retrieves the textual content of the article. Web scraping and content extraction techniques are used to collect the primary text of the article while excluding non-relevant elements such as advertisements, navigation panels, and mul-timedia components. This ensures that only the meaningful textual content is processed by the system.

Once the article content has been collected, the system proceeds to the text preprocessing stage. At this stage, the raw text is cleaned and formatted to prepare it for further analysis. This involves removing HTML tags, punctuation marks, and unnecessary symbols that may interfere with the summarization process. The text is also standardized by converting it to a consistent format, and common stop words that do not contribute significant meaning are eliminated. These preprocessing steps improve the quality of the textual data and enhance the effectiveness of the summarization algorithm.

After preprocessing, the cleaned text is forwarded to the summarization module. In this stage, NLP-based algorithms analyze the structure of the document and evaluate the impor-tance of sentences within the article. Based on this analysis, the system generates a concise summary that highlights the most relevant and informative parts of the news article. This summary enables users to quickly understand the key points of the article without reading the entire document.

Following the summarization stage, the system performs sentiment analysis to determine the emotional tone expressed in the article. The sentiment analysis module examines the textual content and categorizes the sentiment into predefined



classes such as positive, negative, or neutral. This classification helps users quickly understand the overall tone or perspective presented in the news article.

Finally, the generated summary and sentiment analysis results are stored in a database for future reference. The information is then presented to the user through the system dashboard, where visual analytics such as sentiment distribution and summary statistics are displayed. This integrated workflow ensures that users can efficiently access summarized news content along with useful anaAlthough the pro-posed AI-based news aggregation and summarization system demonstrates promising performance in generating concise summaries and performing sentiment analysis on news articles, there are several opportunities for further enhancement. Future developments can focus on improving the accuracy, scalability, and overall functionality of the platform to make it more effective for real-world applications.

One important improvement involves the integration of ad-vanced transformer-based deep learning models for abstractive text summarization. Recent NLP models such as BERT, BART, and T5 have achieved remarkable results in tasks related to language understanding and automatic text generation. By incorporating these advanced architectures, the system could generate summaries that are more coherent, context-aware, and closer to human-written content. This would significantly improve the overall quality of the generated summaries.

Another potential enhancement is the addition of multi-lingual capabilities. The current system primarily processes news articles written in a single language, which limits its usability for international audiences. Future versions of the platform could integrate multilingual Natural Language Processing techniques that allow the system to process and summarize articles written in multiple languages. This feature would expand the reach of the system and enable users from different linguistic backgrounds to benefit from automated news summarization.

The integration of real-time news APIs is another area for improvement. Instead of requiring users to manually provide article URLs, the system could automatically retrieve the latest news from trusted news sources and generate summaries in real time. This capability would transform the system into a dynamic news monitoring platform that continuously provides summarized updates.

Future research may also investigate the use of machine learning techniques for fake news detection. By analyzing patterns in verified datasets, the system could identify misleading or unreliable information within news articles. Additionally, implementing topic classification and personalized recom-mendation mechanisms could allow the system to deliver summaries tailored to individual user interests and reading behavior.

Overall, these future enhancements would increase the in-telligence, scalability, and practical usefulness of the proposed news summarization system, making it a more comprehensive solution for efficient digital news consumption. lytical insights.

VII. FUTURE SCOPE

Although the proposed AI-based news aggregation and summarization system demonstrates promising performance in generating concise summaries and performing sentiment anal-ysis on news articles, there are several opportunities for further enhancement. Future developments can focus on improving the accuracy, scalability, and overall functionality of the platform to make it more effective for real-world applications.

One important improvement involves the integration of ad-vanced transformer-based deep learning models for abstractive text summarization. Recent NLP models such as BERT, BART, and T5 have achieved remarkable results in tasks related to language understanding and automatic text generation. By incorporating these advanced architectures, the system could generate summaries that are more coherent, context-aware, and closer to human-written content. This would significantly improve the overall quality of the generated summaries.

Another potential enhancement is the addition of multi-lingual capabilities. The current system primarily processes news articles written in a single language, which limits its usability for international audiences. Future versions of the platform could integrate multilingual Natural Language Processing techniques that allow the system to process and summarize articles written in multiple languages. This feature would expand the reach of the system and enable users from different linguistic backgrounds to benefit from automated news summarization.



The integration of real-time news APIs is another area for improvement. Instead of requiring users to manually provide article URLs, the system could automatically retrieve the latest news from trusted news sources and generate summaries in real time. This capability would transform the system into a dynamic news monitoring platform that continuously provides summarized updates.

Future research may also investigate the use of machine learning techniques for fake news detection. By analyzing patterns in verified datasets, the system could identify misleading or unreliable information within news articles. Additionally, implementing topic classification and personalized recommendation mechanisms could allow the system to deliver summaries tailored to individual user interests and reading behavior.

Overall, these future enhancements would increase the intelligence, scalability, and practical usefulness of the proposed news summarization system, making it a more comprehensive solution for efficient digital news consumption.

VIII. RESULTS

The developed AI-based news summarization system was successfully implemented and tested using various news article URLs. The system demonstrates the ability to extract textual information from online news sources, generate concise summaries, perform sentiment analysis, and present the results through an interactive web interface.

The Article Summarization Page allows users to paste a news article URL and generate an automatic summary. After processing the URL, the system extracts the article content and applies Natural Language Processing techniques to produce a short and informative summary. This helps users quickly understand the key points of lengthy news articles without reading the full text.

Figure 1 shows the article summarization interface of the system. The interface includes a URL input field, a summarize button, and a display section where the generated summary and article details are presented.

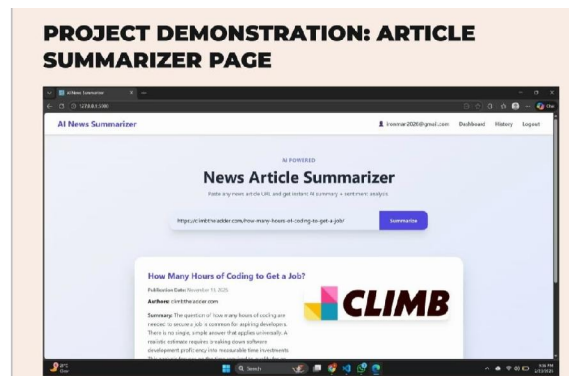


Fig. 2. News Summarizer Interface

The system also includes a Summary History Module that stores previously summarized articles for each user. This feature enables users to review past summaries and track their reading activity. The history page displays article titles, sentiment results, and summarized content in a structured format.

Figure 2 illustrates the summary history page where users can access their previously summarized news articles.

In addition to summarization and history tracking, the system provides an Analytics Dashboard that visualizes information about summarized articles. The dashboard presents statistical insights such as the total number of processed articles, sentiment distribution, and weekly activity trends. These visualizations help users better understand the data generated by the system.



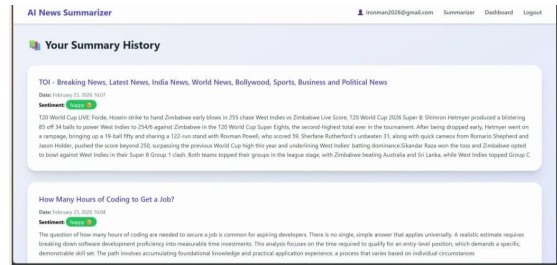


Fig. 3. News Summarizer Interface

Figure 3 shows the analytics dashboard displaying sentiment distribution and user activity graphs. The dashboard includes visual charts that provide an overview of article sentiment classification and system usage statistics.

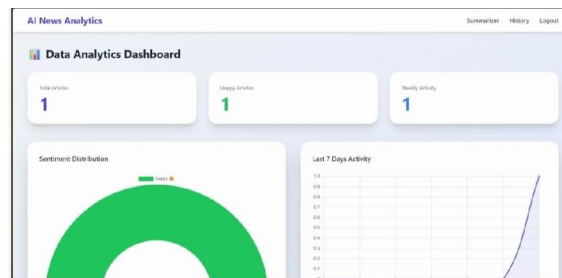


Fig. 4. News Summarizer Interface

IX. CONCLUSION

This research presented the design and implementation of an AI-based news aggregation and automatic summarization system aimed at improving the efficiency of digital news consumption. With the rapid expansion of online news platforms, readers are constantly exposed to large volumes of information from multiple sources. As a result, identifying the most relevant information from lengthy articles has become increasingly challenging. The proposed system addresses this issue by applying Natural Language Processing (NLP) techniques to automatically generate concise summaries from news articles obtained through user-provided URLs.

The developed platform integrates several key modules including article extraction, text preprocessing, automatic summarization, and sentiment analysis. The summarization component processes the textual content of news articles and produces a shorter version that preserves the essential meaning and key ideas of the original content. In addition to summarization, the sentiment analysis module evaluates the emotional tone of the article and classifies it into categories such as positive, negative, or neutral. This functionality allows users to quickly understand the perspective or tone conveyed in the news content.

The system also incorporates additional features that enhance usability and functionality, including user authentication, summary history tracking, and an analytics dashboard. These features enable users to access previously summarized articles and analyze sentiment trends through visual representations. The dashboard provides graphical insights into processed news data, which further improves the user experience and supports efficient information analysis.

Experimental implementation demonstrates that the proposed approach significantly reduces the time required to read and understand news articles while maintaining the essential information contained in the original text. By integrating summarization, sentiment analysis, and visualization capabilities, the system offers a comprehensive platform for efficient news consumption.

Future work may focus on incorporating advanced transformer-based models to improve summarization quality, supporting multiple languages, integrating real-time news APIs, and implementing fake news detection mechanisms. These improvements could further enhance the scalability, intelligence, and practical applicability of the proposed system.



REFERENCES

- [1] A. Vaswani et al., "Attention Is All You Need," Advances in Neural Information Processing Systems, 2017.
- [2] M. Lewis et al., "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension," ACL, 2020.
- [3] C. Raffel et al., "Exploring the Limits of Transfer Learning with a Uni-fied Text-to-Text Transformer," Journal of Machine Learning Research, 2020.
- [4] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," NAACL, 2019.
- [5] R. Nallapati et al., "Abstractive Text Summarization using Sequence-to-Sequence RNNs and Beyond," CoNLL, 2016.
- [6] A. See et al., "Get To The Point: Summarization with Pointer-Generator Networks," ACL, 2017.
- [7] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval, 2008.
- [8] Y. Liu and M. Lapata, "Text Summarization with Pretrained Encoders," EMNLP, 2019.
- [9] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," ICLR, 2013.
- [10] K. Knight and D. Marcu, "Summarization Beyond Sentence Extraction," Artificial Intelligence Journal, 2002.
- [11] S. Chopra, M. Auli, and A. Rush, "Abstractive Sentence Summarization with Attentive Recurrent Neural Networks," NAACL, 2016.
- [12] A. Rush, S. Chopra, and J. Weston, "A Neural Attention Model for Abstractive Sentence Summarization," EMNLP, 2015.
- [13] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," ICLR, 2015.
- [14] J. Pennington, R. Socher, and C. Manning, "GloVe: Global Vectors for Word Representation," EMNLP, 2014.
- [15] I. Goodfellow et al., "Generative Adversarial Nets," NeurIPS, 2014.
- [16] T. Lin et al., "A Structured Self-attentive Sentence Embedding," ICLR, 2017.
- [17] K. Papineni et al., "BLEU: A Method for Automatic Evaluation of Machine Translation," ACL, 2002.
- [18] C. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries," ACL Workshop, 2004.
- [19] S. Bird, E. Klein, and E. Loper, "Natural Language Processing with Python," O'Reilly Media, 2009.
- [20] J. Brownlee, "Deep Learning for Natural Language Processing," Machine Learning Mastery, 2017.
- [21] H. Schütze, C. Manning, and P. Raghavan, "Introduction to Information Retrieval," Cambridge University Press, 2008.
- [22] D. Jurafsky and J. Martin, "Speech and Language Processing," Prentice Hall, 2009.
- [23] A. McCallum and K. Nigam, "A Comparison of Event Models for Naive Bayes Text Classification," AAAI Workshop, 1998.
- [24] F. Chollet, "Deep Learning with Python," Manning Publications, 2017.
- [25] Y. Goldberg, "A Primer on Neural Network Models for Natural Language Processing," Journal of Artificial Intelligence Research, 2016.

