COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING
e-ISSN: 2582-7464

Computational Intelligence and Machine Learning
Vol-4, Issue-2, October 2023

A Review of Text Summarization Techniques Using NLP

Kartik Aggarwal[1]*, Naveen[2]

[1] Dronacharya College of Engineering, Khentawas, Farrukh Nagar, Gurugram, Haryana, India
[2] Assistant Professor, Dronacharya College of Engineering, Khentawas, Farrukh Nagar, Gurugram, Haryana, India
*Corresponding Author Email: kartikaggarwal879@gmail.com

Abstract

Techniques that employ natural language processing (NLP), often known as text summarizing, automatically construct summaries of extensive texts. Extractive and abstractive summarization are two main categories that may be used to classify these methods. In extractive summarizing, the most significant lines or phrases from a text are isolated and used to generate a summary. On the other hand, in abstractive summarization, a summary is generated that is clear, short, and accurate in its representation of the text's primary concepts. NLP methods like sentence segmentation, part-of-speech tagging, named entity recognition, and semantic analysis are used in generating a summary from a text and locating and extracting relevant information from the text. Text summarizing is a subject that has received a significant amount of study and has applications in various fields, including the summation of news articles, documents, and emails, among other things.

Keywords

Deep Learning and Transformer-Based, Extractive and Abstractive Approaches, Natural Language Processing, Text Summarization.

INTRODUCTION

Developing a short and cohesive summary of a given text while maintaining the most significant information from that text is referred to as summarizing the text [1]. Text summarization has emerged as a key component of natural language processing due to the explosion of textual content across various industries, including the media, scientific publications, legal papers, and social media [2].

The ways for summarizing texts may be broken down into two primary categories: extractive and abstractive. Both of these categories can be broken down further into subcategories. Extractive methods entail determining which sentences or phrases from the primary text are the most important and then building a summary by picking those words or phrases and stringing them together [3]. This uncomplicated approach generates outlines that accurately represent the original text; however, it may not successfully convey the bigger picture of the text's context and organization. On the other hand, abstractive techniques entail creating a summary by comprehending the meaning and context of the original text and generating new material that appropriately expresses the primary concepts of the text [4]. This is done to provide a summary of the abstractive approach. This approach is more complex, but it has the potential to generate summaries that are both more useful and more similar to summaries produced by humans [5].

In recent years, due to the fast growth of machine learning and deep learning methods, natural language processing (NLP) strategies for text summarization have garnered substantial interest [6]. Text summarization has seen significant improvements in quality and efficiency because of the work of scholars who have created new methodologies and models in response to the availability of big annotated datasets and powerful computational resources [7].

Text summarization using natural language processing (NLP) techniques offer several advantages, including [8]:

1. Time-saving: Text summarization automates the task of summarizing lengthy texts, saving time for readers who need to quickly understand a document's or article's main ideas.
2. Increased efficiency: Text summarization algorithms can analyze and summarize large volumes of text quickly and accurately, increasing efficiency in tasks such as document processing and information retrieval.
3. Improved comprehension: Summarized texts provide a concise and clear overview of the original text, making it easier for readers to understand complex or technical information.
4. Customization: Text summarization models can be customized to suit specific domains or applications, such as summarizing medical research papers or financial news.
5. Multilingual summarization: NLP techniques can also summarize the text in multiple languages, which can be helpful for multilingual organizations or researchers.
6. Scalability: Text summarization algorithms can be easily scaled to process large amounts of data, making them ideal for summarizing social media feeds or news articles in real-time.

This article provides an overview of text summarization using NLP techniques, including the different approaches, models, and evaluation metrics. We will also discuss some of the applications of text summarization in various domains and future research directions in this field.

BACKGROUND STUDY

The process of constructing a shorter version of a text document while maintaining the content's key meaning and information is referred to as a summary of the text. In text summarizing, methods from the field of Natural Language Processing (NLP) are utilized extensively to automatically produce summaries from vast quantities of text [9].

Extractive and abstract summarization are the two primary methods that may be used to summarize a text. When creating a summary from the original text, extractive summarizing includes picking specific sentences or phrases to rearrange to generate the summary. On the other hand, abstractive summarization requires creating new sentences capable of encapsulating the most important information and the overall meaning of the source text [10]. Natural languages processing methods such as tokenization, part-of-speech tagging, named entity recognition, and syntactic parsing are used to preprocess the text before beginning the summarising process. After the data has been preprocessed, several machine learning methods, including decision trees, neural networks, and support vector machines, are trained on the data and then used to create summaries.

One of the most challenging aspects of text summarizing is ensuring that the resulting summary is not just brief but also instructive, all while ensuring that the text's original meaning and context are preserved in the process. Handling the diversity and complexity of natural language, including synonyms, ambiguity, and sarcasm, which may impact the accuracy and relevancy of the produced summaries, is another problem that must be overcome. Despite these problems, the process of text summary using NLP approaches has several practical uses, such as the summarization of news stories, documents, and social media [11]. Users who need to rapidly comprehend the primary information in vast volumes of text might save time and effort by using the capability to produce summaries of the text automatically.

The natural language processing (NLP) approaches used for text summarization provide several obstacles, some of which include the following [12]

1. Ambiguity and polysemy: Polysemy refers to the fact that the same word or phrase may have numerous interpretations. Natural language is often ambiguous. Since of this, natural language processing-based summarizing systems have a problem since, to provide valuable summaries, they need to comprehend the context and disambiguate the text correctly.
2. Material summarizing entails compressing vast volumes of material into a brief and cohesive summary. This process requires coherence and cohesion. A problem for natural language processing (NLP)-based summarizing systems is ensuring that the summary is grammatically accurate but also coherent and cohesive, which means that the sentences flow logically from one to the next.
3. Realization of the text's context: Realizing the context of the text is necessary to provide accurate and helpful summaries. For NLP-based summarizing systems to provide valuable summaries, they must comprehend the text's significance within its surrounding context.
4. Language particular to a domain: the linguistic qualities and terminology of texts from various fields are distinct. To provide appropriate summaries, systems based on natural language processing (NLP) need to be able to manage domain-specific language.
5. Measures for assessment: When it comes to evaluating the quality of summaries that are created, there are not enough evaluation measures that are consistent and reliable. This is a hurdle for NLP-based summarization systems when evaluating their performance appropriately.
6. Data scarcity is challenging for supervised learning-based summarization systems since having a significant quantity of annotated training data is necessary to produce satisfactory results. On the other hand, annotated data is often lacking or non-existent for many subjects and languages, making it difficult to create successful summarization models.

More research and development has to be done in text summarization using NLP approaches to find solutions to these problems. To overcome these obstacles and produce more efficient summarization systems, improvement is required in several areas, including developing improved algorithms, enhanced training data, and more rigorous assessment measures.

Extracted and abstract summarisation are the most common approaches to summarizing a text [13]

The extractive approach of summarizing involves the computer finding the sentences or phrases that are the most significant from the original text and then extracting them to create a summary. Because it does not result in the creation of new sentences, extractive summarization is considered the more straightforward method of the two. Methods, including clustering algorithms, graph-based models, and sentence ranking, are used in this process. This approach provides a new summary, not just a selection of phrases from the original text. It is referred to as abstractive summarization. It is understanding the original text's meaning and producing a new summary that faithfully conveys that meaning is both required steps in abstract summarization. It is more complicated than extractive summarization because it involves better comprehending the text's context, meaning, and grammar. The abstractive summarisation process uses sequence-to-sequence models, transformers, and reinforcement learning methods.

The extractive and the abstractive approaches of summarization each have their own unique set of benefits and drawbacks. The process of the extractive summary is often less complicated and takes less time than the process of abstractive summarization; nonetheless, it may need to be able to convey the primary ideas expressed in the original text correctly. The abstractive outline has the potential to provide more valuable summaries. Still, it also requires more complex natural language processing methods and has the potential to produce summaries that need to be more accurate to the source text.

In reality, selecting a summarising technique is determined by considerations such as the subject matter, the kind of text being summarized, and the goal of the summary. Some applications may call for high precision and attention to detail, while others emphasize speed and effectiveness. Researchers and industry professionals always develop and refine these methodologies to increase their usefulness and efficiency across various contexts.

Types of Extractive and Abstractive Methods for Text Summarization [14]

Different kinds of extractive techniques may be used to summarise texts, including the following.

1. Methods based on frequency: This approach finds the words or phrases that appear most often in the source material and then creates a summary based on those words or phrases. This method is predicated on the assumption that the words or phrases used most often are also likely to be the ones that best summarise the content of the text.
2. Methods based on graphs: In this technique, graph theory represents the connection between the phrases in the text. The chart is built by considering each sentence as a node and linking the sentences with a high similarity score to those with the same node type. After then, the sentences with the highest significance are determined using centrality metrics like degree centrality, betweenness centrality, or PageRank.
3. Methods based on machine learning: One such technique includes training a machine learning model on a labelled dataset containing texts and summaries. The model is then instructed to recognize the sentences in the text that are the most significant based on the characteristics of those sentences, such as their length, location, and substance.
4. Methods based on clustering entail grouping sentences that are similar and then picking the sentence within each cluster that best exemplifies the approach as a whole. For this particular objective, clustering techniques such as k-means, hierarchical clustering, and density-based clustering are all viable options.
5. One of the ways that sentences may be ranked is by how pertinent and significant they are to the meaning of the whole piece of writing. This is one of the ways that sentences can be ranked. The extraction of keywords, the comparison of phrases, or semantic analysis are all examples of possible approaches that may be used to achieve this goal.

Different kinds of abstractive approaches may be used to summarise texts, such as

1. Sequence-to-sequence models: This approach uses deep learning models such as recurrent neural networks (RNNs) or transformers to develop a mapping between the input text and the summary produced. The summary is produced by the model, which does this by making predictions about the terms that are most likely to appear given the input text.
2. Transformers: This approach generates a summary by concentrating on the most significant aspects of the input text. It does this by using attention processes to do so. Transformers are a specific kind of neural network design that can both consider long-term dependencies and manage inputs of varying lengths.
3. Learning via reinforcement is a technique that includes teaching a model to create summaries by rewarding the model for producing high-quality summaries and punishing the model for making low-quality summaries. The model is trained to maximize a reward function by basing its decisions on the quality of the summaries it generates.
4. Analyzing the semantics of the input text is the first step in the semantic analysis approach. This method aims to provide a summary that accurately conveys the most important ideas and concepts. Techniques such as named entity identification, sentiment analysis, and topic modelling are all viable options for accomplishing this goal.
5. Summarization based on templates: This technique includes generating a summary based on the input text using templates already specified. There is a possibility that the templates may have blanks where vital information or phrases should go. These blanks will be filled in with the material that is provided.

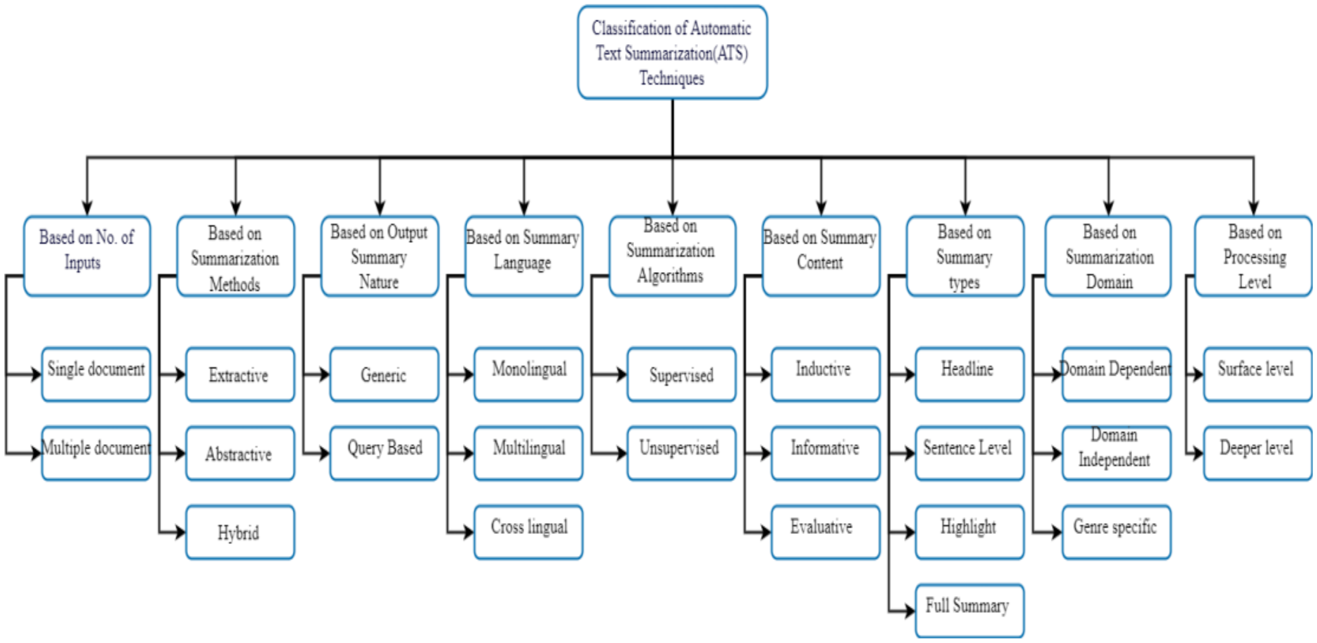

Figure 1. Detailed Automatic Text Summarization System Categorization.

Figure 1 shows that systems that automatically summarise texts may be divided into various categories according to the criteria used to classify them [15]. The following is a list of the several classifications that are used for automated text summarising systems:

1. **Comparison between Extractive and Abstractive Summarization**: An extractive summary is one in which key lines or phrases from the original text are taken and used to construct a summary. An abstractive summary, on the other hand, does not use any of the original text at all. On the other hand, abstractive summarising entails the creation of a summary by the use of natural language generation methods. The extractive process of summarising is more straightforward and more often employed, whereas the abstractive is more complicated, more challenging, needs more complex procedures, and I less prevalent.
2. **Comparison of Summarizing a Single Document to Summarizing Numerous Documents**: A single-document summary seeks to compress a single document, such as an article or report, into a shorter form, while a multi-document summary aims to summarise numerous documents into a single summary.
3. **Domain-Specific Summarizing vs General-Purpose Summarization:** Domain-specific summarising systems are trained on specialized datasets and are targeted to specific domains, such as legal or medical papers. General-purpose summarization systems are designed to summarise information across various topics. On the other hand, systems designed for general-purpose summaries can summarize material from different subject areas.
4. **Learning Methods: Supervised vs Unsupervised Learning**: In supervised learning, a model is trained using a dataset that has been labelled, while in unsupervised learning, a model is trained using a dataset that has not been labelled. The majority of text summarization systems make use of supervised learning approaches; however, unsupervised techniques are also now being investigated.
5. **Comparison of Query-Based and Non-Query-Based Summarizing**: Query-based summarising systems produce summaries based on particular user questions or subjects. In contrast, non-query-based summarisation systems produce summaries without direct input from the user.
6. **Comparison of Sentence Summarization, Paragraph Summarization, and Document Summarization:** The objective of sentence summarisation systems is to condense a document into a set of introductory sentences; the objective of paragraph summarisation systems is to reduce a document into a bunch of essential paragraphs, and the aim to document summarisation systems is to condense an entire document into a shorter version.
7. **Single-Document Compression vs Multiple-Document Fusion:** Techniques for single-document compression try to extract significant lines or phrases from a single document and condense them into a shorter version of themselves. Techniques for multiple-document fusion combine the contents of many documents into a single document. On the other hand, the goal of the approach known as multiple-document fusion is to integrate information from many documents to provide a comprehensive summary that includes all of the essential details from the papers.

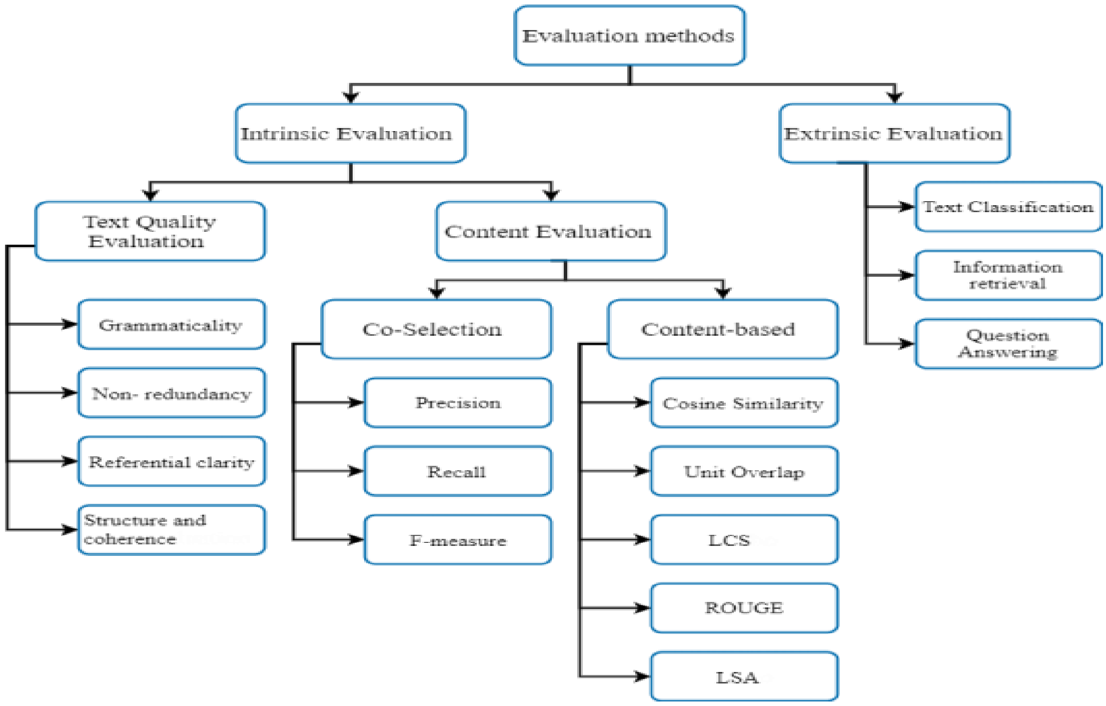

Figure 2. The valuation of Automatic Text Summarization Techniques.

Figure 2, It is vital to analyze several artificial texts summarising approaches to evaluate the quality of the summaries created and compare the various summarisation systems [16]. The following is a list of some of the most frequent strategies used to assess artificial text summarising [17] methods:

1. **ROUGE Metrics ROUGE**, which stands for "Recall-Oriented Understudy for Gisting Evaluation," is a metric collection comparing the amount of overlap between the reference summary and the summary created. ROUGE-1, ROUGE-2, and ROUGE-L are the ROUGE measures used most of the time. These metrics calculate the percentage of overlap between the produced summary and the reference summary based on the number of unigrams, bigrams, and the longest shared subsequence.
2. **Method of the Pyramid**: The Pyramid method includes the creation of various reference summaries via human annotators tasked with writing summaries of the same material. After that, the quality of the produced summary is assessed by contrasting it with the reference summaries using metrics like ROUGE.
3. **Human assessment:** A human assessment consists of requesting human annotators to score the quality of the produced summary based on various criteria, including relevancy, fluency, coherence, and overall quality. Humans perform human evaluations. Human assessment may be time-consuming and costly, but it offers invaluable insights into the quality of the summaries from the users' point of view.
4. **F-Measure:** The F-measure is a statistic that analyses the accuracy and recall of the produced summary. It determines the degree to which the produced summary and the reference summary overlap, in addition to determining whether or not the created summary is comprehensive.
5. **The information content of a summary** is measured against the length of the original text to determine how much of the original material was included in the summary. It evaluates the degree to which the summary length may be reduced while maintaining the integrity of the vital information.
6. **Readability and Coherence:** From a linguistic point of view, the summary's quality may be evaluated based on its readability and degree of readability. The summary's coherence is assessed based on how well it follows a logical flow and is structured. In contrast, the summary's readability is evaluated based on how easily it can be understood and read.

LITERATURE SURVEY

This article comprehensively examines the many approaches to summarising texts, from the simplest to the most sophisticated. This research [18] concludes that improved accuracy may be achieved using the seq2seq model in tandem with the LSTM and the attention mechanism.

This study compares the extractive and abstractive summary methods for Hindi text texts. They start by bringing up the idea of ward hierarchical agglomerative clustering for either type of summary. The PageRank algorithm for extractive summarising follows this. For abstractive summarising, they describe a method based on multi-sentence compression that only needs a POS tagger to make Hindi text summaries [19]. The PageRank method is then used to get an extractive overview. This article discusses modern techniques for text summarising, categorizing them into three categories: extractive, abstractive, and hybrid, and reviewing particular methods that fall under each category. The Extractive method is what most writers emphasize; nonetheless, regardless of the approaches utilized, the produced summaries are very distinct from human summaries [20].

Each word's size changes as you create a word cloud depending on how often it occurs in the original text. With speech recognition enabled and guaranteed semantic coherence and spatial stability of the investigated words, our approach upgrades the user's dynamic data input and the already employed static text word cloud. Our research is represented by an interactive visual study system that helps users analyze text and draw conclusions from a vast collection of documents and voice commands [21]. The effort we have made is personified in this system.

This study investigates the use of machine learning algorithms and classifiers, as well as the usage of extractive text summarisation approaches, to facilitate the retrieval of customer responses from a database containing frequently asked questions by extracting important phrases and semantic information from customer inquiries. To evaluate the relative merits of two approaches to text summarisation, we create a prototype of an automated agent that can respond to customer inquiries in the context of electronic media. (supervised and unsupervised). The study exemplifies machine learning and text summarisation techniques to develop a useful tool that could assist businesses in better managing their customer interactions and introduce robust, efficient, and effective electronic media-enabled customer care mechanisms [22].

To summarise extractive text, first, determine which sections of the source text have the most relevant information, then choose a subset of phrases from that text. Abstractive text summarisation is more analogous to how humans work since it involves replicating or rephrasing the original text based on the interpreter's or understander's interpretation and comprehension of the original text utilizing methods associated with natural language processing. In this work, they give a text summation of the Covid-19 news using an abstractive technique. Their goal was to create something that was as similar to the way humans summarise as possible. They also use data augmentation in the preprocessing section as an example of dealing with data that could be better or varied enough [23]. This is done to illustrate the scenario of working with data.

Reducing a lengthy piece of writing into a more concise version while retaining the essential themes and information presented in the original is known as a summary of the text. The process of text summarising may be broken down into two primary categories. Textual condensation that is both extractive and abstract, Processes of extractive summarising in which the most significant portions of already-existing words, phrases, or sentences from the source material are used, Within the scope of this investigation, a sentence-based mode) using Fuzzy C-Means clustering has been developed. Six essential elements for sentence scoring have been introduced, including a new feature called the "Sentence Highlighter Feature." ROUGE is used to assess the performance of the proposed FCM mode. It has been measured using precision, recall, and f-measure. This FCM model's extraction procedures have been shown to have reduced outline repetition and more data profundity [24], as demonstrated by the results.

The deep learning field methods and various types of neural networks are now the most widely used and effective way for automatically summarising large amounts of text. This work presents a variety of methodologies and different datasets that might be used to produce text summaries automatically. As a result, the time required to transcribe and summarise extensive text documents manually could be reduced [25].

They use a test set to assess the performance of our model, which is based on a BART model that has been fine-tuned. They demonstrate that using this method results in superior summarisation performance measured by the ROUGE score. They achieved a marginal improvement in ROUGE1 F1 (R1 F1), with a score of +0.6 on the development set and +0.5 on the test set, respectively. They contrast it with a model that is independent of adding supplementary data [26].

The secretary problem is a component of the raw text summarising approach; they devise a novel strategy for writing the summary. They create a partition between the printed text. They will alter how we talk to more closely resemble the title's first half. Keep lines in the same group as long as the words match the main title’s terms. There will be various non-title word discussions regarding the secretary issue. When combined with other methods of phrase creation, the Secretary problem guarantees the optimal solution 37% of the time, on average. In this study, we detail how we imagined utilizing a mathematical model to condense the narrative and exclude some details [27]. The outline is created by extracting key phrases from the original text. This study showed a model for describing texts in a general way. The model gets information from the Daily Mail, CNN, and other papers. It then makes two reports of the data. One was from a linguist, and the other was based on a given model. The philologist's summary was saved so that it could be compared to the summary that the machine made. The shown model helped create the report more accurately and efficiently understood [28].

The data extraction method begins by reading the primary source to determine the words containing the relevant information. After finding these phrases, the process then extracts only those phrases. Before going on to the actual creation of the summary, the abstractive approach of summarising starts with an interpretation of the source material. The process of comparing the text and creating the outline is made more difficult by the use of pre-trained models that are based on transformer architecture. The findings acquired from the machine learning models are evaluated and compared with the help of the BBC news dataset for this study investigation [29]. They have completed an analysis of several methods for automatically summarising text. They have presented a model that uses BERT and Group Average Linkage clustering on articles collected from PubMed. This is in consideration of the research gaps that already exist. This effort might benefit people working in biomedicine since it summarises massive amounts of biomedical literature. It would save time by providing a concise summary that condenses the material and delivers it in a nutshell [30].

The study article delves into the debate about whether or not aspect-based summaries and sentiment analysis are necessary. A framework may be constructed by extracting coherent elements from the reviews and then employing the approach of extractive summarisation to create summaries. In addition, aspect-based sentiment analysis provides insights into the evaluations of tourist destinations. Crowdsourcing, Fairsumm, and the Centroid technique are used to analyze the findings. The aspect-based summaries approach that uses crowdsourcing produces the most outstanding results [31].

Several research has looked at the efficacy of various approaches to text summarisation. The techniques discussed in this article often produce summaries or extract summaries of text documents. Additionally, a query-based summarising method is discussed in this article. The summarisation of text documents may be accomplished using these strategies, which are structured and semantically based. Our approach outperforms the state-of-the-art extractive and abstractive baselines on the ROUGE evaluation metric [32], as shown by experimental results from the DUC-2002 dataset.

RESEARCH GAP

Despite the substantial advancements made in text summarisation via the use of NLP, certain research holes still need to be filled. Among these deficiencies are the following:

- **Addressing Bias**: The existence of a possibility for bias in the summaries that are created is one of the most critical research gaps in the field of text summarisation. Bias can originate either in the training data or in the models themselves. For example, models trained on biased datasets might create summaries that propagate preconceptions or engage in discriminatory behaviour. It is essential to address bias in text summarising to guarantee that the summaries created are fair and inclusive.
- **Managing Lengthier Texts**: Although the currently available text summarisation methods are effective when applied to shorter texts, they suffer when applied to larger ones. When attempting to generate a summary for a long text, it is typically necessary to first comprehend the structure of the content and then pick the most critical parts. A research gap that needs to be addressed is the difficulty of summarising lengthy papers using NLP approaches. This obstacle has to be solved.
- **Personalizable Summaries**: One of the difficulties associated with text summarising is the possibility that a single summary may not be appropriate for all consumers or use cases. For instance, a summary written for a researcher may not be suitable for a journalist or someone just reading for pleasure. A research gap that must be addressed is the creation of NLP models that can provide fully configurable summaries that can accommodate different users' various requirements and preferences.
- **Summarizing in Multiple Languages**: The vast majority of the existing text summarising algorithms were trained on datasets in the English language. A hole in the study has to be filled, and one of those holes is making these models applicable to different languages. However, additional obstacles must be addressed to increase the effectiveness of text summarisation in different languages. These issues include variances in grammar, syntax, and lexicon.

CONCLUSION

The job of text summarising via natural language processing (NLP) methods is an important one that may assist users in saving time, increasing their efficiency, and improving their understanding of various fields. Summarising a text may be broken down into two primary categories: extractive summarisation and abstractive summary. Recent studies have concentrated on establishing hybrid strategies that incorporate the benefits of each of these techniques. Recent developments in deep learning and transformer-based models have resulted in significant improvements in the performance of text summarisation. However, there is still an opportunity for further growth in this area. The development of assessment measures and models for text summarisation that is more accurate and efficient should be the focus of future studies. In addition, research should also investigate the difficulties and constraints of text summary, such as how to handle confusing or conflicting material, maintain the text's original context and tone, and generate summaries appropriate for various audiences and purposes.

The application of natural language processing (NLP) to summarising text is quite broad, and there are several fascinating potential directions for future study and development. The following is a list of possible future paths for text summarisation: Enhancement of the performance of abstractive summarising. An abstractive summary still needs to be completed, and there is room for the development of more complex models that can create more accurate and coherent summaries while maintaining the text's original meaning and context. Measures for evaluation: For summarisation models, there is a pressing demand for assessment measures that are both more robust and trustworthy. The current metrics for assessment only sometimes fully represent the quality of the summaries created, and it is necessary to develop new metrics to analyze elements such as coherence, readability, and relevancy. Summarisation models trained on generic corpora may perform less well in specialized domains, such as scientific literature or legal texts. Domain-specific summarisation aims to address this issue. As a result, there is a need for summarisation models that are specialized in specific domains and can provide accurate and relevant summaries. Multimodal summarising: Text is only one communication modality, and there is a rising demand for multimodal summarization models that can produce summaries from a mix of text, voice, pictures, and video. Multimodal summarisation models can generate summaries from all of these different modalities. A summary that protects privacy There is an increasing worry about privacy and data protection today due to the proliferation of digital technology. As a result, there is a need for models of summarising that can provide summaries without compromising the privacy of the persons or organizations involved. The topic of text summarisation with NLP approaches is fast expanding and has substantial untapped potential for future study and growth. Some of the most intriguing possibilities for future research include advancements in abstractive summarisation, assessment metrics, domain-specific summarising, multimodal summarisation, and summarisation that protects individuals' privacy. The continuation of research and development in these areas has the potential to further enhance state of the art in text summarisation and allow information processing and transmission that is both more efficient and effective.

REFERENCES

[1] P. B. Tumpa, S. Yeasmin, A. M. Nitu, M. P. Uddin, M. I. Afjal and M. A. A. Mamun, "An Improved Extractive Summarization Technique for Bengali Text(s)," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2018, pp. 1-4, doi: 10.1109/IC4ME2.2018.8465609.

[2] 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.

[3] M. Afsharizadeh, H. Ebrahimpour-Komleh and A. Bagheri, "Query-oriented text summarization using sentence extraction technique," 2018 4th International Conference on Web Research (ICWR), Tehran, Iran, 2018, pp. 128-132, doi: 10.1109/ICWR.2018.8387248.

[4] W. Palliyali, M. A. Al-Khalifa, S. Farooq, J. Abinahed, A. Al-Ansari and A. Jaoua, "Comparative Study of Extractive Text Summarization Techniques," 2021 IEEE/ACS 18th International Conference on Computer Systems and Applications (AICCSA), Tangier, Morocco, 2021, pp. 1-6, doi: 10.1109/AICCSA53542.2021.9686867.

[5] P. S. Suryadjaja and R. Mandala, "Improving the Performance of the Extractive Text Summarization by a Novel Topic Modeling and Sentence Embedding Technique using SBERT," 2021 8th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA), Bandung, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICAICTA53211.2021.9640295.

[6] K. D. Garg, V. Khullar and A. K. Agarwal, "Unsupervised Machine Learning Approach for Extractive Punjabi Text Summarization," 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2021, pp. 750-754, doi: 10.1109/SPIN52536.2021.9566038.

[7] T. Islam, M. Hossain and M. F. Arefin, "Comparative Analysis of Different Text Summarization Techniques Using Enhanced Tokenization," 2021 3rd International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2021, pp. 1-6, doi: 10.1109/STI53101.2021.9732589.

A. Elsaid, A. Mohammed, L. F. Ibrahim and M. M. Sakre, "A Comprehensive Review of Arabic Text Summarization," in IEEE Access, vol. 10, pp. 38012-38030, 2022, doi: 10.1109/ACCESS.2022.3163292.

[8] R. Sikder, M. M. Hossain and F. M. R. H. Robi, "A Noble Analytical Text Summarization Technique for Natural Bengali Language," 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2018, pp. 1-5, doi: 10.1109/ICCITECHN.2018.8631950.

A. Rahman, F. M. Rafiq, R. Saha, R. Rafian and H. Arif, "Bengali Text Summarization using TextRank, Fuzzy C-Means and Aggregate Scoring methods," 2019 IEEE Region 10 Symposium (TENSYMP), Kolkata, India, 2019, pp. 331-336, doi: 10.1109/TENSYMP46218.2019.8971039.

[9] 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.

[10] Y. Einieh and A. AlMansour, "Deep Learning in Arabic Text Summarization: Approaches, Datasets, and Evaluation Metrics," 2022 20th International Conference on Language Engineering (ESOLEC), Cairo, Egypt, 2022, pp. 45-49, doi: 10.1109/ESOLEC54569.2022.10009528.

A. Farooq, S. Batool and Z. Noreen, "Comparing Different Techniques of Urdu Text Summarization," 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Karachi, Pakistan, 2021, pp. 1-6, doi: 10.1109/MAJICC53071.2021.9526246.

[11] S. Modi and R. Oza, "Review on Abstractive Text Summarization Techniques (ATST) for single and multi documents," 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2018, pp. 1173-1176, doi: 10.1109/GUCON.2018.8674894.

[12] J. Zakraoui, J. M. AlJa'am and I. Salah, "Domain-Specific Text Generation for Arabic Text Summarization," 2022 International Conference on Computer and Applications (ICCA), Cairo, Egypt, 2022, pp. 1-4, doi: 10.1109/ICCA56443.2022.10039630.

[13] J. Chen and H. Zhuge, "Extractive Text-Image Summarization Using Multimodal RNN," 2018 14th International Conference on Semantics, Knowledge and Grids (SKG), Guangzhou, China, 2018, pp. 245-248, doi: 10.1109/SKG.2018.00033.

[14] R. Sheik and S. J. Nirmala, "Deep Learning Techniques for Legal Text Summarization," 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Dehradun, India, 2021, pp. 1-5, doi: 10.1109/UPCON52273.2021.9667640.

[15] R. Boorugu and G. Ramesh, "A Survey on NLP-based Text

Summarization for Summarizing Product Reviews," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 352-356, doi: 10.1109/ICIRCA48905.2020.9183355.

[16] C. Kwatra and K. Gupta, "Extractive and Abstractive Summarization for Hindi Text using Hierarchical Clustering," 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2021, pp. 1-6, doi: 10.1109/ICSES 52305.2021.9633789.

[17] S. Alhojely and J. Kalita, "Recent Progress on Text Summarization," 2020 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2020, pp. 1503-1509, doi: 10.1109/CSCI51800.2020.00278.

[18] K. Padmanandam, S. P. V. D. S. Bheri, L. Vegesna and K. Sruthi, "A Speech Recognized Dynamic Word Cloud Visualization for Text Summarization," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 609-613, doi: 10.1109/ICICT 50816.2021.9358693.

[19] K. Lanyo and A. Wausi, "A Comparative Study of Supervised and Unsupervised Classifiers Utilizing Extractive Text Summarization Techniques to Support Automated Customer Query Question-Answering," 2018 5th International Conference on Soft Computing & Machine Intelligence (ISCMI), Nairobi, Kenya, 2018, pp. 88-92, doi: 10.1109/ISCMI.2018.8703237.

[20] C. Limploypipat and N. Facundes, "Abstractive Text Summarization for Covid-19 News with Data Augmentation," 2022 International Conference on Digital Government Technology and Innovation (DGTi-CON), Bangkok, Thailand, 2022, pp. 56-59, doi: 10.1109/DGTi-CON53875. 2022.9849194.

[21] S. Gani, J. Uddin and M. I. Mobin, "Extractive Text Summarization Technique Using Fuzzy C-Means Clustering Algorithm," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2019, pp. 1-6, doi: 10.1109/IC4ME247184.2019.9036642.

[22] R. S. Shini and V. D. A. Kumar, "Recurrent Neural Network based Text Summarization Techniques by Word Sequence Generation," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 1224-1229, doi: 10.1109/ICICT50816.2021.9358764.

[23] J. Zakraoui, J. M. AlJa'am and I. Salah, "Domain-Specific Text Generation for Arabic Text Summarization," 2022 International Conference on Computer and Applications (ICCA), Cairo, Egypt, 2022, pp. 1-4, doi: 10.1109/ICCA 56443.2022.10039630.

[24] H. Mehta, S. K. Bharti and N. Doshi, "Text summarization using Secretary problem," 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur, Odisha, India, 2022, pp. 1-5, doi: 10.1109/iSSSC56467.2022.10051552.

[25] M. I. Liaqat, I. Hamid, Q. Nawaz and N. Shafique, "Abstractive Text Summarization Using Hybrid Technique of Summarization," 2022 14th International Conference on Communication Software and Networks (ICCSN), Chongqing, China, 2022, pp. 141-144, doi: 10.1109/ICCSN 55126.2022.9817599.

[26] B. N, D. Kumari, B. N, M. N, S. K. P and S. R. A, "Text Summarization using NLP Technique," 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, India, 2022, pp. 30-35, doi: 10.1109/DISCOVER55800.2022. 9974823.

[27] Kuber, S. Kulthe, Y. Kulkarni, R. Pimple and P. Kosamkar, "Extensive Study of Automatic Text Summarization on Biomedical Texts," 2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA, Pune, India, 2022, pp. 1-7, doi: 10.1109/ ICCUBEA54992.2022.10010753.

[28] S. Avasthi, A. Prakash, T. Sanwal, M. Tyagi and S. Yadav, "Tourist reviews summarization and sentiment analysis based on aspects," 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2023, pp. 452-456, doi: 10.1109/Confluence 56041.2023.10048863.

[29] S. -K. Im and K. -H. Chan, "Multilayer CARU Model for Text Summarization," 2023 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, Korea, Republic of, 2023, pp. 399-402, doi: 10.1109/BigComp57234.2023. 00098.