NLP based Text Summarization Techniques for News Articles: Approaches and Challenges

Sara Tarannum[1], Piyush Sonar[2], Aashi Agrawal[3] Krishnai Khairnar[4]

[1]Student, Dept. of Information Technology, SVKMs Institute of Technology, Dhule, Maharashtra, India
[2]Student, Dept. of Information Technology, SVKMs Institute of Technology, Dhule, Maharashtra, India
[3]Student, Dept. of Information Technology, SVKMs Institute of Technology, Dhule, Maharashtra, India
[4]Student, Dept. of Information Technology, SVKMs Institute of Technology, Dhule, Maharashtra, India

---

Abstract - *The amount of text data available has increased dramatically in recent years from a variety of sources. This large volume of literature has a wealth of information and knowledge that must be adequately summarized in order to be useful. Text summarizing is one of the jobs that has received a lot of attention, but there hasn't been any practical application for it aside from tasks based on Extractive summarization, such as news or book summarization. Text summarization creates an automatically generated summary that includes all pertinent significant information from the original material. Observing the summary results, extractive and abstractive techniques are one of the most common. Extractive summarizing is maturing, and research is currently shifting to abstractive summarization and real-time summarization. The results of the analysis provide an in-depth explanation of the trends that are the focus of their research in the field of text summarization, pre-processing and features that have been used, describes the techniques and methods that are often used by researchers as a comparison and means for developing methods. In this study, various recommendations are made regarding opportunities and problems in text summarization research, and why abstraction and hybrid summarization methods produce the best results.*

Key Words*: *News Summarization, Abstractive Summarization, Automatic News Summary, NLP, Hybrid Summarization

1. INTRODUCTION

People have become overwhelmed by the vast amount of information and documents available on the internet as the internet and big data have grown in popularity. Many researchers are motivated to develop a technology solution that can automatically summaries texts as a result of this. The task of compressing a piece of text into a shorter version, lowering the size of the original text while keeping crucial informational aspects and content meaning, is known as summarization. Because manual text summarizing is a time-consuming and inherently tedious activity, automating it is expanding in popularity and so serves as a powerful motivator for academic study.

The purpose of automatic text summarization is to reduce the length of documents or reports while preserving crucial information. [5] Text summarization can be used in a variety of places, including news stories, emails, research papers, and online search engines to get a quick summary of the results.

[1] Text summary is the act of condensing a long text into a smaller, more accurate, and fluent set of phrases that is easy to grasp and provides the reader with the information they need in a limited amount of words. [1] Extractive summarization and abstractive summarization are the two primary text summarizing approaches. The summary is generated using extractive based summarization, which extracts the relevant phrases and sentences from the original text. The statistical properties of the sentences are taken into account when determining the most important sentences from the text. Abstractive summarization recognizes the major ideas in the original text and creates a summary in plain English. The linguistic qualities of the sentences are taken into account here. [2]

Both extractive and abstractive summarization have their own set of advantages and disadvantages. Extractive summarization selects the most important and correct phrases but suffers from incoherency, whereas abstractive summarization provides a word-by-word summary that is crisp and readable but may lose significant data in large manuscripts. [11]

The majority of extant text summary applications are extractive, with only a few producing abstractive summaries. For constructing multi-document summaries, graph-based algorithms for abstractive text summarizing have demonstrated to be superior to previous methods. Knowledge graphs provide a summary that is more in line with human reading patterns and has the logic of human reasoning. As a result, we suggest a possible solution to this problem by gathering crucial information from news websites and extracting only the substance in the form of an automatically created summary, allowing newsreaders to make more informed decisions in less time. [3]

In this paper, we describe a unique approach for extractive summarization with sentiment analysis for two-level text summarizing from online news sources. Important sentences from various news stories pertaining to a topic are extracted first, and individual summaries are prepared. Sentiment analysis is used to go further into the individual summaries for each topic. [4]

2. Background

2.1 Motivation:

The amount of text data available from various sources has exploded in the big data era. This large volume of literature has a wealth of information and knowledge that must be adequately summarized in order to be useful. Because of the growing availability of documents, extensive research in the field of natural language processing (NLP) for automatic text summarization is required. The job of producing a succinct and fluent summary without the assistance of a person while keeping the sense of the original text material is known as automatic text summarization. [20] It's difficult since, in order to summarize a piece of literature, we normally read it completely to gain a better knowledge of it and then write a summary stressing its important points. Because computers lack human language and understanding, automatic text summarization is a complex and time-consuming operation.

Capturing the most relevant components of a document, which serve as a condensed highlight of its content, is critical for saving time and allowing end-users to consume material quickly. There has been an increasing demand for automatic text summarizing solutions as manual summarization becomes un-scalable and hence unfeasible at such a large content influx rate. Because so much of this content is accessed on mobile devices, on-device text summarizing can assist content providers display their information succinctly while limiting cloud involvement to save bandwidth and protect privacy. Because of the limitations of storage space and processing resources, on-device summarization algorithms remain mostly studied. [12]

This large volume of literature contains a wealth of knowledge and facts that must be effectively summarized in order to be useful. This problem's main purpose is to automate text summarizing. As the Internet has developed tremendously, people are becoming overwhelmed by the vast amount of online knowledge and data. The increased availability of papers necessitates substantial research into automatic text summarizing. Text summarization is the subject of a lot of research these days. As the amount of information available on the internet expands, incidents like this are becoming increasingly typical. [13]

The use of an automatic summary technique is a smart way to deal with information overload. It takes a single or a group of papers as input and generates a brief summary that highlights the most relevant details. Personalization, the focus of this study, is both a cause and a reaction to the issues that the traditional media face in relation to their websites. The issues stem in large part from online audience consumption behaviors as well as the economics of advertising, which is the key source of revenue for online news sites. [19]

For this goal, various machine learning methods have been proposed. The majority of these approaches model this topic as a classification problem that determines whether or not a sentence should be included in the summary. Topic information, Latent Semantic Analysis (LSA), Sequence to Sequence models, Reinforcement Learning, and Adversarial processes have all been applied in other techniques.

2.2 Classification of Summarization Techniques:

Summarization is based on numerous factors: news extraction methods, input data, and type of output, summary generation methodology, and similar news recommendations. A combined analysis of these factors determine the priority of the summarization methodology that will be utilized to create the most optimal news summary. The summarization parameters are classified as follows:

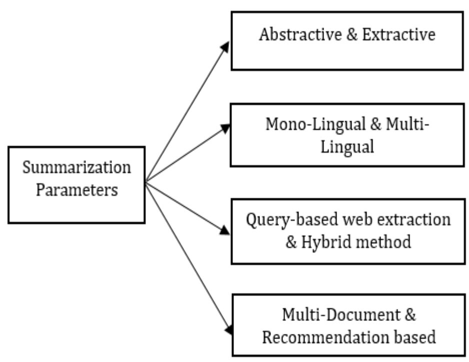

Fig -1: Sample Table format

2.3 Method of Automatic Summarization

2.3.1 Abstractive: Abstractive summarization approaches try to synthesize a new shorter text by analyzing the original text using advanced natural language algorithms. A neural network reads the text, encodes it, and then generates target text in these models, which are made up of an encoder and a decoder. [3]

2.3.2 Multi-Lingual: Not all avid readers are native English speakers. Hence, apart from English, there is a need for inter-translation of each summary for ease of reading. So, multi-lingual translation approach has been chosen.

2.3.3 Recommendation Based: To eliminate the risk of losing out on related news from different sources, provide a broad overview in less time, and reduce resource bias, recommendation-based summarizing was chosen over multi-document summarization as a single summary. Because combining news from several sources into a single summary fails to distinguish which news parameter was collected from which specific news source, the reader has no idea where any given news snippet came from.

2.3.4 Hybrid Method: Each of their flaws may be overcome by each of their summarizing breakthroughs, hence the hybrid method with a combination of query based extraction and web-based extraction was chosen over the individual methods. This method aids in the narrowing down of news phrases for summation to particular criteria in terms of utility, indicative, and informative differentiation for better summary production.

2.4 Detailed walkthrough of survey method

When the paper article search stage is completed, a large number of articles that meet the requirements will be filtered out during the search adjustment procedure. The inclusion and exclusion process provided the basis for determining the article criteria used in the main study. Aside from that, we need a strategy like the one shown in Fig. 1 to produce limited paper articles that meet the research topic for later review. The first paper article found after automatically filtering titles, abstracts, and key terms during a preliminary search. Then select the major paper article that best suits the entire text's contents. The papers' final results will be compiled and assessed for further analysis. Mendeley's software is used to organize search results so that they can be grouped according to specific subjects.

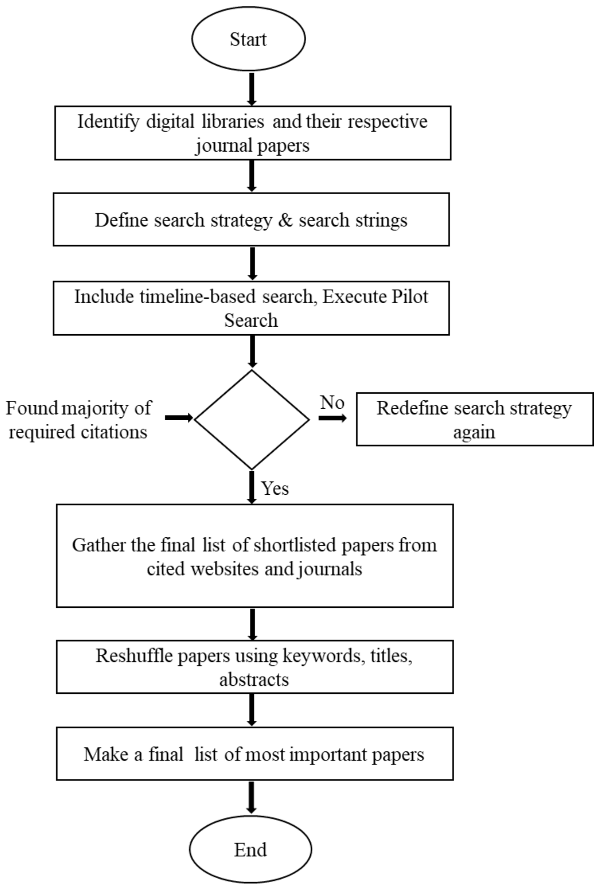

Fig -2: Survey Process

2.4.1 Data abstraction and extraction

Extractive Summarization: Extractive summarization uses a scoring mechanism to extract sentences from a document and combine them into a logical summary. This method works by detecting key chunks of the text, cutting them out, and then stitching them back together to create a shortened form.

The following is a typical extractive summarization system flow:

- Creates an intermediate representation of the input text in order to discover the most important information. TF metrics are computed for each sentence in the provided matrix in most cases.
- Scores the sentences based on the representation, assigning a number to each sentence that indicates the likelihood of it being included in the summary.
- Generates a summary from the top k most important sentences. Latent semantic analysis (LSA) has been utilized in several research to identify semantically significant sentences. [20]

Abstractive Summarization: Abstractive approaches make use of recent deep learning advancements. Abstractive methods take advantage of the recent success of sequence to sequence models since it can be thought of as a sequence mapping task where the source text should be mapped to the target summary. A neural network reads the text, encodes it, and then generates target text in these models, which are made up of an encoder and a decoder. [20]

Alexander et al. [28] suggested a neural attention model for abstractive sentence summary (NAMAS), which investigated a fully data-driven strategy for creating abstractive summaries utilising an attention-based encoder-decoder method. The attention mechanism is widely employed in sequence to sequence models, in which the decoder pulls information from the encoder depending on the source-side information's attention ratings. Other abstractive summarization approaches have incorporated concepts from Vinyals et al [29]'s pointer network to solve the undesired behaviour of sequence to sequence models. The Pointer Network is a sequence-to-sequence architecture based on neuronal attention that learns the conditional probability of an output sequence from discrete tokens corresponding to places in an input sequence.

To improve model accuracy, some abstractive summarization projects have taken principles from the reinforcement learning (RL) field. Chen et al. [30], for example, suggested a hybrid extractive-abstractive architecture that uses two neural networks in a hierarchical manner to choose important lines from a source using an RL guided extractor and then rewrites them abstractly to provide a summary. In other words, the model mimics how humans summarise extensive documents by first utilising an extractor agent to choose salient sentences or highlights, and then rewriting each of these extracted sentences using an abstractor — an encoder-aligner-decoder model — network. To connect both the extractor and abstractor networks and learn sentence saliency, the model leverages policy-based reinforcement learning (RL) with sentence-level metric incentives to train the extractor on accessible document-summary pairings. [20]

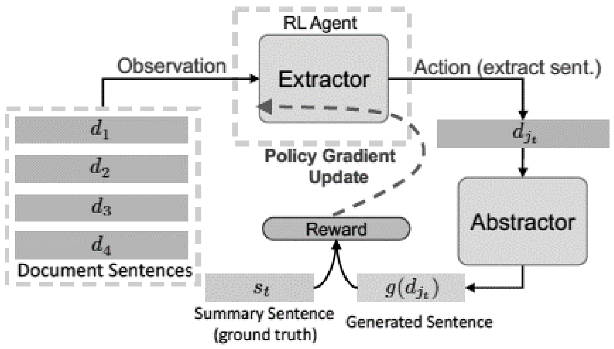

Fig -3: Abstractive Summarization

2.4.2 Processing walkthrough for news summarization

Extraction: This module is in charge of transforming raw data. It entails steps like as discovering the data source, extracting relevant entities from the page, transforming them into usable forms for further processing, and archiving them in an indexed format for future use. Pre-processing is a key stage that includes identifying, removing stop words, language stemming, permitting input in the correct format, and removing duplicate sentences or words. [31]

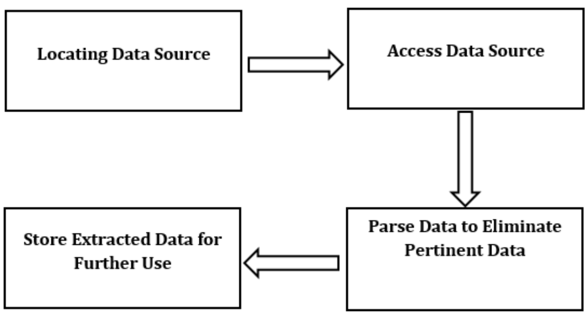

Fig -4: Extraction

Summarization: The summarizing process can be described as reducing the size of a document so that it contains the most relevant information in the fewest possible words. The result can be obtained by filtering out words and phrases and then ranking the relevance of the remaining elements based on the filter scores. These scores, which are based on a threshold ranking methodology, produce a good summary, guaranteeing that only the most important aspects stay in the body of the text.

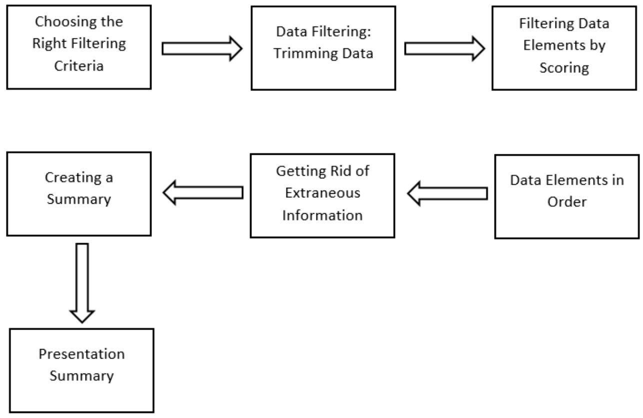

Fig -5: Summarization

3. METHOD

3.1 Systematic Literature Review was used to perform this review research on text summarizing (SLR). SLR (Okoli and Schabram, n.d.) is a method for identifying, evaluating, and interpreting research results that have been conducted as a whole and are relevant to the issue field or research questions that attempt to provide answers to research questions, such as text summarizing study.

ID	Questions	Solutions
R1	Which are the papers related to news summarization	Categorize the most relevant journal/conference papers for news summarization
R2	What kind of dataset is used in news summarization?	Identify the commonly used datasets in news summarization
R3	What is the research trend of news summarization?	Identify the research motive of news summarization
R4	What are the walkthrough methods used in news summarization?	Identify the process of news summarization
R5	What features are included in news summarization?	Identify the topographies taken into consideration for news summarization
R6	What are the common approaches for news summarization?	Collect the common methods of news summarization
R7	What are the average drawbacks	Identify the drawbacks of the existing news summarization

	of news summarization?	methods
R8	What are the evaluation techniques used in news summarization?	Identify how the existing systems evaluate the news summarization methods

Table -1: Research questions and motivations

Data Extraction

R2, R4 are the important research questions while R1 and R3 are designated to get a precise idea of where to begin. R5, R6, R7, R8 are used to evaluate the existing drawbacks of summarization, highlight available features, data extraction and processing techniques.

Classification Criteria	Question Number
Type of publication	R1
Research abstract and trend	R3
Existing Features	R5
Potential summarization drawbacks	R7
Data Extraction	R2, R4
Data Processing	R6, R8

Table -2: Classification Criteria

3.2 Selection criteria for research papers concerning news summarization:

3.2.1 Include*:* Topics, challenges, datasets, methodologies, and methods used are all included in studies that summarize literature. The articles and papers from the conference that particularly explored text summarization were used in this study. The studies that were taken were from 2015 to 2021.

3.2.2 Exclude: Uncertainty exists in studies that employ a dataset but do not have experimental outcomes. Studies that go beyond just summarising texts Non-English-language studies

3.3 Search Strategy*:*

3.3.1 Site Based*:* ieeexplore.ieee.org, sciencedirect.com, some external case studies, reports, and Scopus Journals are some of the data sources used for reference document access.

3.3.2 Search String Based*:* Enter the following keywords or synonyms of the keywords derived from the research topic being done to find papers that fit the topic. The following is a search string that was utilised throughout the paper search process: news summarization, extractive summarization, abstractive summarization, hybrid summarization, and automatic NLP summarization (OR technique OR method). Only English-language papers are included in the collection.

3.3.3 Timeline Based*:* The most recent advances in terms of news summarising approaches and algorithms were examined when searching for publishing years 2015-2021. Look for publications in journals and at conferences.

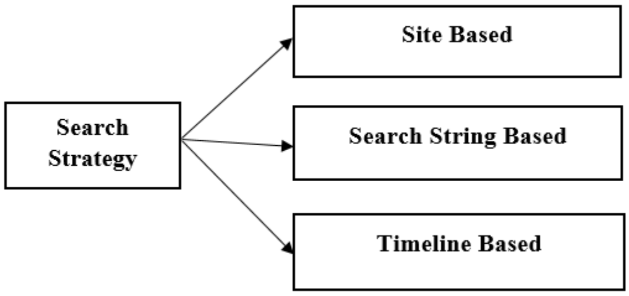

Fig -6: Search Strategy

4. RESULTS:

4.1 Characteristics of improved, efficient summary extraction technique:

- To obtain news stories, multiple sources must be made available. The system must adapt to each source, learning as it encounters new forms and minimising processing time.
- To ensure proper and comprehensive extraction of protected data formats on the Internet, special attention must be taken. The degree of relevancy of the articles to the query that was fired must be determined.
- Before the cleaning process begins, articles extracted may contain random and malicious content, as well as various marketing ads or external links, all of which must be trimmed and removed.

4.2 Methods to navigate around summarization of news articles:

- A user-friendly interface for sorting through queries and filters to aid in the extraction process. The use of a web crawler is required to automate the extraction of online-based news articles.
- In a real-time scenario, a web crawler with parallel vertical search is more efficient and faster than a general search method in retrieving data from that source.
- A mapping mechanism was devised to assess the query's relevancy to the articles.
- A storage solution for indexing and referencing previously sorted files in preparation for future use. Extractions are updated in real time to ensure that real-time data is always available.

4.3 Characteristics of abstractive news summarization

- Include all pertinent information from the unprocessed article. The importance of a sentence must be quantified in terms of the input data file as well as its relation to the query that was executed.
- The extraction process must be in sync in order for the processor to receive real-time data.
- Opinions and indirect semantics that affect news must be handled with efficiently while not being fully eliminated.
- Sentences and paragraphs must be connected, and a logical and grammatical flow must be established. Indicative summaries with references to source articles are preferred.
- Using the pointer-generator network, copying words from the source text is simple.
- The pointer-generator paradigm can even replicate terms from the source text that aren't in the dictionary. This is a significant benefit, as it allows us to deal with invisible words while simultaneously limiting our vocabulary (which requires less computation and storage space).
- The pointer-generator model is easier to train than the sequence-to-sequence attention system, requiring fewer training iterations to reach the same level of performance.

4.4 Ideal method to approach abstractive news summarization

- Encoder — A bi-directional LSTM layer that extracts data from the source text. This is highlighted in red above.
- The bidirectional LSTM reads one word at a time, and because it's an LSTM, it modifies its hidden state based on the current word and previous words.
- Decoder — A one-word-at-a-time LSTM layer that creates summaries. When the decoder LSTM receives the signal that the entire source text has been read, it begins to function.
- It creates the probability distribution over the next word using information from the encoder as well as what it has written before.
- Attention Mechanism — Encoder and Decoder are the building blocks here, however without attention, encoder decoder architecture hasn't been particularly successful in the past.
- Without paying attention, the final concealed state from the encoder is the input to the decoder, which can be a 256 or 512 dimension vector, and since this little vector can't possible contain all of the information, it becomes an information bottleneck.
- The attention method allows the decoder to examine the encoder's intermediate concealed states and use all of that information to determine which word comes next.

5. ANALYSIS:

5.1 Comparison of data-extraction methods:

	Suleiman et al	Adam et al	Ziyi et al
Algorithm	Generative Pre-trained Transformer 2	RSS Feed Crawling	Blocking Tag Tool
Data Source	Dedicated database with Topic-based webpages	RSS News Feeds on the internet	Web Page
Method for accessing data source	Semantic web-based on the queries by the user	Train to add new RSS and interval detection of modification in RSS for updation	HTML code scanned to detect block tag and extract clean page from repository or detection system.
Data Parsing	Weighted query subgraph relations to page subgraph for relevance and sematic scores	Parallel parsing of RSS on the web for regular updation and creation of metadata using counters	Cleaned page is analyzed for news content. If failed the detection process is repeated.
Storage	Page access and ranking using ordered result set to produce search list	Centralized server database to store URL and metadata	The pages are stored in a database using URL and metadata
Advantage	No restriction on pages thus removing bias from the result.	Real-time updated data available at all times	Processed data ready for further use.
Disadvantage	The graphs constructed are concept-based thus restricting the variety of searches	No categorization available for stored data	Block detection mechanism is rudimentary

5.2 Comparison of summarization techniques:

	Mirani and Sasi	Retzaputra and Khotra	Sabha et al
Raw data format	Categorized data sets in corpus with multiple files	Manual multiple document feed into system	Single Document is fed into system

Data Trimming	Tokenization and Regular Expression of HTML files	Tokenization and PoS tagging of txt files	Features like title, length, weight and data points are extracted using fuzzifier
Filtering Criteria	TF-ISF on words	Relationships between sentence structs form clusters	Triangular membership functions prioritize internal sentences
Ranking Data elements	Weighted scoring of sentences based on TF-ISF of words	Sentence similarity score using MMR tree relation	Fuzzy sets classify based on 5 priorities- very low to very high
Eliminating excessive details	Sentimental Analysis on polarity and subjectivity	Sentence similarity defines priority and eliminates copies or irrelevant data	Priorities are used to feed abstractive element that uses vocabulary generation to rebuild sentence structure
Constructing Summary	Sentences ranked and sorted with polarity index	Sentence from each cluster is strung together using NLP until word limit is reached	Attention mechanisms applied on LSTM to form context vector and article distribution
Summary Presentation	Saved as a text file with summaryparagraph	Produces a text file with summary paragraph	Produces an output on system console
Advantage	Sentimental Analysis provides perspective instead of bare facts	The relation deviated from basic words to sentence structure for tree relations	The use of feature extraction as a filter makes the system subjective
Disadvantage	No logical relation for sentence arrangement	The sentence structure identification is not accurate	Article distribution is not relevant in most forms

6. CONCLUSIONS

Automatic text summarization is a fascinating academic topic with a wide range of commercial applications. By reducing huge volumes of information into brief bursts, summarizing can aid with a variety of downstream applications, including news digests, report production, news summarization, and headline construction. The most often used summarizing algorithms are of two types.

Extractive summarizing methods begin by reordering and copying passages from the source material. Second, by rephrasing or inserting terms not contained in the original text, abstractive summarization approaches generate new phrases. The great majority of past work has been extractive due to the difficulty of abstractive summarization.

The extractive approach is more convenient since it assures grammar and accuracy by copying large portions of text from the source document. Advanced abilities like paraphrase, generalization, and assimilation of real-world knowledge, on the other hand, are only possible in an abstractive framework and are required for high-quality summary. Despite the fact that abstractive summarization is a more challenging task, there has been some success thanks to recent advances in the deep learning sector.

7. REFERENCES:

1. S. Singh, A. Singh, S. Majumder, A. Sawhney, D. Krishnan and S. Deshmukh, "Extractive Text Summarization Techniques of News Articles: Issues, Challenges and Approaches," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 2019, pp. 1-7, doi: 10.1109/ViTECoN.2019.8899706.

2. V. Alwis, "Intelligent E-news Summarization," 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer), 2018, pp. 189-195, doi: 10.1109/ICTER.2018.8615590.

3. M. V. P. T. Lakshika, H. A. Caldera and W. V. Welgama, "Abstractive Web News Summarization Using Knowledge Graphs," 2020 20th International Conference on Advances in ICT for Emerging Regions (ICTer), 2020, pp. 300-301, doi: 10.1109/ICTer51097.2020.9325453.

4. T. B. Mirani and S. Sasi, "Two-level text summarization from online news sources with sentiment analysis," 2017 International Conference on Networks & Advances in Computational Technologies (NetACT), 2017, pp. 19-24, doi: 10.1109/NETACT.2017.8076735.

5. P. Sethi, S. Sonawane, S. Khanwalker and R. B. Keskar, "Automatic text summarization of news articles," 2017 International Conference on Big Data, IoT and Data

Science (BID), 2017, pp. 23-29, doi: 10.1109/BID.2017.8336568.

6. J. Wu, "Web News Summarization via Soft Clustering Algorithm," 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009, pp. 618-621, doi: 10.1109/FSKD.2009.838.

7. H. Yu, "Summarization for Internet News Based on Clustering Algorithm," 2009 International Conference on Computational Intelligence and Natural Computing, 2009, pp. 34-37, doi: 10.1109/CINC.2009.194.

8. M. E. İlter, L. Akarun and A. Özgür, "Identifying Image Related Sentences in News Articles," 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019, pp. 1-4, doi: 10.1109/SIU.2019.8806258\

9. K. Chen et al., "Extractive Broadcast News Summarization Leveraging Recurrent Neural Network Language Modeling Techniques," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 8, pp. 1322-1334, Aug. 2015, doi: 10.1109/TASLP.2015.2432578.

10. M. Steinert, R. Granada, J. P. Aires and F. Meneguzzi, "Automating News Summarization with Sentence Vectors Offset," 2019 8th Brazilian Conference on Intelligent Systems (BRACIS), 2019, pp. 102-107, doi: 10.1109/BRACIS.2019.00027.

11. M. S M, R. M P, A. R E and E. S. G SR, "Text Summarization Using Text Frequency Ranking Sentence Prediction," 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP), 2020, pp. 1-5, doi: 10.1109/ICCCSP49186.2020.9315203.

12. M. P. Dhaliwal, R. Kumar, M. Rungta, H. Tiwari and V. Vala, "On-Device Extractive Text Summarization," 2021 IEEE 15th International Conference on Semantic Computing (ICSC), 2021, pp. 347-354, doi: 10.1109/ICSC50631.2021.00065.

13. P. Raundale and H. Shekhar, "Analytical study of Text Summarization Techniques," 2021 Asian Conference on Innovation in Technology (ASIANCON), 2021, pp. 1-4, doi: 10.1109/ASIANCON51346.2021.9544804.

14. R. Rawat, P. Rawat, V. Elahi and A. Elahi, "Abstractive Summarization on Dynamically Changing Text," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1158-1163, doi: 10.1109/ICCMC51019.2021.9418438.

15. H. Tauseef and M. Asfand-E-Yar, "Extractive Summarization of Text Using Supervised and Unsupervised Techniques," 2021 3rd International Conference on Natural Language Processing (ICNLP), 2021, pp. 37-41, doi: 10.1109/ICNLP52887.2021.00012.

16. M. Jain and H. Rastogi, "Automatic Text Summarization using Soft-Cosine Similarity and Centrality Measures," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1021-1028, doi: 10.1109/ICECA49313.2020.9297583.

17. Rahul, S. Adhikari and Monika, "NLP based Machine Learning Approaches for Text Summarization," 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 535-538, doi: 10.1109/ICCMC48092.2020.ICCMC-00099.

18. S. R. K. Harinatha, B. T. Tasara and N. N. Qomariyah, "Evaluating Extractive Summarization Techniques on News Articles," 2021 International Seminar on Intelligent Technology and Its Applications (ISITIA), 2021, pp. 88-94, doi: 10.1109/ISITIA52817.2021.9502230.

19. S. Huang, R. Wang, Q. Xie, L. Li and Y. Liu, "An Extraction-Abstraction Hybrid Approach for Long Document Summarization," 2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC), 2019, pp. 1-6, doi: 10.1109/BESC48373.2019.8962979.

20. T. B. Mirani and S. Sasi, "Two-level text summarization from online news sources with sentiment analysis," 2017 International Conference on Networks & Advances in Computational Technologies (NetACT), 2017, pp. 19-24, doi: 10.1109/NETACT.2017.8076735.

21. E. ShafieiBavani, M. Ebrahimi, R. Wong and F. Chen, "A Query-Based Summarization Service from Multiple News Sources," 2016 IEEE International Conference on Services Computing (SCC), 2016, pp. 42-49, doi: 10.1109/SCC.2016.13.

22. Thurman, N. and Schifferes, S., "The Future of Personalisation at News Websites: Lessons from a Longitudinal Study. Journalism Studies", 13(5-6), doi: 10.1080/1461670X.2012.664341

23. Opgenhaffen, Michael & d'Haenens, Leen "The Impact of Online News Features on Learning from News: A Knowledge Experiment", 2011 The International Journal of Internet Science. 6. 8-28.

24. Adhika Pramita Widyassari, Supriadi Rustad, Guruh Fajar Shidik, Edi Noersasongko, Abdul Syukur, Affandy Affandy, De Rosal Ignatius Moses Setiadi, Review of automatic text summarization techniques & methods, Journal of King Saud University - Computer and Information Sciences, 2020, ISSN 1319-1578.