

Received 31 October 2022, accepted 13 December 2022, date of publication 20 December 2022,  
date of current version 29 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3231016



## SURVEY

# Feature Based Automatic Text Summarization Methods: A Comprehensive State-of-the-Art Survey

DIVAKAR YADAV<sup>ID1</sup>, (Senior Member, IEEE), RISHABH KATNA<sup>2</sup>, ARUN KUMAR YADAV<sup>1</sup>,  
AND JORGE MORATO<sup>ID3</sup>

<sup>1</sup>School of Computer and Information Sciences (SOCIS), Indira Gandhi National Open University (IGNOU), Maidan Garhi, New Delhi 110068 India

<sup>2</sup>Department of Computer Science and Engineering, National Institute of Technology Hamirpur (NIT Hamirpur), Hamirpur, Himachal Pradesh 177005, India

<sup>3</sup>Department of Computer Science, Universidad Carlos III de Madrid, 28911 Leganés, Spain

Corresponding author: Jorge Morato (jmorato@inf.uc3m.es)

**ABSTRACT** With the advent of the World Wide Web, there are numerous online platforms that generate huge amounts of textual material, including social networks, online blogs, magazines, etc. This textual content contains useful information that can be used to advance humanity. Text summarization has been a significant area of research in natural language processing (NLP). With the expansion of the internet, the amount of data in the world has exploded. Large volumes of data make locating the required and best information time-consuming. It is impractical to manually summarize petabytes of data; hence, computerized text summarization is rising in popularity. This study presents a comprehensive overview of the current status of text summarizing approaches, techniques, standard datasets, assessment criteria, and future research directions. The summarizing approaches are assessed based on several characteristics, including approach-based, document-number-based, Summarization domain-based, document-language-based, output summary nature, etc. This study concludes with a discussion of many obstacles and research opportunities linked to text summarizing research that may be relevant for future researchers in this field.

**INDEX TERMS** Abstractive summarization, cosine-similarity, deep learning, extractive summarization, graph-based algorithm, neural networks.

## I. INTRODUCTION

The World Wide Web (WWW) has become an immense information resource. Today, some websites generate more data every day than was produced in the previous ten years combined. However, the majority of the data generated by these websites is irrelevant, redundant, and loud, masking the most pertinent information. In addition, users must explore several files and web pages to find the information they seek. It wastes the time of many users. A strong document summary can fix the aforementioned issue. If every online page provided a concise summary of its content, it would save time for many users and boost website engagement. However, it is not possible to manually summarize each web page on the World Wide Web. Automated text summarization (ATS)

The associate editor coordinating the review of this manuscript and approving it for publication was Seifedine Kadry .

technologies can resolve the issue. Consequently, ATS has become a focus of NLP study.

ATS systems are designed to accomplish objectives like as extracting the most important and relevant information from a document, generating summaries that are much shorter than the original content, etc. The ATS systems can be categorized generally into one of the following categories:

- Single-document summarization system:* This type generates a single summary for a single document.
- Multi-document summarization system:* The generation of a single summary for multiple documents is performed in this type.

These systems are more susceptible to duplication and inaccuracy due to the fact that various documents may contain identical sentences representing different information (inaccuracy) and different sentences representing identical information (redundancy).

There are three primary methods for generating summaries:

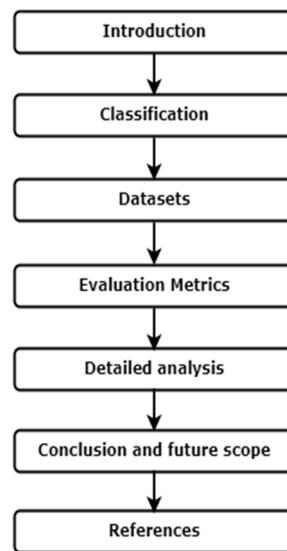
- a. *Extractive approach*: In this approach, important sentences from a document are picked and combined to generate a final summary. Major steps in an extractive approach include:
  - i. Document pre-processing
  - ii. Create a provisional representation of the document
  - iii. Score sentences according to their retrieval value
  - iv. Select the sentences with the highest scores.
- b. *Abstractive approach*: This strategy seeks a much deeper comprehension of the document. Instead than selecting meaningful sentences directly, it generates new sentences that convey the same information using natural language processing algorithms. Important steps in an abstractive approach include:
  - i. Preprocessing the document
  - ii. Making an Intermediate representation of the document
  - iii. Generating new sentences based on IR.
- c. *Hybrid approach*: This approach combines both the abstractive and the extractive approaches to generate the summary.

Automatic Text summarization is one of the most challenging areas of text and data mining. There are numerous obstacles associated with developing high-quality automated summaries mentioned as below:

- (i) *Redundancy*: Most ATS systems generate phrases with similar informational content. Because the size of the summaries is limited, more valuable and diverse information-carrying sentences may not be included in the summary. It may result in the loss of crucial information.
- (ii) *Time-zones for multi-document summarization*: Different documents in a dataset can belong to different time zones. Hence, they might use temporal words to convey different meanings. It is a big challenge in the multi-document summarization.
- (iii) Generating short summaries for very large documents like novels, books, etc.
- (iv) Generated summaries may not maintain a proper flow. This is more significant in extractive text summarization.

These significant challenges in text summarization are the focus of intense research. Nevertheless, certain models perform better than others in certain criteria, such as abstractive summarizers' ability to maintain a decent flow and decrease repetition, but they cannot solve the remaining problems.

Numerous research articles have been published on this subject. Survey papers are vital for imparting concept knowledge to a novice audience and offering information on current trends and future horizons in a single document. Some survey papers covered a specific subdomain of text summarization: Jain et al. [1] surveyed on legal document summarization;



**FIGURE 1.** Paper workflow.

Al-Saleh and Menai [2] on Arabic text summarization techniques; Kumar et al. [3] on multilingual text summarization; and some studies ([4], [5]) attempted to provide an overview of the entire field of text summarization.

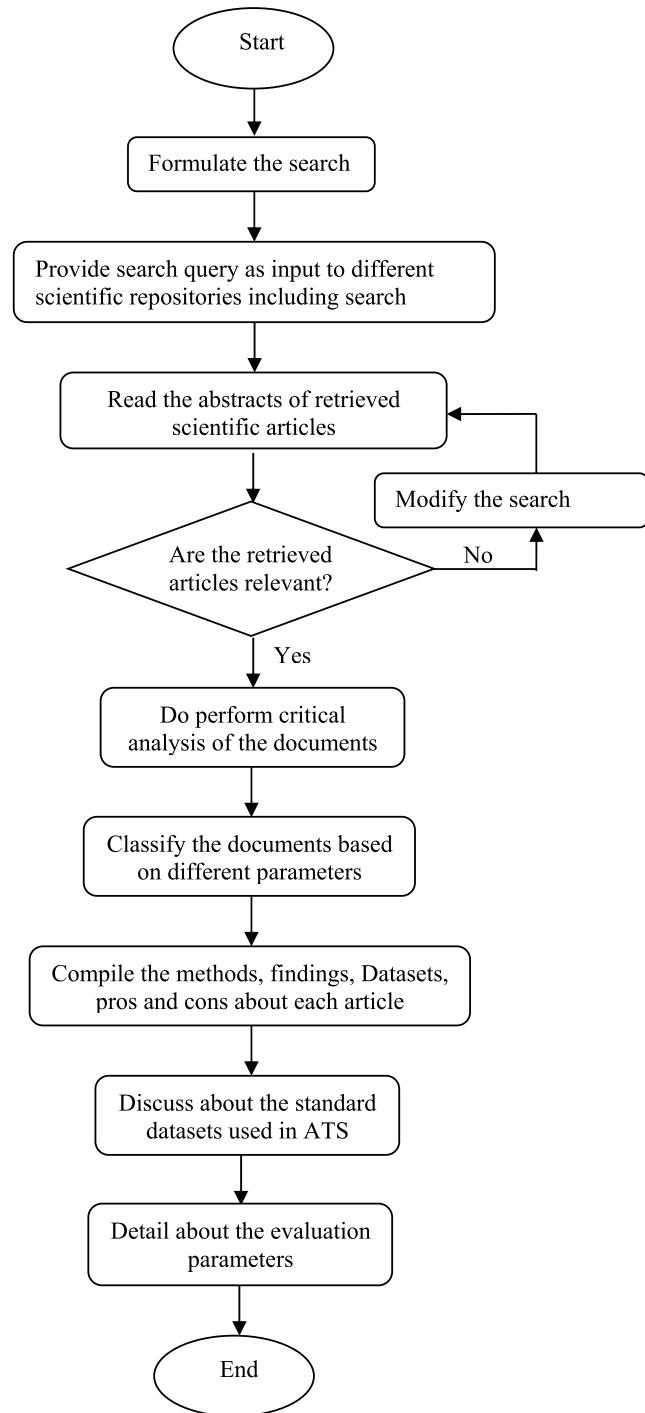
Existing survey articles, however, do have limitations. Either the information covered is minimal ([1], [2], [6]), the articles examined are outdated and do not address the most recent developments in this subject, or the information supplied is difficult to comprehend. By presenting a succinct, up-to-date, and comprehensible overview of the topic of text summarizing, this survey paper overcomes these drawbacks of prior publications.

In this paper, we explore the various classifications of Text summarizing approaches based on several parameters such as methodology, document count, language, etc. We also briefly address investigations undertaken within each classification. We listed the outcomes, benefits, and drawbacks of each study. Finally, we present a comprehensive review of the performance of various approaches on prominent datasets. However, a comprehensive analysis of each study is outside the scope of this work. In addition, this paper discusses the most popular and effective methodologies, as a comprehensive treatment of all approaches would exceed its scope.

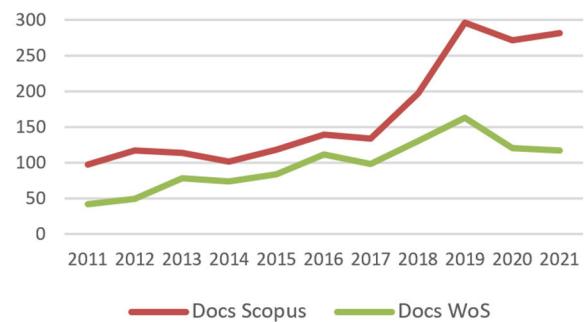
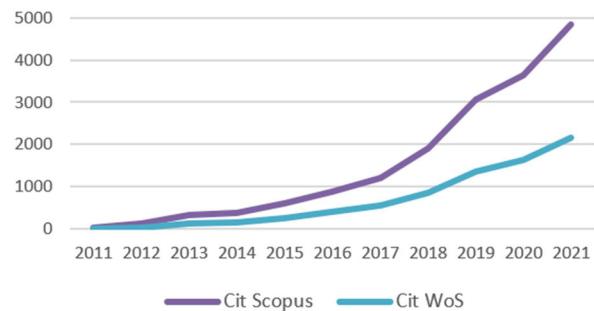
The main contributions of this study are as follows:

- Provides a tabular and comprehensive analysis of different studies, making it easy for the reader to compare and evaluate various methodologies
- Describes the benefits and drawbacks of each study analysed in this paper.
- Offers a comprehensive analysis of numerous strategies and their performance on popular datasets.
- A comprehensive discussion of future horizons, recommended methods, and research directions.

The flow of this paper is explained in the diagram in FIGURE 1.

**FIGURE 2.** Procedure adopted to perform the study of scientific articles.

This paper's body is divided into numerous sections. The first section provides a quick bibliographic analysis of the growing interest in the topic and identified tendencies. The classification of text summarizing algorithms based on various factors is discussed in Section 2. Section 3 enumerates the assessment criteria employed by various studies to compare and contrast their systems with those of others. Section 4 provides a listing of the significant datasets utilized

**FIGURE 3.** Number of documents about ATS in Scopus and WoS in the period 2011-2021.**FIGURE 4.** Number of citations obtained by the documents collected from Scopus and WoS in the period 2011-2021.

in the research described in section 2. Section 5 demonstrates alternative methods of classifying ATS. In section 6, we conduct a comprehensive examination of the prevalent strategies for text summarizing and provide some observations on the enhancements and results gained by other investigations. The seventh segment discusses the difficulties of text summary, followed by a conclusion in the last portion.

## II. A BRIEF BIBLIOMETRIC STUDY ON THE EVOLUTION OF THE FIELD

Following is a brief literature overview demonstrating how interest in the topic has progressed (figure 2). Following this is a classification of approaches according to the approach taken by the various summarization systems.

Regarding the academic interest generated by reputable publications, a study of the works published in the past few years is informative. FIGURE 2 depicts the primary approach used to classify and analyze scientific papers. This diagram's sequence is based on the principles provided in [7] and [8].

A search was conducted in the Web of Science (WoS) and Scopus databases to determine the evolution of the works published in the field. Their selection reflects the fact that they are the data sources with the most extensive coverage and the greatest prevalence in bibliometric research. Both resources are complimentary because their geographical scopes and journal collections are distinct [9]. In addition, the journals included in these databases are chosen based on their quality and influence.

**TABLE 1.** Queries performed in scopus and wos core collection databases.

Resource	Query	# docs
Scopus	SUBJAREA(COMP) PUBYEAR AFT 2010 AND PUBYEAR BEF 2022 AND (TITLE (summar* OR "abstracting") AND TITLE(text OR "document*" OR "multi-document") AND TITLE- ABS-KEY(automatic* OR "extractive" OR "abstractive" OR "hybrid" OR "text summarization" OR "automatic summarization" OR "multi-document" OR document* OR text* OR "ATS"))	1865
WoS Core	WC="computer science" AND PY=(2011-2021) AND (TI= (summar* OR "abstracting") AND TI=(text OR "document*" OR "multi- document") AND AB=(automatic* OR "extractive" OR "abstractive" OR "hybrid" OR "text summarization" OR "automatic summarization" OR "multi-document" OR document* OR text* OR "ATS"))	1064

**TABLE 2.** Languages with more than 5 works in the scopus collection (english excluded).

Language	# docs	Language	# docs
Arabic	65	Spanish	15
Chinese	63	French	9
Bangla/Bengali	28	German	9
Indonesian	21	Japanese	9
Malay	18	Punjabi	8
Vietnamese	18	Russian	8
Turkish	18	Kannada	8
Hindi	17	Portuguese	7
Persian	16	Telugu	7

Given that we aim to study the current trajectory in computing, we restricted our search to the years 2011 through 2021. The executed queries are listed in TABLE 1.

Figures 3 and 4 reveal a strong upward trend that has become more evident since 2018. In the previous two years, there has been a slowdown, although this may be related to the time required to update the database's publications. The trend of citations is quite progressive, indicating a focus on the achievements made in ATS during the period. In fact, the h-index in WoS is 39 while in Scopus it is 56.

Most systems are language-dependent, and the dearth of native speakers or digital resources in certain languages impedes study. Analyzing the summaries, titles and keywords in Scopes show that most of the language's studies are amongst the most spoken languages in the world (TABLE 2).

Several observations can be made regarding the number of works published in the various languages: 1) the number of works does not reflect the number of speakers; for example, Nigerian pidgin is the 14th most spoken language, but it is not mentioned in the results; 2) There are languages among the 30 more spoken that have no study, such as Cantonese, Tagalog, Hausa, Swahili, Nigerian, and Javanese; 3) Indian languages are well represented: Bengali (28), Hindi (17), Punjabi (8), Kannada (8), Telugu (7), KonKonkani (5), Assamese (4), Tamil (2), or Marathi (2). However, the representation of Hindi, the third-most-spoken language, is inadequate, and other languages, such as Nepali, are not mentioned.

### III. EVALUATION METRICS

Automatic text summarizing approaches are evaluated using performance measurement measures, as is the case with all other methods. These metrics are discussed in this section.

#### A. ROUGE (RECALL-ORIENTED UNDERSTUDY OF GISTING EVALUATION)

It is the most popular evaluation metric used in the field of text summarization. ROUGE has four types:

- a) **ROUGE-N:** In this metric, N stands for N-grams co-occurrence statistics. It measures the quality of a summary using n-gram recall between the summary and a set of manually generated summaries as shown in Eq. (1).

$$\text{ROUGH} - N = \frac{\sum_{S \in \{\text{Reference summaries}\}} \sum_{\text{gram}_n \in S} \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S \in \{\text{Reference summaries}\}} \sum_{\text{gram}_n \in S} \text{Count}(\text{gram}_n)} \quad (1)$$

- b) **ROUGE-L:** Here, L stands for longest common substring. A sentence is represented as a set of words. The longer the LCS between our summary and manual summary sentences, the better the quality of the summary.
- c) **ROUGE-W:** Here, W stands for weighted LCS. It tries the limitation of LCS that it cannot differentiate LCSs of different spatial relations within their word embedding.
- d) **ROUGE-S:** S stands for skip-bigrams co-occurrence statistics. Skip-bigrams are bigrams that do not have to appear together in a sentence. For the sentence "I am Ram", the skip-bigrams generated will be {("I", "am"), ("I", "Ram"), ("am", "Ram")}.

ROUGE-S uses skip-bigrams to compute the similarity between our generated summary and manual summaries.

#### B. GENERIC PERFORMANCE METRICS

##### 1) PRECISION

It is computed by dividing the number of sentences common in the Reference and Candidate summary by the number of sentences in the candidate summary as shown in Eq. (2).

$$\text{Precision} = N(S_r \cap S_c) / N(S_c) \quad (2)$$

where,

$Sr := \text{Reference summary}$

$Sc := \text{Candidate summary}$

$N(S) := \text{Number of sentences in summary } S.$

### 2) RECALL

It is computed by dividing the number of sentences common in the Reference and Candidate summary by the number of sentences in the reference summary as shown in Eq (3).

$$\text{Recall} = N(Sr \cap Sc)/N(Sr) \quad (3)$$

where,

$Sr := \text{Reference summary}$

$Sc := \text{Candidate summary}$

$N(S) := \text{Number of sentences in summary } S.$

### 3) F-MEASURE

It is computed by computing the harmonic mean between precision and recall as shown in Eq. (4).

$$F = 2(\text{Precision})(\text{Recall})/(\text{Precision} + \text{Recall}) \quad (4)$$

### C. SUPERT

Summarization evaluation with Pseudo references and BERT (SUPERT) is an un-supervised summary evaluation metric for evaluating multi-document summary by measuring the semantic similarity between the summary and the pseudo reference summary. SUPERT was made by [81]. The limitation of ROUGE is that it needs manual summaries to judge the quality of a summary. SUPERT can be used to summarize a dataset that does not have manual summaries.

## IV. DATASETS FOR TEXT SUMMARIZATION

In this section, we discuss about the popular dataset, used for text summarization methods among researchers.

### A. DOCUMENT UNDERSTANDING CONFERENCES (DUC)

The National Institute of Standards and Technology (NIST) provides these groups of datasets. DUC is part of a Defense Advanced Research Projects Agency (DARPA) program, Translingual Information Detection, Extraction, and Summarization (TIDES), explicitly calling for major advances in summarization technology. The datasets consist of the following parts:

- Documents
- Summaries, results, etc.
  - manually created summaries
  - automatically created baseline summaries
  - submitted summaries created by the participating groups' systems
  - tables with the evaluation results
  - additional supporting data and software

DUC distributed seven datasets from 2001 to 2007. DUC 2002 is the most popular dataset for extractive summarization ([23], [38], [56], [70]). These datasets are available at <https://duc.nist.gov/data.html>.

### B. CNN/DAILY MAIL

It contains over 300,000 articles from CNN and Daily Mail. The dataset is generated using a python script available at CNN [71]. The processed version of this dataset is available on GitHub [72]. It is a very popular dataset among extractive ([58], [73]) and abstractive summarization studies ([44], [66]).

### C. OPINOSIS

It is a dataset constructed from user reviews on a given topic. It is very suitable for semantic analysis and has been used by multiple studies for the same purpose. It consists of 51 topics, with each topic having 100s of review sentences. It also comes with gold standard summaries and some scripts to evaluate the performance of a summarizer using ROUGE metric. The dataset and related material can be downloaded from Opinosis [74]. This dataset was prepared by [45], [75], and [76] for their research.

### D. GIGAWORD

This dataset consists of more than 4 million articles. It is a part of TensorFlow dataset collections and is highly popular among abstractive summarization studies [77]. The source code for this dataset is available at Gigaword [78].

### E. MEDLINE CORPUS

The MEDLINE corpus is provided by NLM (National Library of Medicine). NLM produces this dataset in the form of XML documents on an annual basis. This dataset can be downloaded from [79]. Shang et al. [59] used this dataset to develop an extractive summarizer.

### F. LCSTS (LARGE SCALE CHINESE SHORT TEXT SUMMARIZATION DATASET)

It is a Chinese text summarization dataset. This dataset consists of 2 million short texts from a Chinese microblogging website Sina Weibo. It is also provided with short summaries for each blog, written by the blog authors. It is a very suitable choice for Chinese abstractive summarization systems as the dataset size is large and it can be used to train neural networks efficiently. Li et al. [77] used this dataset to develop an encoder-decoder based abstractive text summarizer.

### G. BC3(BRITISH COLUMBIA UNIVERSITY DATASET)

The corpus is composed of 40 email threads (3222 records) from the W3C corpus. Each thread is commented on by three different commenters. The dataset consists of:

- (i) Extracted abstracts
- (ii) Abstract abstracts with linked sentences

**TABLE 3.** Papers mentioning the ats type in the scopus database (based on title and keyword sections).

2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1	1	2	2	10	10	11	23	51	66	70
12	15	13	12	26	28	38	59	96	98	103
1	1	3	3	2	4	1	7	11	13	8

Yousefi-Azar and Hamey [36] used this dataset to develop a deep learning based extractive text summarizer.

#### H. EASC (ESSEX ARABIC SUMMARY CORPUS)

This dataset consists of Arabic articles and extractive summaries generated for those articles. It is one of the most popular Arabic datasets used in text summarization. Alami et al. [37] and Elayeb et al. [80] used this dataset for Arabic text summarization.

#### I. GEOCLEF

GeoCLEF is used in geographical studies. It consists of 169,447 documents; each document consists of stories and newswires from the Los Angeles Times newspaper (1994) and the Glasgow Herald newspaper (1995). It is used by Perea-Ortega et al. [55] for developing a geographical information retrieval system.

### V. CLASSIFICATION OF SUMMARIZATION APPROACHES

Based on the summarization approach, text summarization can be further divided into 3 main types:

- a. Extractive approach
- b. Abstractive approach
- c. Hybrid approach

The impact of these summarization approaches in the study mentioned above shows a growing of the abstractive types in the last decade (Table 3)

A selection of relevant papers was made based on quality aspects. For each of the approaches to be described below and for each technique applied with that approach, we have selected those articles that, mentioning the technique used in each approach, most clearly and illustratively describe its practical application.

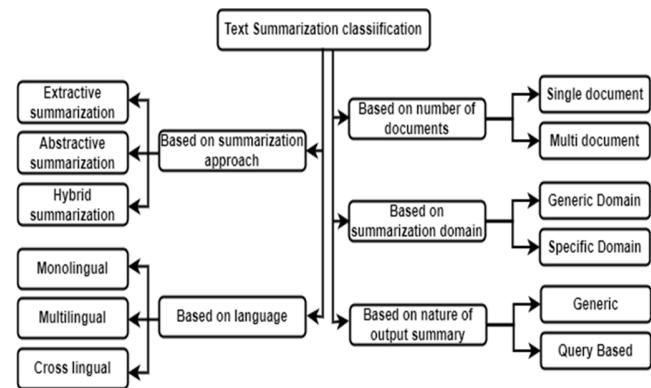
In the remaining of this section, we discuss the classifications of text summarization methods based on different classification parameters. The different classifications of a text summarization system are represented in the FIGURE 5.

In the following subsections, each of these approaches will be discussed.

#### A. EXTRACTIVE TEXT SUMMARIZATION

In this approach, the most important sentences are selected from documents and then assembled to produce the summary. The typical workflow of the extractive-based approach is:

- i. Preprocessing



**FIGURE 5.** Text summarization classification.

- ii. Intermediate representation
- iii. Sentence scoring
- iv. Summary construction and post-processing

The preprocessing and summary construction stages are common for most extractive text summarizers. They are mostly different in terms of techniques for intermediate representation and sentence scoring. Most of the research around extractive text summarization is also focused on these steps. The main extractive text summarization methods are discussed in the following sections.

In the following subsections, we will review the methods employed to each of the main approaches in text summarization. In first place, we will discuss the Extractive Text methods: statistical, topic-based methods, clustering, graph, semantic, machine learning, deep-learning methods, fuzzy-logic techniques, and discourse based (RST). Next, we will discuss the Abstractive Text methods: graph based tree-based, domain specific methods, and deep-learning methods, and finally, the Hybrid Text methods.

#### 1) STATISTICAL-BASED METHODS

In these methods, statistical features are used to compute a sentence's importance. Statistical features may include sentence position [10], sentence length, number of proper nouns in the sentence, term frequency [10], and cosine similarity can be used for computing sentence scores [11] as shown in TABLE 4.

#### 2) TOPIC-BASED METHODS

In this approach, the main topics of a document are extracted. Then the sentences are scored based on their coverage of document topics. TF-IDF [6], Term frequency, Document titles [12] can be used to find document topics. Further, N-gram co-occurrence and semantic sentence similarity can also identify document topics [13] as shown in TABLE 5.

#### 3) CLUSTERING-BASED TECHNIQUES

In this method, the sentences are clustered based on some similarity measure. Then a summarizer extracts the most central sentences from each cluster and processes them to generate a summary. Clustering algorithms like

k-means ([14], [15], [16], [17]), k-medoids [18], etc. are used for sentence clustering as shown in TABLE 6.

#### 4) GRAPH-BASED TECHNIQUES

In these methods, the document is represented as a graph of sentences. The sentences represent the nodes. The edges represent the similarity between the nodes. The similarity between words can be represented using some similarity measures like cosine similarity ([6], [19], [20]). Graph-based techniques are prevalent for extractive summarizers. Popular summarizers such as TextRank [21], LexRank [19] and [22] use a graph-based approach. The sentences are then scored based on the properties of the graph. The summary of such methods is shown in TABLE 7.

#### 5) SEMANTIC-BASED TECHNIQUES

In these methods, sentence semantics are also taken into consideration. LSA (Latent Semantic Analysis), ESA (Explicit Semantic Analysis) and SRL (Semantic Role Labeling) are some ways of doing semantic analysis of textual data. Out of the three, LSA is the most common and is used by most of the studies ([12], [24], [25], [26], [27]) as shown in TABLE 8. Common steps in semantic analysis using LSA is:

- Creating a matrix representation of the input.
- Apply SVD (Singular value decomposition) to capture the relationship between individual terms and sentences.

#### 6) MACHINE-LEARNING-BASED TECHNIQUES

Machine learning approaches have gained popularity in recent years. These techniques convert the text summarization problem into a supervised classification problem, in which each sentence is classified as either a ‘summary’ or ‘non summary’ sentence. In the end, ‘summary’ sentences are collected to generate the summary.

Rather than defining rules manually, the model is trained on a training set. The set consists of documents and their respective human-generated summaries. Various classification techniques like SVM ([27], [28], [29]), Naive-Bayes ([27], [29], [30]), Decision-Trees [30], Ensemble methods ([27], [31], [32]) and neural-network ([33], [34], [35]) have been used for text summarization as shown in TABLE 9.

#### 7) DEEP-LEARNING BASED METHODS

Deep learning techniques are getting more and more popular for text summarization. Seq2seq and encoder-decoder based models [36] are used for extractive text summarization. Alami et al. [37] developed deep learning and clustering-based model for Arabic text summarization. Feed forward neural networks are also being used for extractive summarization [33]. The brief about these methods is shown in TABLE 10.

#### 8) OPTIMIZATION BASED METHODS

In these techniques, the summarization problem is formulated as an optimization problem. The steps involved in an optimization-based technique are as follows:

- Preprocessing and converting the document to an intermediate representation.
- Using an optimization algorithm to extract summary sentences from the IR.

Multi-Objective Artificial Bee Colony algorithm (MOABC) is the most common optimization algorithm used by many studies ([28], [38], [39], [40]) as discussed in TABLE 11.

#### 9) FUZZY-LOGIC BASED TECHNIQUES

In these techniques, Fuzzy-logic based systems are used to compute the sentence scores. Fuzzy-logic techniques are popular because we can represent scores more precisely. The typical workflow of a fuzzy-logic based system is given as under:

- Extracting meaningful features from a sentence. e.g., sentence length, term-weight etc.
- Using a fuzzy system to provide scores to those features. The score ranges between 0 and 1.

Babar and Patil [12], Abbasi-ghalehtaki et al. [28], Azhari and Jaya Kumar [41], and Goularte et al. [42] developed fuzzy-systems based text summarizers. Some studies even integrated different domains like cellular learning algorithms [28] and neural networks [41] with the fuzzy systems to further improve the results as shown in TABLE 12.

#### 10) DISCOURSE BASED

Discourse based studies include analyzing bigger language structures like lexemes, grammar and context and their effect on sentence weights. Rhetorical structure theory (RST) has been used widely by multiple studies ([34], [43]) for discourse analysis and text summarization as shown in TABLE 13.

In recent years, it has been observed that machine-learning, deep-learning, rhetorical structure theory and fuzzy-systems based techniques are getting more popular for extractive text summarization. Hence, for future research, these techniques can be explored extensively.

The main advantages and disadvantages of extractive text summarization are pointed out as below:

- Extractive summarizers are easier to implement than abstractive summarizers.
- Capture more accurate information as sentences are directly extracted from the document without altering the contents.
- Generate more accurate information as this is not how humans generate summaries.
- Multi-document extractive summarization suffers from sentence redundancy.
- Can mix information from different timelines, resulting in wrong summaries.

**TABLE 4.** Studies on statistical-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Kutlu <i>et al.</i> [10]	Term frequency, Keyphrase occurrence, Title similarity and sentence position	Turkish Journal articles	ROUGE-1: 0.369	Developed a useful dataset for Turkish text summarization using Turkish Journal articles.	High redundancy in generated summaries
Bouras & Tsogkas[11]	POS tagger, Web crawler, TFIDF, Cosine similarity	Real-time input from users	Accuracy: 0.654	Enhanced the personalization algorithm of a popular news indexing system, peRSSonal.	The system can incorporate a caching feature to prevent computing summaries repeatedly.

**TABLE 5.** Studies on topic-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Tayal <i>et al.</i> [13]	NLP Parser, POS tagger, Title and topic selection: N-gram co-occurrence, Semantic similarity	DUC 2007	ROUGE-2: 0.13523 ROUGE-SU4: 0.112561	Developed their own Title identification, sentence reduction and combination algorithms.	Cannot control the size of the summary easily

**TABLE 6.** Studies on clustering-based techniques for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Alguliyev et al.[14]	K-means clustering	DUC2002	ROUGE-1: 0.4723 ROUGE-2: 0.2103	Developed a novel mutation strategy for sentence selection.	Did not take the semantics of the sentences into consideration.
Alguliyev et al.[15]	K-means clustering	DUC2002	ROUGE-1: 0.4727 ROUGE-2: 0.2103	Produced better results than the state-of-the-art LexRank method.	Since individual words are used for clustering, lemmatization should have been used rather than stemming for better results
Mohdet al.[16]	K-means	DUC2007	ROUGE -1: 0.34 ROUGE-2: 0.07	Used semantics in integration with statistical features.	Model is computationally expensive and time-consuming
Rouaneet al.[17]	K-means	Binomial central database	ROUGE-1: 0.23840 ROUGE-2: 0.08715 ROUGE-SU4: 0.11456	Proposed novel heuristics for sentence ranking for biomedical texts.	High redundancy in summaries

## B. ABSTRACTIVE TEXT SUMMARIZATION

In this approach, the summary is generated in the same way humans summarize documents. The summary does not consist of sentences from the documents, rather new sentences

are generated by paraphrasing, merging the sentences of the original document. Abstractive text summarization requires a deeper understanding of the input document, the context and the semantics. It also requires some deeper understanding of

**TABLE 7.** Studies on graph-based techniques for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Ferreira <i>et al.</i> [23]	Sentence clustering	DUC 2002	F-measure: 25.4	Designed novel sentence-clustering algorithms to remove redundancy and increase diversity.	Since it treats a document collection as a single input file, different documents written on different dates are combined to form a single file. It can lead to ambiguity.
Erkan &Radev[19]	Connectivity matrix based on intra similarity of sentences. Cosine similarity for measuring sentence similarities	DUC 2003, DUC 2004	ROUGE-1: 0.3666	Showed that the degree-based methods outperform centrality-based methods for sentence ranking for graphs.	Discarded sentences having length < 9, which can discard some important information
Mihalcea&Tarau[21]	TextRank	Read this later	Precision: 31.2 Recall: 43.1 F-measure: 36.2	Gave a state-of-the-art graph-based summarization model.	Low recall scores due to less keywords being selected

NLP concepts. The typical workflow for the abstractive text summarizer is:

- a. Preprocessing
- b. Creating an Intermediate representation
- c. Generating summary from intermediate representation.

In the following subsections, the techniques and methods used in Abstractive Text summarization are discussed.

### 1) GRAPH-BASED METHODS

In these methods, the individual words are taken as the graph's nodes. The edges represent the structure of the sentence. AMR (Abstract Meaning Representation) graphs are popular graph-based text representation methods. Various sentence generators are integrated with AMR graphs for abstractive text summarization [44]. Ganesan et al. [45] developed a popular text summarizer, Opinosis. The brief about these methods is shown in TABLE 14. The processing steps of the OPINOSIS model are as follows:

- The path in the intermediate is considered as the summary.
- The goal is to find the best path.
- To do this, rank all the paths and sort them based on decreasing scores.
- Use a similarity measure metric (e.g., Cosine similarity) to remove redundant paths.
- The best path is chosen for the summary.

### C. TREE-BASED METHODS

In these techniques, parsers convert text documents to parse trees. Then various tree-processing methods like pruning and linearization are used to generate tree summaries. Deep learning models like encoder-decoder neural networks can also be used to generate meaningful information from the parse trees [46]. Techniques like sentence fusion are also used to eliminate redundancy in the generated summary [47]. The further details about these methods are shown in TABLE 15.

### D. DOMAIN-SPECIFIC METHODS

Many studies focus on domain-specific text summarizers. These studies can be benefitted by using knowledge dictionaries unique to each domain. In addition, the sentences that do not hold much importance in normal text summarization can be imperative depending on the domain. Sports news may contain some sport-specific keywords that are important to convey the necessary information about a game, e.g., “out” in cricket is considered an important word that is more significant than other words like “high”. Okumura and Miura [48] developed a sports news summarization system utilizing the above domain characteristics. Lee et al. [49] developed a text summarizer for Chinese news articles. Further details about these methods are shown in TABLE 16.

### E. DEEP-LEARNING BASED METHODS

Advances in deep learning have made abstractive text summarization more approachable. Sequence-to-sequence

**TABLE 8.** Studies on semantic based techniques for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Babar & Patil, [12]	LSA + Fuzzy based system	10 datasets	Precision: 86.91213	Better features were extracted	Use of a small dataset for training
Priya Umamaheswari [24]	LSA + TFIDF	Hotel review and movie review dataset	Precision: 0.58 Recall: 0.83 F-measure: 0.78	Developed a more scalable text summarizer for cloud environments	A better representation of text, like Word2Vec, could have improved the model even further
Mashechkin et al.[25]	LSA Sentence matrix representation	DUC2001, DUC2002	F-measure: 0.18965(LSA square)	Representing the sentences in the topic-space rather than term-space reduced dimensionality	Sentences occurring to topics with small weight might occur in summary.
Yeh et al. [26]	LSA Genetic algorithms	New Taiwan Weekly corpus	Precision: 0.5370 Recall: 0.1964 F-measure: 0.2870	Trained the scoring function using a Genetic algorithm to get the appropriate feature weights	LSA based model is less efficient
Shen et al.[27]	LSA machine learning classifiers like SVM, BM and Ensemble approach	LookSmart web directory	Precision:70.7 Recall: 57.7 F- measure: 63.6	Developed an effective model for noisy web page summarization.	Only evaluated the model on 2 classifiers: NB and SVM.

models is being explored for abstractive text summarization ([50], [51]). Pre-trained transformers are also used for abstractive text summarization [51] as shown in TABLE 17

The main advantages and disadvantages of extractive text summarization are pointed out as below:

- Generate better quality summaries as the sentences are not directly extracted from the document.
- Summaries are safe from plagiarism.
- More complex to implement than extractive summarizers.
- Captures less information as some of the information can be lost while rephrasing the sentences

### 1) HYBRID TEXT SUMMARIZATION

In this approach, a hybrid of extractive and abstractive summarizers generates the summary. Generally, hybrid text summarizers generate better quality summaries than extractive summarizers, and they are less complex than abstractive text summarizers. Lloret et al. [52] developed a hybrid summarization system called Compendium Gupta and Kaur [53] developed a machine learning-based model, and Binwahlan et al. [54] developed a fuzzy-systems based hybrid text summarization model. The details about few of such methods are as shown in TABLE 18. Some of the advantages and

disadvantages of hybrid text summarization are as shown below:

- Generates better quality summaries than pure extractive models.
- It is easier to implement than abstract text summarizers.
- The quality of summaries is less than pure abstractive summarizers.

## VI. OTHER CLASSIFICATION CRITERIA

The following classification shows other criteria for classifying scientific papers:

- a. Classification based on the number of papers: single or multiple.
- b. Classification according to the domain of the abstract
- c. Classification based on the number of languages used.
- d. Classification based on the nature of the output

These classifications are discussed and exemplified below.

### A. BASED ON THE NUMBER OF DOCUMENTS

The text summarization methods based on the number of documents are classified in different categories as discussed in below sections.

**TABLE 9.** Studies on machine learning based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Neto et al.[30]	Naive Bayes + C4.5 Decision Tree	TIPSTER collection	Precision: 34.68 Recall: 1.01	Explained different feature extraction procedures which can be very helpful for different models	Models trained on very small dataset
Bui et al.[31]	BOT Context + Semantic model	Cochrane reviews	Recall: 0912 F-measure: 0.847	Produced more informative summaries at sentence levels than human-generated summaries.	Didn't extract all data elements like age, and sex of the author has to be extracted using multiple parsing algorithms.
Shen et al.[27]	NB, SVM, Random Forest	LookSmart web directory	Precision: 70.7 Recall: 57.7 F- measure: 63.6	Developed an effective model for noisy web page summarization.	Only evaluated the model on 2 classifiers: NB and SVM.
Alami et al.[32]	Word2Vec model	EASC (Essex Arabic summary corpus)	ROUGE-1: 0.1299	Merged Word2Vec and BOW models to improve performance	High implementation cost.

**TABLE 10.** Studies on deep learning-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Alami et al.[37]	NN_Sent2Topic_prob model K-means clustering	ESSEX Arabic summaries corpus	ROUGE-1(Recall): 0.2937	Deep produced better results and solved the inference problem	VAE Didn't take vocalization into consideration. Arabic
Sinha et al.[33]	Feed-forward neural networks	Wikihow	ROUGE-1: .551	Developed a large dataset with large diversity of topics rather than just news articles.	Large topics and writing styles further complicate abstractive text summarization, hence an even larger dataset must be used for deep learning
Yousefi-Azar & Hamey [36]	Ensemble Noisy Auto encoder (ENAE)	SKE, CN3	ROUG-2(Recall): 0.1370	Implemented a noise induction feature in term frequency to further improve recall scores	A word2Vec representation must have produced better results than term-frequency

### 1) SINGLE-DOCUMENT

In this type, the summary is generated for a single document. It is easier than multi-document text summarization as the single document has generally only one topic and is written in a single period. It is less prone to redundancy than multidocument text summarization. Perea-Ortega et al. [55], Sankarasubramaniam et al. [56],

Abbasi-ghalehtaki et al. [28], and Alguliyev et al. [14] developed single document text summarizers, as shown in TABLE 19.

### 2) MULTI-DOCUMENT

In this type, a single summary is generated for multiple documents. It is more complex than single document text

**TABLE 11.** Studies on optimization-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Abbasi-ghalehtaki <i>et al.</i> [28]	MOABC cellular automation theory	+ DUC2002	ROUGE-1: 0.4865	Introduced the GWO (Grey wolf optimization) algorithm.	Too language-dependent, hence porting to another language will become very difficult.
Sanchez-Gomez <i>et al.</i> [38]	Parallel MOABC	DUC2002	An efficiency of 86.72% has been reported for 64 threads	Made a more efficient version of MOABC algorithm	Use of GPU could have improved performance even further
Sanchez-Gomez <i>et al.</i> [40]	MOABC	DUC2002	2.23% improvement for ROUGE-2 metric	Compared different summary quality criteria for text summarization i.e. redundancy, coverage, relevance and coherence.	Did experiments on extractive summarization in which sentences are already picked from the story. A study for abstractive summarizers should have been done also.

**TABLE 12.** Studies on fuzzy systems-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Abbasi-ghalehtaki <i>et al.</i> [28]	Fuzzy systems + cellular learning algorithms	DUC2002	ROUGE-1: 0.4865	Introduced the GWO (Grey wolf optimization) algorithm.	Too language-dependent, hence porting to another language will become very difficult.
Azhari& Jaya Kumar [41]	Fuzzy systems + neural network	DUC2002	Precision: 0.86704 Recall: 0.64000	Integrated neural networks and fuzzy systems to improve model performance.	Did not compare results with state-of-the-art fuzzy logic based methods.
Goularte <i>et al.</i> [42]	Fuzzy systems	DUC2003, DUC2006	Precision:0.417 Recall:0.398 F-measure:0.406	Provided a better f-measure than other methods	Better training dataset could have improved the results even more

**TABLE 13.** Studies on rst-based methods for extractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Xu <i>et al.</i> [34]	RST Parsing and Language generation	CNN/Daily Mail	ROUGE-1:43.77 ROUG-2: 20.85 ROUGE-L:40.67	Produced better rouge scores than the traditional BERT model	Less grammatical sentences compared to original sources.
Hou & Lu [43]	RST + domain knowledge	RST-DT	SGD score: 71.1	Better informative summaries are generated.	Performance not measured using standard metric;hence, results cannot be compared to the state-of-the-art methods directly.

**TABLE 14.** Studies on graph-based methods for abstractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Dohareet al.[44]	AMR (Abstract meaning representation) graph	CNN/Daily mail,	ROUGE-2: 21.5	Listed some limitations of ROUGE metric for abstractive text summarization evaluation.	Model is generalizable to other domains.
		AMR Databank	ROUGE-L: 35.2		
Ganesan et al.[45]	Opinosis graph	Opinosis	F-measure: 0.3271 Precision: 0.4482 Recall: 0.2831	Requires no pretraining and domain knowledge.	Model cannot group sentences on a deeper semantic level.

**TABLE 15.** Studies on tree-based methods for abstractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Song et al.[46]	Shift-reduce parser + encoder-decoder model	Gigawords	Precision: 45.17 Recall: 39.77 F-measure: 41.06	Generated grammatically acceptable sentences and sentences are true to the original content.	Produce only single sentence summaries.
Barzilay & McKeown [47]	Sentence fusion	Newsblaster	Precision: 65 Recall: 72 F-measure: 68	Introduced sentence fusion which is very useful for abstractive text summarization.	Improper placement of punctuation in the generated sentence

**TABLE 16.** Studies on domain-based methods for abstractive text summarization.

Ref. No	Domain	Methods	Dataset	Result	Advantages	Disadvantages
Lee et al. (2003), [49]	Chinese news	NN-based fuzzy model	ChinaTimes corpus	ROUGE-1: 0.3825	Developed a novel fuzzy system and ontology-based method for Chinese news.	Summarizer hard coded for ChinaTimes news.
Okumura et al.(2015), [114]	sports	Mecab, TF-weights, keyGraph weights	Mainichi News Corpus 2012	Label precision: 15.4	Developed a highly scalable document labelling algorithm for cloud systems.	Assumed every word has only one sense.

summarization as the documents may refer to different periods. In addition, different documents may cover different topics, which makes multi-document text summarization more challenging Ferreira et al. [23], Nguyen et al. [57], Barzilay and McKeown [47], Xu and Durrett [58], and Patel et al. [20] developed multi-document text summarizers as discussed in TABLE 20.

#### B. BASED ON THE SUMMARIZATION DOMAIN

Based on summarization domain, text summarization is of two types: generic domain-based text summarization

and specific domain based text summarization as discussed below:

##### 1) GENERIC DOMAIN TEXT SUMMARIZATION

This type of text summarization is based on without having a specific domain. In this type of summarization, the importance of a sentence, keyword or key phrase depends on its grammatical properties, e.g., proper nouns, numerical terms and references can be given higher importance. It is more common than domain-specific summarization as these algorithms tend to perform well in different domains but may end

**TABLE 17.** Studies on deep learning-based methods for abstractive text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Alami <i>et al.</i> [32]	Word2vec rep + neural network	DUC2002, Arabic news dataset	ROUGE-1: 0.5347	Merged Word2Vec and BOW models to improve performance	High implementation cost.
Al-Maleh& Desouki[50]	SSTM encoder	Arabic DS(AHS) Arabic headline summary-300,000entries	ROUGE-1-P: 57.02 ROUGE1-R: 21.05 ROUGE1-F: 30.67	Used a length penalty to improve the performance of the pointer-generator framework.	The dataset size can be expanded for better training of the model.
Khandelwal <i>et al.</i> [51]	Seq2seq framework for encoder-decoder model	CNN/Daily Mail	ROUGE-1: 39.18 ROUGE-2: 17.00 ROUGE-L: 36.33	Very high rouge score (13.1) with using only 1% training data.	High implementation cost.
Alami <i>et al.</i> [37]	NN_Sent2Topic_prob + k means clustering	ESSEX Arabic summaries corpus.	Precision:60.64 Recall:47.20 F-measure:53.08	Deep VAE produced better results and solved the inference problem	Did not take Arabic vocalization into consideration.

**TABLE 18.** Studies on hybrid text summarization.

Ref. No	Methods	Dataset	Result	Advantages	Disadvantages
Gupta & Kaur[53]	SVM	Hindi Unicode Based Corpus, 2. AZIT (Punjabi news corpus)	ROUGE-2:0.83	Location-based features improved the performance of the statistical model.	Punjabi words which contain numeric information are not used as a feature.
Lloret <i>et al.</i> [52]	Word graph representation, Pagerank algorithm,	50 research articles	Recall:38.66 Precision:41.81 F-measure:40.20	Provided a vocabulary independent model for biomedical summarization.	Could not solve Excessive Dereferenciation Timeout Error i.e., queries demanding excessive dereferencing computation.
Binwah lan <i>et al.</i> [54]	MMI (maximum marginal importance), Fuzzy system, Particle swarm-based algorithm	DUC2002	Recall: 0.43962 Precision: 0.49548 F-measure:0.4589	Made a hybrid of swarm algorithms and fuzzy systems to improve results.	Unimportant semantic features were included.

up losing some important domain information in summary Ferreira et al. [23], Babar and Patil [12], and Al-Maleh and

Desouki [50] worked on generic text summarizers as shown in TABLE 21.

**TABLE 19.** Studies on single document text summarization.

Ref. No	Dataset	Results	Advantages	Disadvantages
Perea-Ortega <i>et al.</i> [55]	GeoCLEF	Recall: 0.8803	Improved efficiency of geographical information retrieval task	Less accurate summaries when compared to other state-of-the-art methods
Sankarasubramanian <i>et al.</i> [56]	BMIR-J2	ROUGE-1: 0.46	Less redundancy than other statistical approaches.	Context misinformation due to openness of Wikipedia.
Abbasi-ghalehtaki <i>et al.</i> [28]	DUC2002	ROUGE-1: 0.4865	Introduced the GWO (Grey wolf optimization) algorithm.	Too language-dependent, hence porting to another language will become very difficult.
Alguliyev <i>et al.</i> [14]	DUC2001, DUC2002	ROUGE-1: 0.4723 ROUGE-2: 0.2103	Developed a novel mutation strategy for sentence selection.	Didn't take the semantics of the sentences into consideration.

**TABLE 20.** Studies on multi-document text summarization.

Ref. No	Dataset	Results	Advantages	Disadvantages
Ferreira <i>et al.</i> [23]	DUC2002	F-measure: 25.4	Designed novel sentence-clustering algorithms to remove redundancy and increase diversity.	Since it treats a document collection as a single input file, different documents written on different dates are combined to form a single file. It can lead to ambiguity.
Nguyen <i>et al.</i> [57]	DUC2005	ANOVA: 13.021	Took users' cognitive style into the assessment of summary qualities	Generated summary is too biased towards a user's interests. Since users' interests differ, a different summary will have to be generated for different types of users.
Barzilay & McKeown [47]	Newsblaster	Precision: 65 Recall: 72 F-measure: 68	Introduced sentence fusion which is very useful for abstractive text summarization.	Improper placement of punctuation in the generated sentence
Xu & Durrett [58]	CNN/Daily Mail	ROUGE-1: 41.6	Produced grammatically acceptable sentences.	Removed stop words, which lead to underestimation of some bigrams by the model.

## 2) SPECIFIC DOMAIN TEXT SUMMARIZATION

This type of text summarization is concerned with a specific domain. In this type, the importance of a sentence, keyword or key phrase depends not only on its grammatical properties but also on its relation to the domain of study. This approach can capture better domain-specific summaries as some keywords, key phrases which are important for some domains, may not hold much importance in others. Shang et al. [59], Reeve et al. [60], Rouane et al. [17] or

Moradi et al.[61] worked on biomedical summarization, Farzindar and Lapalme [62] and Jain et al. [1] on legal documents, Perea-Ortega et al. [55] on the geographical study and Lloret et al. [52] on scientific paper summarization as shown in TABLE 22.

### C. BASED ON LANGUAGE

Based on language, the text summarization methods are classified in different categories as discussed in the section below:

**TABLE 21.** Studies on generic domain-based summarization.

Ref. No	Dataset	Results	Advantages	Disadvantages
Ferreira <i>et al.</i> [23]	DUC2002	F-measure: 25.4	Designed novel sentence-clustering algorithms to remove redundancy and increase diversity.	Since it treats a document collection as a single input file, different documents written on different dates are combined to form a single file. It can lead to ambiguity.
Babar & Patil [12]	DUC2001	Precision: 86.91213	Better semantic features were extracted	Use of a small dataset for training
Al-Maleh & Desouki [50]	AHS (Arabic Headline Summary)	ROUGE-1-P: 57.02 ROUGE-1-R: 21.05 ROUGE-1-F: 30.67	Used a length penalty to improve the performance of the pointer-generator framework.	The dataset size can be expanded for better training of the model.

**TABLE 22.** Studies on domain-specific text summarization.

Ref. No	Domain	Dataset	Results	Advantages	Disadvantages
Shang <i>et al.</i> [59]	Biomedical research	MEDLINE corpus	ROUGE-1: 0.3104	Summary includes the “cause”, “treatment” and “pathogenesis” of the diseases.	Sometimes the important sentences are characterized as “too general”.
Reeve <i>et al.</i> [60]	Biomedical Research	24 biomedical papers	ROUGE-2: 0.12653	Integrated statistical features with lexical chain analysis, resulting in better Rouge scores.	The dataset size was very small. A bigger dataset like MEDLINE corpus could have better evaluated the proposed model.
Farzind ar& Lapalm e [62]	Legal documents	Federal court of Canada dataset	F-measure: 0.90	Developed a novel sentence segmentation method for legal document summarization.	High redundancy in extracted summaries.
Perea-Ortega <i>et al.</i> [55]	Geography	GeoCLEF	Recall: 0.8803	Improved efficiency of geographical information retrieval task	Less accurate summaries when compared to other state-of-the-art methods

### 1) MONOLINGUAL

In this type of summarization, the document and the summary are in the same language. Perea-Ortega et al. [55] and Sankarasubramaniam et al. [56] worked on summarizers for the English language Al-Maleh and Desouki [50] worked on the Arabic text summarization, shown in TABLE 23.

### D. MULTILINGUAL

In this type of summarization, the document and the summary are written in multiple languages. Rani and Lobiyal [63] made a summarizer for Hindi and English as shown TABLE 24.

### E. CROSS LINGUAL

In this type of summarization, the document is of one language and the summary is generated in some other language. Linhares Pontes et al. [64] developed a French to English text summarizer as shown Table 25.

### F. BASED ON NATURE OF OUTPUT SUMMARY

Based on the nature of the output summary, the summarization methods are classified in to two categories as discussed below:

**TABLE 23.** Studies on mono-lingual text summarization.

Ref. No	Language	Dataset	Results	Advantages	Disadvantages
Perea-Ortega <i>et al.</i> [55]	English	GeoCLEF	Recall: 0.8803	Improved efficiency of geographical information retrieval task	Less accurate summaries when compared to other state-of-the-art methods
Sankarasubramanian <i>et al.</i> [56]	English	BMIR-J2	ROUGE-2(Recall): 0.1370	Less redundancy than other statistical approaches.	Context misinformation due to openness of Wikipedia.
Al-Maleh & Desouki [50]	Arabic	AHS (Arabic Headline Summary)	ROUGE1-P: 57.02 ROUGE1-R: 21.05 ROUGE1-F: 30.67	Used a length penalty to improve the performance of the pointer-generator framework.	The dataset size can be expanded for better training of the model.

### 1) GENERIC

The output is not influenced by external factors. The generated summary is not controlled by external queries Babar and Patil [12], Gupta and Kaur [53], Sankarasubramanian et al. [56], and Chatterjee and Sahoo [65] developed non-query-based text summarizers as shown in Table 26.

### 2) QUERY-BASED

The summary can be controlled using user-defined queries. The summary is generated based on the user requirements. This approach is prevalent among search engines depending on the query. Some sentences can have more importance than others. Shang et al. [59], He et al. [66], Salton et al. [67], and Van Lierde and Chow ([68], [69]) developed query based models for text summarization as shown Table 27.

## VII. ANALYSIS OF POPULAR TEXT SUMMARIZATION TECHNIQUES

In this section, we are going to perform a detailed analysis of the various popular text summarization techniques. These techniques have always been a popular choice among researchers as they are well researched, efficient, and have the most tendency to be improvised on. We will also analyze studies incorporating these techniques, their results, and enhancement ideas will also be discussed.

### A. K-MEANS CLUSTERING

In this algorithm, an unlabeled dataset is divided into ‘k’ number of clusters. Items in each cluster have properties similar to each other.

For text summarization, k-means can be used to cluster sentences containing similar information. This can be helpful in removing redundant sentences and improving overall summary quality.

Alguliyev et al. [15] used the K-means algorithms on the DUC 2002 dataset and got a ROUGE-1 score of 0.4727 Mohd et al. [16] employed a k-means-based model on the DUC 2007 dataset and got a ROUGE-1 score of 0.34. This clearly indicates that k-means is a promising technique in text summarization and can produce great results.

### B. LSA (LATENT SEMANTIC ANALYSIS)

In this method, a document is first converted into a term-to-sentence matrix. This representation can be then used to collect information about the words that occur commonly together. That information can then be used to generate quality summaries. The performance of LSA-based models is further improved using SVD(Singular Value Decomposition).

Babar and Patil [12] used LSA with a fuzzy system model to get a precision of 0.8654. Priya and Umamaheswari [24] used LSA with TF-IDF on a Hotel review dataset to get an accuracy of 0.54. LSA based models can produce significant results, most modern studies are shifting towards neural network-based models. However, an LSA model alongside a neural network-based model can definitely achieve some interesting results.

### C. TEXTRANK

In this method, a document is represented in the form of a graph. Each node of the graph represents a word, and the edges between two nodes represent the relationship between two words. It also applies a voting mechanism such that nodes having more incoming edges are given higher ranks. Also, while ranking a node the ranks of the nodes casting the vote are taken into consideration [21].

### D. LEXRANK

Like TextRank, it is also a graph-based voting algorithm. In this algorithm, the nodes of the graph are represented by the sentences of the document and the edges

**TABLE 24.** Studies on multilingual text summarization.

Ref. No	Language	Dataset	Results	Advantages	Disadvantages
Rani & Lobiyal[63]	English and Hindi	Hindi and English novels	ROUGE2-P: 0.23	Developed a novel POS tagger for the Hindi language.	Use of only statistical features. Integrating semantic features can improve performance.

**TABLE 25.** Studies on cross-lingual text summarization.

Ref. No	Language	Dataset	Results	Advantages	Disadvantages
Linhares Pontes et al.[64]	French - English	MultiLing pilot 2011 dataset	ROUGE-1: 0.4743	Better grammaticality than other approaches.	Sentence compressor removed too much relevant info, resulting in less informativeness.

**TABLE 26.** Studies on generic output-based text summarization.

Ref. No	Results	Advantages	Disadvantages
Sankarasubramaniam et al.[56]	ROUGE-2(Recall): 0.1370	Less redundancy than other statistical approaches.	Context misinformation due to openness of Wikipedia.
Gupta & Kaur [53]	ROUGE-2: 0.83	Location-based features improved the performance of the statistical model.	Punjabi words which contain numeric information are not used as a feature.
Babar & Patil [12]	Precision: 86.91213	Better semantic features were extracted.	Use of a small dataset for training.

represent the similarity between two sentences. It employs a recommendation-based mechanism to compute sentence ranks [19].

Unlike textRank, the edge weights are computed based on some similarity metric (e.g., Cosine similarity), producing better output in some scenarios.

### E. MOABC

This algorithm is an enhancement of the popular ABC (Artificial Bee Colony) algorithm. The ABC algorithm is inspired by the natural food searching behavior of honeybees. In the ABC algorithm, the optimization is done in three phases:

- Employed bees:* These bees exploit the food source, return to the hive, and report to the onlooker bees.
- Onlooker bees:* These bees gather data from employed bees, then select the food source to gather data from.
- Scout bees:* These bees try to find random food sources for our employed bees to exploit.

This algorithm tries to convert the text summarization problem into an optimization problem, with the best summary representing the global minima.

Sanchez-Gomez et al. [40] used MOABC on DUC 2002 dataset to get a 2.23% improvement on ROUGE-2 scores over state-of-the-art methods. Abbasi-Ghalehtaki et al. [28] implemented a MOABC + cellular automation theory-based algorithm on the DUC 2002 dataset to get significant results.

### F. MACHINE LEARNING TECHNIQUES

#### 1) LOGISTIC REGRESSION

Logistic regression is a classification algorithm, which is very useful in binary classification i.e., whether the gender of the author is male or female. Unlike linear regression, it models the data using a non-linear function like the sigmoid function. It can also be used for classification problems, where the number of classes in the output are more than 2. The mathematical expression for the sigmoid function is given in

**TABLE 27.** Studies on query-based text summarization.

Ref. No	Dataset	Result	Advantages	Disadvantages
Shang <i>et al.</i> [59]	MEDLINE corpus	ROUGE-1: 0.3104	Semantic information like disease cause, treatment and symptoms is retained.	The semantic relation graph does not provide a natural language description of the text. Hence, a better sentence representation could have been used.
He <i>et al.</i> [66]	CNN/Daily Mail	ROUGH-1: 45.65 ROUGH-2: 22.35 ROUGH-L : 42.50	Produces better results than other studies	Requires specific pre-training procedures in high-resource settings
Van Lierde& Chow [68]	DUC2007	ROUGH-2: 0.12997	A better time-complexity than the existing hypergraph based methods.	Did not take polysemy of terms into account.

the following equation.

$$\varnothing_{sig}(z) = \frac{1}{1 + e^{-z}} \quad (5)$$

Neto et al. [30] used the logistic regression classifier on the TIPSTER collection. They got a precision of 0.34 for the model. Neto et al. [32] used logistic regression on the EASC (Essex Arabic Summary Corpus) and got a ROUGE-1 score of 0.129.

## 2) SVM

The main idea behind an SVM classifier is to choose a hyperplane that can segregate n-dimensional data into different classes with minimum overlapping. Support vectors are used to create the hyperplane, hence the name ‘Support vector machines’. In an SVM model, the distance between a point  $x$  and the hyperplane, represented by  $(w, b)$ , where,

$$\|W\| = 1 \text{ if } |w, x| > +b \quad (6)$$

Shen et al. [27] used SVM on the LookSmart web directory along with LSA and achieved significant results. Neto et al. [30] used SVM on the TIPSTER collection to a precision of 0.34.

## 3) RANDOM FOREST

Random forest classifiers are a part of ensemble-based learning methods. Their main features are ease of implementation, efficiency, and great output in a variety of domains. In the Random Forest approach, many decision trees are constructed during the training stage. Then, a majority voting method is used among those decision trees during the classification stage to get the final output. Alami et al. [32] used a Random Forest classifier on the EASC collection and got a ROUGE-1 score of 0.129. John and Wilscy [82] used Random Forest and Maximum Marginal Relevance (MMR), achieving significant results. The MMR coefficient selects

the sentences that have the highest relevance, with the least redundancy with respect to the rest of the sentences generated for the summary.

Machine learning-based methods achieved significant results in the text summarization domain, however, due to limited dataset sizes, the models could not learn that efficiently and thus they could not compete with the state-of-the-art graph-based models. However, neural network-based models overcame the limitations of machine-learning-based models and produced even better results than the state-of-the-art graph-based models.

## G. NEURAL NETWORK-BASED APPROACHES

Task summarization can be formalized as a seq2seq model, where input sequence is the input document, the output sequence is the output summary. Since the input size can keep varying, we cannot use a traditional neural network for this task. These seq2seq models are getting very popular in recent times. The most popular seq2seq models being used in for text summarization are RNN, LSTM and GRU.

### 1) RNN

RNN (Recurrent Neural Networks) belong to a class of neural networks that can use the previous outputs as input for next state. The structure of a basic RNN model is given in FIGURE 6.

The activation vector ( $a$ ) is computed as shown in Eq. (7):

$$a^{<t>} = g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a) \quad (7)$$

The output value ( $y^t$ ) is computed as shown in Eq. (8):

$$y^{<t>} = g_2(W_{ya}a^{<t>} + b_y) \quad (8)$$

### 2) LSTM

Although RNN can generate significant results for text summarization, they suffer from the ‘vanishing gradient’ problem

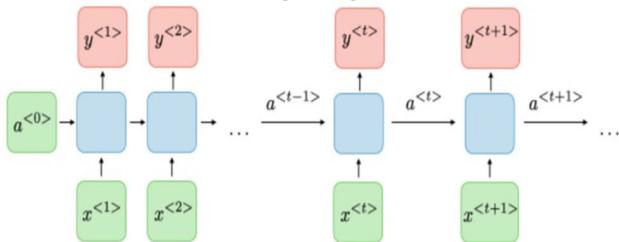


FIGURE 6. RNN.

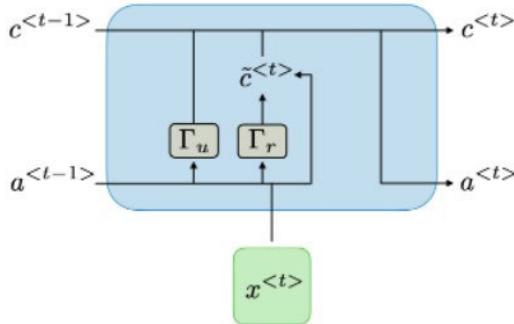


FIGURE 7. LSTM Unit.

while backpropagation. This limits the learning abilities of the model. To counter this, LSTM (Long short-term memory) models were introduced. In an LSTM model, a gate-based mechanism is employed in each LSTM cell that is used to memorize the relevant information. This solves the vanishing gradient problem of RNN. The cell of an LSTM model is shown in *Figure 7*.

### 3) GRU

Gated Recurrent Units (GRU) are another modification over standard RNNs that can solve the vanishing gradient problem. Similar to LSTM units, GRU units have a gate-based mechanism to store the relevant data for backpropagation training. The construction of a GRU cell is given in *FIGURE 8*.

## VIII. CHALLENGES AND FUTURE SCOPES

Even with these advancement in text summarization, multiple challenges still exist, and researchers are working to overcome the challenges. These challenges can also act as future research directions for the new studies. These challenges are in many domains like multi-document summarization, applications of text summarization and some user-specific summarization tasks. Few of the challenges are as discussed below:

### A. CHALLENGES RELATED TO MULTI-DOCUMENT SUMMARIZATION

Multi-document text summarization is more complex than single-document text summarization due to the following issues:

- i. Redundancy
- ii. Temporal dimension
- iii. Co-references
- iv. Sentence reordering

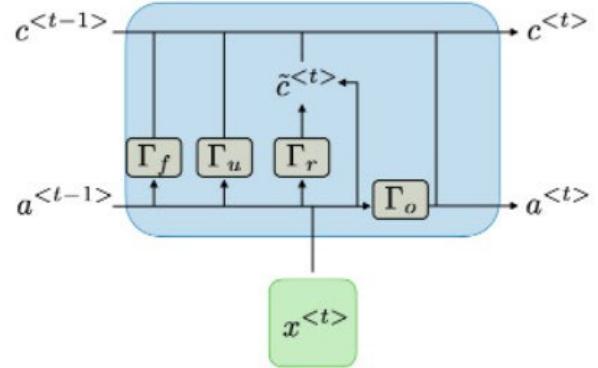


FIGURE 8. GRU Unit.

Some approaches for multi-document summarization can also generate improper references, e.g., assume one sentence in a document contains a proper noun, and the following sentence consists of a reference to the noun. If the summarizer ranks the second sentence higher than the first and does not include it, it will create improper references to other sentences. It is a massive challenge in multi-document summarization.

### B. CHALLENGES RELATED TO APPLICATIONS OF TEXT SUMMARIZATION

Since most current studies focus on a specific text domain, i.e., news, biomedical documents, etc., some of these domains do not have significant economic value. Focusing on a long text, such as an essay, dissertation thesis or reports, may be more economically profitable. However, since the processing of long text requires high computational power, it remains a major challenge.

### C. CHALLENGES RELATED TO USER-SPECIFIC SUMMARIZATION TASKS

Summarizing semi-structured resources like web pages databases is an important application of text summarization since most of the textual data is present in a semi-structured format. This type of summarization is more complex than simple text summarization since there is much more noise in the data. Hence developing efficient summarizers for these domains is a massive challenge.

### D. CHALLENGES RELATED TO FEATURE SELECTION, PREPROCESSING AND DATASETS

For any natural language processing problem, the performance of the selected methods dramatically depends on the selection of the features, so is valid with text summarization techniques. Irrespective of the methods such as machine learning, statistical, fuzzy, deep learning etc. that have been used at a large scale in recent times for such problems, selecting appropriate features for concerning documents to be summarized is still a significant challenge in front of researchers. So, there is much scope in solving the feature selection problem, such as determining the most appropriate features to summarize the dataset, discovering new features,

optimizing the commonly used features, using features for semantic, adding grammatical features, linguistics features etc. Preprocessing a dataset using appropriate methods also affects the performance of the summarization methods, so it also needs attention in the future. One can explore the appropriate stemming approaches, stop word removal techniques, tokenizers, and suitable POS taggers to categorize token classes among nouns, verbs, adjectives, adverbs, etc. The creation of a new dataset is also a demanding task. Many little-explored domains, such as legal, tourism, health, etc., need new datasets to be created and used to expedite the summarization work at a different level.

## IX. CONCLUSION

Text summarization is an exciting research topic among the NLP community that helps to produce concise information. The idea of this study is to present the latest research and progress made in this field with a systematic review of relevant research articles. In this study, we consolidated the research works from different repositories, related to various text summarization methods, datasets, techniques, and evaluation metrics.

We have also added a section on “Analysis of Popular Text Summarization Techniques”, which articulates the most popular techniques in the text summarization domain and gives the strengths and limitations of each technique, and hints at future research directions. We have presented the information in a tabular format, covering the advantages and disadvantages of each research paper, which can make it easier for the readers to use our review paper as a base paper for text summarization domain knowledge. We presented a detailed discussion on the different types of text summarization studies based on approach (extractive, abstractive and hybrid), the number of documents (single-document and multi-document), summarization domain (generic domain and domain-specific summarization), language (monolingual, multilingual, cross-lingual), and nature of the output summary (generic and query-based summarizer). We also presented a detailed analysis of various studies in a tabular format, which will save the readers the hassle of reading through long texts and save their time. We also gave a detailed review of various datasets used in this domain and provided references to the datasets. We discussed various standard evaluation metrics used (ROUGE, F-measure, recall, precision etc.), which can be used to measure the quality of a text summarization model. Finally, we discussed various challenges faced in text summarization that can lead future studies in the domain.

## REFERENCES

- [1] D. Jain, M. D. Borah, and A. Biswas, “Summarization of legal documents: Where are we now and the way forward,” *Comput. Sci. Rev.*, vol. 40, May 2021, Art. no. 100388, doi: [10.1016/j.cosrev.2021.100388](https://doi.org/10.1016/j.cosrev.2021.100388).
- [2] A. B. Al-Saleh and M. E. B. Menai, “Automatic Arabic text summarization: A survey,” *Artif. Intell. Rev.*, vol. 45, no. 2, pp. 203–234, Feb. 2016, doi: [10.1007/s10462-015-9442-x](https://doi.org/10.1007/s10462-015-9442-x).
- [3] Y. Kumar, K. Kaur, and S. Kaur, “Study of automatic text summarization approaches in different languages,” *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 5897–5929, Dec. 2021, doi: [10.1007/s10462-021-09964-4](https://doi.org/10.1007/s10462-021-09964-4).
- [4] A. Nenkova and K. McKeown, “A survey of text summarization techniques,” in *Mining Text Data*, C. C. Aggarwal and C. Zhai, Eds. Boston, MA, USA: Springer, 2012, pp. 43–76, doi: [10.1007/978-1-4614-3223-4\\_3](https://doi.org/10.1007/978-1-4614-3223-4_3).
- [5] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, “Automatic text summarization: A comprehensive survey,” *Expert Syst. Appl.*, vol. 165, Mar. 2021, Art. no. 113679, doi: [10.1016/j.eswa.2020.113679](https://doi.org/10.1016/j.eswa.2020.113679).
- [6] B. Polepalli Ramesh, R. J. Sethi, and H. Yu, “Figure-associated text summarization and evaluation,” *PLoS ONE*, vol. 10, no. 2, Feb. 2015, Art. no. e0115671, doi: [10.1371/journal.pone.0115671](https://doi.org/10.1371/journal.pone.0115671).
- [7] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, “Systematic literature reviews in software engineering—A systematic literature review,” *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, Jan. 2009, doi: [10.1016/j.infsof.2008.09.009](https://doi.org/10.1016/j.infsof.2008.09.009).
- [8] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, “Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement,” *BMJ*, vol. 339, p. b2535, Jul. 2009, doi: [10.1136/bmj.b2535](https://doi.org/10.1136/bmj.b2535).
- [9] P. Mongeon and A. Paul-Hus, “The journal coverage of Web of science and scopus: A comparative analysis,” *Scientometrics*, vol. 106, no. 1, pp. 213–228, Jan. 2016, doi: [10.1007/s11192-015-1765-5](https://doi.org/10.1007/s11192-015-1765-5).
- [10] M. Kutlu, C. Cigir, and I. Cicekli, “Generic text summarization for Turkish,” *Comput. J.*, vol. 53, no. 8, pp. 1315–1323, Jan. 2010, doi: [10.1093/comjnl/bxp124](https://doi.org/10.1093/comjnl/bxp124).
- [11] C. Bouras and V. Tsogkas, “Noun retrieval effect on text summarization and delivery of personalized news articles to the user’s desktop,” *Data Knowl. Eng.*, vol. 69, no. 7, pp. 664–677, Jul. 2010, doi: [10.1016/j.datkat.2010.02.005](https://doi.org/10.1016/j.datkat.2010.02.005).
- [12] S. A. Babar and P. D. Patil, “Improving performance of text summarization,” *Proc. Comput. Sci.*, vol. 46, pp. 354–363, Jan. 2015, doi: [10.1016/j.procs.2015.02.031](https://doi.org/10.1016/j.procs.2015.02.031).
- [13] M. A. Tayal, M. M. Raghuwanshi, and L. G. Malik, “ATSSC: Development of an approach based on soft computing for text summarization,” *Comput. Speech Lang.*, vol. 41, pp. 214–235, Jan. 2017, doi: [10.1016/j.csl.2016.07.002](https://doi.org/10.1016/j.csl.2016.07.002).
- [14] R. M. Alguliyev, R. M. Aliguliyev, N. R. Isazade, A. Abdi, and N. Idris, “A model for text summarization,” *Int. J. Intell. Inf. Technol.*, vol. 13, no. 1, pp. 67–85, Jan. 2017, doi: [10.4018/IJIIT.2017010104](https://doi.org/10.4018/IJIIT.2017010104).
- [15] R. M. Alguliyev, R. M. Aliguliyev, N. R. Isazade, A. Abdi, and N. Idris, “COSUM: Text summarization based on clustering and optimization,” *Expert Syst.*, vol. 36, no. 1, Feb. 2019, Art. no. e12340.
- [16] M. Mohd, R. Jan, and M. Shah, “Text document summarization using word embedding,” *Expert Syst. Appl.*, vol. 143, Apr. 2020, Art. no. 112958, doi: [10.1016/j.eswa.2019.112958](https://doi.org/10.1016/j.eswa.2019.112958).
- [17] O. Rouane, H. Belhadjef, and M. Bouakkaz, “Combine clustering and frequent itemsets mining to enhance biomedical text summarization,” *Expert Syst. Appl.*, vol. 135, pp. 362–373, Nov. 2019, doi: [10.1016/j.eswa.2019.06.002](https://doi.org/10.1016/j.eswa.2019.06.002).
- [18] Y.-H. Hu, Y.-L. Chen, and H.-L. Chou, “Opinion mining from online hotel reviews—A text summarization approach,” *Inf. Process. Manag.*, vol. 53, no. 2, pp. 436–449, Mar. 2017, doi: [10.1016/j.ipm.2016.12.002](https://doi.org/10.1016/j.ipm.2016.12.002).
- [19] G. Erkan and D. R. Radev, “LexRank: Graph-based lexical centrality as salience in text summarization,” *J. Artif. Intell. Res.*, vol. 22, pp. 457–479, Dec. 2004.
- [20] D. Patel, S. Shah, and H. Chhinkaniwala, “Fuzzy logic based multi document summarization with improved sentence scoring and redundancy removal technique,” *Expert Syst. Appl.*, vol. 134, pp. 167–177, Nov. 2019, doi: [10.1016/j.eswa.2019.05.045](https://doi.org/10.1016/j.eswa.2019.05.045).
- [21] R. Mihalcea and P. Tarau, “TextRank: Bringing order into text,” in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Barcelona, Spain, Jul. 2004, pp. 404–411. [Online]. Available: <https://aclanthology.org/W04-3252>
- [22] R. Elbarougy, G. Behery, and A. El Khatib, “Extractive Arabic text summarization using modified PageRank algorithm,” *Egyptian Informat. J.*, vol. 21, no. 2, pp. 73–81, Jul. 2020, doi: [10.1016/j.eij.2019.11.001](https://doi.org/10.1016/j.eij.2019.11.001).
- [23] R. Ferreira, L. de Souza Cabral, F. Freitas, R. D. Lins, G. de França Silva, S. J. Simske, and L. Favaro, “A multi-document summarization system based on statistics and linguistic treatment,” *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5780–5787, 2014, doi: [10.1016/j.eswa.2014.03.023](https://doi.org/10.1016/j.eswa.2014.03.023).
- [24] V. Priya and K. Umamaheswari, “Enhanced continuous and discrete multi objective particle swarm optimization for text summarization,” *Cluster Comput.*, vol. 22, no. S1, pp. 229–240, Jan. 2019, doi: [10.1007/s10586-018-2674-1](https://doi.org/10.1007/s10586-018-2674-1).
- [25] I. V. Mashechkin, M. I. Petrovskiy, D. S. Popov, and D. V. Tsarev, “Automatic text summarization using latent semantic analysis,” *Program. Comput. Softw.*, vol. 37, no. 6, pp. 299–305, Nov. 2011, doi: [10.1134/S0361768811060041](https://doi.org/10.1134/S0361768811060041).

- [26] J.-Y. Yeh, H.-R. Ke, W.-P. Yang, and I.-H. Meng, "Text summarization using a trainable summarizer and latent semantic analysis," *Inf. Process. Manage.*, vol. 41, no. 1, pp. 75–95, Jan. 2005, doi: [10.1016/j.ipm.2004.04.003](https://doi.org/10.1016/j.ipm.2004.04.003).
- [27] D. Shen, Q. Yang, and Z. Chen, "Noise reduction through summarization for web-page classification," *Inf. Process. Manage.*, vol. 43, no. 6, pp. 1735–1747, Nov. 2007, doi: [10.1016/j.ipm.2007.01.013](https://doi.org/10.1016/j.ipm.2007.01.013).
- [28] R. Abbasi-ghalehtaki, H. Khotanlou, and M. Esmailpour, "Fuzzy evolutionary cellular learning automata model for text summarization," *Swarm Evol. Comput.*, vol. 30, pp. 11–26, Oct. 2016, doi: [10.1016/j.swevo.2016.03.004](https://doi.org/10.1016/j.swevo.2016.03.004).
- [29] Y. Ko, J. Park, and J. Seo, "Improving text categorization using the importance of sentences," *Inf. Process. Manage.*, vol. 40, no. 1, pp. 65–79, 2004, doi: [10.1016/S0306-4573\(02\)00056-0](https://doi.org/10.1016/S0306-4573(02)00056-0).
- [30] J. L. Neto, A. A. Freitas, and C. A. A. Kaestner, "Automatic text summarization using a machine learning approach," in *Proc. 16th Brazilian Symp. Artif. Intell., Adv. Artif. Intell.*, Berlin, Germany, 2002, pp. 205–215.
- [31] D. D. A. Bui, G. Del Fiol, J. F. Hurdle, and S. Jonnalagadda, "Extractive text summarization system to aid data extraction from full text in systematic review development," *J. Biomed. Informat.*, vol. 64, pp. 265–272, Dec. 2016, doi: [10.1016/j.jbi.2016.10.014](https://doi.org/10.1016/j.jbi.2016.10.014).
- [32] N. Alami, M. Meknassi, and N. En-Nahnah, "Enhancing unsupervised neural networks based text summarization with word embedding and ensemble learning," *Expert Syst. Appl.*, vol. 123, pp. 195–211, Jun. 2019, doi: [10.1016/j.eswa.2019.01.037](https://doi.org/10.1016/j.eswa.2019.01.037).
- [33] A. Sinha, A. Yadav, and A. Gahlot, "Extractive text summarization using neural networks," 2018, *arXiv:1802.10137*.
- [34] J. Xu, Z. Gan, Y. Cheng, and J. Liu, "Discourse-aware neural extractive text summarization," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, Jul. 2020, pp. 5021–5031, doi: [10.18653/v1/2020.acl-main.451](https://doi.org/10.18653/v1/2020.acl-main.451).
- [35] B. Mutlu, E. A. Sezer, and M. A. Akcayol, "Candidate sentence selection for extractive text summarization," *Inf. Process. Manage.*, vol. 57, no. 6, Nov. 2020, Art. no. 102359, doi: [10.1016/j.ipm.2020.102359](https://doi.org/10.1016/j.ipm.2020.102359).
- [36] M. Yousefi-Azar and L. Hamey, "Text summarization using unsupervised deep learning," *Expert Syst. Appl.*, vol. 68, pp. 93–105, Feb. 2017, doi: [10.1016/j.eswa.2016.10.017](https://doi.org/10.1016/j.eswa.2016.10.017).
- [37] N. Alami, M. E. Mallahi, H. Amakdouf, and H. Qjidaa, "Hybrid method for text summarization based on statistical and semantic treatment," *Multimedia Tools Appl.*, vol. 80, no. 13, pp. 19567–19600, May 2021, doi: [10.1007/s11042-021-10613-9](https://doi.org/10.1007/s11042-021-10613-9).
- [38] J. M. Sanchez-Gomez, M. A. Vega-Rodríguez, and C. J. Pérez, "Comparison of automatic methods for reducing the Pareto front to a single solution applied to multi-document text summarization," *Knowl.-Based Syst.*, vol. 174, pp. 123–136, Jun. 2019, doi: [10.1016/j.knosys.2019.03.002](https://doi.org/10.1016/j.knosys.2019.03.002).
- [39] J. M. Sanchez-Gomez, M. A. Vega-Rodríguez, and C. J. Pérez, "Parallelizing a multi-objective optimization approach for extractive multi-document text summarization," *J. Parallel Distrib. Comput.*, vol. 134, pp. 166–179, Dec. 2019, doi: [10.1016/j.jpdc.2019.09.001](https://doi.org/10.1016/j.jpdc.2019.09.001).
- [40] J. M. Sanchez-Gomez, M. A. Vega-Rodríguez, and C. J. Pérez, "Experimental analysis of multiple criteria for extractive multi-document text summarization," *Expert Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112904, doi: [10.1016/j.eswa.2019.112904](https://doi.org/10.1016/j.eswa.2019.112904).
- [41] M. Azhari and Y. Jaya Kumar, "Improving text summarization using neuro-fuzzy approach," *J. Inf. Telecommun.*, vol. 1, no. 4, pp. 367–379, Oct. 2017, doi: [10.1080/24751839.2017.1364040](https://doi.org/10.1080/24751839.2017.1364040).
- [42] F. B. Goularte, S. M. Nassar, R. Fileto, and H. Saggion, "A text summarization method based on fuzzy rules and applicable to automated assessment," *Expert Syst. Appl.*, vol. 115, pp. 264–275, Jan. 2019, doi: [10.1016/j.eswa.2018.07.047](https://doi.org/10.1016/j.eswa.2018.07.047).
- [43] S. Hou and R. Lu, "Knowledge-guided unsupervised rhetorical parsing for text summarization," *Inf. Syst.*, vol. 94, Dec. 2020, Art. no. 101615, doi: [10.1016/j.is.2020.101615](https://doi.org/10.1016/j.is.2020.101615).
- [44] S. Dohare, H. Karnick, and V. Gupta, "Text summarization using abstract meaning representation," 2017, *arXiv:1706.01678*.
- [45] K. Ganeshan, C. Zhai, and J. Han, "Opinosis: A graph based approach to abstractive summarization of highly redundant opinions," in *Proc. 23rd Int. Conf. Comput. Linguistics (Coling)*, Beijing, China, Aug. 2010, pp. 340–348. [Online]. Available: <https://aclanthology.org/C10-1039>
- [46] K. Song, L. Lebanoff, Q. Guo, X. Qiu, X. Xue, and C. Li, "Joint parsing and generation for abstractive summarization," in *Proc. AAAI*, Apr. 2020, vol. 34, no. 5, pp. 8894–8901, doi: [10.1609/aaai.v34i05.6419](https://doi.org/10.1609/aaai.v34i05.6419).
- [47] R. Barzilay and K. R. McKeown, "Sentence fusion for multidocument news summarization," *Comput. Linguistics*, vol. 31, no. 3, pp. 297–328, Sep. 2005.
- [48] N. Okumura and T. Miura, "Automatic labelling of documents based on ontology," in *Proc. IEEE Pacific Rim Conf. Commun., Comput. Signal Process. (PACRIM)*, Victoria, BC, Canada, Aug. 2015, pp. 34–39, doi: [10.1109/PACRIM.2015.7334805](https://doi.org/10.1109/PACRIM.2015.7334805).
- [49] C.-S. Lee, Y.-J. Chen, and Z.-W. Jian, "Ontology-based fuzzy event extraction agent for Chinese e-news summarization," *Expert Syst. Appl.*, vol. 25, no. 3, pp. 431–447, Oct. 2003, doi: [10.1016/S0957-4174\(03\)00062-9](https://doi.org/10.1016/S0957-4174(03)00062-9).
- [50] M. Al-Maleh and S. Desouki, "Arabic text summarization using deep learning approach," *J. Big Data*, vol. 7, no. 1, p. 109, Dec. 2020, doi: [10.1186/s40537-020-00386-7](https://doi.org/10.1186/s40537-020-00386-7).
- [51] U. KhanDELWAL, K. Clark, D. Jurafsky, and L. Kaiser, "Sample efficient text summarization using a single pre-trained transformer," 2019, *arXiv:1905.08836*.
- [52] E. Lloret, M. T. Romá-Ferri, and M. Palomar, "COMPENDIUM: A text summarization system for generating abstracts of research papers," *Data Knowl. Eng.*, vol. 88, pp. 164–175, Nov. 2013, doi: [10.1016/j.dat.2013.08.005](https://doi.org/10.1016/j.dat.2013.08.005).
- [53] V. Gupta and N. Kaur, "A novel hybrid text summarization system for Punjabi text," *Cognit. Comput.*, vol. 8, no. 2, pp. 261–277, Apr. 2016, doi: [10.1007/s12559-015-9359-3](https://doi.org/10.1007/s12559-015-9359-3).
- [54] M. S. Binwahlan, N. Salim, and L. Suanmali, "Fuzzy swarm diversity hybrid model for text summarization," *Inf. Process. Manage.*, vol. 46, no. 5, pp. 571–588, Sep. 2010, doi: [10.1016/j.ipm.2010.03.004](https://doi.org/10.1016/j.ipm.2010.03.004).
- [55] J. M. Perea-Ortega, E. Lloret, L. Alfonso Ureña-López, and M. Palomar, "Application of text summarization techniques to the geographical information retrieval task," *Expert Syst. Appl.*, vol. 40, no. 8, pp. 2966–2974, Jun. 2013, doi: [10.1016/j.eswa.2012.12.012](https://doi.org/10.1016/j.eswa.2012.12.012).
- [56] Y. Sankarasubramaniam, K. Ramanathan, and S. Ghosh, "Text summarization using Wikipedia," *Inf. Process. Manage.*, vol. 50, no. 3, pp. 443–461, May 2014, doi: [10.1016/j.ipm.2014.02.001](https://doi.org/10.1016/j.ipm.2014.02.001).
- [57] H. Nguyen, E. Santos, and J. Russell, "Evaluation of the impact of user-cognitive styles on the assessment of text summarization," *IEEE Trans. Syst., Man, Cybern., A, Syst. Humans*, vol. 41, no. 6, pp. 1038–1051, Nov. 2011, doi: [10.1109/TSMCA.2011.2116001](https://doi.org/10.1109/TSMCA.2011.2116001).
- [58] J. Xu and G. Durrett, "Neural extractive text summarization with syntactic compression," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process.*, Hong Kong, Nov. 2019, pp. 3292–3303, doi: [10.48550/ARXIV.1902.00863](https://doi.org/10.48550/ARXIV.1902.00863).
- [59] Y. Shang, Y. Li, H. Lin, and Z. Yang, "Enhancing biomedical text summarization using semantic relation extraction," *PLoS ONE*, vol. 6, no. 8, Aug. 2011, Art. no. e23862, doi: [10.1371/journal.pone.0023862](https://doi.org/10.1371/journal.pone.0023862).
- [60] L. H. Reeve, H. Han, and A. D. Brooks, "The use of domain-specific concepts in biomedical text summarization," *Inf. Process. Manage.*, vol. 43, no. 6, pp. 1765–1776, Nov. 2007, doi: [10.1016/j.ipm.2007.01.026](https://doi.org/10.1016/j.ipm.2007.01.026).
- [61] M. Moradi, M. Dashti, and M. Samwald, "Summarization of biomedical articles using domain-specific word embeddings and graph ranking," *J. Biomed. Informat.*, vol. 107, Jul. 2020, Art. no. 103452, doi: [10.1016/j.jbi.2020.103452](https://doi.org/10.1016/j.jbi.2020.103452).
- [62] A. Farzindar and G. Lapalme, "Legal text summarization by exploration of the thematic structure and argumentative roles," in *Proc. Text Summarization Branches Out*, Barcelona, Spain, Jul. 2004, pp. 27–34. [Online]. Available: <https://aclanthology.org/W04-1006>
- [63] R. Rani and D. K. Lobiyal, "An extractive text summarization approach using tagged-LDA based topic modeling," *Multimedia Tools Appl.*, vol. 80, no. 3, pp. 3275–3305, Jan. 2021, doi: [10.1007/s11042-020-09549-3](https://doi.org/10.1007/s11042-020-09549-3).
- [64] E. Linhares Pontes, S. Huet, J.-M. Torres-Moreno, and A. C. Linhares, "Compressive approaches for cross-language multi-document summarization," *Data Knowl. Eng.*, vol. 125, Jan. 2020, Art. no. 101763, doi: [10.1016/j.dat.2019.101763](https://doi.org/10.1016/j.dat.2019.101763).
- [65] N. Chatterjee and P. K. Sahoo, "Random indexing and modified random indexing based approach for extractive text summarization," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 32–44, Jan. 2015, doi: [10.1016/j.csl.2014.07.001](https://doi.org/10.1016/j.csl.2014.07.001).
- [66] J. He, W. Kryscinski, B. McCann, N. Rajani, and C. Xiong, "CTRLsum: Towards generic controllable text summarization," 2020, *arXiv:2012.04281*.
- [67] G. Salton, J. Allan, C. Buckley, and A. Singhal, "Automatic analysis, theme generation, and summarization of machine-readable texts," *Science*, vol. 264, no. 5164, pp. 1421–1426, Jun. 1994, doi: [10.1126/science.264.5164.1421](https://doi.org/10.1126/science.264.5164.1421).
- [68] H. Van Lierde and T. W. S. Chow, "Query-oriented text summarization based on hypergraph transversals," *Inf. Process. Manage.*, vol. 56, no. 4, pp. 1317–1338, Jul. 2019, doi: [10.1016/j.ipm.2019.03.003](https://doi.org/10.1016/j.ipm.2019.03.003).
- [69] H. Van Lierde and T. W. S. Chow, "Learning with fuzzy hypergraphs: A topical approach to query-oriented text summarization," *Inf. Sci.*, vol. 496, pp. 212–224, Sep. 2019, doi: [10.1016/j.ins.2019.05.020](https://doi.org/10.1016/j.ins.2019.05.020).

- [70] A. Joshi, E. Fidalgo, E. Alegre, and L. Fernández-Robles, "SummCoder: An unsupervised framework for extractive text summarization based on deep auto-encoders," *Expert Syst. Appl.*, vol. 129, pp. 200–215, Sep. 2019, doi: [10.1016/j.eswa.2019.03.045](https://doi.org/10.1016/j.eswa.2019.03.045).
- [71] *CNN-Dailymail*. GitHub. Accessed: Jun. 26, 2022. [Online]. Available: <https://github.com/abisee/cnn-dailymail>
- [72] *GitHub—JafferWilson/Process-Data-of-CNN-DailyMail: This Repository Holds the Output of the Repository*. Accessed: Jun. 26, 2022. [Online]. Available: <https://github.com/abisee/cnn-dailymail> and <https://github.com/JafferWilson/Process-Data-of-CNN-DailyMail>
- [73] H. M. Lynn, C. Choi, and P. Kim, "An improved method of automatic text summarization for web contents using lexical chain with semantic-related terms," *Soft Comput.*, vol. 22, no. 12, pp. 4013–4023, Jun. 2018, doi: [10.1007/s00500-017-2612-9](https://doi.org/10.1007/s00500-017-2612-9).
- [74] K. Ganesan. *Opinosis Dataset—Topic Related Review Sentences*. Accessed: Jun. 30, 2022. [Online]. Available: <https://kavita-ganesan.com/opinosis-opinion-dataset/>
- [75] R. C. Belwal, S. Rai, and A. Gupta, "A new graph-based extractive text summarization using keywords or topic modeling," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 10, pp. 8975–8990, Oct. 2021, doi: [10.1007/s12652-020-02591-x](https://doi.org/10.1007/s12652-020-02591-x).
- [76] R. C. Belwal, S. Rai, and A. Gupta, "Text summarization using topic-based vector space model and semantic measure," *Inf. Process. Manage.*, vol. 58, no. 3, May 2021, Art. no. 102536, doi: [10.1016/j.ipm.2021.102536](https://doi.org/10.1016/j.ipm.2021.102536).
- [77] P. Li, W. Lam, L. Bing, and Z. Wang, "Deep recurrent generative decoder for abstractive text summarization," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Copenhagen, Denmark, 2017, pp. 2091–2100, doi: [10.18653/v1/D17-1222](https://doi.org/10.18653/v1/D17-1222).
- [78] *Datasets/Gigaword.Py at Master · Tensorflow/Datasets*. GitHub. Accessed: Jun. 26, 2022. [Online]. Available: <https://github.com/tensorflow/datasets>
- [79] *MEDLINE/PubMed Data*. Accessed: Jun. 26, 2022. [Online]. Available: [https://www.ncbi.nlm.nih.gov/databases/download/pubmed\\_medline.html](https://www.ncbi.nlm.nih.gov/databases/download/pubmed_medline.html)
- [80] B. Elayeb, A. Chouigui, M. Bounhas, and O. B. Khiroun, "Automatic Arabic text summarization using analogical proportions," *Cognit. Comput.*, vol. 12, no. 5, pp. 1043–1069, Sep. 2020, doi: [10.1007/s12559-020-09748-y](https://doi.org/10.1007/s12559-020-09748-y).
- [81] Y. Gao, W. Zhao, and S. Eger, "SUPERT: Towards new frontiers in unsupervised evaluation metrics for multi-document summarization," 2020, *arXiv:2005.03724*.
- [82] A. John and M. Wilscy, "Random forest classifier based multi-document summarization system," in *Proc. IEEE Recent Adv. Intell. Comput. Syst. (RAICS)*, Trivandrum, India, Dec. 2013, pp. 31–36, doi: [10.1109/RAICS.2013.6745442](https://doi.org/10.1109/RAICS.2013.6745442).



**DIVAKAR YADAV** (Senior Member, IEEE) received the bachelor's degree in computer science and engineering, in 1999, the master's degree in information technology, in 2005, and the Ph.D. degree in computer science and engineering, in 2010. He spent one year as a Postdoctoral Fellow at the Universidad Carlos III de Madrid, Spain, from 2011 to 2012. He has more than 22 Years of Teaching and Research Experience. He did his PDF from the University of Carlos-III, in 2011 to 2012, Ph.D. (CSE) in 2010, M.Tech. (IT) from IIIT Allahabad in 2005 and B.Tech. (CSE) from IET Lucknow in 1999. He is working as a Professor at the School of Computer and Information Science, Indira Gandhi National Open University (IGNOU), New Delhi, since September 2022. Prior to join IGNOU, he worked as an Associate Professor and Head,

Department of Computer Science and AMP; Engineering, National Institute of Technology, from 2019 to 2022, Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur (UP) from 2016 to 2019, and Jaypee Institute of Information Technology, Noida (UP) from 2005 to 2016. He is currently an Associate Professor with the Department of Computer Science and Engineering, National Institute of Technology Hamirpur (NIT Hamirpur) (An Institution of National Importance). He supervised eight Ph.D. thesis and 31 M.Tech. dissertations. He has published more than 125 research papers in international journals and conference proceedings of repute. His research interests include machine learning, soft computing, information retrieval, NLP, and e-learning. He is a member of ACM.



**RISHABH KATNA** received the bachelor's and master's degrees in computer science and engineering from the National Institute of Technology Hamirpur (NIT Hamirpur), Hamirpur, in 2021 and 2022, respectively. He has been working with Standard Chartered GBS as an Intern Software Engineer. He is currently working as a Software Engineer at Qualcomm. His research interests include social networking, automation for web developers, and natural language processing.



**ARUN KUMAR YADAV** received the Ph.D. degree in computer science and engineering, in 2016. He is currently an Assistant Professor with the Department of Computer Science and Engineering, National Institute of Technology Hamirpur (NIT Hamirpur). He is also working on government sponsored funded projects and supervised many students. He has published more than 20 research papers in reputed international/national journals and conference proceedings. His research interests include information retrieval, machine learning, and deep learning.



**JORGE MORATO** received the Ph.D. degree in library science from the Universidad Carlos III de Madrid, Spain, on the topic of knowledge information systems and their relationship with linguistics. He is currently a Professor of information science with the Department of Computer Science, Universidad Carlos III de Madrid. His research interests include NLP, information retrieval, web positioning, and knowledge organization systems.