Multimedia Tools and Applications (2023) 82:29135–29197
https://doi.org/10.1007/s11042-023-14613-9

State-of-the-art approach to extractive text summarization: a comprehensive review

Avaneesh Kumar Yadav[1] · Ranvijay[1] · Rama Shankar Yadav[1] · Ashish Kumar Maurya[1]

Received: 1 November 2021 / Revised: 13 June 2022 / Accepted: 3 February 2023 /
Published online: 16 February 2023

Abstract

With the rapid growth of social media platforms, digitization of official records, and digital publication of articles, books, magazines, and newspapers, lots of data are generated every day. This data is a foundation of information and contains a vast amount of text that may be complex, ambiguous, redundant, irrelevant, and unstructured. Therefore, we require tools and methods that can help us understand and automatically summarize the vast amount of generated text. There are mainly two types of approaches to perform text summarization: abstractive and extractive. In Abstractive Text Summarization, a concise summary is generated by including the salient features of the input documents and paraphrasing documents using new sentences and phrases. While in Extractive Text Summarization, a summary is produced by selecting and combining the most significant sentences and phrases from the source documents. The researchers have given numerous techniques for both kinds of text summarization. In this work, we classify Extractive Text Summarization approaches and review them based on their characteristics, techniques, and performance. We have discussed the existing Extractive Text Summarization approaches along with their limitations. We also classify and discuss evaluation measures and provide the research challenges faced in Extractive Text Summarization.

✉ Avaneesh Kumar Yadav
avaneesh17@mnnit.ac.in

Ranvijay
ranvijay@mnnit.ac.in

Rama Shankar Yadav
rsy@mnnit.ac.in

Ashish Kumar Maurya
ashishmaurya@mnnit.ac.in

[1] Department of Computer Science & Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj, India

Keywords Text summarization · Extractive text summarization · Research issues · Graph-based approaches · Machine learning techniques · Clustering-based approaches

1 Introduction

The use of multimedia data such as text, images, audio, and video are increasing day by day. The length and size of the data are very crucial to understand the data. For example, a video clip of a few seconds or 1–2 pages of text documents is easily understandable, but when the length of the document is very large, or the duration of videos and audios are in hours, we need some summarization methods to generate a summary and understand these large multimedia documents. There are several categories of summarization methods based on the types of data, such as text, image, audio, and video summarization, as shown in Fig. 1. Text summarization shortens a set of data to create a summary that comprises the most important or relevant information within the original text document [134]. The image summarization technique effectively reduces extraneous information from the dataset images. It is used in a variety of applications, including effective web-data navigation, rapid surfing, and deep-learning model training. It takes a long time to fine-tune deep networks, and training with a representative summary is more important in these scenarios than training with the original image dataset [147]. In audio summarization, an audio summary aims to extract the most meaningful information from audio and condense it into a format suitable for a specific task. Summarized audio should be easier to understand than a verbatim summary because it eliminates common speech breaks, flaws, corrections, and repetitions [51, 129]. The continuous advancements in audio capture quality, accuracy, and the growing appeal of natural language as a computer interface have spurred the present surge in interest in audio summarizing systems. The audio summary has been used in various situations, including interviews, chats, meetings, TED presentations, lectures, and news broadcasts [87]. Video summarization is used to generate a short summary of the content of a longer video document by selecting and presenting the most informative or interesting materials for potential users. Media companies utilize automatic video summaries to provide suitable indexing, retrieval, browsing, and advertising of diverse media assets. Video sharing services employ it to improve video quality, increase user engagement, and grow online content. Video summarization caters to the needs of specific content presentation situations (viz., to create trailers or teasers for TV shows and movies). It presents the highlights of an event (e.g., a music band accomplishment, a sporting event, or a public debate) [27, 163]. It creates a video summary of the frequent activities, such as the last

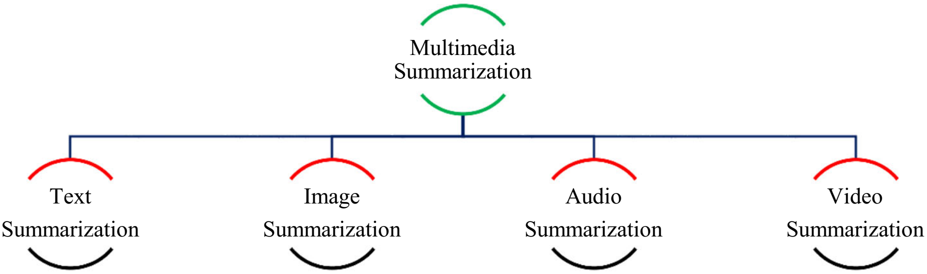

Fig. 1 Multimedia summarization based on the type of data

24 hours of footage from a surveillance camera for time-efficient progress tracking or security purposes. We have focused on text summarization in this article.

The documents mainly involve text data in the form of e-newspapers, blogs, e-books, research articles, official records, personal records, govt. records, stock market data, e-magazines, summaries of the stories (digest), medical reports, and so on. Text data is currently proliferating due to many reasons, such as advancements in technologies, involvement of people in online social activities, keeping the records in the digital form, and publishing the contents online [176]. Text summarization can be used to save time in reading and understanding the whole document. It is the process of extracting the most important information from the given input document (s). It chooses paragraphs and sentences from the given text input and systematizes them into some meaningful order. Producing a summary from the documents is a challenging job. Text summarization includes several issues such as degree of ambiguity, compression ratio, maximum relevance, etc. Currently, many issues have been overcome; for example, common methods are given to identify relevant content or important keywords, and the absence of consistency in summaries. Due to the issues involved in summarization, generating a summary for multiple documents is more complex than a single document. The concept of text summarization was introduced in late 1958 [88]. Since then, exponential progress has been notified in this field, and the researchers have given several methods for text summarization [66]. Text summarization is classified into several categories, such as follows in Fig. 2.

Based on output text summarization, it is of two kinds: (i) Abstractive Text Summarization (ATS) (ii) Extractive Text Summarization (ETS). Extractive summarization selects more relevant sentences and paragraphs from given text input and puts them in summary without any change. It is a faster and easier way of summarization than abstractive text summarization [49]. ETS uses several techniques like Deep Learning Techniques (DLT) [48] (Autoencoder, RBM, RNN [114], Convolutional Neural Network (CNN)), LexRank [41], TextRank [102], Bayesian method, Binary classifier, SVM, Logistic Regression Model (LRM), graphical methods [157], maximal marginal relevance (MMR) algorithm, Hidden Markov Model

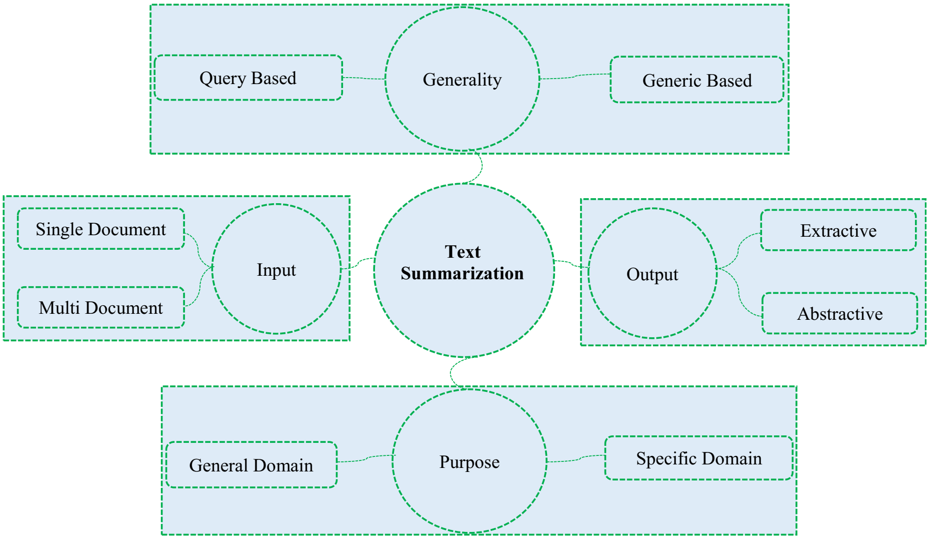

Fig. 2 Classification of Text Summarization

(HMM), Decision Trees (DT), TF-IDF, and Clustering. In ATS, a summary is produced that comprises the significant features of the given text input. The existing sentences and paragraphs of the input text may or may not be present in the summary generated by ATS. Mostly, ATS generates a better summary than ETS. But, producing the summary in ETS takes lesser time and computational steps than ATS. ATS uses several techniques like Naive Bayes decision theory, k-means algorithm, CNN [146], RNN [114], Neural Network (NN) [113], Sequence-to-Sequence model, Singular Vector Decomposition (SVD), WordNet, etc.

Based on the input documents, text summarization is of two kinds: (1) Single Document (SD) and (2) Multiple Documents (MD), as shown in Fig. 2. In SD, there is a summarization based on single documents. SD has extracted the sentences or words from the given text input to generate a summary. In MD summarization, a summary is generated from multiple documents that belong to the same category. MD summarization can create a few problems such as temporal dimension, ambiguity, co-references, order of sentences, etc. It is more difficult to create a summary in MD summarization than in SD summarization [53, 80, 81, 135]. There are many distinct issues in the multiple document summarization that may be solved with the help of a selecting sentence from the initial paragraph or passages in the source documents. There are calculated similarities in the several sentences. Although, it selects the sentences from the input document and maintains the sentence order of new related content of the summary. Few researchers have given distinct kinds of techniques which can calculate the performance in MD text summarization [20, 109, 121, 128, 161].

Based on the generality of document content, text summarization can be categorized as generic-based and query-based summarization [33, 50, 52, 84, 126], shown in Fig. 2. In query-based summarization, a summary of the document is generated by queries of the users. It is also known as topic-based or user-based summarization [24, 38, 71, 151, 159]. In generic-based summarization, the summary is produced by selecting the important data from the source document, which is not specified by the user's summarization [97, 151].

Based on the purpose of document content, it is categorized into general and specific domains shown in Fig. 2. Both domains are independent of the text summarization document. This domain-based text summarization creates a summary from the source of text data like Encyclopedia Articles, Web Pages, Radio News, Newspaper Articles, Transcription Dialogues, Biomedical Domain, Technical Report, Journal Articles, sports documents, legal documents, medical documents, etc., [44, 105].

Based on the language of an input text document (s), text summarization is of several kinds like cross-lingual language summarization, multi-lingual language summarization, and mono-lingual languages summarization. Cross-lingual language summarization generates summaries in languages other than English, for example, Urdu. Multi-lingual language summarization produces summaries in several numbers of languages such as Telegu, Tamil, Punjabi, English, Urdu, and Hindi, Spanish languages (CNN-corpus) [80]. Mono-lingual language summarization generates a summary in the language same as the source language. For example, if the document's source language is Hindi, it generates a summary in the Hindi language.

From the existing definition of text summarization, three foremost aspects are identified to start research in this field. First, how many numbers of documents are considered to generate the summary, for instance, single or multiple documents. Second, how to identify which information is critical in the document so that it must be included in the summary. Third, how the summary can be kept concise that it should give brief idea about the document.

1.1 Motivation

Based on text data analysis, the motivation is (1) Effectively identify relevant sentences from the input documents and surveys to find the more useful data from the text input. (2) ETS summarizes based on previous and recent works and the various techniques used. There are many summaries for English and other foreign languages like Spanish, German, Turkish, French, Chinese, Czech, Rome, Arabic, Indonesia, Greek, Persian, etc. Some ETS systems focus on Indian languages, such as Hindi, Punjabi, Sindhi, Bengali, Sanskrit, Odia, Nepali, Tamil, Assamese, Telegu, Marathi, Konkani, Urdu, and Gujarati. The documents of these languages are very challenging tasks to summarize; hence ETS system is needed. Nowadays, text data is increasing online and offline, so there is no more space for storing the data, and people have less time to read the huge documents, so it is needed to summarize text documents. The benefits of ETS are briefly discussed in the following points:

- Suppose people want to watch movies, then they decide on the basis of reviews.
- Summaries reduce the amount of time it takes to scan any document.
- According to the summaries, people can make decisions in less or sudden time.
- It aims to shorten search time, especially when the user's purpose is to gather as much information as possible on a specific issue.
- It's about writing better reports and obtaining more relevant information in a shorter amount of time.
- ETS algorithms are less partial than summaries created by humans.
- An article's unique information is summarized in a single document.
- It helps to select and sort research papers while searching existing research papers.
- It helps in pruning relevant sentences from the input document and saves reading time of the documents.

ETS is particularly beneficial to a variety of users because it saves time that is wasted when summarized manually. It also aids in the retrieval of summaries of one or numerous documents about the same subject. The final summary can also provide a broad overview of many themes covered in the input documents, allowing the user to understand deeper into the source papers for more information.

1.2 Related surveys

In literature, the researchers have performed various surveys on extractive text summarization. In [40], the authors have surveyed unsupervised based learning techniques of extractive text summarization. Moratanch and Chitrakala [107] discussed the literature survey about ETS summarization based on both supervised and unsupervised techniques. In [12], a review of the extractive text summarization for Bengali languages is presented. In [48], the authors have discussed the literature review of recent automatic text summarization in the combination of ATS and ETS reviews, but it has not focused completely on ETS literature reviews. Gholamrezazadeh et al. [50] discussed a literature review of text summarization for a few ETS and ATS summarization approaches. In [71, 126], the query-based text summarization approaches are given, including Artificial Neural Networks (ANN) and machine learning techniques. Kumar et al. [77] described the review of text summarization for different Indian languages like (i.e., Malayalam, Tamil, Marathi, Bengali, Punjabi, Urdu, Kannada, Assamese, Konkani, Sanskrit, Odia, etc.). In this paper, we concentrate on presenting state-of-the-art extractive text summarization (ETS) techniques.

1.3 Our contribution

The existing surveys discussed the various ETS approaches given in many research papers. In this work, we study the state-of-the-art ETS approaches and perform a survey of those approaches on various parameters. The contribution of this paper can be summarized as follows:

- We give a classification of ETS approaches and discuss them along with highlighting their limitations.
- We review the existing ETS approaches based on their characteristics, techniques, and performance.
- A classification of evaluation measures with respect to mathematical formulation is also given.
- We discuss the dataset (standard or non-standard datasets) in detail used in the ETS approaches from the existing research papers.
- We discuss the research gap, open issues, and challenges faced during the extractive text summarization process.

These contribution points to be covered in this survey paper are discussed in detail in terms of theoretical and tabular form.

1.4 Paper organization

The rest of the paper's structure is explained as follows: In Section 2, we have explained the review process for our survey. The background of the ETS is described in Section 3. The classification of evaluation measures is given and discussed in Section 4. Section 5 performs the literature review of existing ETS approaches. The detail provided the sources of datasets related to ETS in Section 6. Further, Section 7 discusses various research gaps, open issues, and research challenges in ETS. Section 8 concludes the work and gives future scope. A description of frequently used words is listed in Table 1.

2 The review process

There are many research papers available in the literature regarding extractive text summarization. To identify appropriate summarization technique, we require to search for research papers from the well-known databases manually, which is a bit challenging. Therefore, a systematic approach is needed to review the existing literature on extractive text summarization to know about the state-of-the-art techniques. The proposed review methodology contains two steps, as described below.

Step 1: Identification of research questions (RQs) and their motivation for state-of-the-art

In this step, we have created the six RQs and their motivations, as shown in Table 2. It helped us to collect the required information, statistics, and guidance in our survey. The formulated RQs are very clear, definite, and well-structured.

Table 1 Frequently used words

ATS	Abstractive Text Summarization	NLP	Natural Languages Processing
BNC	British National Corpus	NN	Neural Networks
CNN	Convolution Neural Network	RBM	Restricted Boltzmann Machine
CAST	Computer-Aided Summarization Tool Corpus	RNN	Recurrent Neural Network
DUC	Document Understanding Conference	ROUGE	Recall-Oriented Understudy for Gisting Evaluation
ETS	Extractive Text Summarization	SVD	Singular Value Decomposition
EASC	Essex Arabic Summaries Corpus	SVM	Support Vector Machine
GRU	Gated Recurrent Unit	TAC	Text Analysis Conferences
HAN	Hierarchical Attention Network	TF-ISF	Term Frequency - Inverse Sentence Frequency
HITS	Hyperlink-Induced Topic Search	TS	Text Summarization
LCS	Longest Common Sub-sequence	NER	Named Entity Recognition
LCSTS	Large-scale Chinese Short Text Summarization	LSTM	Long-Short Term Memory
LSA	Latent Semantic Analysis	IDF	Inverse Document Frequency

Step 2: Identification of related work

In this step, research papers are chosen from the year 2011 to 2022 based on the RQs. The papers are included as per their titles, abstract, and keywords regarding extractive text summarization. We used the following sub-steps to identify the related works:

(i) The following search engines are used to find the research papers related to text summarization:

- Springer (https://www.springer.com/in),
- ACM digital library (https://dl.acm.org/),
- Elsevier (https://www.elsevier.com/en-in)
- Scopus (https://www.scopus.com/home.uri),
- Google Scholar (https://scholar.google.com/)
- ScienceDirect (https://www.sciencedirect.com/),
- IEEE Xplore (https://ieeexplore.ieee.org/Xplore/home.jsp),
- Web of Science (https://clarivate.com/webofsciencegroup/solutions/web-of-science/)

Table 2 RQs and their motivation for the text summarization

RQs for text summarization	
RQ$_1$	Which approaches are more appropriate for text summarization?
RQ$_2$	Whether approaches are document-dependent?
RQ$_3$	Can text summarization be applied to all datasets?
RQ$_4$	What is the current status of extractive text summarization?
RQ$_5$	How can a quality summary be achieved?
RQ$_6$	What are the limitations of the existing research papers?
Motivations of the above RQs	
M$_1$	To identify approaches that are often in the ETS.
M$_2$	To do exhaustive text summarization for a variety of documents.
M$_3$	It helps to find the appropriate evaluation measures to ensure quality.
M$_4$	To identify datasets commonly used in the ETS.
M$_5$	To identify the limitation of ETS approaches from existing research papers.

(ii) We have used the following search strings to find the papers: (extractive OR text OR multi-document OR single-document OR sentence OR document OR summarization OR graph OR machine learning OR clustering OR categorization OR question OR answering OR learning) in title only AND (text summarization OR single-document OR multi-document) anywhere in the paper.

(iii) The manual reading of the papers based on their titles from journals, transactions, and conference proceedings has also been done. Very pertinent research papers have been chosen on text summarization. Figure 3 shows the selection process of the research papers for literature review.

In this step, we have identified 1282 research papers, in which 627 are not considered on the basis of their title at first glance. Further, 379 research papers are discarded on the basis of abstract, keywords, and conclusions at a second glance. At last, we have a total of 276 research papers in hand. After reading these 276 research papers and performing in-depth scanning based on the content, we have 93 core papers providing different aspects of the work.

Step 3: Quality Assessment

In quality assessment, there is focused on evaluation, those selected research papers around our related work-study of our research question to be answered. Some of the quality parameters are continually, conceptually, relevancy, mathematical modeling, and correlation of approaches used to the issues involved in text mining. These state-of-the-art summaries are being fabricated around these quality parameters and ensure the lucid presentation of state-of-the-art through the use of tables related to different sub-strings. The table itself is self-explanatory in terms of the contribution of the core selected article, along with remarks for demerit/improvement possible with 179 research articles. This research article includes six handbooks, one guide, eleven reports, and 75 auxiliary papers that were distinguished from references and, on account of being perused, complete knowledge put into the set.

3 Extractive text summarization

Text Summarization is a process that extracts important salient features from an original text document and assembles them into a meaningful summary. It originated in the late fifties and is used till now. Text summarization is essential as it provided meaningful information as per individual needs from the occasion of data available in different diversify domains. This is the

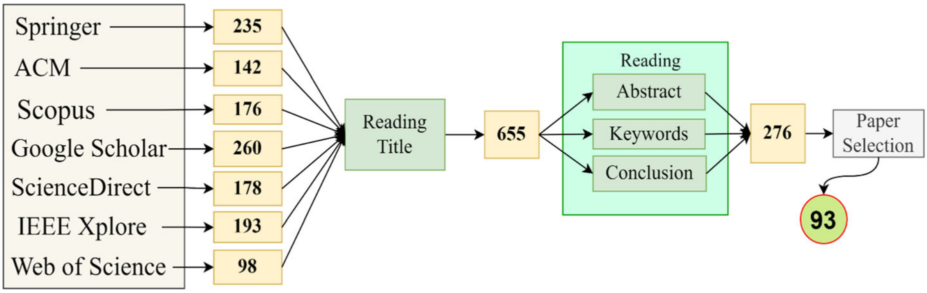

Fig. 3 Shortlisting process of existing research papers for this literature review

ultimate need of today's mobile-packed working environment. There are many approaches to text summarization that create the summary of the documents. The summary consists of the document's most relevant, complete, compact, and a lot for individual requirements and requiring lessor time and space. There are many summaries generation issues, such as effective sentence ordering, co-reference, and redundancy, that different text summarization approaches can address. The text summarization approaches are mainly classified into ATS and ETS. In this research paper, we have focused on ETS summarization.

Extractive text summarization is a snippet of the crucial sentences of the document, and it concatenates the important and meaningful sentences of a document into a summary without any changes. The summary length depends on the compression rate [44, 49]. This paper focuses on ETS approaches that commonly select sentences that contain the most significant concepts in the document. The relevance of sentence features may include a syntactic word, event, frequency, importance sentence relevance, and many more. The text summarization in ETS can be performed in four major steps, as shown in Fig. 4: (i) pre-processing of text, (ii) feature extraction, (iii) sentence selection and assembly, and (iv) generation of the summary [77, 127].

3.1 Preprocessing

It is a text pre-processing consisting of different types of operations like removing stop-words, segmentation, word tokenization, part of speech (POS) tagging, word stemming, etc. These operations are shown in Fig. 5.

Segmentation In segmentation, the document is broken into words, sentences, and paragraphs from the given input documents [53, 121, 135].

Word tokenization It is similar to segmentation. It breaks the sentences into words or a group of words identified as uni-grams, bi-grams, tri-grams words, etc. [53, 121, 135]. Single-word tokenization is known as uni-grams. Two words and three words tokenization are bi-grams and tri-grams, respectively. Similarly, others grams are defined. For example, consider the sentence "Ravi Dahiya Moves to Gold Medal Match." This sentence has seven unigram words such as "Ravi", "Dahiya", "Moves", "to", "Gold", "Medal", and "Match". Similarly, we can find bigrams in the given sentence. There is a total of six bigrams' tokens such as "Ravi Dahiya",

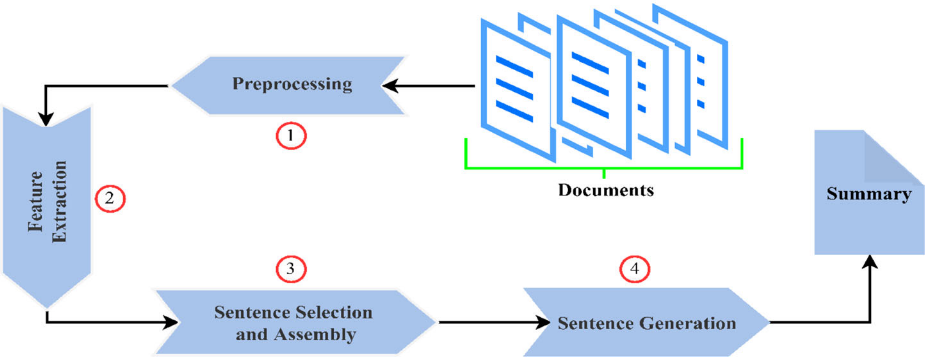

Fig. 4 Process of Extractive Text Summarization

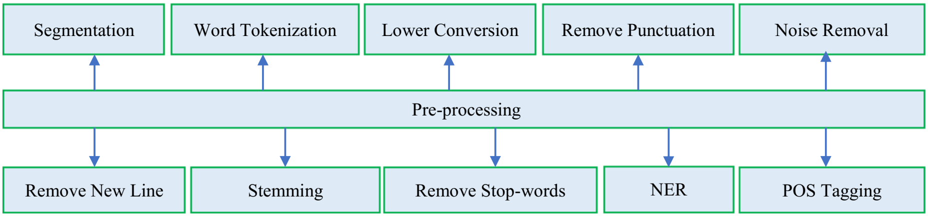

Fig. 5 Various Types of Text Pre-processing in ETS

"Dahiya Moves", "Moves to," "to Gold", "Gold Medal", and "Medal Match" in the given sentence.

Word stemming It refers to the process of converting each word into a "basic root" and finding the word in different forms such as different POS tags, tenses, singular, and plural [53, 121, 135]. Example: The words "connection", "connected", "connecting", "connections", and "connect" have only a single root word, such as "connect".

Lower conversion It is used for text conversion from upper to lower text form in given input documents, and it helps during the removal of the stop-word. Example: "Ravi Dahiya Moves to Gold Medal Match". It is converted into lower cases like "ravi dahiya moves to gold medal match".

Remove new line If the given document contains a new line, it is removed from the given input document.

Remove punctuation It removes the symbol as well as punctuation which are available in the input documents (i.e., Like; &, *, @, #, −, $, etc.) [121, 139, 169].

Noise removal It removes the unnecessary text from given dataset/documents, such as the footer, header, etc.

Removing stop-words There are many stop-words present in the given input document that should be ignored while summarizing the document [53, 121, 135]. By default, a total of 180 stop-words are available in the English language. Examples of stop-word are "is", "our", "should", "would", "he", "she", "they", "there", "had", "have", "may", "can", and so on. Now, take the following sentence as an example to understand the stop words: "We are writing a letter to our parents." This sentence has four stop words like "we", "are", "for", and "our". After removing all stop words from the existing sentence, the newer sentence will become "writing letter parents."

NER It is used for identifying the words of the input text, known as the name of things such as company name, location name, people name, name of the organization, etc., [171].

POS tagging Tagging the part of speech (POS) in a large document is a fast and efficient way to reduce the document during the pre-processing of text. The most relevant POS tags in the document are Adjective, Adverb, Pronouns, Noun, and Verbs [53]. It is a very important task in summarization to extract the most relevant linguistic terms from the input text documents. Using the NLTK POS tagger, we can easily extract the POS word from the document [109, 127].

3.2 Feature extraction

This step acquires the features from the sentences of the document to find the relevant information from the document. Some common features are Term Frequency2 (TF), Similarity to Keywords, Sentence to Centroid Similarity, Proper Noun (PN), Cue Methods (CM), Word Frequency, Sentence Position, and Sentence Length. These features are described in detail and shown in Fig. 6.

Term frequency (TF) The term frequency is an important measure that gives information about a term or word and the number of times it appears in the document [96, 140]. With the help of TF, other measures such as TF-IDF and TF-ISF [97, 117, 143] are calculated. It is denoted in the form of tf (t, d), which is the frequency of the term t. It is evaluated by Eq. (1), where $f_{t,\,d}$ is the number of times term t occurs in document (d) where $\sum_{t^{\cdot}\epsilon d} f_{(t^{\cdot},d)}$ is the total number of terms in document (d).

$$\text{tf }(t,d)=\frac{f_{t,d}}{\sum_{t^{\cdot}\epsilon d}}$$

$f_{(t^{\cdot},d)}$ The idf is the ratio of the total number of documents (N), contains the term (t) from the number of documents, and takes the logarithmic quotient. Idf is denoted as idf $(t,\ D)$ as in Eq. (2).

$$\text{idf }(t,D)=\log_e\frac{N}{|\{d\epsilon D:t\epsilon d\}|} \tag{2}$$

In Eq. 2, number of documents is $|\{d\epsilon D\ :\ t\epsilon d\}|$, and when the term (t) appears, then tf$(t,\ d)$ $\neq$ 0. Here, it calculates the tf-idf weight values with the help of tf $(t,\ d)$ and idf $(t,\ D)$ values as in Eq. (3).

$$\text{tf}-\text{idf }(t,d,D)=\text{tf}(\text{t, d})\times\text{idf}(\text{t, D}) \tag{3}$$

TF-ISF (Term Frequency-Inverse Sentence Frequency) is evaluated by Eq. (4), and it is denoted by tf-isf (t, s, D) [139].

$$\text{tf}-\text{isf }(t,s,D)=\frac{\log(isf)\times tf(t,s)}{L(s)} \tag{4}$$

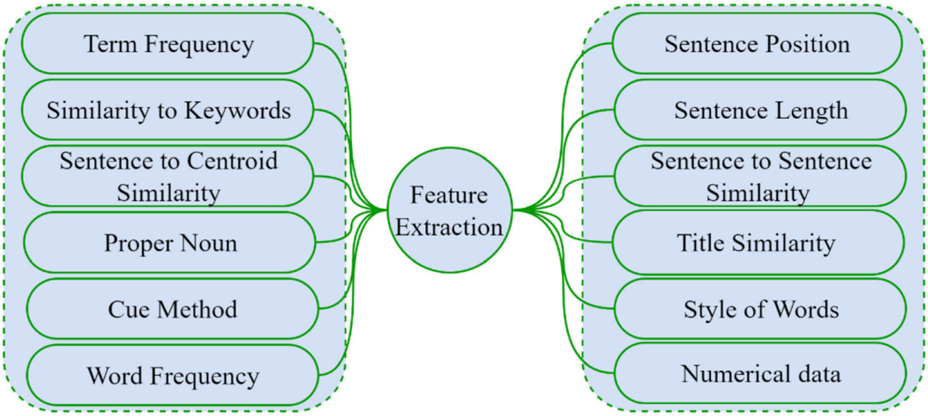

Fig. 6 Taxonomy of feature extraction of ETS

where $L(s)$ is the total length in kth sentences, $tf(t, s)$ is a term (t) in kth sentences, and isf is the total number of occurrences term (t) in kth sentences.

Similarity to keywords (SK) This feature is defined in terms of cosine similarity, which finds the similarity between two keywords and input sentences. It is mainly used in query-based summarization.

Sentence to centroid similarity (SCS) It measures the similarity between each sentence and the centroid of the sentences [40, 143]. The sentence having a maximum TF-ISF value is selected as a centroid among all the sentences.

Proper noun (PN) The sentences written in proper nouns are considered as most essential sentences of the document. If a sentence contains proper nouns, it should be given priority to other sentences [44, 97, 139]. The sentences for example, may include the name of the person, organization, places, and so on. It is denoted as p(n) and evaluated by Eq. (5):

$$\mathrm{p(n)} = \frac{\mathrm{N}}{\mathrm{L}_s} \tag{5}$$

where N is the number of proper nouns in kth sentences and L_s is kth sentences length [139].

Cue method (CM) In a document, sentences may be written using cue words or weighted words. These words may be positive words or negative words. The cue method calculates the significance of a sentence on the basis of these cue words in the cue dictionary. Cue words are present in the summary section as well as conclusions [97].

Word frequency (WF) It is the number of times a word is repeated in a pre-processed document [53, 77]. It is evaluated by Eq. (6) and denoted as wf(t).

$$\mathrm{wf(t)} = \frac{X}{Y} \tag{6}$$

where X is the number of times a word available in the input documents, and Y is the total number of words in the documents.

Sentence position (SP) It gives the location of an important sentence in the given document or paragraph. Positional values of sentences are calculated, and the highest value is assigned to the first sentence of the document. Sentences are sorted based on their positions [40, 44, 53, 77, 139].

Assume ‘S’ is a sentence of the document or paragraph, ‘m’ is a sentence position in the paragraph, and ‘n’ is the sentence number in the paragraph, then the feature score ($sp(f)$) is evaluated by Eq. (7):

$$sp(f) = \frac{m}{n} \tag{7}$$

Sentence length (SL) The length of a sentence resembles the importance and meaning of the sentence in summarization. Generally, very long or very short sentences are not suitable for a summary. Very long sentences may contain unnecessary text data, but it is not useful for the summarization of the original document. Whereas sentences that are too short do not provide more text information from the text document [44, 77, 97, 127, 139], and sentence length (*sl*(*f*)) is evaluated by Eq. (8):

$$sl(f) = \frac{\Psi}{\overline{\Psi}} \tag{8}$$

where Ψ is the count of the number in the sentences, and $\overline{\Psi}$ is the word count from the longest sentence.

Sentence to sentence similarity (SSS) It is a complex type of phenomenon which is not word-dependent. Sentences may be either similar to others or the opposite. It is based on text coherences [117].

Title similarity (ST) This feature finds the similarity between the words present in a sentence and the title of the document and assigns a score to the sentence based on this similarity. Cosine similarity can be used to assess the title similarity [53, 97]. In the absence of a document's title, title similarity cannot be applied.

Style of words (SW) Capital letters or emphasis of words, i.e., underline, bold, and italic, are essential [53].

Numerical data (ND) Sentences in the document which have several numbers of ND need to be included in the summary.

3.3 Sentence selection and assembly

This step finds more relevant sentences available in the given document using methods such as PageRank, LexRank, TextRank, etc. [67, 172]. These relevant sentences are arranged (assembled) in descending order of sentence rank. Choosing more important sentences for the relevant summary is based on the percentage rate of sentences. The user is required to choose the percentage rate of summary (say *p*), and then the first *p* percent of total sentences are considered, and the rest of them are discarded. Next, selected sentences are assembled as per their order in the original document.

3.4 Summary generation

It places the meaningful sentences into the proper position and generates the summary of the given document [127].

4 Evaluation measures (EM)

EM is a difficult assignment for authors because of the various types of summaries available in the papers. It is not an easy task to collect information about summaries. The taxonomy of summary measures and the determination of summary evaluation can be categorized into two kinds of ETS performance: intrinsic and task-based (Extrinsic based) evaluations. It is shown in Fig. 7.

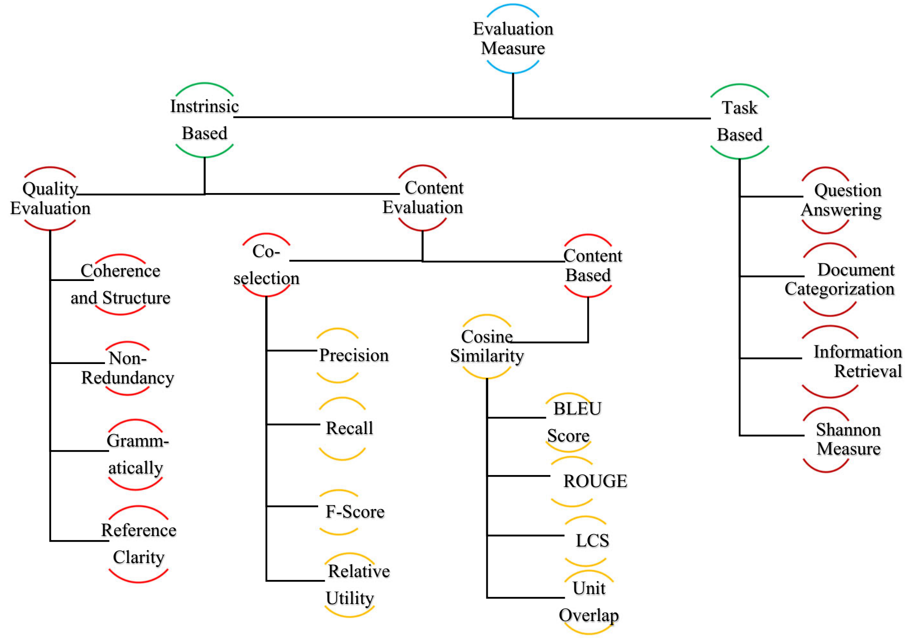

Fig. 7 Taxonomy of the evaluation measures

4.1 Intrinsic evaluation measure

There is a need to compare the quality of machine-generated and human-generated summaries, which are separated into the quality and content evaluation, as illustrated in Fig. 7.

4.1.1 Quality evaluation (QE)

QE measures the quality of the text, and it is categorized into four types, which are described below:

(a). Coherence and Structure (C & S): Coherence is a bit similar to cohesion. It is used for the logical link between paragraphs, sentences, and words in text documents. When the document’s sentences are well-structured, the document’s summary should be well-structured as well.

(b). Non-Redundancy: Repetition of words in a single sentence or phrase in the document is known as redundancy. However, if they do not include duplicate words or sentences during the evaluation of the performance of the document, then it is known as non-redundancy.

(c). Grammatically: The document’s performance does not consider a non-textual item or punctuation error and incorrect words.

(d). Reference clarity: It always consider the noun and pronouns in the summary of the documents

4.1.2 Content evaluation (CE)

CE is part of intrinsic-based evaluation measures, and it is also categorized into two types: Co-selection and content-based evaluation measures.

Co-selection There are numerous forms of it, including (a) Precision (b) Recall (c) F-score (d) Relative utility.

Where, T_P = True Positive, F_P=False Positive, T_N = True Negative, F_N=False Negative. Sometimes confusion matrix is also known as an error matrix.

(a) Precision ($\mathcal{P}_r$): It is determined by the ratio of true positive classes of prediction to total positive classes value from the positive column of the prediction section of the confusion matrix [7, 96, 109, 155, 172] (Table 3) and its mathematical representation is shown in Eq. (9).

$$\mathcal{P}_r = \frac{T_P}{T_P + F_P} \tag{9}$$

(b) Recall ($\mathcal{R}_e$): It is determined by the ratio of true positive prediction classes to the total positive class value and the false-negative class value from the positive row of the actual section of the confusion matrix [7, 96, 109, 155, 172] (Table 3) and its mathematical representation is shown in Eq. (10).

$$\mathcal{R}_e = \frac{T_P}{T_P + F_N} \tag{10}$$

(c) $\mathcal{F}$-Score: It is used to find the unique score with the help of precision and recall, which is twice the multiplication of precision and recall divided by the summation of precision and recall (using Eqs. (9) and (10)) as in Eqs. (11) and (12) [7, 96, 109, 155, 171, 172].

$$\mathcal{F}\text{−score} = \frac{\mathcal{R}_e.\mathcal{P}_r.(\alpha^2 + 1)}{R_e + \alpha^2.\mathcal{P}_r} \tag{11}$$

$$\text{when } \alpha = 1, \text{then } \mathcal{F}\text{−score} = \frac{2.\mathcal{P}_r.\mathcal{R}_e}{\mathcal{P}_r + \mathcal{R}_e} \tag{12}$$

Where the weighting factor is α, when α >1, it is a favors precision case, and α <1, it is favor recall case.

Table 3 Confusion matrix used for precision, recall, and f-score

Confusion Matrix		Prediction	
		Positive	Negative
Actual	Positive	T_P	F_N
	Negative	F_P	T_N

(d) Relative Utility (RU): It is more applicable in reference summaries, which are extracted from a unit of the paragraph, sentences, and words from the document. These are used in fuzzy membership values for reference summaries. Every sentence of the given document contains a reference summary with a trustable number of the values and is used in the additional summary [56, 172]. The RU can be used in more than one summary of documents because it has constructed one sentence from the redundant document. However, it is automatically penalized for the rank of evaluations. Systems can extract more than two sentences; however, the system has only focused on extracting the foregoing sentences, although it has generated less sentence information. Its mathematical formulation is represented in Eq. (13).

$$\mathrm{RU}=\frac{\sum_{i=1}^{k}\rho_i\cdot\sum_{j=1}^{K}\omega_{ji}}{\sum_{i=1}^{k}\varphi_i\cdot\sum_{j=1}^{K}\omega_{ji}} \tag{13}$$

Where ω_{ji} is a sentence utility score i from annotator j and ρ_i=1 for extracted few top sentences by the system, otherwise $\rho_i = 0$. Similar $\varphi_i = 1$ use for top few sentences due to total scores of utilities from all judges, otherwise $\varphi_i = 0$.

Content-based evaluation measure (CBEM) CBEM is a part of intrinsic evaluation measure that has multiple forms, including (a) Cosine Similarity, (b) BLEU Scores, (c) ROUGE, (d) LCS (Longest Common Sub-sequence), and (e) Unit Overlap, all of them are described with mathematical formulations.

(a) Cosine Similarity: Cosine Similarity is a part of content-based evaluations. It depends on reference documents and system summary and contains vector-space models (VSM) [4, 65, 67, 121, 135, 150, 151]. Its mathematical formulation is represented in Eq. (14).

$$\cos(\mathrm{M},\mathrm{N})=\frac{\sum_i M_i.N_i}{\sqrt{\sum_i M_i^2}.\sqrt{\sum_i N_i^2}} \tag{14}$$

Where M and N are both system summaries. $\sum_i M_i.\ N_i$ is the summation of sentences of summary, $\sqrt{\sum_i M_i^2}$ is the square root of the summation of squares of $\mathrm{M_i}$ sentences and $\sqrt{\sum_i N_i^2}$ is the square root of the summation of squares of $\mathrm{N_i}$ sentences.

(b) Bilingual Evaluation Understudy (BLEU) Score: It is used to evaluate text documents that evaluate a produced sentence to a reference sentence. It counts the number of matches by comparing the n-grams of the produced sentences with the n-grams of the reference sentence. Here, the matches do not depend on the locations where they take place [25, 63, 119]. BLEU metrics are used in many applications such as text summarization, image caption generation, machine translation, and speech recognition. However, it offers few advantages as it is easy to understand and highly correlates with the evaluation of human summary. BLEU is widely adopted as it is language-independent. BLEU Score can be computed using Eqs. (15), (16), and (17).

$$\mathrm{BLEU}=\mathrm{BP}.e^{\sum_{n=1}^{N}W_n\log P_n} \tag{15}$$

$$BP = \begin{cases} e^{\left(1-\frac{r}{c}\right)}, & if\ r \geq c \\ 1, & if\ r < c \end{cases} \quad (16)$$

$$\text{logBLEU} = \sum_{n=1}^{N} W_n \log P_n + \min\left(\left(1-\frac{r}{c}\right), 0\right) \quad (17)$$

Where P_n is the geometric average of the modified n-gram precision, W_n is $_{\text{the}}$ uniform weight ($W_n = \frac{1}{N}$) and uses the value N is 4 (by default), N is the length of n-grams, r is an effective length of reference corpus, and c is a candidate translation length.

(c) ROUGE: ROUGE is similar to n-grams co-occurrence because it is mostly applicable in automatic evaluation measures [4, 7, 25, 48, 67, 146]. It behaves like n-grams and finds the n-gram subsequences from the document sentences. There are many n-grams words present in the sentences, and these words are converted into 1-g, 2-g, 3-g, etc. Its mathematical formulation is represented in Eq. (18).

$$\text{ROUGE}-\text{n} = \frac{\sum_{\text{C}\in\text{X}}\sum_{\text{P}\in\text{C}}\text{COUNT}_{\text{match}}(\text{P})}{\sum_{\text{C}\in\text{X}}\sum_{\text{P}\in\text{C}}\text{COUN}(\text{P})} \quad (18)$$

where P is gram$_{\text{n,}}$ COUNT$_{\text{match}}$(P) is a candidate summary and a reference summary, and this is the maximum number of n-grams co-occurs. COUNT(P) is the reference summary for the number of n-grams., X = A collection of reference summaries–a collection of reference summaries. ROUGE metric has several types, such as:

ROUGE-1: ROUGE-1 is based on the 1-g score in between reference summary and candidate summary [4, 61, 76, 85, 130, 155, 169]. However, n-gram recall has ROUGE-N measures in between reference summary and candidate summaries.

ROUGE-2: It works for the bi-gram measures in between reference summary and candidate summary [4, 76, 99, 130, 136, 155, 169, 176].

ROUGE-L: It works for the longest common subsequences between a reference summary and candidate summary [135, 136, 155].

ROUGE-S*: It is used in overlap ratio in terms of "skip-bigrams" in-between reference summary and a gold summary [155].

ROUGE-SU*: It is an extension of ROUGE-S*, where the number of skip words denotes "*". ROUGE-SU* uses unigram and bigram as a counting unit i.e., ROUGE-SU4 allows bigrams that are not adjacent words with the highest of four words in between two bigrams [85, 99].

(d) LCS: LCS is the longest common subsequence dependent on two-word string lengths [76]. Where LCS uses edit of distances, its mathematical formulation is represented in Eq. (19).

$$\text{LCS}(\text{M}, \text{N}) = \frac{\text{len}(\text{M}) + \text{len}(\text{N}) - \text{edit}_{\text{di}}(\text{M}, \text{N})}{2} \quad (19)$$

where, M and N (both) are word sequences, LCS (M, N) denotes the length of the LCS between M and N. len(M) = string length (M), len (N) = string length (N), and edit$_{\text{di}}$ (M, N) = edit-distance of M and N.

(e) Unit Overlap: It is almost similar to the cosine similarity method. This similarity method is based on the set of words. Its mathematical formulation is represented in Eq. (20).

$$\mathrm{Overlap}(\mathrm{M},\mathrm{N}) = \frac{\|M\cap N\|}{\|M\| + \|N\| - \|M\cap N\|} \tag{20}$$

Where M and N are set of words. $\|M\|$=size of set M, and $\|N\|$=size of set N.

4.2 Task-based evaluation measure (TBEM)

TBEM has focused on the quality of the summary from the given documents. It performs many tasks like Question Answering, Shannon Measure, Document Categorization, and Information Retrieval.

4.2.1 Question answering (QA)

QA is a part of an extrinsic evaluation or task-based evaluation. The summarization in ETS can be carried out with the help of QA. Four Combined Pre-Medical Test (CPMT) is collected for reading comprehensive exercise. This comprehensive exam exercise has a number of multiple choices of QA available. Although, there is only one option that is correct, and that option is chosen from each question. So that correct QA can be evaluated from a different situation. Consequently, it creates a summary of the main QA passages.

4.2.2 Document categorization (DC)

Text classification is the foundation of DC, which includes sentiment analysis, subject labeling, and intent detection. It performs the analysis of the text on different levels such as level of documents, level of the paragraph, level of sentences, and level of sub-sentence [172]. It measures the evaluation result for an inaccurate summary of the document into all different types of levels.

4.2.3 Information retrieval (IR)

ETS uses it to remove redundancy and unimportant words or sentences in the texts [172]. It is used to improve the quality of the text summary and helps to choose the document from the database.

4.2.4 Shannon measure

It is mainly used in information theory and measures the information of document content by concentrating on document tokens like words, letters, etc., from the given document text [172]. It measures the three groups, such as relevant paragraphs from the given document, does not consider text at all from the given documents, and creates a summary from the documents.

5 Sources of datasets

Dataset provides an overview of many resources used to compare the ETS system and evaluate the summary. These resources include the standard datasets, automatic summary and manual criteria evaluation tools.

This state-of-art review presents the most common datasets used in ETS summarization evaluation [19, 31, 39]. These datasets are described in Table 4:

TAC dataset It was started in 2008, named TAC2008, and lasts till 2015 as TAC2015. TAC has total four datasets such as TAC2008 [174], TAC2009 [174], TAC2010, TAC2011 [7], TAC2013, TAC2014, and TAC2015. These datasets are needed to fill a few application forms on the TAC website (https://tac.nist.gov/data/forms/index.html), and their respective details are listed in Table 4.

DUC dataset It is provided by the NIST organization and is the most common dataset [135]. These datasets are frequently used in ETS research. Its corpora were liberated as the summarization segment organized at the DUC conference [31, 39]. The first DUC challenge was DUC2001, and the last DUC challenge was DUC2007, held in 2007. DUC website contains a total of seven DUC datasets such as DUC2001 [7, 142, 157], DUC2002 [7, 33, 67, 135, 142, 157, 169], DUC2003 [33], DUC2004 [33, 99, 121], DUC2005 [29, 30, 99, 173], DUC2006 [99, 173], and DUC2007. These datasets are available on the DUC website (https://www-nlpir.nist.gov/projects/duc/data/). When these datasets get access, there is a need to fill out few application forms on the DUC website. Every DUC dataset contains the documents, and it has three types of summaries such as (a) automatically created, (b) automatically created base-line, and (c) manually created that were created summaries by challenging participant systems. This dataset is mainly used in the ETS system for evaluation, but it does not supply more data to train NNs-models.

Table 4 Standard and other datasets of ETS

Datasets	Single	Multi	Domain	Languages	Dataset Size
TAC2008 [174]	–	✓	News	English	48 × 20
TAC2009 [174]	–	✓	News	English	44 × 20
TAC2010	–	✓	News	English	46 × 20
TAC2011 [7]	–	✓	News	English	44 × 20
TAC2013	–	✓	News	English	48 × 20
TAC2014	–	✓	News	English	–
TAC2015	–	✓	News	English	44 × 20
DUC2001 [142, 157]	✓	✓	News	English	60 × 10
DUC2002 [33, 67, 142, 157]	✓	✓	News	English	60 × 10
DUC2003 [33]	✓	✓	News	English	30 × 25, 60 × 10
DUC2004 [33, 121]	✓	✓	News	English, Arabic	100 × 10
DUC2005 [29, 30, 173]	–	✓	News	English	50 × 32
DUC2006 [30]	–	✓	News	English	50 × 25
DUC2007	–	✓	News	English	25 × 10
CAST	✓	–	News	English	147
EASC [4, 25, 37]	✓	–	Wikipedia News, News	Arabic	153
CNN/DailyMail [8, 65]	✓	–	News	English	312,084
CNN Corpus [80]	✓	–	News	English	3000
SummBank	✓	✓	News	English, Chinese	40 × 10
LCSTS	✓	–	Blogs	English	2,400,591
Opinosis	–	✓	Reviews	English	51 × 100
ACL-all [21]	–	✓	News	English	173
AI-all [21]	–	✓	News	English	372
SIGIR 2018 [110]	–	✓	Conference Preceding	English	125

CAST dataset This dataset contains a group of newswire texts, and it is taken from the Reuters Corpus website (http://about.reuters.com/researchandstandards/corpus/) and some science text from the BNC website (http://www.natcorp.ox.ac.uk/), and provides three types of corpus information annotation such as text fragments, links between sentences, and essential sentences which is isolated from marked sentences. It finds the unessential sentence which is not annotated. These datasets could be more beneficial for the development of reduction of sentences and selection of sentences algorithms.

EASC dataset It is a type of NLP resource, and this dataset contains 153-Arabic articles and 765-human generated ETS summaries of this dataset [4, 25, 37, 63]. It is taken from the Arabic website (https://www.lancaster.ac.uk/staff/elhaj/corpora).

These datasets need copyright materials. There are responsible datasets users to follow the relevant rule and regulations of copyright materials.

CNN/DailyMail dataset This type of dataset is used for the passage-based question-answering task. It is being used for the evaluation of ETS summarization systems. It is taken from this website (https://github.com/deepmind/rc-data/). However, it is a modified version of the corpus in [11, 113], which contains multiple sentence summaries for ATS summarization and in [8, 65], this dataset is used for ETS summarization.

CNN-corpus dataset This is widely used in single document ETS. This dataset has original texts, gold standard summaries, and highlights. It was recently used in the ETS competition DocEng'2020 [83], and DocEng'2019 [82]. When needed for research purposes, CNN-corpus with entire its annotated text is easily obtainable by (By email request through [80, 81]) requesting from its authors.

SummBank dataset This dataset contains 360 multi-documents, human-written non-ETS summaries, and 40 clusters of news. It is taken from the website (https://catalog.ldc.upenn.edu/LDC2003T16). It is near about 2-million SD and MD extracts, generated by two methods such as manual and automatic methods.

LCSTS dataset This dataset contains over 2-million short texts with short summaries and is taken from the website (http://icrc.hitsz.edu.cn/Article/show/139.html) [31, 39]. It is constructed from the SinaWeibo website (such as the Chinese-microblogging website).

Opinosis dataset This dataset contains 51 topics and each topic has approximately 100 sentences. This dataset contains five manually produced "Gold standard Summaries" for each topic. It is taken from the website (http://kavita-ganesan.com/opinosis-opinion-dataset/).

Association for Computational Linguistics (ACL) dataset This dataset contains 173 ACL 2014 conference papers and ACL-all divided into 85 short papers and 88 long papers [21].

Artificial intelligence (AI-all) journal dataset It contains 372 papers from artificial intelligent journals. AI-all is divided into six subparts: AI-less10, AI-1013, AI-1316, AI-1619, AI-1925, and AI-more25 [21].

SIGIR 2018 corpus dataset SIGIR 2018 corpus taken from 41st International ACM SIGIR Conference. It contains 125 research papers [110], and this dataset was taken from Research and Development in Information Retrieval, known as SIGIR-2018.

The following features of all datasets are discussed in Table 4: (1) the name of the dataset, (2) the type of document dataset: single or multi-document (3) The domain data: blogs, news, conference preceding, and reviews (4) the languages of dataset: English, Chinese, and others, (5) the dataset size.

Table 4 includes a detailed description of 24 features and these features were used in corresponding papers. These datasets are extracted from the existing research paper and websites. The multi-document summarization dataset was mentioned in terms of "40 × 15", which means the dataset includes a total of 40-cluster of documents, and every cluster has nearly about 15-documents.

6 Literature review of ETS approaches

In the literature, numerous authors have focused on extractive text summarization and given various approaches to generate a summary of the given document (s). Here, we have categorized ETS approaches as shown in Fig. 8.

6.1 Fuzzy logic approach

Using fuzzy logic, we can extract the most essential and relevant sentences from the given document. We can apply fuzzy measures and inference to make reasonable decisions in scenarios where flexibility and tolerance are required. This approach demonstrates how a number of title sentences, first sentence, last sentence, number of words in a sentence, etc., can be utilized for making summaries [137]. Some method uses triangular fuzzy to gain performance as the fuzzy rules on triangular fuzzy sets are more straightforward than other shapes of fuzzy membership functions and less expensive computationally. Kiyomarsi and Esfahani [73] utilized fuzzy inferences to rank and extract the sentences and cover the documents' main content by choosing top-scoring sentences. Dixit and Apte [32] described ETS summarization with the help of fuzzy logic system and found feature scores after feature extraction using a fuzzy system. It consists of three steps such as fuzzifier, fuzzy rule, and defuzzification. It produces a relevant summary of the given documents. Irfan and Zulfikar [60] discussed the fuzzy c-means (FCM) algorithm and TF-IDF methods in text summarization. TF-IDF model evaluated the weight of sentences, and the FCM method was used for resultant weight sentences. Authors have categorized fuzzy into two parts: the important group (sentence score

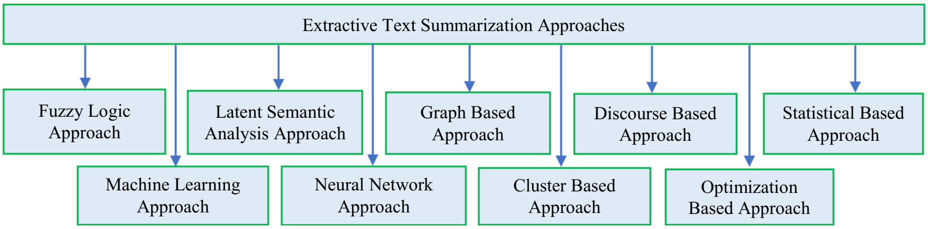
Extractive Text Summarization Approaches
Fuzzy Logic Approach
Latent Semantic Analysis Approach
Graph Based Approach
Discourse Based Approach
Statistical Based Approach
Machine Learning Approach
Neural Network Approach
Cluster Based Approach
Optimization Based Approach

Fig. 8 Categorization of ETS Approaches

is high) and the unimportant group (sentence score is low). Sahba et al. [132] proposed automatically summarized complex articles such as technical documents and news articles. They presented two models, i.e., Sequence-to-Sequence (S2S) model, and fuzzy logic systems. Kumar et al. [75] presented and proposed Genetic Case Base Reasoning (GCBR) to identify cross-document relationships from unannotated texts. It identified the cross-document relations with the help of the new sentence scoring model based on fuzzy reasoning. Abhiman and Hiraman [3] discussed the feature of multi-linguistic and fuzzy logic approaches to sentences. Yadav et al. [169] described extractive text summarization (ETS) using three different methods: WordNet synonyms methods, bushy path methods, and fuzzy logic-based methods. In ETS, WordNet synonyms and fuzzy logic-based methods handled ambiguity issues and estimated values while generating the summary. The authors have used DUC 2002 dataset to evaluate the performance of given methods using ROUGE metrics. Among the given three methods, fuzzy logic-based provides better performance than the other two methods. Shirwandkar and Kulkarni [139] generated a short-term summary with the help of ETS and deep learning. This model calculates many features such as sentence position, sentence length, numerical token, TF-ISF, RBM, fuzzy logic, etc. Authors have categorized sentence feature extraction into RBM and fuzzy logic for creating a summary. RBM is a generative model that uses fuzzy logic and unsupervised learning algorithms and upgrades the summary's accuracy. RBM is a neural network using in-text classification that is a random probability distribution. RBM contains input neurons and hidden neurons for a summary generation. Patel et al. [121] proposed the three main aspects of text summarization using fuzzy logic, i.e., information richness, minimum diversity (means the content has minimum similarity), and decided compression ratio for final summary length. Kumar et al. [76] proposed four component models as LexRank, TextRank, Fuzzy logic, and LSA. All models work based on the graph system, where keyword extraction is done using the text summarization technique and these keywords are converted into rank. Keywords like noun, centrality, recurrence, etc., occurred in the final results. All four-component models were generated to the keyword list, and other keywords were placed into decreasing orders of scores. The top few keywords were selected for a summary generation from the final keyword list.

Limitations:

- Fuzzy logic-based ETS summarization has minimal scope, but it may or may not eliminate redundancy.
- It uses k-means clustering and faces difficulties during the creation of cluster groups for multiple languages.
- It involves an issue of content ambiguity.
- It is language-independent and does not support massive datasets.

6.2 Latent semantic analysis approach

Latent Semantic Indexing (LSI) or Latent Semantic Analysis (LSA) is one of the extractive summarization methods. LSA analyzes the relationship between a set of documents and the terms contained in the document [151]. It uses a mathematical concept called Singular Value Decomposition (SVD) to compute a set of matrices that gives the similarity between the documents [48]. LSA provides the weightage to the documents belonging to different topics.

Summarization has been done by selecting the top N documents from each topic depending on the weightage. Semantic Role Labeling (SRL) has been used as an LSA extractive technique in this paper. In [105], the authors used SRL to design a semantic representation for input documents. To produce a relevant summary, a semantic feature was used for ETS summarization with the help of another feature. Ozsoy et al. [118] proposed the LSA-based methods such as cross and topic for ETS to summarize the Turkish and English text. Moiyadi et al. [106] discussed the LSA-based ETS summarization to summarize documents from the user, and it uses the semantic and SVD summarizers. Chowdhury et al. [26] described the generic Bengali language text summarization using the LSA approach and generated relevant information from the given input document. Merchant and Pande [101] proposed the automated text summarization system, concisely summarized the legal text document, and generated relevant information with the help of the LSA approach. Suleman and Korkontzelos [152] described the LSA as a frequently used corpus-based approach for assessing text similarity based on semantic relationships between words. LSA has been utilized successfully in various language systems to calculate the semantic similarity of texts. However, because LSA overlooks the structural composition of sentences, it suffers from syntactic blindness. LSA cannot discriminate between sentences with semantically similar terms but radically different meanings. LSA is similarly blind to a sentence's grammatical structure, making it impossible to distinguish between sentences and keyword lists [153]. Lwin and Nwet [90] proposed Myanmar language text summarization using the LSA approach. The LSA approach is an NLP method that belongs to an unsupervised approach that generates a set of concepts related to the terms and document.

Limitations:

- The LSA-based approach does not work with large sentence index values.
- It is incapable of dealing with polysemy (a word with various meanings).
- It is not suitable for nonlinear dependencies as it is a linear model.
- Large amounts of storage are required for LSA vectors. Despite significant advancements in the electronic storage medium, the loss of sparsity due to massive data is a serious concern.

6.3 Graph based approach

Graph-based summarization generates the summary from the given text documents. The graph $G = (V, E)$; where $v_1, v_2, \ldots., v_n \in V, e_1, e_2, \ldots\ldots, e_n \in E$, consists of a set of vertices and edges [167, 170] where vertices represent a set of sentences, and edges represent a set of the words from the given documents. Here, two sentences are connected to each other through vertices such that the first sentence is denoted by vertex V_1, and the second sentence is denoted by vertex V_2. The connection between two sentences is represented by an edge that shares the common words of the two sentences. Each sentence of the graph produces a rank of the sentences. Rank-based algorithms are used to do searching and scoring in the documents.

In this approach, several sentence ranking algorithms are used in ETS, such as PageRank, TextRank, LexRank, TOPSIS, SimRank, and CoSimRank. These algorithms are shown in Fig. 9.

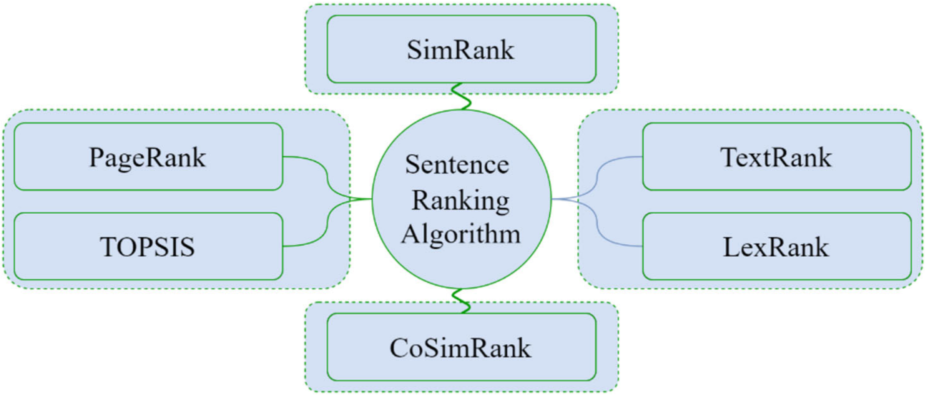

Fig. 9 Various Sentence Ranking Algorithms used in ETS

PageRank It is generally used in graph-based techniques. It helps in calculating the rank of each node from the graph, and this rank decides the summary of texts [37, 140]. Page rank can be represented mathematically as shown in Eq. (21).

$$R(P_k) = \mathrm{d}^*\sum_{V\in \boldsymbol{B}_{\boldsymbol{P}_k}} \frac{\boldsymbol{PR}(\boldsymbol{v})}{C(v)} + (1-d) \tag{21}$$

$R(P_k)$ is page rank of the graph, and d is the damping factor ($0 \leq d \leq 1$), and by default, the value of d is 0.85. C(v) is the number of backlinks from page C(v), and PR(v) is the sum of all PageRank P_k.

TextRank It is used to rank the graph's vertices to be further used for a summary generation [102]. It is similar to page ranking algorithms and can be represented mathematically as shown in Eq. (22).

$$S(V_i) = (1-d) + d^*\sum_{j\in \boldsymbol{V}_j} \frac{S(V_j)}{|V_j|} \tag{22}$$

$S(V_i)$ is TextRank of document sentences and $|\mathrm{Out}(V_j)|$ is out degree of V_i.

LexRank It computes the LexRank of important sentences of the documents based on the eigen vector centrality from the representation of graph sentences [41]. It behaves similar to Page Ranking, and LexRank L(m) can be represented mathematically as shown in Eq. (23).

$$L(m) = \frac{d}{M} + (1-d)^*\sum_{n\in \boldsymbol{adj}(\boldsymbol{m})} \frac{idf_{\ modified_{cosine(m,n)}}}{\sum\limits_{l\in \boldsymbol{adj}(\boldsymbol{n})} idf_{\ modified_{cosine(l,n)}}} L(n) \tag{23}$$

M is the total number of sentences or vertices in the graph, and adj[m] is the set of vertices adjacent to m.

TOPSIS This method makes a decision based on the matrix. The matrix value is converted into graphs then the TOPSIS method decides the rank of graph vertices [43]. TOPSIS stands for "Technique for order preference by similarity to ideal solution" and where decision matrix

D=Y_{mn} holds only for square matrix like the adjacency matrix. It is normalized into an alternative decision matrix represented in Eqs. (24) and (25).

$$R_{ij} = \frac{Y_{ij}}{\sqrt{\sum_{j=1} m^*Y_{ij}^2}}; \mathrm{i} = 1, 2, 3, \ldots\ldots, \mathrm{m} \text{ and } \mathrm{j} = 1, 2, 3, \ldots\ldots, \mathrm{n} \tag{24}$$

$$V_{ij} = R_{ij}{}^*W_j; \mathrm{i} = 1, 2, 3, \ldots\ldots\ldots, \mathrm{n} \text{ and } \mathrm{j} = 1, 2, 3, \ldots\ldots., \mathrm{n} \tag{25}$$

where W_j is the weight decision matrix multiplied by the column of the normalized decision matrix. The rank of each vertex (V_{ij}) is calculated with respect to the decision weight matrix.

SimRank It matches the similarity between two objected (sentences) and similar work as PageRank Algorithms. Yu et al. [175] discussed SimRank computation for large networks such as SimRank in the form of Matrix [140], which is represented by the Eq. (26).

$$S = (1{-}C).I_n + C\left(P^T.S.P\right) \tag{26}$$

where S is similarity matrix and I_n is square identity matrix such as n × n, C ∈ (0, 1) is a damping factor, and P is a column normalized adjacency matrix i.e. $P_{ij} = \frac{1}{|I(i)|}$, here edges from j to i; otherwise, its values are zeros (0).

CoSimRank It is personalized PageRank algorithms derivatives [131]. Its behavior of SimRank algorithms, such as matrix formulation of CoSimRank. This is represented by Eq. (27).

$$S = I_n + C\left(P^T.S.P\right) \tag{27}$$

These methods are used in several existing research papers related to graph-based approaches. Many authors focused mainly on only PageRank, TextRank, and LexRank. Herskovic et al. [58] presented MEDRank, a graph-based ranking method (page ranking algorithms) for biomedical indexing content. MEDRank uses textual concepts to find the most important terms, a useful complement to indexing systems like Medical Text Indexer (MTI). Li et al. [79] proposed graph-based methods to update summarization (update the content of the previous document) such as salience, relevance, and non-redundancy terms. It uses the PNR^2 (Positive and Negative Reinforcement Ranking) ranking algorithm for updating the previous document summarization, and it is an extension of TextRank algorithms. Al-Taani and Al-Omour [9] proposed graph-based Arabic ATS summarization and concentrated on semantic relation between two sentences. The authors generated the summary with the help of page ranking algorithms and evaluated the performance of the three basic units such as stem, word, and n-grams. Fang et al. [42] addressed the sentence scoring method and proposed a new word-sentence co-ranking model known as the CoRank model. It uses to join the relationship between the graph-based ranking model and word sentences. CoRank is generally used in view of matrix operation, and it generates quality of the summary. Azadani et al. [14] proposed the biomedical summarization techniques and combined the GBA with itemset mining techniques. It considered the relationships between numerous concepts while assessing the similarity of the source text's sentences. In [15], Baralis et al. proposed a new technique called GRAPHSUM, which discussed the association rule to represent correlations terms dependent on graphical models. Both TextRank and LexRank are fully unsupervised learning algorithms.

In unsupervised learning, there is no requirement for training, but it depends on the text. Parveen and Strube [120] proposed an unsupervised extractive GBA for summarizing single texts that considered three important summary properties: importance, non-redundancy, and local coherence [141]. Authors represented the input document as a bipartite graph which consists of nodes denoting words and entities. Vale et al. [158] gave a GBA ranking method that first translated texts into a graph model, then used that graph to rate the sentences based on their importance. The summary becomes non-redundant and locally coherent after a process of optimization. Uçkan and Karci [157] proposed a method for ETS that used the Maximum Independent Set (MIS) and textual graph. The authors have used a text processing tool to preserve the semantic cohesion among sentences in the demonstration phase of preliminary tests. The text processing tool first ensures the efficient transformation of relationships between sentences into graphs in this method. Then MIS is computed on the obtained graph. After that, the nodes that formed the Independent Set (IS) are identified and deleted from the graph. The remaining nodes are connected to generate the summary from the remaining graph. The performance of the proposed method was evaluated on DUC2002 and DUC2004 datasets using the ROUGE metrics. Elbarougy et al. [30] gave a method for Arabic text summarization. Due to the complex morphological structure of the Arabic language, extraction of nouns is not very easy, which is utilized as a feature in summarization. The authors used Al-Khalil morphological analyzer to extract nouns from the sentences. The given method is based on the graphical system in which a graph is represented on the basis of a document. In the graph, sentences denote vertices, and words of sentences denote the edges. In the final step of the given method, the authors applied the Modified PageRank algorithm that uses the nouns present in the sentences as the initial score or rank for the nodes. At last, they sorted the sentences based on their final score, removed redundancy, and generated a summary. The performance of the method was calculated using EASC Corpus datasets. Cao and Zhuge [21] proposed a group-based method for text summarization. The method performed the grouping of semantically associated sentences based on the semantic features of the network. It ranks the groups based on four types of semantic network links: cause-effect link, similar-to link, is-part-of link, and sequential link. After that, it concatenated the topmost ranked groups for a summary generation. The authors concluded that the summaries produced by paragraphs or groups comprise more keywords in comparison to the summaries generated by sentences. Dutta et al. [34] focused on GBA summarization. Authors removed the named entity from given document sentences and measured the similarity for each pair of given sentences. Important words from news article datasets were selected to find the interesting result from it. The proposed method used the ROUGE metric for the evaluation of summarized documents, and it generated the quality of the summary. Awan and Beg [13] proposed TOP-Rank as a key extraction and key classification method. For key extraction, the authors used a method based on the position of keywords in the document. The key extraction technique, in particular, examines documents and extracts essential phrases from them by assigning topical terms a higher priority. After extracting the keys, the authors have divided them into three categories: process, material, and task. On a variety of datasets, it is found that TOP-Rank achieves an F1-score of 0.73 for key classification, outperforming state-of-the-art approaches by a large margin. Joshi et al. [68] presented Ranksum, an approach to ETS of SD based on the rank fusion of sentences. The Ranksum uses an unsupervised approach to estimate the sentence saliency rankings correlating to each attribute, then uses a weighted fusion of the four scores to order the phrases according to overall importance. The fusion weights are learned from a labeled document collection because the scores are created fully unsupervised.

The fusion weights are then applied to a variety of DUC2002 and CNN datasets. The document's important keywords and associated sentence ranks were determined using a graph-based technique. Using a sentence uniqueness metric based on bigrams, trigrams, and phrase embedding removes duplicate sentences from the summary.

Limitations:

- The graph-based approach is used in lesser semantics sentences, and it is more complex in the linguistic process.
- It does not work properly in multiple languages.
- It creates a problem for large graphs due to insufficient memory.
- The summary generated by graph-based methods may involve ambiguity and local coherences.

6.4 Discourse based approach

This approach is used in ETS and is also referred to as linguistic techniques/methods. In this approach, the relation of two sentences is called discourse relations. It uses several types of resources like e-dictionary, tree tagger, POS pattern, n-grams, and WordNet for lexical analysis. In [93], the authors discussed Rhetorical Structure Theory (RST), which increased the quality of extracting the semantic process text. The paper presented a framework to use RST for rhetorically parsing, understanding, and summarizing Arabic texts in the Arabic language. Further, the authors have given a three-phase approach to extract the Arabic rhetorical relations: 1. Study of English relations, 2. Analysis of Arabic corpus 3. Understand and use the Arabic cue phrases to facilitate given Arabic texts. Arabic text summarization through this approach uses a few features such as regular expression, action, relation, status, and position. Existing works mainly focused on sentence-level syntactic tasks to evaluate the ability of pre-trained language models (LMs). Koto et al. [74] introduced document-level discourse probing and used them to find document-level relations and calculate the ability of pre-trained LMs. They also utilized BERT, BART, RoBERT as the model for evaluation of the results. Nawaz et al. [116] proposed two frameworks such as the Global Weight (GW) approach and the Local Weight (LW) approach. Weights always depend on the content of the text. One word may have distinct weights called LW in different documents or articles, and an independent dataset is computed for the words' weight. All word weight remains equal in the distinct documents or articles known as GW. Both approaches are used in languages like Urdu, English, etc., but the authors used LW-based approaches here for a generated summary. GW-based approaches are widely applicable in English languages for text classification and information retrieval applications. Finally, ETS generates the summary by integrating both approaches (LW and GW). Gupta et al. [55] suggested that text summarization is based on the sentence ranking method to use linguistics features. It used the lexical chain and WordNet to find the lexicographical relationship and also used a vector space model to measure the similarity of the sentences based on which it produced a relevant summary. Amarappa and Sathyanarayana [10] discussed the NER and classification in the languages of Kannada. In this language, the main focus is on NE tags and their meaning, which are more affected to generate the summary from input documents. Abdi et al. [1] proposed query-based multi-document summarization with the help of linguistic techniques. It is executed by computing the syntactic and semantic similarity of sentence to query and sentence to sentence. The proposed method needs to minimize the redundancy of generated summary. Zopf et al. [178] discussed the linguistic annotation for the summarization of ETS documents. It focused on bigram representation to be a snippet of more important estimation of sentences in the documents. In this method, text annotations like name entities and frame, i.e., replacing bigrams with several types of linguistic annotations such as connotation frame, verb stem, n-gram, discourses relation sense type, entity type, etc., helps in generating a summary. Baruah et al. [17] focused on text summarization in Assamese languages. It is the most challenging task to summarize the Indian language and generate good accuracy. It summarized the document by WordNet techniques for Assamese languages and produced a relevant summary. Yogatama et al. [174] discussed ETS using Maximizing Semantic Volume (MSV), which maximized bigram coverage for subject constraint. Each sentence is embedded in the semantic space for document summarization, and a summary was generated by selecting sentences from documents. The authors also focused on the greedy algorithms that efficiently perform volume maximization by the Gram-Schmidt process on the budget constraint.

Limitations:

- The discourse-based approaches use domain-specific datasets**.**
- It lacks coherence and cohesion during text summarization of the documents.
- It is not suitable for volume maximization problems.
- It has the possibility of redundant sentences.

6.5 Statistical based approach

Many studies have performed ETS using statistical approaches, which indeed used frequency and centrality. Also, both frequency and centrality are more useful in unsupervised learning. Instead of linguistic features of the document, this approach concentrates on other features such as the position of sentences and the position of words within the document [48]. The researchers used n-grams for extracting information and then applied a statistical approach to fetch the collection of keywords from the given documents. Several kinds of measures are used in this approach, such as TF-IDF, word frequency, term frequency [141], word co-occurrences, etc. Statistical-based approaches (SBA) have many drawbacks, such as ambiguous references, misapplied rhetoric, synonyms, other context-dependent terms, etc. To solve these problems, we require some domain knowledge and linguistic analysis. In such cases, SBA cannot go beyond a certain limit. Naik and Gaonkar [112] proposed the rule-based summarizer, which generated better information from the given documents. The author took the DUC 2002 dataset for rule-based summarizer to many features like sentence position, numerical value, title, and keyword weight which is selected as an essential sentence along with the dataset. Rule-based summarizer compared to the existing GSM summarizer. The rule-based summarizer obtained better average precision, recall, and f-measure than the GSM summarizer. Ravinuthala and Chinnam [130] proposed the Keyword Extraction Approach (KEA). KEA focused on graph themes, where the keyword of the document defines graph representation that is put off into a thematic text graph. This type of graph defines the maximum number of incoming edges. Here, the theme is mentioned as graph representations and topic centrality. KPA was used to select relevant information from the DUC 2002 datasets. KPA improved the quality of ETS summary information. Barrera and Verma [16] described single-document summarization with the semantic and syntactic techniques for sentence scoring, where text, position, and WordNet scores were used. Sentence scoring is based on the mentioned topic's results heading relevant to its sentence position and generates an effective summary based on scientific magazine articles and DUC2002 newswire datasets. Lwin and Nwet [91] discussed the extractive Myanmar News summarization with centroid-based techniques. This technique calculated the rank of sentences with the help of their similarity to the centroid. It created a meaningful summary and centroid based on the word embedding technique to produce a good performance of the summary. Daiya et al. [28] discussed the semantic and statistical methods for multi-document ETS summarization. Here statistical methods used similarity matrices such as TF-IDF, n-gram model. Word co-occurrence semantic methods used the Facebook InferSent model, WordNet, and Glove models. These two methods help to generate a summary from multi-documents. Hernández-Castañeda et al. [57] gave ETS which was independent of language and domain, and it was based on a clustering scheme with the help of GA to get a proper group of sentences. The topic modeling algorithm was then used to extract important sentences from these clusters with the assistance of automatically generated keywords. This results in a relevant summary of the document. Elayeb et al. [36] described the Arabic text summarization using analogical proportions. Analogical proportion represented the relationship between summaries and documents. At the same time, an analogical proportion used independent summaries generated in a few languages such as LSA, Luhn, TextRank, and LexRank using a large corpus of Arabic News Text (ANT), small test collection EASC by computing BLEU and ROUGE metric. Alami et al. [5] proposed the hybrid method of graph-based Arabic text summarization system, which is a combination of semantic and statistical approaches. In contrast, semantic similarity computed the semantic relation and statistical similarity based on the content overlap of two sentences. The PageRank evaluated the rank of sentences from the graph to generate a relevant summary of the documents.

Limitations:

- It has no semantic meaning and mapping the text documents for summarization.
- It is more suitable for a single document and moderate datasets.
- This approach has a problem in finding the position and weight of sentences while summarizing the documents.

6.6 Machine learning approach

Machine learning is a part of artificial intelligence (AI). It utilizes data analysis methods and applies computer algorithms. It uses many approaches or techniques such as SVM [95], k-NN, random forest, Naïve Bayes, CNN, RNN, RBM, NN, HMM, Regression, and Genetic Algorithms (GA), etc. Machine learning is categorized into three types such as semi-supervised, unsupervised, and supervised learning approaches.

Unsupervised learning approach (ULA) It generates text summaries without training data. It neither classifies nor labels the data. The algorithms used in ULAs extract information without guidance. It focuses on a group of CBA summarization according to the patterns, similarities, and without training data. It consists of many techniques such as centrality, similarities, frequency, statistical techniques, and machine learning approaches like HMM, and deep learning techniques (Autoencoder, RBM, CNN and RNN, GRU, LSTM [8]), association rule, and k -mean clustering.

Supervised learning approaches (SLA) It needs labeled data or training datasets to represent the collection of the documents to detect and learn the most important sentences from the documents. It consists of many techniques like SVM [95], Decision Tree (DT), Multilayer Neural Network (MNN), Regression, GA, Naïve Bayes Classifier (NBC), CRF, Linear Classifier, Bagging Models, Boosting Models, Deep Neural Network (DNN), i.e., Recurrent CNN, GRU, BRNN, LSTM, and k-NN [35].

Semi-supervised learning (SSL) approaches It combines labeled and unlabeled data to create a classifier of a suitable summary of the documents. This approach consists of several techniques like NBC, SVM, Logistic Regression, Linear Regression, etc.

Recently, Machine Learning Approach (MLA) is being used in HMM, and it has two classifiers, i.e., the binary classifier and the Bayesian classifier. The Bayes rule validates binary classification. It generates the summary of the documents by calculating the probability of the document. Shen and Li [138] discussed the supervised learning methods to rank for query-focused summarization of multi documents. The authors used two aspects: sentence-to-sentence and training data. These two aspects help to generate the text summary from given input documents. Liu et al. [86] discussed multi-document summarization via unsupervised deep learning (UDL). The authors presented three phases of a uniform framework: concepts extraction, generation of summary, and validation of reconstruction of the final relevant summary. Castillo et al. [22] employed the NER and SVM technique to train the documents for ranking of sentences using multiple features such as sentence location, semantic characteristics, word and phrase features, and named entities. However, SVM is utilized to find the relevant summary of texts in query-based summarization. Mutlu et al. [110] discussed the methods, datasets, and features for sentence selection. The authors focused on supervised learning and used two features such as semantic and syntactic. Glove and Word2vec are two semantic features that fall under embedding. Syntactic features create informative, meaningful summaries based on datasets. Both features (semantic and syntactic) are integrated with the LSTM-NN model. They used SIGIR 2018 corpus datasets and generated the summary. Qaroush et al. [125] described Arabic single document text summarization using semantic and statistical features. According to them, these features find the important sentence, diversity, and coverage of the documents. There used two summary techniques, such as machine learning and rank-based, to create a relevant summary. The performance of the given method was evaluated on the EASC corpus dataset using ROUGE metrics. Arimae and Liu [11] proposed a method for text summarization based on supervised learning. The authors focused on the quality of the summary using question-answers from the given documents. The given method substitutes important sentences of the documents in the form of question-answers. It learns to promote fluent summaries and executes useful question-answers. The performance of the given method was evaluated on CNN/Daily Mail dataset. Luo et al. [89] proposed an approach for ETS that is a contextual bandit problem solved by reinforcement learning. The authors categorized the given framework into careful reading and rough reading. Two NNs are taken from the encoded document used for sentence embedding and producing local and global features. But careful reading has document-level features and sentence vectors provided by rough reading. These types of reading select sentences from documents one by one to form a summary. Sirohi et al. [145] proposed machine learning with LSTM summarization (which is used in both ETS and ATS summarization). The given approach used the word2vec for word embedding to combine the sentences for the relevant summary from the given documents. It generates good summary accuracy. Mendoza et al. [100] proposed the memetic algorithms for single document summarization (MA-SingleDocSum). This method combined the population-based global search with memetic algorithms (Local Search Heuristic (LSH)). LSH uses the problem knowledge to search for the best solutions. The local optimization algorithm is applied in the MA-SingleDocSum, which has GLS (Guided Local Search). MA-SingleDocSum has feature title, length, coverage, and cohesion. These algorithms select suitable sentences from the document and calculates the best results obtained by MA-SingleDocSum. It measures the result using ROUGE-1 and ROUGE-2 metrics based on DUC2001 and DUC2002 datasets. Tarnpradab et al. [154] discussed a HAN model used to create thread and sentence representation. These are performed to the sentence encoder and thread encoder, where the sentence encoder is read as an input sentence and vector as the output sentence. Vector sentences are the weighted sum of the word. Thread encoder is the input of a sequence of sentence vectors so that these models perform for unsupervised and supervised baselines other than MEAD. MEAD is used in terms of sentence prediction. Khurana and Bhatnagar [72] presented Shannon's entropy to capture the informativeness of phrases. A non-negative matrix factorization factor was used to calculate the entropy of nouns, phrases, and themes. The E-Summ algorithm is an unsupervised ETS approach that understands the generated entropy using information theory. By using the information-theoretic approach in a systematic method, the computer selects significant sentences from essential subjects in the document. The technique described is both generic and rapid, making it is ideal for real-time document summarization. It also has no preference for domains, collections, or languages. It just needs positively-oriented non-negative matrix factorization factor matrices and the E-Summ algorithm. The proposed strategy is tested on four well-known public datasets using the standard ROUGE toolset. The authors additionally compute the semantic similarity of the E-Summ summary to the original document to measure its quality quantitatively. Singh et al. [144] proposed a novel method termed as SHEG (Summarization and Headline Generation), which uses the concept of both extractive and the abstractive methodology of summary extraction. This hybrid model was done using a pipeline approach. As the name suggests, this method produces a brief summary of the news and which in turn was used to produce headlines of the news. Srivastava et al. [149] discussed the unsupervised ETS approach. The cluster is paired with topic modeling (TM) to minimize topic bias. For TM techniques, LDA is used, which means the k-medoids cluster is used for summary creation. The ROUGE-metric was used to assess the summary's performance.

Limitations:

- The machine learning approach works in terms of a "bag of word" that is more time-consuming due to a large number of features and less memory.
- When it is used in both unsupervised and supervised techniques, there is a massive training corpus requirement.
- Its dominant by fitness agent over population-based summarization with the help of a memetic algorithm.
- It lacks the ability to provide coverage, diversity, and important sentences and generates redundancy problems.
- It has taken the high cost of computational complexity.

6.7 Neural network approach

A Neural Network (NN) is a processing system that attempts to replicate the learning process of the human brain. NN is a collection of artificial neurons that are linked together. For data processing, a numerical model of computation is used. Each sentence in the test paragraph is reviewed to verify if it should be included in the summary. The user's needs are taken into account when training the model. The classification of summary sentences using NN is accurate, but it takes a long time to train the model. According to the features in the feature vector, Kaikhah in [70] assigned a score to each sentence. Small weights are trimmed during neural network training to remove unusual features. With high-scoring sentences, a summary has been created. In Fig. 10, multi-document is passed as an input. These documents of sentences produce features of sentences at multiple levels using RNN [113]. All levels generate new sentence IDs, such as *Sentence ID$_1$, Sentence ID$_2$,* and *Sentence ID$_3$.* Each sentence ID generated in the previous level is passed to the next level after matching the original document sentences. Finally, the sentences for the summary are generated by matching them with the original document.

SoftMax function (SF) It is used in the N-dimensional sentence vector of the real values. SF produces a sentence in terms of the real value in the range of "0 and 1" in N-dimension, and it is denoted as S(k), where $k\epsilon k_1, k_2, \ldots, k_n$. The function S(k) is mapped as $\mathfrak{R}^N \rightarrow \mathfrak{R}^N$. It is represented by Eq. (28).

$$S(k):\begin{bmatrix} k_1 \\ k_2 \\ . \\ . \\ k_N \end{bmatrix} \rightarrow \begin{bmatrix} S_1 \\ S_2 \\ . \\ . \\ S_N \end{bmatrix} \tag{28}$$

The SF uses the sum of the weight value of sentences from different sentence features [8]. In [156], the authors suggested a neural network-based Vietnamese text summarizing system that employs semi-supervised learning to decrease computation. Here, the sentences are rated using the word set to generate a summary. Chen et al. [23] discussed an RNN language model based on word co-event auxiliary data. In order to construct a summary, the language model uses a

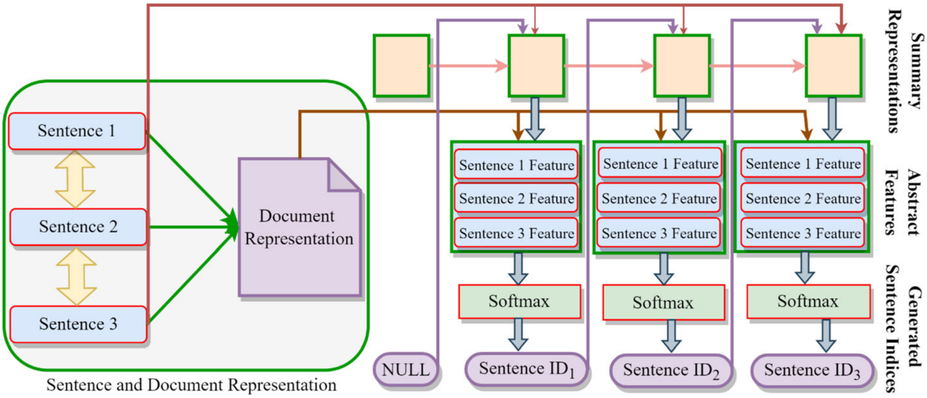

Fig. 10 Extracting the sentences from the document based on Neural Network ETS

probabilistic generative paradigm to rank sentences based on the frequency of each unique term. Wang et al. [164] proposed a heterogeneous graph-based NN for ETS summarization. It produces relevant information document summary based on HETERSUMGRAPH model. Kågebäck et al. [69] suggested an autoencoder to extract phrase embedding, a simple sum of words. Summarization is accomplished by comparing the similarity of phases. LeClair et al. [78] discussed the improved code summarization techniques using Graph Neural Network (GNN). GNN is divided into four parts Spatial-Temporal Graph Neural Networks (STGNNs), Graph Autoencoders (GAEs), Convolutional Graph Neural Networks (ConvGNNs), and Recurrent Graph Neural Networks (RecGNNs). It is helpful for generating meaningful summary from the input document. Al-sabahi et al. [8] proposed a hierarchical model for extractive summarization and used the neural network architecture. The given model includes the information in document structure and utilizes a hierarchical-structured self-attention mechanism for creating the sentences and document embedding. The attention mechanism gives additional knowledge for the summary process. The authors considered the summarization problem as a classification problem and computed the probability for each sentence corresponding to their summary membership through the given model. The model predictions are divided via numerous features such as positional representation, novelty, salience, content richness, information content, and redundancy. The performance of the given model was evaluated on two datasets, DUC2002 and CNN/Daily Mail, using ROUGE metrics. Nallapati et al. [114] presented an extractive model for text summarization based on RNN. The presented model breaks the visualized prediction by the abstract features, for instance, novelty, salience, and information content. The given model eliminates the requirement for sentence-level and extractive labels and then trains model on human-produced summaries. The performance of the given method was evaluated on two datasets such as DUC2002 and CNN/Daily Mail corpus, using ROUGE metrics. Isonuma et al. [61] proposed Multi-Task Learning (MTL) framework which extracts sentences using classification and LSTM-RNN. MTL obtains better features by extracting sentences from the document and evaluating the summary on two datasets: News corpus and financial report. MTL is less useful for the news corpus dataset. Xu and Durrett [168] have presented the NN framework for ETS and suppressible summarization. It uses bidirectional LSTM to encode words for each sentence in the given document, and it generates a meaningful summary. Narayan et al. [115] discussed the sentence ranking model with ETS and summarized the document with the help of reinforcement learning. The training algorithm presented by the authors for document encoder and decoder uses the RNN with LSTM to avoid gradient problems due to long-term training. Finally, the relevant information was generated from the input document. Abdi et al. [2] described a novel deep-learning-based strategy for extracting summarization of multi-documents that is generically opinion-oriented. The sentiment analysis embedding space (SAS), text summarization embedding spaces (TSS), and opinion summarizer module (OSM) are all part of this technique. SAS uses a recurrent neural network (RNN) built of long short-term memory (LSTM) to take advantage of sequential processing and overcome various shortcomings in older approaches, such as the loss of word order and information. TSS uses a variety of statistical and linguistic knowledge factors to improve word-level embedding and extract an appropriate representation. Fitrianah and Jauhari [46] discussed ETS summarization based on LSTM and GRU. Authors used feature engineering to fetch meaning from the given document to produce a meaningful summary of the original document. Hin et al. [59] presented a DL architecture, termed LineVD, which focused on formulating vulnerability detection at the statement level as the node classification task. LineVD uses Graph Neural Network (GNN). It enhances prediction performance, excluding vulnerability status. Gambhir and Gupta [49] proposed a completely data-driven based on the NN model for ETS single-document summarization. They have termed this as a Word-level Attention mechanism (WL-AttenSumm). Authors have focused on selecting salient words in the sentence so that the sentence's meaning is conveyed. Alami et al. [6] described an unsupervised NN technique for Arabic language text summarization. To tackle issues with Arabic text corpus, a new approach based on document clustering, unsupervised NN, and TM was presented to develop an efficient document representation model. The proposed method was tested against other Arabic text summaries algorithms using the ROUGE metric on the EASC dataset. Muthu et al. [108] developed the ETS algorithm for text documents using the Deep Learning Modifier NN (DLMNN) classifier. Based on the entropy levels, it created a valuable summary of the texts. Six steps make up the DLMNN framework. The developed method framework (DLMNN) compared several steps. It is a pre-processed text document and extracted feature of processing text data and selected the most appropriate feature that improves the fruit fly optimization algorithm (IFFOA). This method has been classified into two classes: lowest entropy and highest entropy values. To find the relevant summary of the document, pick the highest entropy values. The developed methods gave an accuracy that is 83.53%.

Limitations:

- Text summarization with the neural network has more complicated computation.
- The neural network's training takes a long time due to the large dataset and uses the multiple number of layers. It also uses a lot of resources.
- NN (deep learning) needs multiple number of levels for a relevant summary.
- There are limitations of memory because datasets need training and testing for the document summarization.
- In this approach, generated summary may contain ambiguous sentences.

6.8 Cluster-based approach

Generally, documents are organized in various sections to address related ideas in one place. It is normal to assume that summaries should cover various topics covered in different sections of the article. If the document covers a wide range of topics, the summarizer takes this into account by grouping the information. According to Fig. 11, a large sentence corpus is generated from the different types of documents, which produces a textual graph with the help of cosine similarity to find the relation (common word) between two sentences. When these sentences have high similarity, they are grouped into clusters (CoRank) [142] as shown in Fig. 11 which includes three clustered groups $CH_1,$ $CH_2,$ and CH_3.

Each cluster group consists of different number of sentences. A cluster has the highest sentence rank values (using TextRank, LexRank, and other ranking methods), and these rank values of sentences are arranged in descending order of the sentences. Sentences are selected for the generation of summary which depends on the compression ratio or the number of words so that relevant summary is generated from the cluster group of sentences. In cluster-based approaches, a TF-IDF score of words is used to represent the document. A high-frequency word denotes a cluster's topic. Cluster-based methods produce high-relevance

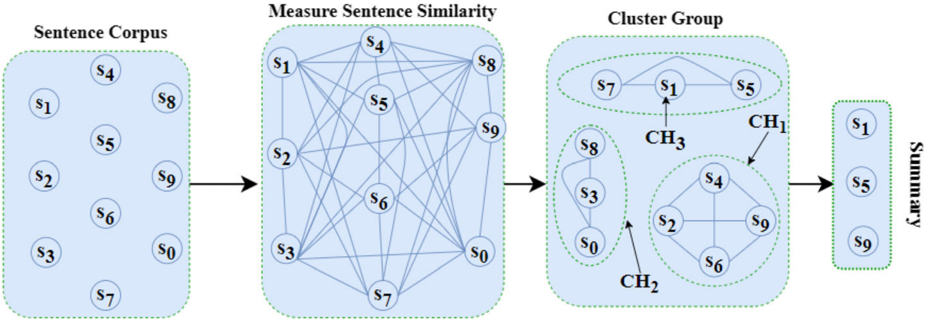

Fig. 11 Summarize the text document using clustered group formation

summaries for a given query or document topic. Jain et al. [62] used the k-means clustering algorithm [94] to create a cluster of texts. The central sentence of the cluster was included in the summary based on its sentence features. Patil and Mahajan [122] took sample sentences from a research paper and grouped them. They created a summary using local and global search strategies. Wu et al. [165] suggested a spectral clustering and LexRank strategy that achieves maximum coverage with the least amount of redundancy. The authors used the k-NN method to create a sparse matrix of comparable texts and derived the LexRank score based on common features to generate a summary. Ferreira et al. [45] offered a network model-based clustering technique that uses TextRank to recognize the important sentences and generated a cluster based on sentence similarity. Zhang et al. [177] suggested a clustering and labeling strategy that constructs a summary by sorting semi-structured connected elements into semantic groups and assigning a label to them. Mei and Chen [99] presented the core SumCR algorithm for multi-document summarization and calculated the sentence similarity, exemplar score, and position score. After exemplar and position scores are calculated, it produces a final summary of the given input document. Mehta and Majumder [98] presented three new ETS summarization techniques: ranking of sentence algorithm, the similarity of sentence metric, and representation of text scheme where numerous different sentence similarities are compared with other existing algorithms. It produces better summary performance. Verma et al. [160] proposed a mixed strategy including cluster and fuzzy logic. The authors focused on cluster-based summarization mainly. The presented method used a partitional CBA to use the cluster sentence of document similarity. The author employs a linear mix of the normalized word mover and Google distances to distinguish the two sentences. The partitional clustering technique determined the partitions for each document. It isolates the meaningful sentences from each cluster and recognizes them based on their updated text feature ratings. It employed a teaching learning-based optimization strategy to determine the best weights for the text features in another way. A fuzzy inference system with a full-fledged knowledge base supplied by humans is used to establish the final sentence score. Three datasets are used in the suggested methods: DUC2001, DUC2002, and CNN. It produced a relevant summary.

Limitations:

- The efficient formation of a cluster is a problem in this approach.
- The resulting summary may lack an appropriate flow when the sentences from distinct group are not related.
- It faces a problem of sentence ordering during the summary generation.

6.9 Optimization-based approach

The text summarization problems are generally addressed by single-objective or multi-objective optimization-based approaches (OBA). For example, Sanchez-Gomez et al. [135, 136] performed an experimental study by comparing the numerous criteria appropriate for standard extractive multi-document text summarization. They considered every probable combination of two, three, and four objectives: redundancy reduction, content coverage, coherence, and relevance. The authors evaluated the performance and compared all the combinations of objective functions on DUC datasets using ROUGE metrics. An OBA is presented which summarizes documents by scoring in two phases. In the first phase, it generates an appropriate representation for the text input, such as the widely used vector representation. In the second phase, there are one or more optimization criteria such as redundancy reduction, content coverage, coherence, and relevance when an optimization algorithm is utilized to pick the summary sentences based on the needed summary length. Furthermore, the capability of genetic algorithms in modifying weights could be applied to ETS. For an ETS genetic algorithm based on a summarizer, the process of sentence scoring can be performed in two steps: (1) recognizing text properties such as sentence length, sentence position, and so on from the input text, and (2) adjusting the weights of these factors when computing the sentence scores using the genetic algorithm [97]. Jain et al. [64] proposed the PSO (Particle swarm optimization) algorithm text summarization based on the Punjabi language. It summarized bilingual Punjab-Hindi text corpus and monolingual Punjabi text corpus summarization. Gamal et al. [47] focused on Chicken Swarm Optimization (CSO) and the Genetic Algorithm (GA) of summarization. Both these methods obtained optimal solutions and generated the effective summary with the help of hybrid techniques using CSO and GA algorithms. Mirshojaei and Masoomi [104] used the CSO (Cuckoo Search Optimization) algorithm and increased ETS approach performance. ETS approach is based on the standard DUC2002 dataset and analyzes the result using the ROUGE metric. MirShojaee et al. [103] discussed the BBO (biography-based optimization) algorithm applied in the domain of ETS summarization, and it produces relevant information summary from standard text datasets (DUC2002). It measures the performance of the DUC2002 dataset by ROUGE software. Mandal et al. [92] proposed summarizing the document in an extractive way using the PSO algorithm. The PSO algorithm is also named as population-based stochastic method, and it has several similarities with evolutionary methods like GA. The dimension and large document term have been calculated through the term-document matrix. The presented algorithm uses K-mean clustering with PSO techniques, which obtain the optimal number of concept clusters. For text summarizing, these key concepts were employed to determine the relevant summary in documents. Srivastava et al. [148] discussed a new extractive-based approach where the document's redundant parts combine to form a single text file from many text files. Word Mover Distance (WMD) and Modified Normalized Google Distance (M-NGD) (WM) Hybrid Weight Method-based similarity techniques accomplish content coverage and non-redundancy features. Authors have applied the Dolphin swarm optimization (DSO), a metaheuristic technique, for feature weight optimization. The suggested technique is evaluated using ROUGE and AutoSummENG metrics and is tested in python using the Multiling 2013 dataset. Potnurwar et al. [123] discussed the binary PSO based on ETS summarization to use multi-Hindi text documents. It has many features for finding the result like TF-IDF, Thematic word, Numerical Data (ND), sentence length, sentence position, and title features. However, binary PSO generated the optimal value of features. Nagalla and Kumar [111] presented the OLOA and DNN-based multi-document summarization. The selection of optimal sentences from the document is done using the OLOA techniques and calculation of the rank of sentences through DNN techniques. Based on the sentence score, it generates a relevant summary.

Limitations:

- Optimization-based approach for ETS summarization takes a long time and a lot of costs to compute the sentence's score.
- The number of iterations for optimization-based approach techniques must be specified.
- The generated summary may contain ambiguous sentences.

Table 5 shows the selection of Paper ID. Table 6 shows the summary of existing ETS approaches based on their characteristics, techniques or approaches, and performance. Table 7 shows the comparison of current ETS approaches based on the number of documents, domain types, and datasets. Table 8 lists the abbreviations used in Table 7.

6.10 Discussion and analysis

We have discussed many extractive text summarization approaches such as fuzzy-logic, latent semantic analysis, graph-based, discourses-based, statistical-based, machine learning, neural network, cluster-based, and optimization-based approaches. Each approach has some set of constraints with it. For example, the fuzzy-logic technique enhances sentence ranking concerns, but the fuzzy rule has limited scopes. The LSA technique has a polysemous problem. It gives semantic relationships and produces more knowledge with little noise. Graph approaches generate summaries using ranking algorithms and determine the accuracy of these summaries with the help of mathematical functions. There is a lack of linguistic semantics in graph approaches. The discourse approach uses languages to construct a summary, but it requires a domain-based dataset to overcome the cohesiveness and coherence. The statistical method doesn't require a training dataset; it processes data quickly and generates a summary without comprehending semantics or discourses. Machine learning techniques yield a relevant summary based on the features extracted from the documents and use a large set of test datasets. The neural network approach reduces redundancy and produces a relevant summary by taking care of sentence flow, but it requires a large data corpus. Cluster-based approaches avoid the repetition of sentences during the summarization. In the optimization-based approach, the weights of sentences are used to generate a meaningful summary from the given input documents.

We have compared all the discussed approaches based on their characteristics, performance, type of documents, and datasets from the twelve-year research done in extractive text summarization. A pie-chart comparison of ETS approaches is shown in Fig. 12. According to the chart, neural networks, machine learning, and graph-based techniques are the favorite ETS approaches, with 17%, 14%, and 15% of total research done in ETS.

7 Research gap, open issues, and research challenges in ETS

This survey addressed various research gaps, issues, and challenges from several existing research articles.

The research gaps based on existing ETS approaches as shown in Fig. 13:

Table 5 Paper IDs corresponding to their references

Paper ID	Authors name and Ref.	Paper ID	Authors name and Ref.
1	Kiyomarsi and Esfahani [73]	48	Shen and Li [138]
2	Dixit and Apte [32]	49	Liu et al. [86]
3	Irfan and Zulfikar [60]	50	Castillo et al. [22]
4	Sahba et al. [132]	51	Mutlu et al. [110]
5	Kumar et al. [75]	52	Qaroush et al. [125]
6	Abhiman and Hiraman [3]	53	Arumae and Liu [11]
7	Yadav et al. [169]	54	Luo et al. [89]
8	Shirwandkar and Kulkarni [139]	55	Sirohi et al. [145]
9	Patel et al. [121]	56	Mendoza et al. [100]
10	Kumar et al. [76]	57	Tarnpradab et al. [154]
11	Ozsoy et al. [118]	58	Khurana and Bhatnagar [72]
12	Moiyadi et al. [106]	59	Singh et al. [144]
13	Chowdhury et al. [26]	60	Srivastava et al. [149]
14	Merchant and Pande [101]	61	Chen et al. [23]
15	Suleman and Korkontzelos [152, 153]	62	Wang et al. [164]
16	Lwin and Nwet [90]	63	Kågebäck et al. [69]
17	Mohamed et al. [105]	64	LeClair et al. [78]
18	Herskovic et al. [58]	65	Al-Sabahi et al. [8]
19	Li et al. [79]	66	Nallapati et al. [114]
20	Al-Taani and Al-Omour [9]	67	Nallapati et al. [113]
21	Fang et al. [42]	68	Isonuma et al. [61]
22	Azadani et al. [14]	69	Xu and Durrett [168]
23	Baralis et al. [15]	70	Narayan et al. [115]
24	Parveen and Strube [120]	71	Abdi et al. [2]
25	Vale et al. [158]	72	Fitrianah and Jauhari [46]
26	Uçkan and Karci [157]	73	Hin et al. [59]
27	Elbarougy et al. [37]	74	Gambhir and Gupta [49]
28	Cao and Zhuge [21]	75	Alami et al. [6]
29	Dutta et al. [34]	76	Muthu et al. [108]
30	Awan and Beg [13]	77	Jain et al. [62]
31	Joshi et al. [68]	78	Patil and Mahajan [122]
32	Koto et al. [74]	79	Wu et al. [165]
33	Nawaz et al. [116]	80	Ferreira et al. [45]
34	Gupta et al. [55]	81	Zhang et al. [177]
35	Amarappa and Sathyanarayana [10]	82	Mehta and Majumder [98]
36	Abdi et al. [1]	83	Mei and Chen [99]
37	Zopf et al. [178]	84	Verma et al. [160]
38	Baruah et al. [17]	85	Sanchez-Gomez et al. [136]
39	Yogatama et al. [174]	86	Jain et al. [64]
40	Naik and Gaonkar [112]	87	Gamal et al. [47]
41	Ravinuthala and Chinnam [130]	88	Mirshojaei and Masoomi [104]
42	Barrera and Verma [16]	89	MirShojaee et al. [103]
43	Lwin and Nwet [91]	90	Mandal et al. [92]
44	Daiya et al. [28]	91	Srivastava et al. [148]
45	Hernández-Castañeda et al. [57]	92	Potnurwar et al. [123]
46	Elayeb et al. [36]	93	Nagalla and Kumar [111]
47	Alami et al. [5]		

- **Sentence Sequencing Problem:** Many research articles lack proper sentence arrangement. While creating meaningful summaries in numerous research papers, sentence sequencing is more crucial issue in ETS.

Table 6 Summary of existing ETS approaches based on their characteristics, techniques/approaches, and performance

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
1	Similarity to Title, Similarity to Keywords, Sentence length, Sentence position	Precision, Recall	Fuzzy Logic Approach	Evaluated the performance of summary through compression rate (CR). When CR = 10%, precision = 0.3788, recall = 0.4352; CR = 20%, precision = 0.4101, recall = 0.4467; CR = 30%, precision = 0.4336, recall = 0.4505.
2	Sentence to Sentence Similarity, Term weight, Title, Sentence Length, etc.	Cosine similarity	Fuzzy Logic Approach	Calculated the similarity of fuzzy summarizer, Copernic summarizer, and MS-Word 2007 summarizer. Compared the similarity between reference and candidate summary of fuzzy summarizer, Copernic summarizer, and MS-Word 2007 summarizer. The fuzzy summarizer has an average 79% resemblance to the reference summary, and it is highest among the other two summarizers.
3	Centroid, Membership value	Precision, Recall, F-score	Fuzzy Logic Approach	The results of system summaries and manual summaries are compared resulting in highest recall value of 65% at the evaluation stage, and precision and f-score as 30% and 41.05%, respectively.
4	Term weight, Proper noun, sentence length, etc.	ROUGE-1 ROUGE-2 ROUGE-L	Fuzzy Logic Approach	Evaluated the ROUGE scores of various models such as fuzzy logic system (ROUGE-1 = 25.73, ROUGE-2 = 25.04, ROUGE-L = 24.52) and sequence-to-sequence (S2S) ((ROUGE-1 = 28.12, ROUGE-2 = 16.20, ROUGE-L = 26.79)).
5	Sentence length, Sentence Similarity	Precision, Recall, F-score	Fuzzy Logic Approach	Calculated the performance of the DUC 2002 datasets summary using ROUGE metrics. For ROUGE-1 (FUZZY CST with COM) value of precision, recal1 and f-score is 0.33206, 0.34101, and 0.33568, and for ROUGE-1 (FUZZY CST with COM) value of precision, recal1 and f-score is 0.12806, 0.12986, and 0.1287.
6	Thematic features, Term weight, etc.	Precision, Recall	Fuzzy Logic Approach	Modern-featured basic text summarization achieved precision in performance and a recall score of 90%.
7	Title similarity, sentence length, sentence centrality, etc.	ROUGE-1 ROUGE-2	Fuzzy Logic Approach	For ROUGE-1: precision=0.48514, recall=0.52895, f-measure=0.50408. ROUGE2 gives the average value of precision, recall, f-measure for all the summarizer.
8	Sentence Position, Sentence length	Precision, Recall F-measure	Fuzzy Logic Approach	Precision = 88%, Recall = 80%, F-measure = 84%.

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
9	Title-word, Thematic-word, Proper-noun word, sentence position, etc.	ROUGE-2 ROUGE-4	Fuzzy Logic Approach	Evaluated the performance of ROUGE-2 and ROUGE-4. MDS (ROUGE-2) for precision, recall and f-measure value was 0.155, 0.073, and 0.099 respectively. MDS (ROUGE-4) for precision, recall and f-measure value was 0.068, 0.026, and 0.038 respectively.
10	Sentence position, sentence length, title, etc.	ROUGE-1	Fuzzy Logic Approach	ROUGE-1 = 0.87634 by using hybrid summarizers.
11	Sentence Position, Sentence Order	ROUGE-1 ROUGE-L	Latent Semantic Analysis Approach	For DUC2002 dataset: ROUGE-1: Cross = 0.333, Topic=0.347; ROUGE-L: Cross=0.453, Topic=0.428 and for DUC 2003: ROUGE-1: Cross=0.085, Topic=0.083; ROUGE-L: Cross=0.072, Topic=0.075.
12	Positive Keyword. Negative Keyword, Centrality	ROUGE-1	Latent Semantic Analysis Approach	It has only summarized the text document but does not calculate the summary results.
13	Sentence position, Sentence length	F-score	Latent Semantic Analysis Approach	The performance of the LSI model is 0.324347 (F-score).
14	Sentence extraction	ROUGE-1 ROUGE-2 ROUGE-L	Latent Semantic Analysis Approach	Using DUC Dataset: ROGUE-1 = 0.58, ROGUE-2 = 0.15, ROGUE-L = 0.35.
15	Sentence Dependency Structure	F-score	Latent Semantic Analysis Approach	A comparison between standard LSA and xLSA has been made. Standard LSA has 100% semantic similarity but xLSA has much lesser semantic similarity.
16	Word Segmentation, Stop words removal etc.	Precision Recall F-score	Latent Semantic Analysis Approach	Measured the performance of ROUGE-2. Cross Method: Precision= 0.71, Recall = 0.60, F-score = 0.65. Steinberger and Jezek's Approach: Precision = 0.66, Recall = 0.58, F-score = 0.61.
17	Characteristic features, Attributes	ROUGE-1 ROUGE-2 ROUGE-L	Latent Semantic Analysis Approach	For SDS (ROUGE-1 = 0.504, ROUGE-2 = 0.235, ROUGE-L = 0.335) values were obtained; For MDS (ROUGE-1 = 0.474, ROUGE-2 = 0.212, ROUGE-SU4 = 0.246) obtained.
18	Sentence Length	Recall Precision	Graph Based Approach	

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
				MEDRank boosted major heading recall by 26% from MTI’s 0.376–0.475. MEDLINE indexing boosted recall of essential concepts by 30% while lowering overall precision by 4%.
19	Sentence Position	ROUGE-1 ROUGE-2	Graph Based Approach	The performance of TAC 2008 and 2009 datasets: TAC 2008: ROUGE-1 = 0.37002, ROUGE-2 = 0.08812, BewT = 0.20143; TAC 2009: ROUGE-1 = 0.36247, ROUGE-2 = 0.08554, BewT = 0.19255.
20	Term frequency Title similarity Sentence position Sentence length	Precision	Graph Based Approach	Evaluated the performance of n-grams compression ratio (CR) summary as: For 25% CR summary: precision = 0.502, and for 40% CR summary: precision = 0.492.
21	Sentence Rank	ROUGE-1 ROUGE-2 ROUGE-L	Graph Based Approach	Calculated the performance of CoRank+ Method based on two datasets Chinese News and DUC 2002. For Chinese News: F-score=0.594, ROUGE-1=0.697, ROUGE-2=0.606, ROUGE-L=0.661, and For DUC 2002: F-score=0.524, ROUGE-1=0.526, ROUGE-2=0.258, ROUGE-L=0.451.
22	Sentence selection	ROUGE-1 ROUGE-2 ROUGE-W-1.2 ROUGE-SU4	Graph Based Approach	The performance of ROUGE-1=0.7648, ROUGE-2=0.3524, ROUGE-W-1.2=0.0913, ROUGE-SU4=0.4090.
23	Term relevance, Correlation	Recall Precision F-score	Graph Based Approach	GRAPHSUM: ROUGE-2 (recall = 0.093, precision = 0.099, f-score = 0.097); ROUGE-SU4 (recall = 0.015, precision = 0.021, f-score = 0.015).
24	Sentence ranking	Hyperlink Induced Topic Search (HITS), ROUGE	Graph Based Approach	The performance of (1) PLOS Medicine: System Coherence and Position using R-SU = 0.224 and R-2 = 0.189; (2) DUC 2002: Single document summarization using R-SU=0.253, R-2=0.230, and R-1 = 0.485.
25	Sentence location, Word rank Lexical similarity	Precision Recall F-score	Graph Based Approach	Evaluated the result of WORDFREQ based on MUSST: Precision = 0.381, recall = 0.502, and F-score = 0.424.
26	Sentence Ranking, Sentence Position,	ROUGE	Graph Based Approach	The performance of DUC document (DUC2002, DUC2004) datasets: For 100-word Summary: ROUGE =0.38072; For

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
	Sentence Order			140-word Summary: ROUGE =0.59208; For 200-word Summary: ROUGE =0.51954.
27	Sentence Position, Title, keyword, numerical value	Precision, Recall, F-measure	Graph Based Approach	Recall=72.94, Precision=68.75, and F-measure=67.99.
28	Sentence length	ROUGE-N (N ∈ {1, 2, 3, 4}), ROUGE-L	Graph Based Approach	Measured the performance of sentence Group: R1=75%, R2=77.78%, R3=83.33%, R4=77.78%.; Paragraph: R1=63.89%%, R2=66.67%, R3=77.78%, R4=69.44% etc.
29	Sentence length, Sentence score	ROUGE-1, ROUGE-2, ROUGE-L	Graph Based Approach	For 25% of summary rate; ROUGE-1: Precision=0.6125, Recall= 0.57198, F-measure=0.59155); ROUGE-2: (Precision=0.28033, Recall=0.26172, F-measure=0.27071); ROUGE-L: (Precision= 0.450, Recall=0.42023, F-measure=0.43461).
30	Phrase position ranking TopicalRank Keyphrase selection	F-score	Graph Based Approach	Material – 0.778, Process – 0.742, Task – 0.420, F1-score – 0.730
31	Sentence rank, Sentence Position	ROUGE (R-1, R-2, R-L)	Graph-based Approach	SoA utilized ROUGE measures on the DUC2002 dataset. It scored 53.2, 27.9, and 49.3 on R-1, R-2, and R-L. Accuracy improved for R-1, R-2, and R-L scores by 44.5, 24.0, and 41.0 in the CNN/Daily Mail dataset.
32	Sentence order	Spearman correlation	Discourse Based Approach	It compared seven pre-trained language models like English, Chinese, German, and Spanish by spearman correlation.
33	Sentence weight, word frequency	ROUGE-1, ROUGE-2	Discourse Based Approach	The performance of SWA dataset, UCE-ROUGE. ROUGE-1: (UCE) Precision=0.78, Recall=0.81, F-measure=0.80; ROUGE-2: (UCE) Precision=0.47, Recall=0.61, F-measure=0.53.
34	Term frequency, Sentence Score	Precision Recall	Discourse Based Approach	Calculated the summary for 20% and 30% number of sentences. For 20%: precision=81, recall=86; and for 30%: precision=83, recall=89.
35	Annotate Corpora, Tagging	Summary Comparison	Discourse Based Approach	Generated summary of Kannada language.
36	Sentence rank	ROUGE-1 ROUGE-2	Discourse Based Approach	

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
		ROUGE-SU4		The performance of DUC2006 datasets: For DUC2006 (ROUGE-1=0.4287, ROUGE-2=0.0968, ROUGE-SU4=01673, ARS=0.2310).
37	Verb stems Bigrams Chunks	Evaluate the value of Verb Stem.	Discourse Based Approach	DUC 2003 (Verb-stem=0.497), DUC 2004 (Verb-stem=0.517), TAC 2008 (Verb-stem=0.515), and TAC 2009 (Verb-stem=0.500), etc.
38	Cue words Word Frequency, Assamese WordNet	ROUGE-1 ROUGE-2	Discourse Based Approach	The document summary's two performance compression ratios as 30% and 50%, and used ROUGE-1 and ROUGE-2 metrics
39	Sentence Selection	ROUGE ($R_{1,}$ R_2)	Discourse Based Approach	Determined the result on dataset TAC 2008/09: TAC 2008 (Volume 500): R_1=37.40, R_2=9.17; TAC 2009 (Volume 500): R_1=34.08, R_2=8.91; TAC 2008 (Volume 600): R_1=37.50, R_2=9.58; TAC 2009 (Volume 600): R_1=34.37, R_2=8.76.
40	Sentence position, title, numerical value, and keyword weight	Recall, Precision, F-measure	Statistical Based Approach	Calculated the value of Precision, Recall, F-Measure for Rule-Based Summarizer:0.784, 0.424, and 0.5068.GSM Summarizer: 0.578, 0.295, 0.358 from fifteen set of documents from the DUC dataset.
41	Topical word, keyword extraction	ROUGE-1, ROUGE-2	Statistical Based Approach	ROUGE-1: (ES-KWI) Precision=0.51257, Recall=0.61380, F-measure=0.55838; ROUGE-2: (ES-KWI) Precision=0.40122, Recall=0.48129, F-measure=0.43741.
42	TextRank score, Position Score, WordNet Score, Sentence score.	Precision Recall F-score	Statistical Based Approach	Computed the performance of DUC2002 dataset. SynSem: Precision=0.45130, Recall=0.48258, F-score=0.46631.
43	Sentence Selection, Sentence Embedding, etc.	Precision Recall F-score	Statistical Based Approach	Measured the performance of two summarizer: Bag of words (BOW) and centroid based words (CBW) embedding summarizer. ROUGE-2 (BOW): R=46%, P=41%, F=43%; ROUGE-2 (CBW): R=74%, P=76%, F=75%.
44	Sentence Rank, Sentence pair Similarity	ROUGE-2	Statistical Based Approach	For DUC2004 used ROUGE–2: Glove-vec Model=0.03054; WordNet Model=0.03354; InferSent Model=0.0381.
45	One-hot encoding TF-IDF Doc2Vec	ROUGE-N (R1, R2, RL, R-SU)	Statistical Based Approach	For CNN Dataset: (R1–41.4, R2–18.4, RL – 37.6) For DUC2002 Dataset (R1–0.48681, R2–0.23334, RSU – 0.24954)
46	Diversity, Sentence order	ROUGE–1	Statistical Based Approach	

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
		BLEU–1		ANT corpus v2.1: AATS1 (ROUGE–1 = 0.83, BLEU–1 = 0.54) and AATS2 (ROUGE–1 = 0.80, BLEU–1 = 049); EASC collection: AATS1 (ROUGE–1 = 0.75, BLEU–1 = 0.47) and AATS2 (ROUGE–1 = 0.74, BLEU–1 = 0.49).
47	Semantic Similarity, Statistic Similarity	Precision Recall F-score	Statistical Based Approach	Compression ratio summary size was 30% for Dataset-1: Precision = 56.89, Recall = 58.93, F1-measure = 57.89 and Dataset-2: Precision = 68.49, recall = 59.03, F1-measure = 63.41.
48	Sentence weight, Sentence position	ROUGE (R_1, R_2, R-W, R-SU)	Machine Learning Approach	Evaluated the performance of SVR based on DUC datasets, used ROUGE metric (R1, R2, R-W, R-SU): For DUC2006 (R1= 0.4166, R20.0952, R-W=0.1106, R-SU=0.1517), For DUC2006 (R1 = 0.4395, R2 = 01179, R-W = 0.1163, R-SU = 0.1652).
49	Sentence Position, Sentence Length	ROUGE (R_1, R_2, R-W, R-SU)	Machine Learning Approach	Evaluated using ROUGE metric over DUC Datasets such as DUC2005 (R-1 = 0.3751, R-2 = 0.0775, R-SU4 = 0.1341), DUC2006 (R-1 = 0.4015, R-2 = 0.0928, R-SU4 = 0.1479), DUC2007 (R-1 = 0.4295, R-2 = 0.1163, R-SU4 = 0.1685).
50	Sentence Location, Name entity, Semantic word, etc.	Precision Recall F-measure	Machine Learning Approach	Precision = 85.29%, Recall = 85.83%, and F-Measure = 85.56%.
51	Sentence Ranking, Syntactic, Semantic.	ROUGE-1, ROUGE-2	Machine Learning Approach	Summarization feature: ROUGE (Ensemble feature of ROUGE-1 = 0.65, Semantic features of ROUGE-1 = 0.60); Summarization methods: LSTM-NN based on SummaRuNNe modified the ROUGE-1 by three percentages.
52	Sentence Centrality, Sentence Length etc.	ROUGE-1, ROUGE-2	Machine Learning Approach	Calculated the performance of two datasets using ROUGE-2. F-scores of 0.617 and 0.524 for both datasets.
53	Sentence Position	ROUGE (R_1, R_2, R_L)	Machine Learning Approach	Evaluated results over two datasets, CNN and Daily Mail. For CNN: QASumm+NER: R_1 = 27.38, R_2 = 9.38, R_L = 19.02; For Daily Mail: QASumm+NER R_1 = 25.74, R_2 = 11.89, R_L = 22.38.
54	Sentence Decoder, Bandit Policy, Termination Mechanism	ROUGE-1, ROUGE-2, ROUGE-L	Machine Learning Approach	Calculated HER model f1-score through ROUGE metric (R_1 = 42.3, R_2 = 18.9, and R_L = 37.9).
55		ROUGE-1	Machine Learning Approach	Generated a relevant summary with the help of LSTM techniques.

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
	Sentence similarity to the title, Sentence length			
56	Position, Length, Title, Cohesion	ROUGE-1, ROUGE-2	Machine Learning Approach	Evaluated MA-SingleDocSum on two datasets DUC 2001 and DUC 2002. For Dataset DUC 2001: ROUGE-1 = 0.44862, ROUGE-2 = 0.20142; For Dataset DUC 2002: ROUGE-1 = 0.48280, ROUGE-2 = 0.22840.
57	Sentence Encoder Thread Decoder	ROUGE-1, ROUGE-2	Machine Learning Approach	Measured the performance of HAN model. ROUGE-1: P=42.8%, R=31.0%, F=33.7±0.7%; ROUGE-2: P=17.8%, R=11.2%, F=12.7±0.5%; Sentence-level: P=34.1%, R=26.9%, F=32.4±0.5.
58	Sentence score, Sentence Selection	ROUGE-1, ROUGE-2, ROUGE-L	Machine Learning Approach	The accuracy of E-Summ summaries was 89% semantically comparable to the source material. The best and worst-performing Marathi documents was summarized in E-Summ summaries.
59	Headline length, Dependency features	ROUGE-1, ROUGE-2, ROUGE-L	Machine Learning Approach	R1–31.82 R2–13.2 RL – 28.80 on Gigaword dataset R1–40.67 R2–17.74 RL – 36.69 on CNN/Daily Mail dataset
60	Sentence Extraction	ROUGE (R-1, R-2, R-L)	Machine Learning Approach	The R-1, R-2, and R-L equivalent values for the CNN/Daily Mail Dataset are 43.90%, 19.01%, and 41.50% respectively. For the DUC dataset: R-1, R-2, and R-L were 49.35%, 31.53%, and 41.72% respectively.
61	Acoustic feature, Sentence selection	ROUGE (R-1, R-2, R-L)	Neural Network Approach	Calculated the F-scores of ROUGE metrics (R-1 = 0.600, R-2 = 0.523, R-L = 0.527).
62	Sentence Selection, Sentence Score	ROUGE (R-1, R-2, R-L)	Neural Network Approach	For HDSG model: R-1 = 46.05, R-2 = 16.35, and R-L = 42.08

63	Sentence Position	ROUGE (R-1, R-2, R-SU4)	Neural Network Approach	Calculated Recall ®, Precision (P), and F-measure (F) with respect to ROUGE.
				R-1(OPTR): R = 57.86, P = 21.96, F = 30.28;
				R-2(OPTR): R = 22.96, P = 12.31, F = 15.33;
				R-SU4(OPTR): R = 29.50, P = 13.53, F = 17.70.

64	Sentence Sequence	BLEU Score	Neural Network Approach	

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
				The ConvGNN improved aggregate BLEU scores (BLEU-A) by 4.6% over other graph-based techniques and by 5.7% over flattened AST approaches when used to encode the AST.
65	Novelty, Salience, Information Content, Sentences Position	ROUGE (for n-grams)	Neural Network Approach	Evaluated the performance over two datasets DUC2002 and CNN/Daily Mail with the help of ROUGE-N metric (N=1, 2 and L): (1) R_1=0.521, R_2=0.245 and R_L=0.488 (2) R_1=0.423, R_2=0.178 and R_L=0.376.
66	Novelty, Salience, Information Content, etc.	ROUGE (R-N, R-L), where N=1,2	Neural Network Approach	SummaRuNNer for Daily Mail datasets (limited length recall is 75 byte): ROUGE-1=26.2±0.4, ROUGE-2=10.8±0.3, ROUGE-L=14.4±0.3. SummaRuNNer for Daily Mail datasets (limited length recall is 275 byte): ROUGE-1=42.0±0.2, ROUGE-2=16.9±0.4, ROUGE-L=34.1±0.3. SummaRuNNer for DUC2002 datasets (limited length recall is 75 byte): ROUGE-1=46.6±0.8, ROUGE-2=23.1±0.9, ROUGE-L=43.03±0.8.
67	Sentence position, Sentence Order	ROUGE (R_1, R_2, R_L score)	Neural Network Approach	Evaluated the performance of Recall at 75 bytes: Deep model classifier gave result better than shallow model classifier. Rouge-1=26.2+±0.4, Rouge-2=10.7±0.4, Rouge-L=14.4±0.4 and recall at 275-byte, Rouge-1=42.2±0.2 and Rouge-L=35.0±0.2.
68	Document Classification, Sentence extraction	ROUGE (R_1, R_2)	Neural Network Approach	Evaluated ROUGE score from NYTAC datasets. NN-ML-CL: for 100-word reference summary: $R_{1=}$18.1, R_2=12.7; for 250-word reference summary: $R_{1=}$18.3, R_2=12.7; for 500-word reference summary: $R_{1=}$18.5, R_2=12.9.
69	Sentence length, Sentence location etc.	ROUGE (R_1, R_2, R_L)	Neural Network Approach	For performance, two datasets were used: CNN and CNNDM. For CNN: R_1=29.1, R_2=11.1, R_L=25.8; For CNNDM: R_1=40.3, R_2=17.6, R_L=36.4.
70	Sentence Ranking, Max Pooling	ROUGE (R_1, R_2, R_L)	Neural Network Approach	For performance, three datasets were used: CNN, Daily Mail, and CNN/Daily Mail. For CNN: R_1=29.1, R_2=11.1, R_L=25.9 For Daily Mail: R_1=40.7, R_2=18.3, R_L=37.2; For CNN/Daily Mail: R_1=39.6, R_2=17.7, R_L=36.2.
71	Sentence location		Neural Network Approach	For DUC2001 Dataset: R1–0.3857

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
	Sentence Length Title Cue-phrases	ROUGE-1, ROUGE-2, ROUGE-L		R2–0.0864 Average ROUGE Score (ARS) – 0.2361
72	Paragraph Location, Fixed Phrase feature etc.	ROUGE-N (R1, R2, RL)	Neural Network Approach	(R1–76.25%, R2–59.49%, RL – 72.72%)
73	Classifier Learning, Graph Construction	F-score, Precision, Recall	Neural Network Approach	F1–0.36, Recall – 0.533, Precision – 0.271
74	Sentence length Sentence Position	ROUGE-N (R1, R2, RL)	Neural Network Approach	For Daily Mail Corpus (R-1=32.8%, R-2=11.0%, R-L=27.5%) For DUC2002 Dataset: (R-1=55.9%, R-2=24.8%, R-L=53.9%) For CNN/Dail Mail Dataset: (R-1=42.9%, R-2=19.7%, R-L= 39.3%)
75	AE, VAE, ELM-AE	Precision, Recall, F-score	Neural Network Approach	The execution of the summarization task has the highest accuracy compared to various current approaches for summarizing Arabic documents.
76	Similarity with the Title, Sentence similarity, Proper noun	F-measure Sensitivity Specificity Accuracy	Neural Network Approach	With the help of the DLMNN classifier, the system's performance results F-measure=89.72, sensitivity =81.56, accuracy =91.21, and specificity =83.53.
77	Sentence order	Threshold value	Cluster-Based Approach	Calculated the edge weight of two sentences (two vertexes) with the help of a threshold and find the summary from the given documents.
78	Centroid Sentence, Term-frequency	Comparison classifier value	Cluster-Based Approach	Evaluated the four-classifier method such as SID2, SIS2, DIS2, and DID2.
79	Sentence Rank	Precision Recall F-score	Cluster-Based Approach	The performance of LexRank of the 20% compression ratio: precision=0.3240, recall=0.2209, and f-score=0.2541.
80	Sentence order	F-measure	Cluster-Based Approach	Calculated the performance of the 200 and 400-word summary using the DUC 2002 dataset for the new system. For 200-word summary (F-measure =30) and 400-word summary (F-measure =25.4).
81	Category as Label, Word as Label	Graph model Tree model	Cluster-Based Approach	Calculated cluster group by WiiCluster techniques.

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
82	Centroid, Sentence similarity	Recall	Cluster-Based Approach	The performance of sentence similarity was improved by about 5% to 10% in ROUGE-1 (recall values).
83	Key words, title words, cue words, and sentence location	ROUGE-N (R_2, R_{SU4})	Cluster-Based Approach	Result of SumCR-Q were (DUC2004: R_2=0.0965, R_{SU4}=0.1364; DUC2005: R_2=0.0700, R_{SU4}=0.1251; DUC2006: R_2=0.0906, R_{SU4}=0.1437, etc.).
84	Sentence Position, Sentence length, etc.	ROUGE-1, ROUGE-2	Cluster-Based Approach	The fuzzy logic approach provided a 50% increase in the accuracy of the generated summary.
85	Relevance, Content Coverage	ROUGE-2, ROUGE-L	Optimization Based Approach	Using dataset DUC2002, calculated the results for ROUGE-2 and ROUGE-L. Three parameters were compared: average, range, and CV. For ROUGE-2: Average=0.339, Range=0.001, CV=0.23%; For ROUGE-L: Average=0.567, Range=0.000, CV=0.000%.
86	Title Similarity Sentence Length, etc.	Precision Recall F-measure	Optimization Based Approach	Calculated the performance of Punjabi dataset: ROUGE-1 Precision=0.6904 Recall=0.6731 F-Measure=0.6816; and ROUGE-2 Precision=0.5892 Recall=0.5442 F-Measure=0.565.
87	Title Feature Sentence Length Sentence Position ND, TW	ROUGE-1, ROUGE-2	Optimization Based Approach	Measured the performance of CNN/DM datasets, and provided high accuracy of ROUGE-1=4.4% and ROUGE-2=12.01%.
88	Sentence Location	Evaluate the value of documents by the CSO method	Optimization Based Approach	Evaluated the result using DUC2002 datasets: CSO Algorithm (d061j=0.49761, d067f=0.46476, d070f=0.47126, and d105g=0.42391).
89	Sentence weight	Precision Recall F-score	Optimization Based Approach	Compared 400-words document to calculate the value of d105g using BBO algorithms: precision=0.43130, recall=0.66488, and F-score=0.52320.
90	Weight feature, Matrix Factorization	Precision Recall F-measure	Optimization Based Approach	ROUGE-1: Precision=48.16, Recall=43.15 F-Measure=45.48, and ROUGE-2: Precision=21.841, Recall=17.12 F-Measure=19.12.
91	Sentence Position Sentence Length	ROUGE-1, ROUGE-2	Optimization Based NoneApproach	R1 (P - 0.351, R - 0.127, F1–0.201) R2 (P - 0.160, R - 0.069, F1–0.091)
92	Sentence position, Sentence length,	Precision Recall	Optimization Based NoneApproach	

Table 6 (continued)

Paper ID	Features	Evaluation	Techniques/ Approaches	Performance
	Thematic word, Numerical None Data, TF-IDF	F-measure		Measured the performance of news articles by precision, recall, and f-measure. News number has good performance than others. Precision=0.83, recall=0.82, and F-measure=0.824.
93	Sentence Score, Sentence length	F- measure	Optimization Based NoneApproach	Improved the performance using OLOA and DNN (F-measure value is OLOA=0.64% and DNN=1.14% uses multi-documents) than existing PSO-DNN and DNN algorithm.

Table 7 Comparison of current ETS approaches based on the number of documents, domain types, and datasets

Paper ID	Type of Document		Domain Types					Dataset
	S_D	M_D	TR	EA	WP	NA	JA	
1	✓	–	–	–	–	–	–	OT
2	✓	–	–	–	–	–	–	OT
3	✓	–	–	–	–	–	–	OT
4	✓	–	–	–	–	–	–	C_1
5	–	✓	–	–	–	✓	–	D_2
6	–	✓	–	–	–	–	–	OT
7	✓	–	–	–	–	–	–	D_2
8	✓	–	–	–	–	✓	–	–
9	–	✓	–	–	–	–	–	D_5
10	✓	–	–	–	–	–	–	OT
11	✓	–	–	–	–	–	–	D_2, D_5
12	✓	–	–	–	–	–	–	OT
13	✓	–	–	–	–	–	–	OT
14	✓	✓	–	–	–	–	–	OT
15	✓	–	–	–	–	–	–	OT
16	✓	–	–	–	–	✓	–	–
17	✓	✓	–	–	–	–	–	D_2
18	✓	–	–	–	–	–	–	OT
19	–	✓	–	–	–	–	–	T_1
20	✓	–	–	–	–	–	–	E_1
21	✓	–	–	–	–	✓	–	D_2
22	✓	–	–	–	–	–	–	OT
23	–	✓	–	–	–	–	–	D_5
24	✓	✓	–	–	–	–	✓	D_2
25	–	✓	–	–	–	✓	–	–
26	–	✓	–	–	–	✓	–	D_2, D_5
27	✓	–	–	–	–	–	–	O_1
28	✓	–	–	–	–	–	–	O_1
29	✓	–	–	–	–	–	–	OT
30	✓	–	–	–	–	–	–	OT
31	✓	–	–	–	–	–	–	D_2, C_1
32	✓	–	–	–	–	–	–	OT
33	✓	–	–	–	–	–	–	OT
34	✓	–	–	–	–	–	–	OT
35	✓	–	–	–	–	–	–	OT
36	–	✓	–	–	–	–	–	D_6, D_7
37	–	✓	–	–	–	–	–	D_5, T_1
38	✓	–	–	–	–	–	–	OT
39	–	✓	–	–	–	–	–	T_1
40	✓	–	–	–	–	–	–	D_2
41	✓	–	–	–	–	–	–	D_2
42	–	✓	–	–	–	–	–	D_2
43	✓	–	–	✓	–	✓	–	–
44	–	✓	–	–	–	–	–	D_5
45	✓	–	–	–	–	–	–	D_5, C_5, OT
46	✓	–	–	–	–	–	–	OT
47	✓	–	–	–	–	–	–	OT
48	–	✓	–	–	–	–	–	D_6, D_7, D_8
49	–	✓	–	–	–	–	–	D_6, D_7, D_8
50	✓	–	–	–	–	–	–	OT
51	✓	–	–	–	–	–	–	OT
52	✓	–	–	–	–	–	–	O_1
53	✓	–	–	–	–	–	–	C_1
54	✓	–	–	–	–	–	–	C_1

Table 7 (continued)

Paper ID	Type of Document		Domain Types					Dataset
	S_D	M_D	TR	EA	WP	NA	JA	
55	✓	–	–	–	–	–	–	OT
56	✓	✓	–	–	–	✓	–	D_2, D_4
57	–	✓	–	–	–	–	–	OT
58	–	✓	–	–	–	–	–	D_3, D_4, C_1
59	–	✓	–	–	–	–	–	C_1, OT
60	✓	–	–	–	–	–	–	D_2, C_1
61	–	✓	–	–	–	✓	–	–
62	✓	–	–	–	–	–	–	$C_{1,}$ OT
63	–	✓	✓	–	–	–	–	–
64	✓	–	–	–	–	–	–	OT
65	✓	–	–	–	–	–	–	C_1
66	✓	–	–	–	–	–	–	D_3
67	✓	–	–	–	–	–	–	D_3
68	✓	–	✓	–	–	–	–	–
69	✓	–	–	–	–	✓	–	C_1
70	✓	–	–	–	–	✓	–	C_1
71	–	✓	–	–	–	–	–	D_4
72	✓	–	–	–	–	–	–	OT
73	✓	–	–	–	–	–	–	OT
74	✓	–	–	–	–	–	–	D_2, C_1, OT
75	✓	–	–	–	–	–	–	E_1
76	✓	✓	–	–	–	–	–	D_1
77	✓	–	–	–	–	–	–	OT
78	–	✓	–	–	–	–	–	OT
79	✓	–	–	–	–	–	–	D_1
80	–	✓	–	–	–	✓	–	–
81	✓	–	–	–	–	–	–	OT
82	–	✓	–	–	–	–	–	D_2, D_5, D_8
83	✓	–	–	–	–	–	–	OT
84	–	✓	–	–	–	–	–	$D_{3,}$ D_4
85	✓	–	–	–	–	–	–	D_2
86	–	✓	–	–	–	–	–	OT
87	✓	–	–	–	–	–	–	C_1
88	✓	–	–	–	–	–	–	D_2
89	✓	–	–	–	–	–	–	D_2
90	✓	–	–	–	–	–	–	OT
91	–	✓	–	–	–	✓	–	OT
92	–	✓	–	–	–	✓	–	–
93	–	✓	–	–	–	–	–	OT

- **Computation cost**: The major problem with ETS input document summarization is computation cost. Several research papers have low accuracy, and the cost of computation is extremely high. There is a need to improve the cost of document summarization.
- **Reliability of generated summary:** Various approaches were used to construct the quality of summary from multi-document datasets, which hamper reliability.
- **Grammatical Inconsistency:** The resulting summary must not contain any inappropriate sentences, capitalization errors, or grammar rules.
- **Formation of Clusters Groups:** Creating numerous cluster groups from a multi-document input is a major difficulty in cluster-based ETS summarization.

Table 8 List of abbreviations used in Table 7

Type of Domain							Type of Document				
EA	Encyclopedia Articles		WP		Web Pages		S_D	Single Document	M_D	Multiple Document	
NA	Newspaper Articles		TR		Technical Report						
JA	Journal Articles										
Datasets											
D_1	DUC	D_2	DUC2002	D_3	DUC2002 and DailyMail			D_4	DUC2001	D_5	DUC2004
D_6	DUC2005	D_7	DUC2006		D_8	DUC2007	D_9	DUC2003	O_1	Opinions Opinion	
OT	Other datasets		C_1	CNN/DM		E_1	EASC corpus	T_1	TAC-2008 and TAC-2009		

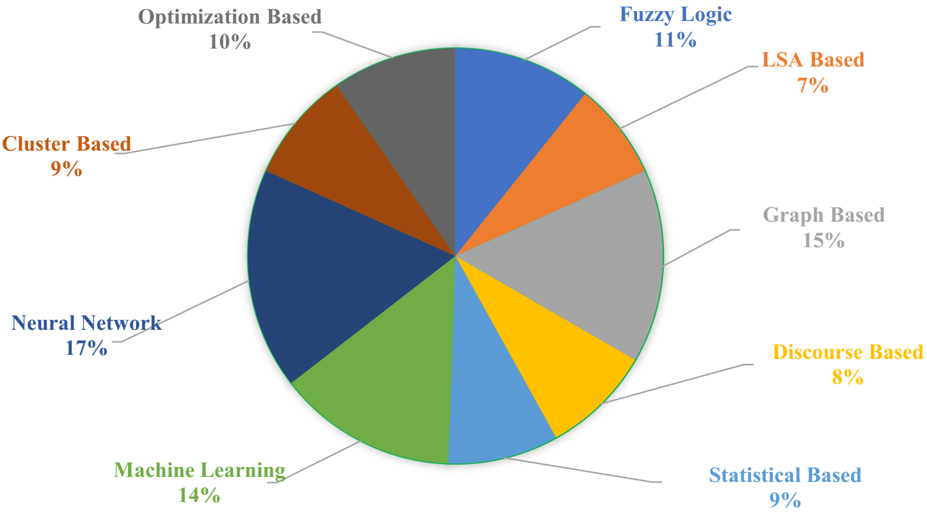

Fig. 12 Comparison of existing ETS approaches in the 12-year research

The issues based on existing ETS approaches as shown in Fig. 13:

- **Hindi Summarization using ETS:** In Hindi, collecting nouns and pronouns in extractive text summarization is a significant issue, and Hindi language datasets are hard to find out.
- **Multiple Languages:** Using multiple languages is a huge problem when it comes to summarizing numerous documents simultaneously. Summary generated using this leads to various other issues like improper readability, poor understanding, and meaninglessness (lacking any significance).

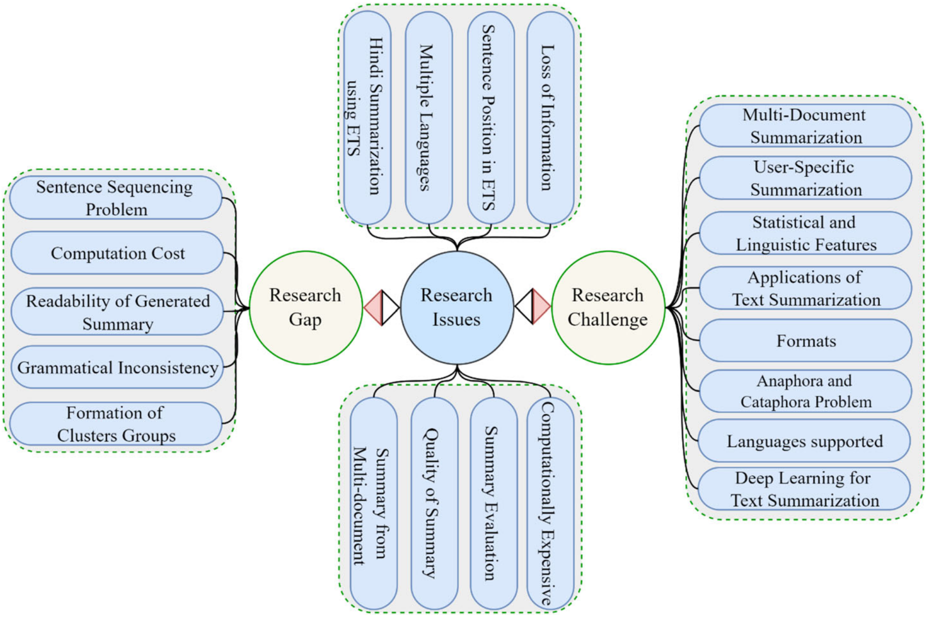
Hindi Summarization using ETS
Multiple Languages
Sentence Position in ETS
Loss of Information
Sentence Sequencing Problem
Computation Cost
Readability of Generated Summary
Grammatical Inconsistency
Formation of Clusters Groups
Research Gap
Research Issues
Research Challenge
Multi-Document Summarization
User-Specific Summarization
Statistical and Linguistic Features
Applications of Text Summarization
Formats
Anaphora and Cataphora Problem
Languages supported
Deep Learning for Text Summarization
Summary from Multi-document
Quality of Summary
Summary Evaluation
Computationally Expensive

Fig. 13 Several research gaps, issues, and challenges of ETS

- **Sentence Position in ETS:** Sentence position is a major issue in multi-document summarization for the relevant summary from the input documents. This is due to summary formation based on the rank of the documents' sentences. The position of the sentences should be correct for the generation of a relevant summary.
- **Loss of Information:** Graph-based approach help to remove the independent node from the graph in the document. So, the chances of losing some critical information increases, which affects the generated summary.
- **Summary from Multi-document:** There are many issues related to multi-document summarization. Some issues arise during the assessment of summaries, such as a co-reference, redundancy, temporal dimension, sentence position, etc.
- **Quality of Summary:** The quality of summaries varies from person to person or from one system to another. Some people believe that some sentences are vital for summary, while others believe that other sentences are necessary for the required summary.
- **Summary Evaluation:** The summary evaluation is a major issue in ETS summarization. Selecting a correct evaluation method for the generation of a relevant summary is a significant issue.
- **Computationally Expensive:** The cost is a major problem in the summary generation. Reducing the cost for the relevant summary is challenging.

It is challenging to arrange sentences and find the relevant summary of the documents. However, more concerns arise, such as cohesion, coherence, and summary construction while generating a summary. The challenges based on existing ETS approaches as shown in Fig. 13:

- **Multi-Document Summarization:** The most important task in creating text summarization in a multi-document is co-referencing multiple sentences [48]. It might generate improper reference if the pronoun is handled without taking care of the proper noun. Other major things which can produce improper sentences are temporal dimension and redundancy [133].
- **User-Specific Summarization:** User-specific summaries depend on various factors like languages, multiple text sources, semi-structured data, etc. It also requires focusing on applying sentiment analysis and user-specific personalized summaries [54].
- **Statistical and Linguistic Features:** To further improve extractive text summarization, it is required to work on the linguistic features. Using hybrid approaches to summarize text can also build newer statistical models to extract key sentences from text to improve quality [54].
- **Applications of Text Summarization:** Most of the research in text summarization is dedicated to online news, reviews, online pages, and multi-documents, but still, many domains are left out and are more challenging such as summarizing very large-scale text, which contains legal books, medical textbooks, novels, etc. Therefore, they require different strategies to handle [166].
- **Formats:** Visualization is less time-consuming than reading and contains more details in a small space. Moreover, there are different multi-media formats where the summaries are explained. In addition, vice versa can summarize other media formats like video, audio, etc., which are attached to online meetings, etc. The summaries should also focus on readability [22]. Removing ambiguous sentences and confusing text when summaries have multiple topics and documents improves readability.
- **Anaphora and Cataphora Problem:** Furthermore, people frequently introduce the challenges and then attempt to explain them later in the text using synonyms or pronouns.

Understanding which pronoun replaces which previously provided phrase is referred to as an "anaphora difficulty" in texts. Similarly, researchers confront the opposite difficulty (when ambiguous phrases and descriptions addressing a specific term are used in the text before the term itself), which is known as the "cataphora problem" [124].

- **Languages supported:** Expanding the text summaries to other more popular languages is also an important and challenging task [18]. It requires developing and improving syntactical and semantics parsing of non-English languages.
- **Deep Learning for Text Summarization:** Deep neural networks have been used for both abstractive summarization and extractive summarization. RNN and CNN have been used successfully by many researchers and industries. These algorithms require large-scale data to train and validate the model. Sometimes a particular domain dataset may not have such a scale that it is required to build deep learning models on smaller datasets [162].

8 Conclusion and future scope

Extractive text summarization is an extremely interesting topic in the area of NLP, with an in-depth examination of different ways that provide non-redundant, short, logical, and useful information of documents in the form of summaries. Enough work has been done in literature in this area, and still, the researchers are exploring this area to give efficient methods to generate summaries from the given documents. We have discussed the review methodology of the selection of papers for this survey and explained the ETS summarization steps for all ETS approaches. We have given a classification of extractive text summarization approaches and reviewed them based on their characteristics, techniques, and performance. We have explained various techniques, starting from a basic statistical-based approach upto difficult approaches like the NN approach, optimization-based approach, etc. All approaches are dependent on single or multi-documents. We have also discussed the limitation of all ETS approaches, and several ETS datasets are discussed in detail, such as DUC, TAC, and others, which are used in the current status of ETS approaches of the documents. The quality of the summary is achieved by using evaluation measures techniques. Further, we have categorized and explained evaluation measures utilized in ETS approaches. Moreover, we have discussed the research gaps, issues, and challenges encountered in extractive text summarization.

In the future, we'll look at abstractive text summarizing methods and discuss their taxonomy, benefits, and drawbacks. We'll also look at the research gaps, issues and challenges, and difficulties that come with abstractive text summarization based on deep learning. Further, we focus on the other categories of multimedia summarization methods such as image, audio, and video summarization.

Acknowledgments We value the opinions of innominate reviewers.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest in this manuscript.

References

1. Abdi A, Idris N, Alguliyev RM, Aliguliyev RM (2017) Query-based multi-documents summarization using linguistic knowledge and content word expansion. Soft Comput 21(7):1785–1801. https://doi.org/10.1007/s00500-015-1881-4
2. Abdi A, Hasan S, Shamsuddin SM, Idris N, Piran J (2021) A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion. Knowl-Based Syst 213:106658. https://doi.org/10.1016/j.knosys.2020.106658
3. Abhiman, BD, Hiraman, PY (2021) A text summarization using multi linguistic features and fuzzy logic technique of sentences
4. Alami N, Meknassi M, En-nahnahi N (2019) Enhancing unsupervised neural networks-based text summarization with word embedding and ensemble learning. Expert Syst Appl 123:195–211. https://doi.org/10.1016/j.eswa.2019.01.037
5. Alami N, Mallahi ME, Amakdouf H, Qjidaa H (2021) Hybrid method for text summarization based on statistical and semantic treatment. Multimed Tools Appl 80(13):19567–19600. https://doi.org/10.1007/s11042-021-10613-9
6. Alami N, Meknassi M, En-nahnahi N, El Adlouni Y, Ammor O (2021) Unsupervised neural networks for automatic Arabic text summarization using document clustering and topic modeling. Expert Syst Appl 172:114652. https://doi.org/10.1016/j.eswa.2021.114652
7. Ali ZH, Hussein AK, Abass HK, Fadel E (2021) Extractive multi document summarization using harmony search algorithm. Telkomnika 19(1):89–95. https://doi.org/10.12928/TELKOMNIKA.v19i1.15766
8. Al-Sabahi K, Zuping Z, Nadher M (2018) A hierarchical structured self-attentive model for extractive document summarization (HSSAS). IEEE Access 6:24205–24212. https://doi.org/10.1109/ACCESS.2018.2829199
9. Al-Taani, AT, Al-Omour, MM (2014) An extractive graph-based Arabic text summarization approach. In The International Arab Conference on Information Technology
10. Amarappa S, Sathyanarayana SV (2013) Named entity recognition and classification in kannada language. Int J Electron Comput Sci Eng 2(1):281–289
11. Arumae K, Liu F (2019) Guiding extractive summarization with question-answering rewards. arXiv preprint arXiv:1904.02321. https://doi.org/10.48550/arXiv.1904.02321
12. Asa AS, Akter S, Uddin MP, Hossain MD, Roy SK, Afjal MI (2017) A comprehensive survey on extractive text summarization techniques. Am J Eng Res 6(1):226–239
13. Awan MN, Beg MO (2021) Top-rank: a topicalpostionrank for extraction and classification of keyphrases in text. Comput Speech Lang 65:101116. https://doi.org/10.1016/j.csl.2020.101116
14. Azadani MN, Ghadiri N, Davoodijam E (2018) Graph-based biomedical text summarization: an itemset mining and sentence clustering approach. J Biomed Inform 84:42–58. https://doi.org/10.1016/j.jbi.2018.06.005
15. Baralis E, Cagliero L, Mahoto N, Fiori A (2013) GRAPHSUM: discovering correlations among multiple terms for graph-based summarization. Inf Sci 249:96–109. https://doi.org/10.1016/j.ins.2013.06.046
16. Barrera A, Verma R (2011) Automated extractive single-document summarization: beating the baselines with a new approach. In proceedings of the 2011 ACM symposium on applied computing (pp. 268-269). https://doi.org/10.1145/1982185.1982247
17. Baruah N, Sarma SK, Borkotokey S (2019) A novel approach of text summarization using Assamese WordNet. In 2019 4th international conference on information systems and computer networks (ISCON) (pp. 305-310). IEEE. https://doi.org/10.1109/ISCON47742.2019.9036285
18. Belkebir R, Guessoum A (2018) TALAA-ATSF: a global operation-based Arabic text summarization framework. In intelligent natural language processing: trends and applications (pp. 435–459). Springer, Cham. https://doi.org/10.1007/978-3-319-67056-0_21
19. Bommasani R, Cardie C (2020) Intrinsic evaluation of summarization datasets. In proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP) (pp. 8075-8096). https://doi.org/10.18653/v1/2020.emnlp-main.649
20. Cai H, Zheng VW, Chang KCC (2018) A comprehensive survey of graph embedding: problems, techniques, and applications. IEEE Trans Knowl Data Eng 30(9):1616–1637. https://doi.org/10.1109/TKDE.2018.2807452
21. Cao M, Zhuge H (2020) Grouping sentences as better language unit for extractive text summarization. Futur Gener Comput Syst 109:331–359. https://doi.org/10.1016/j.future.2020.03.046
22. Castillo JM, Mateo MAL, Paras AD, Sagum RA, Santos VDF (2013) Named entity recognition using support vector machine for Filipino text documents. Int J Future Comput Commun 2(5):530–532. https://doi.org/10.7763/IJFCC.2013.V2.220
23. Chen KY, Liu SH, Chen B, Wang HM, Jan EE, Hsu WL, Chen HH (2015) Extractive broadcast news summarization leveraging recurrent neural network language modeling techniques. IEEE/ACM Transact Audio, Speech, Lang Process 23(8):1322–1334. https://doi.org/10.1109/TASLP.2015.2432578

24. Chieu HL, Lee YK (2004) Query based event extraction along a timeline. In proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval (pp. 425-432). https://doi.org/10.1145/1008992.1009065
25. Chouigui A, Ben Khiroun O, Elayeb B (2021) An arabic multi-source news corpus: experimenting on single-document extractive summarization. Arab J Sci Eng 46(4):3925–3938. https://doi.org/10.1007/s13369-020-05258-z
26. Chowdhury SR, Sarkar K, Dam S (2017) An approach to generic Bengali text summarization using latent semantic analysis. In 2017 international conference on information technology (ICIT) (pp. 11-16). IEEE. https://doi.org/10.1109/ICIT.2017.12
27. Cizmeciler K, Erdem E, Erdem A (2022) Leveraging semantic saliency maps for query-specific video summarization. Multimed Tools Appl 81(12):17457–17482. https://doi.org/10.1007/s11042-022-12442-w
28. Daiya D, Singh A, Jadon M (2018) Using statistical and semantic models for multi-document summarization. arXiv preprint arXiv:1805.04579. https://doi.org/10.48550/arXiv.1805.04579
29. Dang HT (2005) Overview of DUC 2005. In proceedings of the document understanding conference (Vol. 2005, pp. 1-12)
30. Dang HT (2006) DUC 2005: evaluation of question-focused summarization systems. In proceedings of the workshop on task-focused summarization and question answering (pp. 48-55). https://aclanthology.org/W06-0707.pdf
31. Dernoncourt F, Ghassemi M, Chang W (2018) A repository of corpora for summarization. In proceedings of the eleventh international conference on language resources and evaluation (LREC 2018). https://aclanthology.org/L18-1509.pdf
32. Dixit RS, Apte SS (2012) Improvement of text summarization using fuzzy logic-based method. IOSR J Comput Eng (IOSRJCE) 5(6):5–10 http://www.iosrjournals.org/
33. Dunlavy DM, O'Leary DP, Conroy JM, Schlesinger JD (2007) QCS: a system for querying, clustering and summarizing documents. Inf Process Manag 43(6):1588–1605. https://doi.org/10.1016/j.ipm.2007.01.003
34. Dutta M, Das AK, Mallick C, Sarkar A, Das AK (2019) A graph-based approach on extractive summarization. In emerging Technologies in Data Mining and Information Security (pp. 179–187). Springer, Singapore. https://doi.org/10.1007/978-981-13-1498-8_16
35. Dwivedi V, Ghosh S (2022) Classification of Hindi compound nouns using machine learning. SN Comput Sci 3(1):1–5. https://doi.org/10.1007/s42979-021-00895-z
36. Elayeb B, Chouigui A, Bounhas M, Khiroun OB (2020) Automatic arabic text summarization using analogical proportions. Cogn Comput 12(5):1043–1069. https://doi.org/10.1007/s12559-020-09748-y
37. Elbarougy R, Behery G, El Khatib A (2020) Extractive Arabic text summarization using modified PageRank algorithm. Egypt Inf J 21(2):73–81. https://doi.org/10.1016/j.eij.2019.11.001
38. El-Haj MO, Hammo BH (2008) Evaluation of query-based Arabic text summarization system. In 2008 international conference on natural language processing and knowledge engineering (pp. 1-7). IEEE. https://doi.org/10.1109/NLPKE.2008.4906790
39. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK (2021) Automatic text summarization: a comprehensive survey. Expert Syst Appl 165:113679. https://doi.org/10.1016/j.eswa.2020.113679
40. Elrefaiy A, Abas AR, Elhenawy I (2018) Review of recent techniques for extractive text summarization. J Theor Appl Inf Technol 96(23):7739–7759
41. Erkan G, Radev DR (2004) Lexrank: graph-based lexical centrality as salience in text summarization. J Artif Intell Res 22:457–479. https://doi.org/10.1613/jair.1523
42. Fang C, Mu D, Deng Z, Wu Z (2017) Word-sentence co-ranking for automatic extractive text summarization. Expert Syst Appl 72:189–195. https://doi.org/10.1016/j.eswa.2016.12.021
43. Fei L, Hu Y, Xiao F, Chen L, Deng Y (2016) A modified topsis method based on numbers and its applications in human resources selection Mathematical Problems in Engineering, 2016. https://doi.org/10.1155/2016/6145196
44. Ferreira R, de Souza Cabral L, Lins RD, e Silva GP, Freitas F, Cavalcanti GD, Favaro L (2013) Assessing sentence scoring techniques for extractive text summarization. Expert Syst Appl 40(14):5755–5764. https://doi.org/10.1016/j.eswa.2013.04.023
45. Ferreira R, de Souza Cabral L, Freitas F, Lins RD, de França Silva G, Simske SJ, Favaro L (2014) A multi-document summarization system based on statistics and linguistic treatment. Expert Syst Appl 41(13): 5780–5787. https://doi.org/10.1016/j.eswa.2014.03.023
46. Fitrianah D, Jauhari RN (2022) Extractive text summarization for scientific journal articles using long short-term memory and gated recurrent units. Bullet Electr Eng Inf 11(1). https://doi.org/10.11591/eei.v11i1.3278
47. Gamal M, El-Sawy A, AbuEl-Atta AH (2021) Hybrid Algorithm Based on Chicken Swarm Optimization and Genetic Algorithm for Text Summarization. Int J Intell Eng Syst, Vol.14, No.3, https://doi.org/10.22266/ijies2021.0630.27

48. Gambhir M, Gupta V (2017) Recent automatic text summarization techniques: a survey. Artif Intell Rev 47(1):1–66. https://doi.org/10.1007/s10462-016-9475-9
49. Gambhir M, Gupta V (2022) Deep learning-based extractive text summarization with word-level attention mechanism. Multimed Tools Appl, 1-24. https://doi.org/10.1007/s11042-022-12729-y
50. Gholamrezazadeh S, Salehi MA, Gholamzadeh B (2009) A comprehensive survey on text summarization systems. In: 2009 2nd international conference on computer science and its applications. IEEE, pp 1–6. https://doi.org/10.1109/CSA.2009.5404226
51. Goldman J, Renals S, Bird S, De Jong F, Federico M, Fleischhauer C, Wright R (2005) Accessing the spoken word. Int J Digit Libr 5(4):287–298. https://doi.org/10.1007/s00799-004-0101-0
52. Gong Y, Liu X (2001) Generic text summarization using relevance measure and latent semantic analysis. In proceedings of the 24th annual international ACM SIGIR conference on research and development in information retrieval (pp. 19-25). https://doi.org/10.1145/383952.383955
53. Goularte FB, Nassar SM, Fileto R, Saggion H (2019) A text summarization method based on fuzzy rules and applicable to automated assessment. Expert Syst Appl 115:264–275. https://doi.org/10.1016/j.eswa.2018.07.047
54. Gupta V, Lehal GS (2010) A survey of text summarization extractive techniques. J Emerg Technol Web Intell 2(3):258–268. https://doi.org/10.4304/jetwi.2.3.258-268
55. Gupta P, Pendluri VS, Vats I (2011) Summarizing text by ranking text units according to shallow linguistic features. In 13th international conference on advanced communication technology (ICACT2011) (pp. 1620-1625). IEEE
56. Hassel M (2004) Evaluation of automatic text summarization. Licentiate Thesis, Stockholm, Sweden, pp 1–75
57. Hernández-Castañeda Á, García-Hernández RA, Ledeneva Y, Millán-Hernández CE (2022) Language-independent extractive automatic text summarization based on automatic keyword extraction. Comput Speech Lang 71:101267. https://doi.org/10.1016/j.csl.2021.101267
58. Herskovic JR, Cohen T, Subramanian D, Iyengar MS, Smith JW, Bernstam EV (2011) MEDRank: using graph-based concept ranking to index biomedical texts. Int J Med Inform 80(6):431–441. https://doi.org/10.1016/j.ijmedinf.2011.02.008
59. Hin D, Kan A, Chen H, Babar MA (2022) LineVD: statement-level vulnerability detection using graph neural networks. arXiv preprint arXiv:2203.05181.https://doi.org/10.48550/arXiv.2203.05181
60. Irfan M, Zulfikar WB (2017) Implementation of fuzzy C-means algorithm and TF-IDF on English journal summary. In 2017 second international conference on informatics and computing (ICIC) (pp. 1-5). IEEE. https://doi.org/10.1109/IAC.2017.8280646
61. Isonuma M, Fujino T, Mori J, Matsuo Y, Sakata I (2017) Extractive summarization using multi-task learning with document classification. In proceedings of the 2017 conference on empirical methods in natural language processing (pp. 2101-2110). https://doi.org/10.18653/v1/D17-1223
62. Jain HJ, Bewoor MS, Patil SH (2012) Context sensitive text summarization using k means clustering algorithm. Int J Soft Comput Eng 2(2):301–304
63. Jain D, Borah MD, Biswas A (2021) Automatic summarization of legal bills: a comparative analysis of classical extractive approaches. In 2021 international conference on computing, communication, and intelligent systems (ICCCIS) (pp. 394-400). IEEE. https://doi.org/10.1109/ICCCIS51004.2021.9397119
64. Jain A, Yadav D, Arora A (2021) Particle swarm optimization for Punjabi text summarization. Int J Oper Res Inf Syst (IJORIS) 12(3):1–17. https://doi.org/10.4018/IJORIS.20210701.oa1
65. Jang M, Kang P (2021) Learning-free unsupervised extractive summarization model. IEEE Access 9: 14358–14368. https://doi.org/10.1109/ACCESS.2021.3051237
66. Jones KS (2007) Automatic summarising: the state of the art. Inf Process Manag 43(6):1449–1481. https://doi.org/10.1016/j.ipm.2007.03.009
67. Joshi A, Fidalgo E, Alegre E, Fernández-Robles L (2019) SummCoder: an unsupervised framework for extractive text summarization based on deep auto-encoders. Expert Syst Appl 129:200–215. https://doi.org/10.1016/j.eswa.2019.03.045
68. Joshi A, Fidalgo E, Alegre E, Alaiz-Rodriguez R (2022) RankSum—an unsupervised extractive text summarization based on rank fusion. Expert Syst Appl 200:116846. https://doi.org/10.1016/j.eswa.2022.116846
69. Kågebäck M, Mogren O, Tahmasebi N, Dubhashi D (2014) Extractive summarization using continuous vector space models. In proceedings of the 2nd workshop on continuous vector space models and their compositionality (CVSC) (pp. 31-39). https://aclanthology.org/W14-1504.pdf
70. Kaikhah K (2004) Automatic text summarization with neural networks. In 2004 2nd international IEEE conference on'Intelligent Systems'. Proceedings (IEEE cat. No. 04EX791) (Vol. 1, pp. 40-44). IEEE. https://doi.org/10.1109/IS.2004.1344634

71. Keyvanpour MR, Shirzad MB, Rashidghalam H (2019) Elts: a brief review for extractive learning-based text summarizatoin algorithms. In 2019 5th international conference on web research (ICWR) (pp. 234-239). IEEE. https://doi.org/10.1109/ICWR.2019.8765294
72. Khurana A, Bhatnagar V (2022) Investigating entropy for extractive document summarization. Expert Syst Appl 187:115820. https://doi.org/10.1016/j.eswa.2021.115820
73. Kiyomarsi F, Esfahani FR (2011) Optimizing persian text summarization based on fuzzy logic approach. In 2011 international conference on intelligent building and management
74. Koto F, Lau JH, Baldwin T (2021) Discourse probing of pretrained language models. arXiv preprint arXiv: 2104.05882. https://doi.org/10.48550/arXiv.2104.05882
75. Kumar YJ, Salim N, Abuobieda A, Albaham AT (2014) Multi document summarization based on news components using fuzzy cross-document relations. Appl Soft Comput 21:265–279. https://doi.org/10.1016/j.asoc.2014.03.041
76. Kumar A, Sharma A, Nayyar A (2020) Fuzzy logic-based hybrid model for automatic extractive text summarization. In proceedings of the 2020 5th international conference on intelligent information technology (pp. 7-15). https://doi.org/10.1145/3385209.3385235
77. Kumar Y, Kaur K, Kaur S (2021) Study of automatic text summarization approaches in different languages. Artif Intell Rev 54(8):5897–5929. https://doi.org/10.1007/s10462-021-09964-4
78. LeClair A, Haque S, Wu L, McMillan C (2020) Improved code summarization via a graph neural network. In proceedings of the 28th international conference on program comprehension (pp. 184-195). https://doi.org/10.1145/3387904.3389268
79. Li X, Du L, Shen YD (2012) Update summarization via graph-based sentence ranking. IEEE Trans Knowl Data Eng 25(5):1162–1174. https://doi.org/10.1109/TKDE.2012.42
80. Lins RD, Oliveira H, Cabral L, Batista J, Tenorio B, Salcedo DA, Simske SJ (2019) The CNN-Corpus in Spanish: a large Corpus for extractive text summarization in the Spanish language. In proceedings of the ACM symposium on document engineering 2019 (pp. 1-4). https://doi.org/10.1145/3342558.3345423
81. Lins RD, Oliveira H, Cabral L, Batista J, Tenorio B, Ferreira R, Simske SJ (2019) The cnn-corpus: A large textual corpus for single-document extractive summarization. In Proceedings of the ACM Symposium on Document Engineering 2019 (pp. 1–10). https://doi.org/10.1145/3342558.3345388
82. Lins RD, Mello RF, Simske S (2019) DocEng'19 competition on extractive text summarization. In proceedings of the ACM symposium on document engineering 2019 (pp. 1-2). https://doi.org/10.1145/3342558.3351874
83. Lins RD, de Mello RF, Simske SJ (2020) DocEng'2020 competition on extractive text summarization. In proceedings of the ACM symposium on document engineering 2020 (pp. 1-4). https://doi.org/10.1145/3395027.3419579
84. Liu B (2012) Sentiment analysis and opinion mining. Synth Lectures Human Lang Technol 5(1):1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016
85. Liu F, Liu Y (2008) Correlation between rouge and human evaluation of extractive meeting summaries. In proceedings of ACL-08: HLT, short papers (pp. 201-204). https://aclanthology.org/P08-2051.pdf
86. Liu Y, Zhong SH, Li W (2012) Query-oriented multi-document summarization via unsupervised deep learning. In: Proceedings of the AAAI conference on artificial intelligence, vol 26, no 1, pp 1699–1705. https://doi.org/10.1609/aaai.v26i1.8352
87. Liu SH, Chen KY, Chen B, Wang HM, Yen HC, Hsu WL (2015) Combining relevance language modeling and clarity measure for extractive speech summarization. IEEE/ACM Transact Audio, Speech, Lang Process 23(6):957–969. https://doi.org/10.1109/TASLP.2015.2414820
88. Luhn HP (1958) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165. https://doi.org/10.1147/rd.22.0159
89. Luo L, Ao X, Song Y, Pan F, Yang M, He Q (2019) Reading like HER: human reading inspired extractive summarization. In proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP) (pp. 3033-3043). https://doi.org/10.18653/v1/D19-1300
90. Lwin SS, Nwet KT (2018) Extractive summarization for Myanmar language. In 2018 international joint symposium on artificial intelligence and natural language processing (iSAI-NLP) (pp. 1-6). IEEE. https://doi.org/10.1109/iSAI-NLP.2018.8692976
91. Lwin SS, Nwet KT (2019) Extractive Myanmar news summarization using centroid based word embedding. In: 2019 international conference on advanced information technologies (ICAIT). IEEE, pp 200–205. https://doi.org/10.1109/AITC.2019.8921386
92. Mandal S, Singh GK, Pal A (2018) A constraints driven PSO based approach for text summarization. J Inf Math Sci 10(4):703–714. https://doi.org/10.26713/jims.v10i4.891
93. Mathkour HI, Touir AA, Al-Sanea WA (2008) Parsing Arabic texts using rhetorical structure theory. J Comput Sci 4(9):713–720

94. Maurya AK (2020) Resource and task clustering based scheduling algorithm for workflow applications in cloud computing environment. In 2020 sixth international conference on parallel, distributed and grid computing (PDGC) (pp. 566-570). IEEE. https://doi.org/10.1109/PDGC50313.2020.9315806
95. Maurya R, Singh SK, Maurya AK, Kumar A (2014) GLCM and multi class support vector machine based automated skin cancer classification. In 2014 international conference on computing for sustainable global development (INDIACom) (pp. 444-447). IEEE. https://doi.org/10.1109/IndiaCom.2014.6828177
96. Maurya SK, Singh D, Maurya AK (2022) Deceptive opinion spam detection approaches: a literature survey. Applied intelligence, 1-46. https://doi.org/10.1007/s10489-022-03427-1
97. Meena YK, Gopalani D (2015) Evolutionary algorithms for extractive automatic text summarization. Proced Comput Sci 48:244–249. https://doi.org/10.1016/j.procs.2015.04.177
98. Mehta P, Majumder P (2018) Effective aggregation of various summarization techniques. Inf Process Manag 54(2):145–158. https://doi.org/10.1016/j.ipm.2017.11.002
99. Mei JP, Chen L (2012) SumCR: a new subtopic-based extractive approach for text summarization. Knowl Inf Syst 31(3):527–545. https://doi.org/10.1007/s10115-011-0437-x
100. Mendoza M, Bonilla S, Noguera C, Cobos C, León E (2014) Extractive single-document summarization based on genetic operators and guided local search. Expert Syst Appl 41(9):4158–4169. https://doi.org/10.1016/j.eswa.2013.12.042
101. Merchant K, Pande Y (2018) Nlp based latent semantic analysis for legal text summarization. In 2018 international conference on advances in computing, communications and informatics (ICACCI) (pp. 1803-1807). IEEE. https://doi.org/10.1109/ICACCI.2018.8554831
102. Mihalcea R, Tarau P (2004) Textrank: bringing order into text. In proceedings of the 2004 conference on empirical methods in natural language processing (pp. 404-411)
103. MirShojaee H, Masoumi B, Zeinali E (2017) Biogeography-based optimization algorithm for automatic extractive text summarization. Int J Indust Eng Product Res 28(1):75–84 http://ijiepr.iust.ac.ir/article-1-722-en.html
104. Mirshojaei SH, Masoomi B (2015) Text summarization using cuckoo search optimization algorithm. J Comput Robot 8(2):19–24 http://www.qjcr.ir/article_683.html
105. Mohamed M, Oussalah M (2019) SRL-ESA-TextSum: a text summarization approach based on semantic role labeling and explicit semantic analysis. Inf Process Manag 56(4):1356–1372. https://doi.org/10.1016/j.ipm.2019.04.003
106. Moiyadi HS, Desai H, Pawar D, Agrawal G, Patil NM (2016) NLP based text summarization using semantic analysis. Int J Adv Eng Manag Sci 2(10):239678
107. Moratanch N, Chitrakala S (2017) A survey on extractive text summarization. In: 2017 international conference on computer, communication and signal processing (ICCCSP). IEEE, pp 1–6. https://doi.org/10.1109/ICCCSP.2017.7944061
108. Muthu B, Cb S, Kumar PM, Kadry SN, Hsu CH, Sanjuan O, Crespo RG (2021) A framework for extractive text summarization based on deep learning modified neural network classifier. Trans Asian Low-Resource Lang Inf Process 20(3):1–20. https://doi.org/10.1145/3392048
109. Mutlu B, Sezer EA, Akcayol MA (2019) Multi-document extractive text summarization: a comparative assessment on features. Knowl-Based Syst 183:104848. https://doi.org/10.1016/j.knosys.2019.07.019
110. Mutlu B, Sezer EA, Akcayol MA (2020) Candidate sentence selection for extractive text summarization. Inf Process Manag 57(6):102359. https://doi.org/10.1016/j.ipm.2020.102359
111. Nagalla S, Kumar KC (2021) Oppositional lion optimization algorithm and deep neural network based multi-document summarization from large-scale documents. Eur J Mol Clin Med 7(10):1991–2009 https://www.ejmcm.com/article_6857.html
112. Naik SS, Gaonkar MN (2017) Extractive text summarization by feature-based sentence extraction using rule-based concept. In 2017 2nd IEEE international conference on recent trends in electronics, Information & Communication Technology (RTEICT) (pp. 1364-1368). IEEE. https://doi.org/10.1109/RTEICT.2017.8256821
113. Nallapati R, Zhou B, Ma M (2016) Classify or select: neural architectures for extractive document summarization. arXiv preprint arXiv:1611.04244. https://doi.org/10.48550/arXiv.1611.04244
114. Nallapati R, Zhai F, Zhou B (2017) Summarunner: a recurrent neural network-based sequence model for extractive summarization of documents. In Thirty-first AAAI conference on artificial intelligence https://doi.org/10.48550/arXiv.1611.04230, 31
115. Narayan S, Cohen SB, Lapata M (2018) Ranking sentences for extractive summarization with reinforcement learning. arXiv preprint arXiv:1802.08636. https://doi.org/10.48550/arXiv.1802.08636
116. Nawaz A, Bakhtyar M, Baber J, Ullah I, Noor W, Basit A (2020) Extractive text summarization models for Urdu language. Inf Process Manag 57(6):102383. https://doi.org/10.1016/j.ipm.2020.102383

117. Neto JL, Freitas AA, Kaestner CA (2002) Automatic text summarization using a machine learning approach. In Brazilian symposium on artificial intelligence (pp. 205-215). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-36127-8_20
118. Ozsoy MG, Alpaslan FN, Cicekli I (2011) Text summarization using latent semantic analysis. J Inf Sci 37(4):405–417. https://doi.org/10.1177/0165551511408848
119. Papineni K, Roukos S, Ward T, Zhu WJ (2002) Bleu: a method for automatic evaluation of machine translation. In proceedings of the 40th annual meeting of the Association for Computational Linguistics (pp. 311-318). https://aclanthology.org/P02-1040.pdf
120. Parveen D, Strube M (2015) Integrating importance, non-redundancy and coherence in graph-based extractive summarization. In: IJCAI'15: Proceedings of the 24th International Conference on Artificial Intelligence, pp 1298–1304
121. Patel D, Shah S, Chhinkaniwala H (2019) Fuzzy logic-based multi-document summarization with improved sentence scoring and redundancy removal technique. Expert Syst Appl 134:167–177. https://doi.org/10.1016/j.eswa.2019.05.045
122. Patil SR, Mahajan SM (2011) A novel approach for research paper abstracts summarization using cluster-based sentence extraction. In proceedings of the International Conference & Workshop on emerging trends in technology (pp. 583-586). https://doi.org/10.1145/1980022.1980150
123. Potnurwar A, Pimpalshende A, Aote SS, Bongirwar V (2020) Extractive multi-document text summarization by using binary particle swarm optimization. Helix 10(04):263–265. https://doi.org/10.21786/bbrc/13.14/8
124. Prasad SN, Narsimha VB, Reddy PV, Babu AV (2015) Influence of lexical, syntactic and structural features and their combination on authorship attribution for Telugu text. Proced Comput Sci 48:58–64. https://doi.org/10.1016/j.procs.2015.04.110
125. Qaroush A, Farha IA, Ghanem W, Washaha M, Maali E (2021) An efficient single document Arabic text summarization using a combination of statistical and semantic features. J King Saud Univ Comput Inf Sci 33(6):677–692. https://doi.org/10.1016/j.jksuci.2019.03.010
126. Rahman N, Borah B (2015) A survey on existing extractive techniques for query-based text summarization. In 2015 international symposium on advanced computing and communication (ISACC) (pp. 98-102). IEEE. https://doi.org/10.1109/ISACC.2015.7377323
127. Rani R, Lobiyal DK (2021) An extractive text summarization approach using tagged-LDA based topic modeling. Multimed Tools Appl 80(3):3275–3305. https://doi.org/10.1007/s11042-020-09549-3
128. Rautray R, Balabantaray RC (2017) Cat swarm optimization-based evolutionary framework for multi-document summarization. Physica A: Stat Mech Appl 477:174–186. https://doi.org/10.1016/j.physa.2017.02.056
129. Raval KR, Goyani MM (2022) A survey on event detection-based video summarization for cricket. Multimed Tools Appl, 1-29. https://doi.org/10.1007/s11042-022-12834-y
130. Ravinuthala VVMK, Chinnam SR (2017) A keyword extraction approach for single document extractive summarization based on topic centrality. Int J Intell Eng Syst https://doi.org/10.22266/ijies2017.1031.17
131. Rothe S, Schütze H (2014) Cosimrank: a flexible & efficient graph-theoretic similarity measure. In proceedings of the 52nd annual meeting of the Association for Computational Linguistics (volume 1: long papers) (pp. 1392-1402). https://aclanthology.org/P14-1131.pdf
132. Sahba R, Ebadi N, Jamshidi M, Rad P (2018) Automatic text summarization using customizable fuzzy features and attention on the context and vocabulary. In 2018 world automation congress (WAC) (pp. 1-5). IEEE. https://doi.org/10.23919/WAC.2018.8430483
133. Sahoo D, Balabantaray R, Phukon M, Saikia S (2016) Aspect-based multi-document summarization. In 2016 international conference on computing, communication and automation (ICCCA) (pp. 873-877). IEEE. https://doi.org/10.1109/CCAA.2016.7813838
134. Salton G, Singhal A, Mitra M, Buckley C (1997) Automatic text structuring and summarization. Inf Process Manag 33(2):193–207. https://doi.org/10.1016/S0306-4573(96)00062-3
135. Sanchez-Gomez JM, Vega-Rodríguez MA, Pérez CJ (2018) Extractive multi-document text summarization using a multi-objective artificial bee colony optimization approach. Knowl-Based Syst 159:1–8. https://doi.org/10.1016/j.knosys.2017.11.029
136. Sanchez-Gomez JM, Vega-Rodriguez MA, Perez CJ (2020) Experimental analysis of multiple criteria for extractive multi-document text summarization. Expert Syst Appl 140:112904. https://doi.org/10.1016/j.eswa.2019.112904
137. Shaymal AK, Pal M (2007) Triangular fuzzy matrices. Iran J Fuzzy Syst 4(1):75–87 https://www.sid.ir/en/Journal/ViewPaper.aspx?ID=67072
138. Shen C, Li T (2011) Learning to rank for query-focused multi-document summarization. In 2011 IEEE 11th international conference on data mining (pp. 626-634). IEEE. https://doi.org/10.1109/ICDM.2011.91
139. Shirwandkar NS, Kulkarni S (2018) Extractive text summarization using deep learning. In 2018 fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1-5). IEEE. https://doi.org/10.1109/ICCUBEA.2018.8697465

140. Shoaib M, Maurya AK (2014) Comparative study of different web mining algorithms to discover knowledge on the web. In proceedings of Elsevier second international conference on emerging research in computing, information, communication and application (ERCICA-2014) (Vol. 3, pp. 648-654)
141. Shoaib M, Maurya AK (2014) URL ordering-based performance evaluation of web crawler. In 2014 international conference on advances in Engineering & Technology Research (ICAETR-2014) (pp. 1-7). IEEE. https://doi.org/10.1109/ICAETR.2014.7012962
142. Siddiqui MK, Ahmad A, Pal O, Ahmad T (2021) CoRank: a clustering cum graph ranking approach for extractive summarization. arXiv preprint arXiv:2106.00619. https://doi.org/10.48550/arXiv.2106.00619
143. Singh SP, Kumar A, Mangal A, Singhal S (2016) Bilingual automatic text summarization using unsupervised deep learning. In 2016 international conference on electrical, electronics, and optimization techniques (ICEEOT) (pp. 1195-1200). IEEE. https://doi.org/10.1109/ICEEOT.2016.7754874
144. Singh RK, Khetarpaul S, Gorantla R, Allada SG (2021) SHEG: summarization and headline generation of news articles using deep learning. Neural Comput & Applic 33(8):3251–3265. https://doi.org/10.1007/s00521-020-05188-9
145. Sirohi NK, Bansal M, Rajan SN (2021) Recent approaches for text summarization using machine learning & LSTM0. J Big Data 3(1):35. https://doi.org/10.32604/jbd.2021.015954
146. Song S, Huang H, Ruan T (2019) Abstractive text summarization using LSTM-CNN based deep learning. Multimed Tools Appl 78(1):857–875. https://doi.org/10.1007/s11042-018-5749-3
147. Sreelakshmi PR, Manmadhan S (2021) Image summarization using unsupervised learning. In 2021 7th international conference on advanced computing and communication systems (ICACCS) (Vol. 1, pp. 100-103). IEEE. https://doi.org/10.1109/ICACCS51430.2021.9441682
148. Srivastava AK, Pandey D, Agarwal A (2021) Extractive multi-document text summarization using dolphin swarm optimization approach. Multimed Tools Appl 80(7):11273–11290. https://doi.org/10.1007/s11042-020-10176-1
149. Srivastava R, Singh P, Rana KPS, Kumar V (2022) A topic modeled unsupervised approach to single document extractive text summarization. Knowl-Based Syst 246:108636. https://doi.org/10.1016/j.knosys.2022.108636
150. Steinberger J (2009) Evaluation measures for text summarization. Comput Inf 28(2):251–275 http://147.213.75.17/ojs/index.php/cai/article/view/37
151. Steinberger J, Jezek K (2004) Using latent semantic analysis in text summarization and summary evaluation. Proc ISIM 4(93-100):8
152. Suleman RM, Korkontzelos I (2020) Managing the syntactic blindness of latent semantic analysis. In CS & IT conference proceedings (Vol. 10, no. 4). CS & IT conference proceedings. https://doi.org/10.5121/csit.2020.100401
153. Suleman RM, Korkontzelos I (2021) Extending latent semantic analysis to manage its syntactic blindness. Expert Syst Appl 165:114130. https://doi.org/10.1016/j.eswa.2020.114130
154. Tarnpradab S, Liu F, Hua KA (2017) Toward extractive summarization of online forum discussions via hierarchical attention networks. Thirtieth Int Flairs Conf, 288-292. https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS17/paper/view/15500
155. Thakkar HK, Sahoo PK, Mohanty P (2021) DOFM: domain feature miner for robust extractive summarization. Inf Process Manag 58(3):102474. https://doi.org/10.1016/j.ipm.2020.102474
156. Thu HNT, Huu QN, Ngoc TNT (2013) A supervised learning method combine with dimensionality reduction in Vietnamese text summarization. In 2013 computing, communications and IT applications conference (ComComAp) (pp. 69-73). IEEE. https://doi.org/10.1109/ComComAp.2013.6533611
157. Uçkan T, Karcı A (2020) Extractive multi-document text summarization based on graph independent sets. Egypt Inf J 21(3):145–157. https://doi.org/10.1016/j.eij.2019.12.002
158. Vale R, Lins RD, Ferreira R (2020) An assessment of sentence simplification methods in extractive text summarization. In proceedings of the ACM symposium on document engineering 2020 (pp. 1-9). https://doi.org/10.1145/3395027.3419588
159. Van Lierde H, Chow TW (2019) Query-oriented text summarization based on hypergraph transversals. Inf Process Manag 56(4):1317–1338. https://doi.org/10.1016/j.ipm.2019.03.003
160. Verma P, Verma A, Pal S (2022) An approach for extractive text summarization using fuzzy evolutionary and clustering algorithms. Appl Soft Comput 120:108670. https://doi.org/10.1016/j.asoc.2022.108670
161. Wang D, Zhu S, Li T, Chi Y, Gong Y (2011) Integrating document clustering and multi-document summarization. ACM Trans Knowl Discov Data (TKDD) 5(3):1–26. https://doi.org/10.1145/1993077.1993078
162. Wang S, Zhao X, Li B, Ge B, Tang D (2017) Integrating extractive and abstractive models for long text summarization. In 2017 IEEE international congress on big data (BigData congress) (pp. 305-312). IEEE. https://doi.org/10.1109/BigDataCongress.2017.46
163. Wang X, Nie X, Liu X, Wang B, Yin Y (2020) Modality correlation-based video summarization. Multimed Tools Appl 79(45):33875–33890. https://doi.org/10.1007/s11042-020-08690-3

164. Wang D, Liu P, Zheng Y, Qiu X, Huang X (2020) Heterogeneous graph neural networks for extractive document summarization. arXiv preprint arXiv:2004.12393. https://doi.org/10.48550/arXiv.2004.12393

165. Wu K, Shi P, Pan D (2015) An approach to automatic summarization for chinese text based on the combination of spectral clustering and LexRank. In 2015 12th international conference on fuzzy systems and knowledge discovery (FSKD) (pp. 1350-1354). IEEE. https://doi.org/10.1109/FSKD.2015.7382140

166. Wu Z, Lei L, Li G, Huang H, Zheng C, Chen E, Xu G (2017) A topic modeling-based approach to novel document automatic summarization. Expert Syst Appl 84:12–23. https://doi.org/10.1016/j.eswa.2017.04.054

167. Wu M, Pan S, Zhou C, Chang X, Zhu X (2020) Unsupervised domain adaptive graph convolutional networks. In proceedings of the web conference 2020 (pp. 1457-1467). https://doi.org/10.1145/3366423.3380219

168. Xu J, Durrett G (2019) Neural extractive text summarization with syntactic compression. arXiv preprint arXiv:1902.00863. https://doi.org/10.48550/arXiv.1902.00863

169. Yadav J, Meena YK (2016) Use of fuzzy logic and WordNet for improving performance of extractive automatic text summarization. In 2016 international conference on advances in computing, communications and informatics (ICACCI) (pp. 2071-2077). IEEE. https://doi.org/10.1109/ICACCI.2016.7732356

170. Yadav AK, Saxena S (2016) A new conception of information requisition in web of things. Indian journal of science and technology, 9(44). https://doi.org/10.17485/ijst/2016/v9i44/105143

171. Yadav H, Ghosh S, Yu Y, Shah RR (2020) End-to-end named entity recognition from English speech. arXivpreprintarXiv:2005.11184. https://doi.org/10.48550/arXiv.2005.11184

172. Yadav AK, Maurya AK, Yadav RS (2021) Extractive text summarization using recent approaches: a survey. Ingénierie des Systèmes d'Information, 26(1). https://doi.org/10.18280/isi.260112

173. Ye S, Chua TS, Kan MY, Qiu L (2007) Document concept lattice for text understanding and summarization. Inf Process Manag 43(6):1643–1662. https://doi.org/10.1016/j.ipm.2007.03.010

174. Yogatama D, Liu F, Smith NA (2015) Extractive summarization by maximizing semantic volume. In proceedings of the 2015 conference on empirical methods in natural language processing (pp. 1961-1966). https://aclanthology.org/D15-1228.pdf

175. Yu W, Lin X, Zhang W (2013) Towards efficient SimRank computation on large networks. In 2013 IEEE 29th international conference on data engineering (ICDE) (pp. 601-612). IEEE. https://doi.org/10.1109/ICDE.2013.6544859

176. Zajic DM, Dorr BJ, Lin J (2008) Single-document and multi-document summarization techniques for email threads using sentence compression. Inf Process Manag 44(4):1600–1610. https://doi.org/10.1016/j.ipm.2007.09.007

177. Zhang K, Xiao Y, Tong H, Wang H, Wang W (2014) WiiCluster: a platform for wikipedia infobox generation. In proceedings of the 23rd ACM international conference on conference on information and knowledge management (pp. 2033-2035). https://doi.org/10.1145/2661829.2661840

178. Zopf M, Botschen T, Falke T, Heinzerling B, Marasovic A, Mihaylov T, Frank A (2018) What's important in a text? An extensive evaluation of linguistic annotations for summarization. In 2018 fifth international conference on social networks analysis, management and security (SNAMS) (pp. 272-277). IEEE. https://doi.org/10.1109/SNAMS.2018.8554853

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.