ISSN: 0974-5823 Vol. 7 No. 8 August, 2022

International Journal of Mechanical Engineering

A Comprehensive Survey of Different Approach for Text Summarization

***Neeraj Kumar Sirohi**

Research Scholar, Shobhit Institute of Engineering & Technology, Meerut,

Dr Mamta Bansal

Shobhit Institute of Engineering & Technology, Meerut,

Dr S.N Rajan

IMS Engineering College, GhaziabadAbstract

These days, data is expanding incredibly in all domains such as mass media, research, banking, sports, etc. Due to the explosion of data, data have lost information but very only some as treated as valuable knowledge. To obtain this important data automatically from this document we need a text summarizer that is competent to obtain this valuable information automatically and reduce the length of the document specifically for textual data without dropping any necessary information. The text summarization is focused on the summary produced. The approaches for summary generation are divided into two groups generally known as extractive and abstractive summarization

The extractive method is based on extracting the maximum number of rank sentences from the data assembled through important words and sentences then putting them together to form a summary. Further, abstractive summarization is dependent on knowing the important facts about the text and later expressing those concepts, in other words, sometimes the word is not in the e-mail text. This is the newest research field for NLP, ML, and NN. This article investigates and reviews the numerous methods for summarization and explains the usefulness and weaknesses of the different methods. Also discuss the latest approach by using NN based on LSTM this structural design is known as Encoder-Decoder, underneath the ML methodology.

Keywords: Text Summarization. Extractive Summary. Abstractive Summary, NLP, Neural Network, LSTM

1. Introduction:

Summarization is a method of compiling and assembling key information from a document into a short summary of the main content [1]. According to Mani and Maybury [2], text summarization is the most common way of refining the most significant data from a source (or on the other hand sources) to generate a compressed rendition for a particular client (or user) and assignment (or errands). Before discussing summarization first understand the meaning of summary. In 1950[3] describe summary as “A summary is a reduced form of an original document, generally a complete article or book. Summaries are mostly around a passage long, and might be only some paragraphs long subject to the size of the work being compressed”. After that in 2002 [4] re-defined the summary as “ a text that is generated after considering one or more articles, that communicates crucial knowledge of the initial document(s), that is not more than one-half of the initial document(s) and frequently appreciably fewer than that”. Summarization is a powerful and effective way to create a summary of a complex document. The summarization task is divided into two groups, extractive brief, and abstractive brief. The abstractive brief is a current research area where lots of research is going on; but unfortunately, even so, no system (Method) has been achieved an excellent result yet. These summaries were created using data from the document after studying what was expressed in the document and then modifying it into a form expressed through the system. It is accomplished much like humans generate summaries after analyzing the document. while extractive summaries are generated after selecting the essential idioms and sentences from the initial text itself re-organized them and showing them to the user.

The summary is achieved by picking significant keywords which illustrate the text. Computerized keyword selection is the method of picking words and sentences from a text file that, depending on the model, can best describe the document's fundamental sentiment without requiring human intervention [5]. The goal of automatic keyword selection is to use the strength and speed of today's calculation capabilities to the challenge of access and retrieval, with a focus on information structure and without the additional costs of human annotators.

This paper consists study of various ML methodology that utilizes ANN (artificial neural networks) to generate summaries of arbitrary size text are discussed.

The majority of Avant-grade surveys emphasize a content matter of ATS elements, identical to extractive summarising methods, another technique like abstractive method, one particular-area ATS approach (e.g. legal text summarization), and so on. Furthermore, as [5] points out, different ATS approaches provide various summaries belonging to the same input texts, combining outputs as of numerous summarization methods to build superior summaries is quite promising. This article reviews a complete framework of the methods of text summarization at first and then details the study of the Encoder-Decoder method for ML techniques along with implementation using TensorFlow in Keras.

2. SUMMARIZATION, THE IMPACT OF CONTEXT

Additional evidence is frequently available for summarization systems to use to determine the most essential document themes (s). While reviewing blogs, for example, there are debates or statements which follow that record information post for useful resources to determine which wedge of the blog are critical and interesting. There is a sufficient amount of informative pieces is available in scientific paper summaries like mentioned in various articles and seminar reports, that are able to be used for highlighting essential data in the unique work. The following sections go over various situations in greater depth.

2.1 Summarizing of Scientific Articles

Finding more articles that mention the objective document and extracting the phrases where the mentions occur in require to ascertain the relevant attributes of the goal chronicle is a valuable source of knowledge for summarizing a technical paper (i.e. citation-based summarization). [6] offer a language model in which every term in the reference setting phrases which assigned a probability. They then use the KL divergence approach to rate the relevance of sentences in the original manuscript (i.e. discovering the equivalence among the sentence and the language model).

2.2 Summarization Of Email's

An email has certain distinguishing traits that reflect aspects of both verbal and written communication. Summarization strategies, for example, must take into account the interactive aspect of the dialogue, like in spoken discussions. Early research in this area was given by [7], who proposed a method for generating a review for the initial 2 stages of the idea discussion. An idea is made up of one or more talks involving two or additional people throughout the moment. They choose a note from the source note and every answer to the source, taking into account the root context similarity. [8] employed an ML method and integrated threat-related information along with as well as email structure features such as sentence location in the thread, number of recipients, and so on.[9] present a system that clusters mails into thematic classes and then extracts reviews for each group to summarize an entire mailbox rather than a single thread.

2.3 Summarization Of Web

Pictures, for example, are among the many aspects of a web page that cannot be summarized. Because the textual information they have is typically limited, text summarizing approaches are constrained. Nevertheless, the milieu of a web page, i.e. data gathered on or after the subject of entire written content connecting to the same, can be used like supplementary information that improves the feature of summaries. [10] proposed the first study in this area, in which they use web search engines to find pages with links to a certain web page. Then heuristically assess the potential contacts and choose the finest sentences with connections to the web page. [11] built on this method by employing an algorithm that attempts to find a decree almost the same that encompasses as numerous characteristics of the internet page as feasible.

3. Related Works:

Even though the research on Automatic Text summarization (ATS) has been there for a long, initial work has been initiated [3]] around the 1950s in IBM Research Lab. this methodology is based on selecting critical sentences from the document and merging them collectively. To determine the significance of a sentence term frequency is used. Sentences are to be part of the summary of the phrase frequencies of that specific sentence are above average. Soon after this a new algorithm graph-based ranking proposed by [5] for summarization provides enhanced and more prominent results. After that, some important approaches for summarization based on abstractive methods are introduced by various researchers like[12] [13][14].

3.1 EXTRACTIVE TEXT SUMMARIZATION APPROACHES

Depending on the literature, ATS techniques can be categorized into various classifications, like statistical-based, ML dependent, logical based, graph-based, statistical-based as shown in Figure

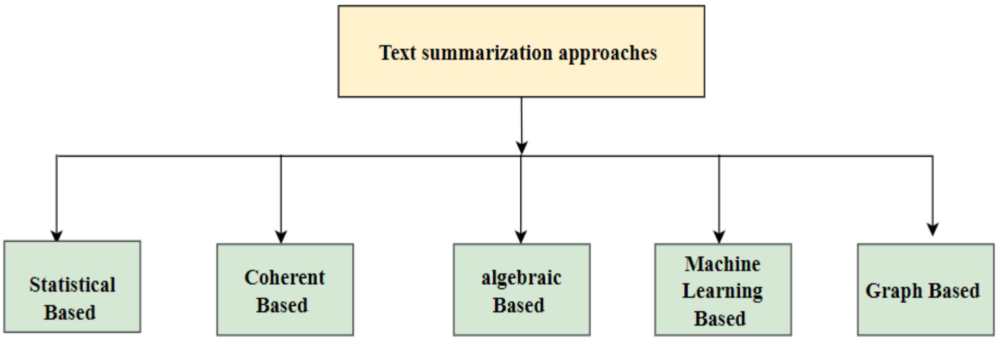

Fig 1. Text Summarization classifications

There are numerous methods in each category the ML method is based on a quality determined and to train and test the model needs to gloss corpus. Here are several promising ML methods, some of which are discussed below.

3.1.1 Decision Tree:

Decision tree processes defined by [15][16] are the prime method that extensively applied inductive understanding techniques. Amongst the numerous D.T techniques selecting C4.5 techniques [16] for summarization is the best option. Understanding a feature that returns the largest amount of information yields a decision tree. A defined set of rules is linked to the attributes is also used to create a new node. This procedure is continued for another feature in sequence till no further info advantage is available. A design is continuously reviewed in testing along with a fork of a decision, tree planting from the root and progressing through the proper nodes actualizing on the circumstance and trait significance till the final branch is zenith. The arrangement afterwards is supposed to fit into the category in which the final node symbolizes. C4.5 was recognized as an incredibly quick and effective technique with excellent simplification ability.

The SUMMARIST is established on a decision tree that utilized the following 'equation':

summarization = topic detection + topic understanding + production.

There are 3 stages in process of summarization are:

Topic Detection: Recognize the extremely critical (significant) themes of the document. SUMMARIST depends on positional significance[17][18], key phrases [18] [19][20], and term frequency. Significance depends on the discussion will be included later [21]. This is the truly improved stage in SUMMARIST.

Topic Interpretation: To combine the notions such as an attendant, menu, and food items into one comprehensive idea that is the restaurant, then if need more information about topic used uncomplicated word cluster which is used in conventional information extraction. We have explored topic calculating[22] and theme signatures to confrontation the combination problem.

Creation of Summary: SUMMARIST is intelligent enough to create summaries in numerous presentations such as essential words (vital noun expressions), access key information (crucial sentences from text), pattern-dependent summaries [23](generated from pre-definite examples), and sophisticated summaries (produced by a sentence manager and realizer) [24] [25].

3.1.2 An approach based on Support Vector Machine

For two-class issues, SVM is a supervised understanding procedure is used. The summary generation procedure of SVM is depicted in Figure 2. Preparation information is provided as (A_1, B_1), (Au, Au), A_j Є Rn, B_{j+1} Є {-1,+1}. Where Aj is the j[th] section's characteristic vector, and class is Aj with a label of +ve (+1) or -ve (-1).

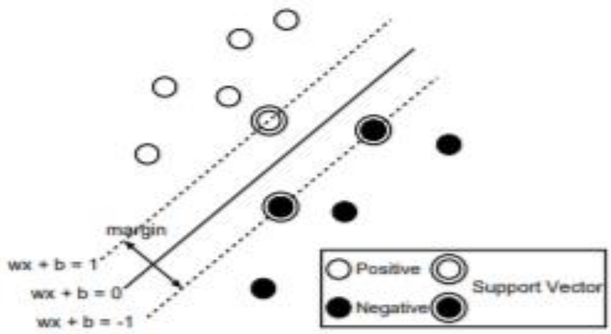

Fig.2 SVM model for summarization of text [37]

Support Vector Machine depends upon the attributes selection which is related to the sentence of Si. Several other crucial characteristics are involved which are as follows

Location of sentences: The position of a sentence plays a significant role in determining which sentence has to be included in a conclusive sentence. Understanding of the sentences at the introduction transmits the fundamental aim of the text, even if the last sentence may be an ended or summary sentence, the collection of the sentence is quite important.

The sentence is judged by using its place in the text. Location of sentence offers the importance of sentence to be a part of the first 5 sentences of summary. This feature result is assessed by [26] [27] and consider the initial 5 sentences in the article

Score = 1[st] if the score is 5/5

2[nd] has a score of 4/5, 3[rd] when the score is 3/5, and so on. And score for the remaining sentences is 0/5. Soon after this [28] proposed the distinct and moderate approach which is utilized while calculating the location of the sentence.

Assign 1 for both sentences whether they are at the beginning or end of the document. Otherwise, assign 0(zero) to all other sentences.

Sentence size: This feature was used to castigate sentences that were too short[29]. If a threshold is set, such as five words, the feature is real if it is exceeded and false otherwise. Sentences with a little number of keywords are unlikely to be significant, so they should be ignored. The two methods provided by Nobata and Sekine[30] express a more complex formula for scoring a sentence. The first technique allocates a single value for a particular sentence based on its length and a maximum number called Lmax. The second, which yields superior results, allows for a negative score to penalize sentences that are less than a predetermined minimum length Lmin. Equation 5 shows a more modern formula developed by Fattah and Ren[31]. It allocates a score to a sentence si depending on the total number of words in the sentence |si |, the text's word count |dn|, and the number of sentences in the file denoted as|{s: s ∈ dn}|.

$\text{Score}_{\text{length(si)}}$ = |si | ∗ |{s : s ∈ dn}| |dn|

Sentence Weight: weighting of a sentence is another important method to identify important sentences from the text this is achieved in two steps. Initially, the text is cleaned using normalization (i.e stop word removal, lemmatization, lowering) after that an individual score value is transferred to each word. The weight is assessed by using:

wt=amount of the word / Total no. of words in the document

After attaching the weight to an individual words. The second step is all about designating a rank to all sentences depending on their weight worth. Then compute the phrase's weight is reckon by adding the scores of each of the words in the sentence and by division of it by the whole number of terms in the sentence, i.e.

wts= (wti)ni=1/n

Where wts = sentence weight.

wt1, wt2,wt3 ,.........wtn = score of individual words .

n = total quantity of words in that sentence.

Sentence Relationship with Label: in this step, a sentence assigns more weight if it has words that are present in the title and have more opportunities to be a part of the summary. These sentences are picked by utilizing the document's label as an "objection" against each of the text's sentences and then determining the alikeness of the tag, each sentence taken from the document using cosine similarity [32].

Sentence-to-Sentence Organization: This attribute is achieved as follows: individual sentences S initial computes the similarity amongst the S and other sentences of the document; then combine these similar ideals and find the primary value of that S; the method is simulated for each sentence. The stabilized score (series of [1,0]) for a phrase S, this attribute is achieved through determining the amount of S's unprocessed parameter over the sentence with the highest unprocessed attribute score in the phrase Closer to 1.0 indicate that sentences with maximum cohesion.

3.1.3 Summarization methods based on Bayesian

In extractive summarization, the importance of sentences in the document depends on how valuable/important they are as a portion of the summary, Bayesian consider a specific grading system that depends on the sentence likelihood.

i.e., P(S|V) (i)

S represents a specified phare, and V = (V_1, . . . , Vn) stands for a DOV(Distribution of votes), an array of monitored vote calculations for the text's words; V_1 indicates the importance of words at the documents primary position, V_2 suggest sentence that holding the second rank, and so on.

Both BIC (Bayesian Information Criterion) and MC (Monte Carlo integration method (MacKay, 1998)) are used to solve problems constructing a summarizer is an uncomplicated task for a Given document D with compression rate R, what a summarizer would do is just abundant the sentences in D depending on P(S_i |V) and choose an R segment of the maximum-position phares

offering a training set of text with humans-generated text summaries, build up a categorization function those estimations the likelihood of a given sentence. By rating a sentence depending on likelihood and scoring a new summary can be generated by picking the sentences having maximum likelihood and top scorer. Calculate the probability of each phrase which is a part of summary S considering the k characteristics fj; j = 1…k, which can be communicated as described by using Bayes' assumption [29]:

P(s ϵ S | $f1$, $f2$, …, fk) = $P(f1,f2,…,fks\epsilon S)$ $P(s\epsilon S)$ $P(f1,f2,…,fk)$

Guessing statistical autonomy of the features:

P(s ϵ S | $f1$, $f2$, …, fk) =

$\prod^{k}{}_{j=1}\ P(fjs \in S)\ P(s \in S)$

$\prod^{k}{}_{j=1}\ P(fj)$

3.1.4 Hidden Markov (HMM) Model

Another important approach for the selection of important sentences for summary is Hidden Markov Model (HMM)[23]. HMM has fewer rules for sentence picking compared to a naive Bayesian technique. Actual, The HMM does not presume that the probability of phrase x is the same as the likelihood of phrase i^{th} position of the summary is fundamentally not- dependent on the i-1 phrase of the synopsis. There are mainly 3 strategies worked in HMM to select a sentence these are as follows:

· sentence location in the text,

· quantity of words in the phrase,

· probability of the phrase relations because of the text terms

The HM Model consists of 2x + 1 states, varying x as a review and x+1 as a non-summary. Below Fig. 3. Shows HMM with seven connections, deciding to x = 3.

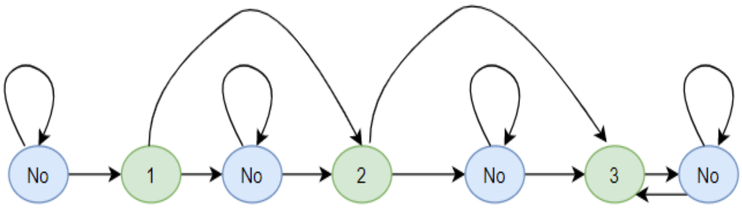

Fig 3. HMM summarization prototype [23]

This series will choose up to x-1 summary statements and a random list of supporting statements. How the chain picks the important sentence note that every single way across the string, view first x-1 summaries conditions. The initial two positions in the series permit random numbers of non-synopsis and summaries sentences. This Markov series has whole 2s requirements which support calculating the likelihood of various progress between sets of states. These boundaries are surveyed by the foundations of training sets. For instance, the assessment of probability among summaries levels 2z and 2z+2 number of periods the summaries phrase z+1 is repeated is the summaries state. in the training sets, immediately followed by summaries sentence z. What's more, the likelihood of changes between summaries states 2z and non-summary state 2z+1 is characterized to have one fewer this probability.

Then, at that point, discover the most extreme likelihood for every and specific by utilizing this calculation, and afterward to make an exchange matrix Markov series known as M, in which [a,b] indicate the anticipated likelihood of altered phrase ***a*** to ***b***. Similarly determining p_a is the most extreme probability assessment of the essential appropriation for the series by utilizing the accompanying condition

P(a) =pri (primary phares relating to state 1)

Where p(a) = 0 for a > 2 then, at that point, the critical sentence is moreover the key sentence of outline (state 2) or an express that drives the essential synopses Phares (state 1). In the wake of doing a few alterations in the series that grants us precisely minimize the S summaries sentence. This further developed chain is displayed beneath below fig, display contrast from of chain shown in above fig. further developed chain eliminates the sequence that arises between last and first Summaries and non-summaries states.

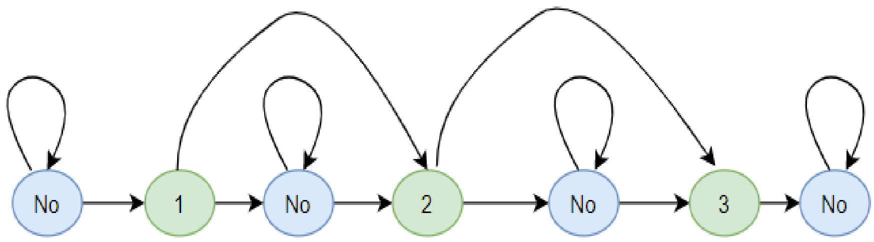

Fig 4. HMM summarization process

This chain is highest appropriate for handling immobile sizes of summaries. There are 2s available for prediction from training data ***a*** as an output function

Bi (o) = pr (0| state a)

Here o is the tentative vector of features associated with a sentence.

3.1.5 Summarization Based on Neural Networks

ANN is a type of machine learning technique that uses artificial neural networks. that is highly regarded and significant. ANN is utilized to provide summaries of news stories of varying lengths. A corpus of articles is applied to train a neural network. The neural network is then improved by combining it with other neural networks to provide a summary of the article's highest-ranking sentences. The network estimates the significance of several elements used to fix the Summary-value of each expression utilizing highlight combination [33]. Training and testing phases are the two steps of ANN. The neural framework examines the plans and properties of sentences that should be considered summary sentences and those that should not be considered summary sentences during the training phase. Three feed-forward layers make up the Neural Network structure. Seven information layer neurons, six secret layer neurons, and one result layer neuron make up this organization. Each statement is addressed by seven-highlighted vector [f1, f2,..., f7]. The highlights are selected with care based on the text's location or the sentence's position.

f1 = Following the title is a paragraph (Paragraph Position)

f2 = text paragraph position

f3 = paragraph sentence position

f4 = the initial sentence of the paragraph

f5 = The length of the sentence

f6 = In a sentence, the total counting of key terms

f7 = Total count of sentence heading terms

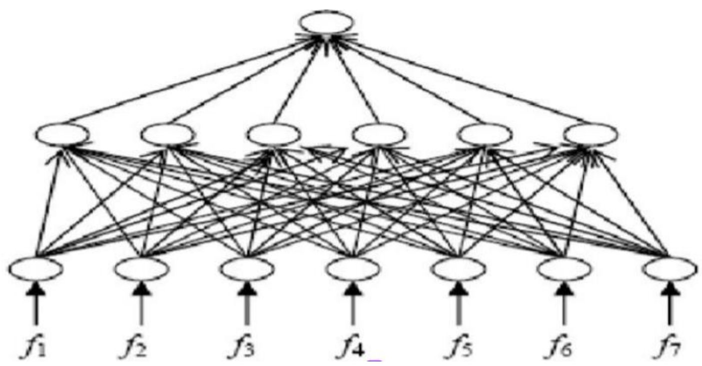

Fig 5. The Neural Network before Pruning [12]

The Summarization procedure has three stages: preparing, highlighting combination, and sentence determination. The underlying advance incorporates preparing a neural organization to distinguish the sort of sentences that should be chosen as an outline sentence. Next phase, try to minimize the neural system and crash down the hidden level item creations into distinct values of occurrences. After that picking of sentences is covered in the third step, which utilizes this trained neural system to sort the document and pick the highest graded sentences [12]. Moreover, any secret layer neuron having no adjoining associations can be eliminated.

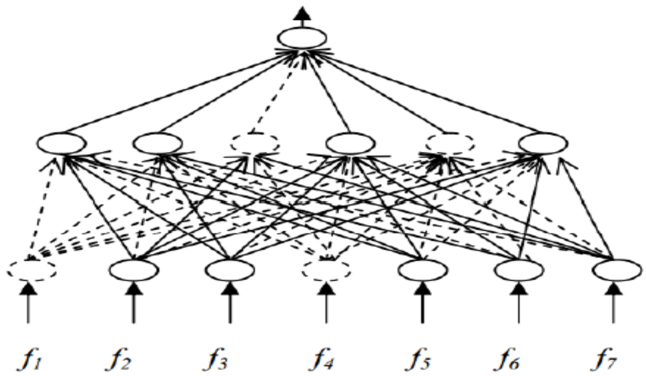

Fig 6. Neural System after Lopping [12]

Using a suitable clustering algorithm, the activation values of the hidden layer and each children layer neuron are clustered. The centroid and frequency of each cluster are used to identify it. The cluster's centroid exchanges the activation rate for every hidden layer cell. The combination of these two processes resulted in inappropriate parameters for sentence ranking by simplifying the effects of features. Another method[34] for document summarization uses a Neural Network with "Numerical Data Feature" as an input list, with the network using 8 neurons as input. After the neural network has discovered high-graded summary sentences, the sentences are passed to the rhetorical structure, which then discovers the language design and observes logical relations in sentences, which may aid the discovery of improved summary judgments, This information can then be utilized to create better summaries. [35] suggested a language model based on word co-events in a recurrent neural system auxiliary data. To provide a summary, To classify sentences, the linguistic methodology uses a Bayesian generating framework based on the frequency of every single exceptional term. [36] studied the results of neural networks and various feature-based summarizing approaches. Based on a binary analysis tree built by a recursive neural system, [37] suggested an autoencoder to separate expression implanting, which is a straightforward amount of words. Summarization is accomplished by comparing the relationship between sentences. Part of speech disambiguation was proposed by Prasad et al.[22] using a recurrent neural network. To construct a summary, a piece vector of grammatical features is taken care of to a neural organization that characterizes sentences.

3.1.6 Summarization using Fuzzy Logic

The fuzzy logic set and fuzzy logic instruction are employed. The objective of this strategy is to observe the main sentences in view of their properties. Skillful techniques and decision-making assistance with significant intellectual component capacities are delivered via fuzzy logic methodologies. Zadeh [27] proposed a fuzzy logic set as a mathematical instrument for dealing with ambiguity, inaccuracy, vagueness, and uncertainty. In the area of text summarization, a little study was conducted using Fuzzy. Witte and Beagle [38] suggested a text summarizing technique depends on fuzzy theory and its application to coreference proposals. There is a lot of ambiguity in automatic coreference fixation between noun phrases. They demonstrate how fuzzy collections may be applied to create a novel coreference technique that explicitly resolves the ambiguity and permits us to specify different levels of coreference. Patil and Kulkarni[39] utilize Fuzzy Logic for ratting the statement as well. after selection of Features selection and a step of pre-processing includes Sentence length, the weight of terms, the similarity between sentences, title term, the position of the sentence, topical words, statistical values, and Appropriate Nouns the eight features they employ for text summary. The system is made up of the following steps:

- The system selects individual sentences from the original text in the preprocessing step after examining the source collection. Isolated input text into distinct terms after that. Eliminate any stop keywords after that. Word stemming is the final step in the preprocessing process.
- Every phrase is linked to one of the eight properties via a vector discussed above, whose values are derived from the content of the sentence;
- According to frequency, a collection of the top-scoring phrases is picked as the document summary

Fuzzifier, Interpretation Engine, Defuzzifier, and Fuzzy Intelligence Center are the four main factors of the Fuzzy Logic Method. To convert difficult arguments into soft inputs, an association function is utilized in the fuzzifier. To obtain the grammatical ideal, the inference mechanism alludes to the fuzzy rules that contain IF-THEN clauses after fuzzification.

In the concluding phase, The de-fuzzifier uses the correlation function to recommend the final paragraph score by transforming the result verbal factors from the research into a conclusive target score[27]. Individual sentences are coupled with 8 feature vectors to develop text summarization based on fuzzy logic. The score for every sentence is calculated using a fuzzy logic technique that incorporates all eight feature scores. The deltoid association function and fuzzy rules are used in the fuzzy logic technique.

. Individual scores are fuzzified using the triangle association function, which has three values: LOW, MEDIUM, and HIGH. The IF-THEN expressions are used to write a fuzzy set of statements. After that, fuzzy rules are applied to decide whether the phrases are irrelevant, frequent, or substantial. Defuzzification is another term for this. For instance, if (F1 is M), (F2 is H), (F3 is H), (F4 is H), (F5 is H), (F6 is H), (F7 is M), and (F8 is M), THEN (the sentence is vital). All phrases are evaluated in descending order depending on their grade in the phrase selection procedure. Based on compression rate, the topmost **n** phrases having maximum grade are chosen as summaries text. Finally, the summary sentences are arranged in the same order as they appear in the original text.'

3.1.7 Latent Semantic Analysis

In the handling of natural languages, this method has emerged to evaluate the relationships between texts through creating relevant terms and the vocabulary used in those texts [40]. LSA implies that phrases with comparable sentences will exist in related sections of the document. The structure of a great A matrix holding word counts each paragraph is a piece of text (Unique terms are denoted by rows, while each paragraph is denoted by columns), SVD (Singular value decomposition) is a scientific method used for reducing the number of rows while keeping the similarity structure among columns. When comparing words, the cosine angle created by any two rows is employed. Values near to 1 correspond to relatively related words, whereas values near to 0 reflect extremely different words.

In LSA, 3 main steps are involved. Which are as follows :

1) The creation of an input matrix.

2) Decomposition of Singular Values.

3) Sentence choice.

3.2 Abstractive Text Summarization Approaches

Structure-based approaches (e.g. trees, graphs, laws based, and patterns), semantic-based methods (e.g. depends on data items, establish arguments, and semantic graphs), and deep-learning-based ways are the three main categories of abstractive text summarization described by[41]. [42] divide abstractive approaches into two categories: neural-centered and traditional, which generally indicates any technique which is not neural-centered. Structure-centered approaches locate the best essential material in the input document, then provide abstractive summaries using graphs, trees, rules, templates, or ontologies [41]. Semantic-based approaches use knowledge details, predicate opinions, or semantic graphs to construct a semantic illustration of the input document, which is later on used to generate abstractive summaries using a natural language generation system.

We discuss some important methods in this article which is as follows:

3.2.1 Methods Based on Graphs:

[43] present an abstractive summarizer called "Opinosis" which utilizes a graph prototype. Each point refers to a term, and nodes are connected by positional information. Sentence structure is expressed by pointed edges. The graph-based technique [43] includes the following processing steps:

1) creating a textual graph to describe the source material, and

2) producing the desired abstractive summaries to do so, the following sub-routes in the graph are investigated and recorded:

1. Pathway's scores are sorted in descending order after rating them. The failed pathways are also included in the ranking.

2. Using a similarity measure, eliminate duplicated (or extremely similar) pathways.

3. Choose the highest waiting routes for the created summary, with a parameter controlling the number of edges that define summary volume.

3.2.2 Methods Based on Tree Approach

These algorithms find similar sentences that contain common information, then put them together to make an abstracted summary [44]. A shape like a tree is used to represent comparable sentences. The most widely used tree-form structures for the data are dependency trees document, which is built using parsers. Some operations are accomplished in the processing of the trees, such as lopping, linearization (i.e. turning trees into strings), and so on, to construct the final summary [41]. [45] suggest the following multi-document abstractive summarizer:

1) analyses the corpus's input texts to create a list of all syntactic dependence trees,

2) From the syntactic dependence trees, select a series of partial dependency trees (of varying amounts).

3) To ensure topical diversity, group the extracted limited dependence trees, and

4) use the restricted trees in each group to build a specific phrase that expresses

3.2.3 Methods Based on Rules

These methods necessitate the definition of rules and classifications to identify the key ideas in the documents, which are again used to generate the summary. This method's steps are as follows:

1) Sort the input text into categories based on phrases and topics found there,

2) design problems depend on the input text's area,

3) respond to the questions by discovering textual terms and notions, and

4) feed the responses into certain formats to get the abstractive summary "What is the event?" "Who did the event?" "When did the event happen?" "Where did the event happen?" "What was the impact of the event?" and so on are examples of inquiries [41].

[46]offer a structural design depends on abstraction strategies each abstraction structure is tailored to a specific subgroup or idea and includes substance Information Extraction (IE), algorithms for selecting principles, and simple production examples. All these guidelines are made by hand. A generalization plan aims to address one or additional qualities, and there may be multiple schemes that are connected to the same aspect. The subject collection component can pick the finest candidates to submit to the generating unit centered on the IE policies that can identify numerous nominees for each position.

3.2.4 Methods Based on Semantic Representation

These methods create a semantic representation of the input document(s) (e.g., knowledge items, predicate-reason arrangements, or semantic diagrams), which is then given to an NLG (natural language generation) method, to generate a final abstractive summary using a verb and noun phrases [41]. suggest an abstractive summarizer for multi-text that is:

1. Uses a semantic similarity measure to cluster throughout the text, there are semantically identical syntactic structures,

2. Rank predicate-disagreement organizations depend on characteristics biased and enhanced through a Genetic System.
3. Uses a semantic likeness determines.

3.2.5 Methods Based on Deep-Learning

Abstractive summarization is now likely to be grateful to the recent achievement of sequence-to-sequence understanding (seq2seq) [47]. Seq2seq has excelled at a variety of NLP tasks, including machine interpretation, speech detection, and conversation techniques [48]. For short text summarization, a collection of RNN versions centered on attention encoder-decoder gets encouraging outcomes; however, deep understanding techniques yet to have several issues, such as

1. Inability to handle out-of-vocabulary (OOV) terms (i.e. rare and limited-occurrence terms) and
2. Incapacity to produce frequent words or idioms

The summarization system in [47] includes the following steps:

1. transforming the corpus to clear text and keeping the initial texts (for example, news items) and their outlines individually
2. processing the data using word separation and a sub-word method, and
3. adjusting the word vectors using the Genism toolkit [49], which will be more skilled in the proposed standard
4. Tensorflow [50] was utilized for execution, with individual bidirectional and single unidirectional Long Short-Term Memory (LSTM) levels for encoder and decoder, respectively.

The loss is calculated using cross-entropy, and the loss is optimized using the Adam optimizer.

4. INTERPRETATIONS OF SUMMARIZATION APPROACHES

The combined literature suggests the following key view:

- The fundamental work before discovering extractive summarization is to uncover substantial knowledge included in summaries, according to the synthesized literature.
- Some sentences are longer than others, which is why they may not always hold.
- In the summary, there is a lot of material that isn't necessary.
- Important information is preserved in separate areas of the text; extraction techniques are occasionally used.
- It's possible that the summary won't find all of the document's useful information.
- The summary may contain redundant information.
- Why Summaries based on extraction are unappealing to read.
- In a summary text, there is a lack of flow because mined items are taken from various portions of the text, resulting in unexpected topic shifts.
- Abstractive summaries can sometimes miss the semantic relationship between key terms in a text.
- Natural Language Generation rules are highly required for creating thorough summaries.
- Abstractive summaries can sometimes be incomprehensible.
- Abstractive summarization necessitates a semantic comprehension of the text.
- The quality of abstractive summaries is determined by extensive linguistic knowledge.

5. Conclusion

Manual text summarising is a time-consuming and expensive process with numerous processes. To manually summarise a single document, for example, the stages are as follows [51]:

1) attempting to comprehend the contents of the document,

2) extracting the "most important" elements from the text.

3) attempting to produce a summary that meets the criteria discussed below [52]:

- The legibility and semantic value of the summaries.
- The content coverage and consistency of the summary.
- The produced summary's non-redundancy.

Because instruction manual text summarization of the vast quantity of documented matter on Cyberspace or in other documents is impossible, ATS methods have emerged as the primary solution to address this critical problem. Although there are several artificial document summarizers in the literature, their results are even much away from those of human document summaries.

There is substantially less literature available for abstractive summary than for extractive document summarization. Since the abstractive methodology is far more difficult and less strong than the extractive methods, most survey papers [53][54][55][56][57][58], The purpose of this analysis is to provide a thorough examination and comprehensive summary of the many components of summarization. Following are the survey's primary contributions:

- The ATS systems' various classifications and functions are discussed.
- Conducting a comprehensive examination of the literature on ATS strategies (particularly extractive, abstractive, and hybrid), as well as the methodology used to apply these methods.
- Classifying and describing the various developing sections and strategies used to create and execute ATS systems, such as

1) Document summarising procedures,

2). Statistical and grammatical characteristics, and

3) Document summarization structure modules (namely the text interpretation standards, the grammatical evaluation and processing methods, and the soft processing methods).

- Offering a high-level overview of the regular datasets, manual assessment standards, and computerized assessment approaches that are commonly used to evaluate computer-generated summaries.
- Providing the ATS research community with a list and classification of upcoming exploration ways. The rest of this section will go over these study directions.

There are some issues with using ATS systems, such as.

User-Specific Summarization Challenges: The main question is to summarise matter from a

variety of documented and semi-organized resources (e.g. files and web pages) in the appropriate manner (language, structure, size of it, and time) for each user [58]. Due to the vast amount of data available in many forms and languages, it is necessary to focus more investigation attempts on multi-document, multi-linguistic, and multimedia summaries. It's also necessary to create summaries with a certain emphasis, such as sentiment-centered, customized summaries, and so on.

Issues with Input and Output Designs: Most summarization systems work with textual (written) input. It is necessary to suggest novel summarizers using inputs such as meetings, videos, audio, and other media and outputs additional than text. For instance, the input could be text, and the output could be tables, numbers, visuals, graphical ranking scales, and so on. Users will benefit from ATS methods that allow for the visualization of summaries since they will be able to receive the information they need in less time[58].

Issues with Input Document Length: Most ATS systems are designed to work with short text

documents. A news story, for example, is briefer than a novel unit (around 650 words against 4,864 words) [59]. The existing ATS approaches may perform well when summarising small texts, but they perform poorly when summarising long texts[48].

Issues with Supported Languages: The majority of ATS systems concentrate on English

language material. The quality of current ATS systems for many more languages requires to be enhanced. It is necessary to expand and enhance NLP technologies such as POS, NER tagging syntactic and semantic parsing, and others that are used to generate summaries for non-English languages. [60]. Other issues associated with ATS systems' approaches and tactics include 1) content summarization methods, 2) arithmetical and grammatical features, and 3) text summarization utilizing deep learning

Issue Using Deep-Learning: In seq2seq method depends on deep learning needs enormous managed data for training during the summary generating phase. In actual NLP applications, the requisite training data is not always available. Building a summarization model with a little number of training records utilizing a mixture of classic NLP approaches such as syntactical evaluation, grammatical evaluation, semantic evaluation, and so on is an interesting research issue.[48]. Finally, there are certain issues with the production and created summary from the summarization model such as 1) the summarising method's stop criteria, 2) the excellence of the produced summary, and 3) the assessment of the produced summary.

6. References

[1] E. Hovy and D. Marcu, "Automated text summarization," Oxford Handb. Comput. Linguist., vol. 583598, 2005.

[2] I. Mani and M. T. Maybury, "Advances in automatic text summarization, vol. 293," Camb MA, 1999.

[3] H. P. Luhn, "The automatic creation of literature abstracts," IBM J. Res. Dev., vol. 2, no. 2, pp. 159–165, 1958.

[4] D. R. Radev, E. Hovy, and K. McKeown, "Introduction to the special issue on summarization," Comput. Linguist., vol. 28, no. 4, pp. 399–408, 2002.

[5] S. Dutta, V. Chandra, K. Mehra, S. Ghatak, A. K. Das, and S. Ghosh, "Summarizing microblogs during emergency events: A comparison of extractive summarization algorithms," in Emerging Technologies in Data Mining and Information Security, Springer, 2019, pp. 859–872.

[6] Q. Mei and C. Zhai, "Generating impact-based summaries for scientific literature," in Proceedings of ACL-08: HLT, 2008, pp. 816–824.

[7] N. Nicolov, K. Bontcheva, G. Angelova, and R. Mitkov, "Recent Advances in Natural Language Processing III: Selected Papers from RANLP 2003," 2004.

[8] O. C. Rambow, L. Shrestha, J. Chen, and C. Lauridsen, "Summarizing email threads," 2004.

[9] P. S. Newman and J. C. Blitzer, "Summarizing archived discussions: a beginning," in Proceedings of the 8th international conference on Intelligent user interfaces, 2003, pp. 273–276.

[10] E. Amitay and C. Paris, "Automatically summarising web sites: is there a way around it?," in Proceedings of the ninth international conference on Information and knowledge management, 2000, pp. 173–179.

[11] J.-Y. Delort, B. Bouchon-Meunier, and M. Rifqi, "Enhanced web document summarization using hyperlinks," in Proceedings of the fourteenth ACM conference on Hypertext and hypermedia, 2003, pp. 208–215.

[12] R. Nallapati, B. Zhou, C. dos Santos, Ç. Gulçehre, and B. Xiang, "Abstractive text summarization using sequence-to-sequence RNNs and beyond," CoNLL 2016 - 20th SIGNLL Conf. Comput. Nat. Lang. Learn. Proc., pp. 280–290, 2016, doi: 10.18653/v1/k16-1028.

[13] A. See, P. J. Liu, and C. D. Manning, "Get to the point: Summarization with pointer-generator networks," ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap., vol. 1, pp. 1073–1083, 2017, doi: 10.18653/v1/P17-1099.

[14] R. Paulus, C. Xiong, and R. Socher, “A deep reinforced model for abstractive summarization,” arXiv Prepr. arXiv1705.04304, 2017.

[15] J. R. Quinlan, “Induction of decision trees,” Mach. Learn., vol. 1, no. 1, pp. 81–106, 1986.

[16] C. Quinlan, “5: programs for machine learning Morgan Kaufmann,” San Fr. CA, 1993.

[17] C.-Y. Lin and E. Hovy, “Identifying topics by position,” in Fifth Conference on Applied Natural Language Processing, 1997, pp. 283–290.

[18] H. P. Edmundson, “New methods in automatic extracting,” J. ACM, vol. 16, no. 2, pp. 264–285, 1969.

[19] C. D. Paice, “Constructing literature abstracts by computer: techniques and prospects,” Inf. Process. Manag., vol. 26, no. 1, pp. 171–186, 1990.

[20] C.-Y. Lin, “Training a selection function for extraction,” in Proceedings of the eighth international conference on Information and knowledge management, 1999, pp. 55–62.

[21] D. Marcu, “The rhetorical parsing of unrestricted natural language texts,” in 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics, 1997, pp. 96–103.

[22] C.-Y. Lin, “Topic identification by concept generalization,” in Proceedings of the Thirtythird Conference of the Association of Computational Linguistics (ACL-95), 1995, pp. 308–310.

[23] K. McKeown and D. R. Radev, “Generating summaries of multiple news articles,” in Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, 1995, pp. 74–82.

[24] E. H. Hovy and L. Wanner, “Managing sentence planning requirements,” in Proceedings, ECAI-96 Workshop on Gaps and Bridges: New Directions in Planning and Natural Language Generation, 1996, pp. 53–58.

[25] I. Langkilde and K. Knight, “Generation that exploits corpus-based statistical knowledge,” 1998.

[26] S. A. Babar and P. D. Patil, “Improving performance of text summarization,” Procedia Comput. Sci., vol. 46, pp. 354–363, 2015.

[27] L. Suanmali, N. Salim, and M. S. Binwahlan, “Fuzzy logic based method for improving text summarization,” arXiv Prepr. arXiv0906.4690, 2009.

[28] G. PadmaPriya, “An approach for text summarization using deep learning algorithm,” Int. J. trends Comput. Sci., no. 1, 2014.

[29] J. Kupiec, J. Pedersen, and F. Chen, “A trainable document summarizer,” in Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, 1995, pp. 68–73.

[30] C. Nobata and S. Sekine, “CRL/NYU summarization system at DUC-2004,” 2004.

[31] M. A. Fattah and F. Ren, “GA, MR, FFNN, PNN and GMM based models for automatic text summarization,” Comput. Speech Lang., vol. 23, no. 1, pp. 126–144, 2009.

[32] F. Kyoomarsi, H. Khosravi, E. Eslami, P. K. Dehkordy, and A. Tajoddin, “Optimizing text summarization based on fuzzy logic,” in Seventh IEEE/ACIS International Conference on Computer and Information Science (icis 2008), 2008, pp. 347–352.

[33] K. Kaikhah, “Automatic text summarization with neural networks,” in 2004 2nd International IEEE Conference on’Intelligent Systems’. Proceedings (IEEE Cat. No. 04EX791), 2004, vol. 1, pp. 40–44.

[34] A. T. Sarda and A. R. Kulkarni, “Text summarization using neural networks and rhetorical structure theory,” Int. J. Adv. Res. Comput. Commun. Eng., vol. 4, no. 6, pp. 49–52, 2015.

[35] K.-Y. Chen et al., “Extractive broadcast news summarization leveraging recurrent neural network language modeling techniques,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 8, pp. 1322–1334, 2015.

[36] K. Kianmehr et al., “Text summarization techniques: SVM versus neural networks,” in Proceedings of the

11th International Conference on Information Integration and Web-based Applications & Services, 2009, pp. 487–491.

[37] M. Kågebäck, O. Mogren, N. Tahmasebi, and D. Dubhashi, "Extractive summarization using continuous vector space models," in Proceedings of the 2nd Workshop on Continuous Vector Space Models and their Compositionality (CVSC), 2014, pp. 31–39.

[38] R. Witte and S. Bergler, "Fuzzy coreference resolution for summarization," in Proceedings of 2003 International Symposium on Reference Resolution and Its Applications to Question Answering and Summarization (ARQAS), 2003, pp. 43–50.

[39] P. D. Patil and N. J. Kulkarni, "Text summarization using fuzzy logic," Int. J. Innov. Res. Adv. Eng., vol. 1, no. 3, pp. 42–45, 2014.

[40] G. Swathi and D. Ruby, "Enhancement of Quality of Service in VANET'S By Using Reliable Routing Scheme," Int. J. Recent Dev. Eng. Technol., vol. 2, no. 3, 2014.

[41] S. Gupta and S. K. Gupta, "Abstractive summarization: An overview of the state of the art," Expert Syst. Appl., vol. 121, pp. 49–65, 2019.

[42] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: A comprehensive survey," Expert Syst. Appl., vol. 165, p. 113679, 2021.

[43] K. Ganesan, C. Zhai, and J. Han, "Opinosis: A graph based approach to abstractive summarization of highly redundant opinions," 2010.

[44] V. Gupta, N. Bansal, and A. Sharma, "Text summarization for big data: A comprehensive survey," in International Conference on Innovative Computing and Communications, 2019, pp. 503–516.

[45] L. J. Kurisinkel, Y. Zhang, and V. Varma, "Abstractive Multi-document Summarization by Partial Tree Extraction, Recombination and Linearization," in Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers), 2017, pp. 812–821.

[46] P.-E. Genest and G. Lapalme, "Fully abstractive approach to guided summarization," in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2012, pp. 354–358.

[47] L. Hou, P. Hu, and C. Bei, "Abstractive document summarization via neural model with joint attention," in National CCF Conference on Natural Language Processing and Chinese Computing, 2017, pp. 329–338.

[48] S. Wang, X. Zhao, B. Li, B. Ge, and D. Tang, "Integrating extractive and abstractive models for long text summarization," in 2017 IEEE International Congress on Big Data (BigData Congress), 2017, pp. 305–312.

[49] R. Rehurek and P. Sojka, "Software framework for topic modelling with large corpora," 2010.

[50] M. Abadi et al., "Tensorflow: A system for large-scale machine learning," in 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16), 2016, pp. 265–283.

[51] K. Takeuchi, "A study on operations used in text summarization," 2002.

[52] E. Lloret, L. Plaza, and A. Aker, "The challenging task of summary evaluation: an overview," Lang. Resour. Eval., vol. 52, no. 1, pp. 101–148, 2018.

[53] A. B. Al-Saleh and M. E. B. Menai, "Automatic Arabic text summarization: a survey," Artif. Intell. Rev., vol. 45, no. 2, pp. 203–234, 2016.

[54] L. M. Al Qassem, D. Wang, Z. Al Mahmoud, H. Barada, A. Al-Rubaie, and N. I. Almoosa, "Automatic Arabic summarization: a survey of methodologies and systems," Procedia Comput. Sci., vol. 117, pp. 10–18, 2017.

[55] V. Dalal and L. Malik, "A survey of extractive and abstractive text summarization techniques," in 2013 6th International Conference on Emerging Trends in Engineering and Technology, 2013, pp. 109–110.

[56] P. Gupta, R. Tiwari, and N. Robert, "Sentiment analysis and text summarization of online reviews: A survey," in 2016 International Conference on Communication and Signal Processing (ICCSP), 2016, pp. 241–245.

[57] N. Moratanch and S. Chitrakala, “A survey on extractive text summarization,” in 2017 international conference on computer, communication and signal processing (ICCCSP), 2017, pp. 1–6.

[58] P. Agarwal and S. Mehta, “Empirical analysis of five nature-inspired algorithms on real parameter optimization problems,” Artif. Intell. Rev., vol. 50, no. 3, pp. 383–439, 2018.

[59] Z. Wu et al., “A topic modeling based approach to novel document automatic summarization,” Expert Syst. Appl., vol. 84, pp. 12–23, 2017.

[60] R. Belkebir and A. Guessoum, “TALAA-ATSF: a global operation-based arabic text summarization framework,” in Intelligent Natural Language Processing: Trends and Applications, Springer, 2018, pp. 435–459.