OPEN ACCESS JOURNALS
Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21559113

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Review Article

Volume-06|Issue-04|2026

An Insights into NLP based Abstractive Text Summarization with Modern Approach, Methods, Datasets and its Challenges.

Prafull S. Mankar[1], A. B. Manwar[2]

[1]Research Scholar, Department of Computer Science, S.G.B.Amravati University, Amravati, Maharashtra, India.
[2]Associate Professor, Department of Computer Science, S.G.B. Amravati University, Amravati, Maharashtra, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Mankar, P. S. & Manwar, A. B. (2026). An Insights into NLP based Abstractive Text Summarization with Modern Approach, Methods, Datasets and its Challenges. *Indiana Journal of Multidisciplinary Research, 6*(4), 52-61.

Abstract: In the research area of NLP and deep learning, generating abstractive text summary exactly and consistent enough is significant challenges for text summarization. In abstractive text summarization issues such as processing a text, counting vocabulary, missing words and its embedding, efficiency of models and reaction time of machine etc. had been occurred. Deep learning takes part in NLP task mainly a text-to-text conversion, an image to text transformation tasks. Thus, deep learning leads to an improvement in abstractive text summarization especially by the transformers-based text summarization methods. Transformer models like encoder-decoder based architecture are found to be prominently modern state of the art. For obtaining accuracy and improve efficiency of recent deep learning methods such as transformers with attention mechanism using encoder decoder architecture gaining more popularity in abstractive text summarization. This research paper provides insights of deep learning-based methods for text summarization.

Keywords: Transformers, Deep Learning, Text Summarization, Seq2Seq, Attention mechanism, Encoder-Decoder.

INTRODUCTION

There is demands for to make summary of growing quantity of online text available in the form of scientific papers, books, novels, news articles, legal documents, various fields related online content blogs articles etc., So various methods and techniques are evolved over time for text summarization of such online content to reduce it in some way and generate summary of those content and compress it. Redundant information always leads to decline in efficiency. Also surfing all online content from browser is not favorable and summarization of it creates a problem of loss of essential information during text summarization process. Automatic text summarization is quite useful to resolve the problem by generating a concise and readable summary for obtaining necessary information in reliable, efficient and rapid way [1]. Text summarization instinctively generates a quality based summing up which contains vital sentences and pertinent content from source text document. There are mainly two categorization of text based summarization mainly an extractive and abstractive based summarization. In Extractive type of text summarization, important sentences or words are included in final summary as against to that of Abstractive summarization where it include only meaning and alternative words or sentences instead of original text document and create brief, coherent, consistent summary from source documents. Now it is very noteworthy to produce real time text summarization for events such as live sports events, breaking news etc. Abstractive text summarization is useful to extract appropriate information from research papers and generates a meaningful and useful summary as well. For summarizing Corpus based methods prominently use for semantic checking with the help of semantic similarity where knowledge based only manually check meaning in semantic database. Abstractive text summarization provides summary that is divided into two categories namely semantic and structure driven methods. The primary goal of abstractive based summarization is to produce summary for small documents rather than large amount of text document [2]. Deep learning is efficiently used to take decision by attempting and extracting features at different level of abstraction, imitating what human brain acquire. Deep learning use in NLP tasks for giving the learning facilities of multilevel hierarchical by using several layers of non linear units to represent data. The deep learning model like RNNs, CNNs and Seq2Seq which can be used for abstractive based text summarization types [3].

An Encoder-Decoder with neural network is important especially for attention driven model in abstractive text summarization. As there are difficulties to extract vital information from source document with the help of abstractive summarization for semantic representation, the transformer mod l plays a vital role for summarization using deep learning. Problem solving issues arises in seq2seq task and the transformer tackle long collection dependency efficiently. There are two

***Corresponding Author:** Prafull S. Mankar

layers of transformer model, in which first encoder and second decoder. Then attention layer integrate feed forward based network layers were linked to these layers. The accurate position and chain of works to create positional encoding the transformer model utilizes cosine functions as well as sine functions. Mechanisms like self attention term intra attention were helpful to compute the chain design by comparing numerous positions with solitary pattern. The decoder and encoder model with attention mechanism is available in transforming network by using multi-head attention layer [4]. Since machine lacks human understanding, the complexity level increase and time consumption factors required for text summarization process also get affected. To deal with problem such as classification problem various techniques evolved such as topic in sequence, latent semantic analysis, Seq2Seq models, reinforcement learning, an adversarial processes. These methods then determine effectively and efficiently whether to include the important sentences at the time of generation of concise summary [5]. Instead of using RNN cells, the transformer based architecture with self attention mechanism was used. The point wise fully connected layers computes easier, convenient and parallelizable manner. Also when transformer get combine with positional encoding captures relative token position which ware unclear and have long range relationships [6].

AUTOMATIC TEXT SUMMARIZATION

For abstractive text summary generation the language understanding tools useful for extracting phrases and lexical chains. For deriving vital features, identifying specific information, for illuminate and compress information an abstractive text summarization is used. As it relies on deep linguistic knowledge, abstractive based text summarization was very difficult and it cannot mathematically or logically devise. These techniques can be broadly categories into structured driven approach, semantic approach along with deep learning. The structured approach consists of interpretation of crucial data in the form of reasonable compositions like script and templates. In semantic approaches, a natural language construction structure represent sentence which takes key in either semantic based or linguistic kind of data. Main objective is to produce phrases based on noun and verb to handle linguistics data. The outputted sentences connected with ideas it depends on the ontology driven annotation and clustering techniques which can be crucial part of document such as paragraph and sentences.

Extractive method is easier as compare to abstractive text summarization which produces quality summary with accuracy, complete and informative manners [7].Various pre-built approaches based on statistics, topic, graph were used. After that, the model based on neural with supervised and unsupervised machine learning techniques arises. Automatic text summarization performs better with time by the use of seq2seq methods of deep learning. By implementing and using various techniques or mechanisms just like attention mechanism, pointer generator, coverage method, bidirectional based architectures as well as knowledge improved method, the process of generating text summarization done in excellent manner. Finally, the transformer and pre-trained language model produces the quality results to improve best summary closer to human generated summary [8].

COMPARATIVE STUDY OF ATS TYPES, METHODS AND MODELS

Automatic text-based summarization separated into various category depends mainly on its classification, certain criteria along with various approaches and their associated methods, operations and models with the help of tabular representation as shown in Table-1.

Table1- Various ATS classification with types.[9][10]

Automatic text summarization (ATS) classification							
Depends on document Input Size	Depends on Summarization Approach	Depends on Nature of Output Summary	Depends on Summary Language	Depends on Summarization Algorithm	Depends on Summary Content	Depends on Summary Type	Depends on Summarization Domain
Single	1.Extractive	1.Generic	1.Mono lingual	1.Supervised	1.Indicative	1.Headline	1.General
Multi	2.Abstractive	2.Query based	2.Multi lingual	2.Unsupervised	2.Informative	2.Sentence Level	2.Domain Specific
	3.Hybrid (Combination of Extractive and Abstractive)	--	3.Cross-Lingual			3.Full Summary	
ATS Approaches and their associated methods.							
Extractive	Statistical	Concept	Topic	Sentence Centrality	Graph Based	Semantic Based	SC Based
	Machine-Learning	Deep Learning	Optimization	Fuzzy Logic	Others		

***Corresponding Author:** Prafull S. Mankar

Abstractive	Structure based	Graph based	Tree based	Rule based	Template based	Ontology based	Semantic based
	Information Item	Predicate Argument	Semantic Graph	Deep learning based			
Hybrid	Extractive method to Abstractive method	Extractive method to Shallow Abstractive					
ATS Summarization Operations							
Single Sentence Operation	Sentence Compression	Syntactic Transformation	Lexical Paraphrasing	Generalization	Specification		
Multi Sentence Operation	Sentence Combination	Sentence Reordering	Sentence Selection	Sentence Clustering			
ATS Models							
Graph based Model	Lexical based Graph	Semantic based Graph		Vector based Model	Bag of the Words	Vector Space based Model	Word Vector
N Gram Model	Bi grams	Tri grams	Quad grams	Meaning Representation	Lambda Calculus	AMR	Meaning Representation

STATE OF THE ARTS ABSTRACTIVE TEXT SUMMARIZATION:

The state-of-the-art with abstractive based text summarization can consists of following: Traditional Seq2Seq based Models, Pre-trained based LLM, approaches based on reinforcement Learning (RL), Hierarchical, and Multi-modal type of Summarization [11].

Sequences-to-Sequence Models:
Seq2Seq models: Few applications like machine translation, headline creation, text based summarization along with speech recognition were NLP tasks where seq2seq models can be efficiently used [12].

Figure1 shows each seq2seq model contains an LSTM cell, encoder-decoder utilizes a word embedding file in text summarizer. The vocabulary size of the files which utilized as input for model has been counted then.

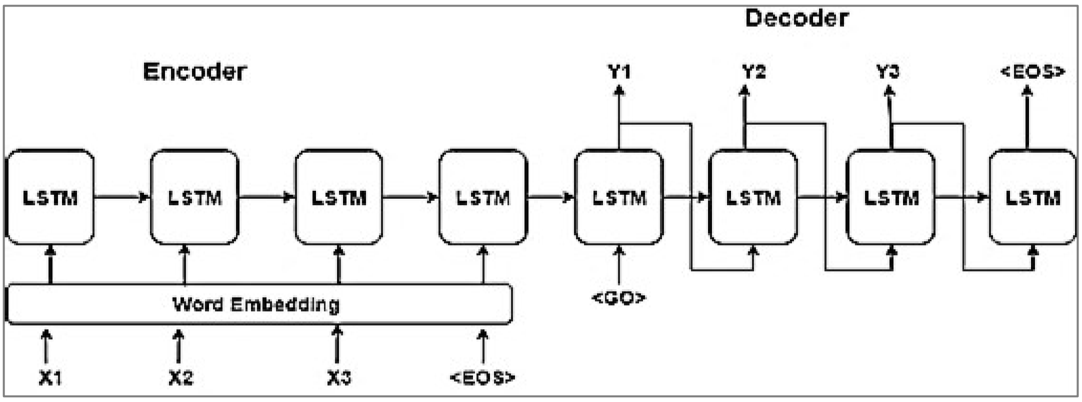

Figure1: Model of Sequence to Sequence [12]

Some factors provided with the controlled vocabulary and few words were not been altered. The end of the sequence that notifies the encoder upon receipt of the input is incorporated in the EOS token. The GO token instructs the decoder to commence handing out the output sequence [13]. Neural network designs known as traditional Seq2Seq models were created to transform input sequences into output sequences. The machine translation and text summarization mostly use Seq2Seq mechanism. The basic idea is to make use of a decoder to produce the yield sequence from the vector based representation. An encoder has altered the input series into an explicit length vector based depiction. Seq2Seq models further categories into four subclasses in the context of abstractive type of text based summarization includes fundamental Seq-2Seq models, attention mechanisms, pointer networks and copy mechanisms [11].

Attention Mechanisms: NLP applications like machine translation, captioning of image, neural abstractive method based text summarization comprises

***Corresponding Author:** Prafull S. Mankar

of the attention mechanism. The attention mechanism has proven to be exceedingly efficient and is persistently make use of seq2seq models. Attention driven encoder decoder structural design consists of the decoder which preferentially spotlight on certain portion of the article at each decoding stage in addition to receiving the encode representation [12]. Prior to the NLP tasks just like text based summarization, the attention mechanism for neural machine translation were used. Since the encoding size is permanent for the source string, the essential encoder-decoder based architecture unable to hold extended sentence. It is unable to take into account all of the components of a lengthy input [13].

Copy Mechanism: Copy mechanism is an exclusive method for abstractive text summarization which tackles the complexity by managing out-of-vocabulary i.e. OOV terms. Since the OOV terms were frequently included in actual world text facts in conventional manner. To solve this issue a method which allows the neural based network to directly duplicate certain terms from the source text to the summary produced. The copying method was integrated into the attention technique of the pre-existing Seq2Seq framework. This approach enhances model's capability afforded more exact and consistent summaries by eliminating challenges cause by exceptional OOV terms [11].

Pointer Generator Networks: A text summarizer produced the output or result using a pointer generator method. A dataset of news items with summaries of few sentences was utilized. Since out-of-vocabulary (OOV) terms with ease and copied them, makes it possible to create good summaries. BERT a pre trained model which improves NLP and whole functions of BERT in summarization process [14].

Pre-trained Large Language Models

Pre-trained language models (PTLMs) very much developed significantly in diversity of NLP based tasks as well as abstractive type of summarization techniques. The prime goal is to examine the associations between large scale corpora at the sentence/token level. Two fundamental approaches are used to accomplish are feature based and fine tuning. Task explicit architectures were additional features for the feature based method. The model has classification layer added in order to use the fine-tuning techniques [15].

PEGASUS: Similar to extractive summarization, the Pegasus pre training task involve deleting that is mask vital sentence from original document by merge the remaining phrases into a single output sequence. The T5 base model is an encoder-decoder which transforms each work to text to text. Because T5 base prefixes each input by a unique prefix, it performs effectively on a range of jobs. BART is a transformer which comprises of GPT and BERT. The pre training task comprise of applying a original in satisfying technique that substitutes a solitary mask token for text spans and randomly rearranging the original phrases order [16] [17].

BERT: The primary responsibility of encoder is to provide or to preserve syntactic and semantic information from the source text to produce required output needed by the decoder. Since the pre trained language model uses a far better corpus in its pre training phase than it does in the text summarization training phase. The knowledge transfer from the pre-trained language model should get better for important information recollection potentially. Therefore it utilizes BERT language representation to train an encoder. One of main benefits is that the methods ability to get around BERT's extent restrictions and hold large documents. The mainstream of summarization datasets supports the longest sequence length that BERT support which is less than the distinctive text extent found. The primary distinction is to use simply the word embedding matrix of BERT rather than its final output [18]. BERT is a deep, bidirectional transformer encoder that does not have a decoder connected. In order to produce a matching series of refined, semantically encoded output vectors is used [19].

GPT: The GPT Generative Transformers OpenAI launched a line of transformer-based models called Pre-trained Transformers (GPT). One well-known autoregressive aspect of GPT models is the ability to forecast the subsequent token in a succession depends on the prior context. Architectural design consists of a load up with transformer layers that comprise a location wise feed forward network and multi head attention. With diverse transformer layers, attention heads and model parameters GPT has different variant [20].

T5: T5 is one particularly functional transformers model created by Google Research. A unified text based structure made each NLP job as text to text issues which are its distinctive feature. Generally, encoder processes key in text and appropriate yield text by decoder respectively. To process key in text efficiently and spawn prediction, encoder and decoder stack uses feed forward based neural network layer along with self- attention mechanism [20]. When completing text summary tasks, the T5 model will be evaluated using ROUGE. The consequence of the T5 capability to execute text summary depends on textual data wh ch went through training. ROUGE technique was then use to assess the T5 model's presentation on text summary tasks [21].

Long T5: To manage lengthy inputs, add global, local attention sparsity to the original T5 encoder. The attention mechanism is where T5 and LongT5 deviate most architecturally. Two different attention based mechanisms of LongT5 are Transient Global attention and Local attention. These versions preserve a number of T5 characteristics, including compatibility with T5

***Corresponding Author:** Prafull S. Mankar

checkpoints which is used to support for packing and relative position representations [22].

Text-to-Text Transformer Transfer: A progressive language representation built on transformer structural design is the T5. By transforming the key in and yield into natural language texts, it uses a single text to text structural design that can supervise any natural language processing operation. T5 uses a full-attention technique that enables it to be acquainted with intricate semantic links and long-range dependency. In natural language text documents, numerous NLP tasks like machine translation, summarization of text, problem answer, and sentiment investigation have seen accomplishment with T5. Figure 2 depicts the structural design of the T5 model, which adheres to standard encoder-decoder topology [23].

The Integrated Text-to-Text Transformer: Examining boundaries of transfer learning, T5 is an back-to-back transformer representation as opposed to BERT models. T5 trained using text data key in and transformed text output instead of only a group label or an expansion of the input. T5 an encoder-decoder paradigm is used to encode and decode problems associated with natural language processing. It is trained by instructor intimidation. Therefore, it always requires a set of inputs and a set of outputs in order to do training. The source sequence is abounding to the representation through key in. A start sequence is appended to the target sequence using the decoder input token prior to the decoder receiving it. T5 may be trained and optimized in both supervised and unsupervised settings [5].

BART: The Facebook AI published transformer based BART model. The standard architecture used by BART process information in a seq2seq approach. A Design comprises of bi-directional encoder as well as distinctive bi-directional decoder of autoregressive type. The auto regressive character of BART in which the model is taught to construct a target sequences from distorted input sequences. To disguise precise mechanism for forecast, the decoder get source tokens while an encoder receive a degraded edition of a token [20]. BART does exceptionally well at comprehending messages and collecting context using bidirectional feature. Based on the original content, it can produce flowing and logical summaries using auto-regressive function. Pre training done on considerable dataset and then excellent modification on the task explicit information wer the methods used to train the BART model. It can therefore produce outstanding outcomes in the translation and summarization fields [24]. In NLP, BART is a strong translation and summarization paradigm. Transformer based model BART performs exceptionally well on linguistic problems. It is the blend of auto regressive and bidirectional transformers or BART. It works well for both translation and summarization jobs because bidirectional nature along with auto regressive characteristics [25].

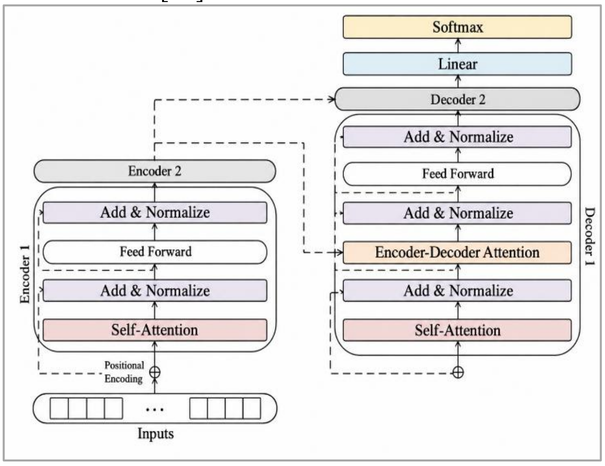

Figure 2: Architecture of the T5 model [23]

Large Bart: The BART model has been the potent language and BART Large is a variation of it. In contrast to the standard BART model, the BART Large is bigger representation size and particularly trained on a extensive dataset. Large Bart improved size enables it to take hold of more complex language patterns thus improving its capacity for language creation and comprehension [24].
Bart Model: The innovative BART transformer concept was developed by Facebook AI. The transformer architecture's central component of BART is made up of feed forward neural networks and several levels of self attention processes. It can effective y handle a range of NLP problems since it has both bidirectional and auto regressive training objectives. BART can gather rich semantic representations of text and produce coherent outputs for a variety of downstream tasks. BART may be optimized for text summarizing tasks using summarization datasets like CNN/Daily Mail to provide excellent abstractive summaries of news items [26]. To generate the summary, extractive summarization chooses input phrases from supply text. Avoiding repetition as well as identifying the most useful sentences are challenges [27]. Common strategies include clustering, TF-IDF, graph-based techniques, and others. By creating new phrases abstractive summarization uses techniques like BERT, BART, neural networks, and others to create more succinct generalized summaries [28].

EGASUS: A sequence-to-sequence based transformer exclusively made for abstractive summarization called as PEGASUS. The gap sentence generation approach which was developed by Google Research pre trains PEGASUS using a large corpus by eliminating few sentences from the given key in sequences. The transformer then has to restructure the output sequence from the given input tokens in the form of a degraded sequence [20]. Abstractive summarization pre-training along with extract break sentence was released on eliminating or masking key sentence from supply text and then constructing a single output sequence from the remaining phrases. It is a pre training exercise for Pegasus that

*Corresponding Author: Prafull S. Mankar

creates a task similar to an extractive summary [5]. After extracting the most significant lines from the input data, they are assembled into distinct outputs. Additionally, picking the most pertinent lines is preferable to picking them at random. CNN/ Daily Mail newspaper datasets were used to pre train the model [29].

LSTM: It is a unique variation of RNN addresses the issue of extended term dependency. Recurrent neural network are not proficient of storing information for extended periods of time. For instance, if the next word in a sequence has anticipated and the word that is successfully forecasted depends on information which has spoken a few sentences. That's why LSTMs (long-term gaps/dependencies) are used and longer learning times are possible with LSTM [17].

DEEP LEARNING BASED METHODS AND MODEL

A abstractive text summarization deep learning approach uses an LSTM-CNN based structural design. For faster processing in the forward training technique the LSTM model is combined with an encoder based on convolutional neural networks. The ROUGE parameter is used to evaluate the approach, and the average number of distinct phrases is used generated the abstractive summaries using the seq2seq encoder-decoder paradigm. The relevance model and generic summarization are the foundations of the query-focused summarization that is used to remove redundancies and provide an effective summary. Each word in the seq2seq model is given a predetermined relevance score in order to assess its weight [30]. In the past, people have been manually condensing text manuscripts for a long time. However, consumers are finding it difficult to efficiently manage the huge amount of information available over internet includes news, blogs, articles with lengthy written materials. Text summaries have gained significance and are crucial for addressing this problem. The importance of text based summary is found its competence to quickly and effectively dig up significant information from a extensive text [12].

Deep neural networks: Deep learning is based on DNN which train diverse models using complex mathematical techniques. It is rarely referred to as a multi layer perceptron due to the large number of unknown layers it has.

RNN: The idea behind the recurrent neural network is the intuitive realization that, there is a chronological association inside the series and neighboring items rely on one another. Network uses properties of input at the previous and current time steps to predict the yield of the succeeding instance step. In particular, the RNN's concealed layer nodes are interconnected, the input layer's output and the preceding hidden layer make up hidden layer input [31].

GNN: A neural network with a focus on processing graph data called as graph neural network GNN. GNN's fundamental concept is to entrench nodes based on their local neighborhoods. It makes instinctive sense that a neural network would aggregate the properties of every node and nodes that are linked to it [1].

SUMMARIZATION METHODS

The Bayesian Classifier: It can create a classification function which calculates the likelihood that a certain sentence is contained in an extract from a training set of documents containing hand-picked document extracts. After evaluating sentences based on a user-specified number of the peak scoring sentences can be chosen to create new extracts.

Markov Hidden Model: Another technique for extracting a phrase from a text is the hidden markov model HMM. The Hidden Markov model makes fewer assumptions about independence than a naïve Bayesian method. Three characteristics were employed in the HMM Model which are sentence's position within the document, the amount of terms it contains and the likelihood of sentence requisites specified the document vocabulary.

Text Summary Based on Neural Networks: One of most popular and potent categories of machine learning algorithms is artificial neural networks ANN. News stories of any length may be summarized using artificial neural networks. A summary of the article's top ranked sentences is then generated by fusing the neural network.

Text Summarization Based on Fuzzy Logic: This approach uses a fuzzy logic set and uzzy logic rule to identify the key sentences based on their characteristics. Fuzzy logic approaches provide expert and decision-support systems controlling reasoning power. A mathematical technique for dealing with ambiguity, imprecision, vagueness, and uncertainty is the fuzzy logic set [32].

Deep Learning Based T5 Model
Transformer Model T5

The transformer that can create the input and output pictures without the need of convolution or sequence recurrence was the first sequence transduction model. Completely connected and shown, are the various normalize and multi-head attention layers mound self attention layer and point wise layers for the encoder-decoder [30]. A versatile language model is the text to text transfer transformer T5 model. It is capable of handling a broad range of natural language processing tasks such as text to text difficulties. It goes through a pre training phase in which it uses a sizable corpus of text data to learn how to understand and generate text broadly [26]. Google pioneered the use of Transformers for succession knowledge problems. Transformers don't need recurring or convolutional unit since depend on attention mechanism. Encoders-decoders should be layered in the transformer outline. Attention units and feed forward units construct the encoder and decoder blocks. Six identical decoder units construct the decoder portion, while six encoder units construct the encoder portion. Every encoder unit has a feed forward item and a multi- head based attention item [28]. The feed forward as well as multi-head attention unit, every decoder item also has a covered multi-head attention item. The word embeddings of the key in series initiate transformer's operation [29]. The transformer-based models are often used for text summarization. By comprehending intricate word and sentence associations, transformers were capable of capturing the main ideas of the text further accurately. These models may be adjusted to do certain tasks, such as summarizing scientific publications or news stories. Transformers are effective in management of longer texts, which makes them appropriate for summarizing news items or research papers [33].

TEXT SUMMARIZATION DATASETS AND EVALUATION METRICS

More than 300,000 specifically unique news stories collected in CNN/Daily Mail dataset by the journalist in English language. Hugging face available where CNN/Daily Mail dataset is collected [26].

ROUGE: To assess the computerized text based summarization, ROUGE metrics have been used. One of the most popular methods for assessing summaries is ROUGE. It contrasts a set of human generated reference summaries with an automatically produced summary.

BLEU Score: The BLEU score can be used for automatically evaluate machine generated summary. The score always range from 0 to 1, where 0 denotes no overlap and 1 denotes perfect overlap between the machine-generated summary and the ground truth [6].

BLEU-Bilingual Evaluation Understudy: This assessment tool was first created to evaluate the quality of material that has been machine based translation also applied to task of text summarization. The summary generated system contrast with suggestion summary by means of these precision based criteria. BLEU measures overlap of n-grams involving the reference texts as well as produced output at the n-gram level. It determines the accuracy for every n-gram size typically ranging from 1 to 4 grams then computes the final score by taking a weighted geometric mean. **METEOR** an evaluation metric originally designed for machine translation daily jobs and also it could be use in text based summarization assessment process [34]. Comparison of various ROUGE Scores on the CNN/Daily Mail Dataset for Text Summarization is as shown in Table 2.

APPLICATION OF ATS AND ITS CHALLENGES

Table 2 Comparison of various ROUGE Scores on the CNN/Daily Mail Dataset for Text Summarization.

Paper Name	Model	Dataset	R1	R2	R L
Neural Abstractive Text Summarization with Sequence-to-Sequence Models [2].	NATS (Neural Abstractive Text Summarization) [2]	CNN/Daily Mail Dataset [2]	39.11	27.54	35.99
Abstractive Type Text Summarization on Language Model Conditioning and Locality Modelling [10]	Bert-transformer Encoder + Bert-transformer Decoder [10]	CNN/Daily Mail dataset	33.23	14.99	31.26
	Transformer	CNN/Daily Mail dataset	24.82	6.27	22.99
	BERT + Copy Transformer	CNN/Daily Mail dataset	40	18.42	37.15
An Analysis of Abstractive Text Summarization Using Pre-trained Models [16]	Google/Pegasus–cnn-daily mail	CNN/Daily Mail dataset	40.22	17.64	37.51
	T5-base	CNN/Daily Mail dataset	33.96	13.37	31.43
	Facebook / bart-large-cnn	CNN/Daily Mail dataset	41.76	18.67	38.84
Survey on abstractive text summarization: dataset, Models, and metrics [19]	BART large: facebook/BART-large-cnn	CNN/Daily Mail dataset	22.92	5.0	13.18
	PEGASUS large: google / PEGASUS-large-cnn daily mail	CNN/Daily Mail dataset	23.68	6.52	15.24
	BART large: datien228 / distil BART-cnn – 12 – 6 – ftn -multi news; [19]	CNN/Daily Mail dataset	32.0	11.75	19.25
Text based Summarization using diverse Methods for Deep Learning [25]	Reinforcement learning	CNN/Daily Mail dataset	34.11	12.78	22.5
	LSTM-CNN	CNN/Daily Mail dataset	33.26	10.97	23.28
	Seq2Seq model equipped with attention mechanisms, Scheduled Sampling	CNN/Daily Mail dataset	34.18	12.47	23.68

APPLICATION OF ATS AND ITS CHALLENGES

The text document information needs suitable summarization- based technique to decrease upholding attempt. Also require methods to pack in information into noteworthy summary exclusive of trailing the unique situation. The significance of ATS has grown over time

*Corresponding Author: Prafull S. Mankar

owing to the large volume of textual information produce over the Internet. One of the main areas of research has been coming up with ways to distill or summarize large amounts of textual data with text summarization serving as a classic example. Applications for ATS may also be found in a variety of knowledge areas where search kinds of algorithms can be greatly enhanced by substituting a textual summary for the novel data. Creating a textual summary in a certain amount of time is known as real time automatic text summarization [35].

CHALLENGES: Problems in summarizing multiple documents which includes redundancy, sequential dimension, co reference, and sentence reorder are only a few of the many problems that arise while summarizing several documents. Challenges concerning to the user specific summarization includes condensing information from several textual and semi- structured sources. Such as databases, most difficult applications such as book, novel, and long text summaries. Challenges arise with formats for input and output consist of text input and outputs are handling by the majority of ATS systems. New summarizers must be proposed with the yield in a layout other than text and the key in in the form of meetings, videos, sounds, etc. For instance, the yield may be displayed as tabular information, statistics and images, visual rating scales, etc., whereas the input could be text. Additional difficulties with automatic text summarization system methodology and techniques include text summarizing strategies, linguistic and statistical aspects, and text summarization utilized deep learning approach. Also problems occurred with Statistical and Linguistic characteristics include so as to semantically extracting the significant sentences from the source text or documents. Its necessary to find a quantity of innovative linguistic as well as statistical characteristics for words and sentences. Furthermore, choosing the appropriate weights for each characteristic is crucial as it affects the final summary's quality. The difficulty arises in applying deep learning to text summarization since RNN in seq2seq arrangement needs a lot of planned training information during summary generating stage. In real world NLP applications the necessary training data is not always accessible [10][36].

CONCLUSION AND FUTURE SCOPE:

There is vital significance of automatic text summarization as it grows over time owing to the large volume of textual information produced over Internet. Various techniques and mechanism like attention mechanism, pointer generator, coverage method, bidirectional architectures and knowledge enhanced mechanism etc. the text summarization has improved time to time. In NLP, by using the transformer and pre-trained language model which produces and improve the quality of summary closer to human generated summary in more convenient and efficient ways. The Transformer is works well with large volume of text summarizing news items and research papers etc. T5 model capable of handling a wide variety of NLP tasks by using just one framework. Pre-training phases in which T5 uses a sizable quantity of text information to learn how to understand and generate text broadly. This research paper gives insights into the various Deep learning based text summarization methods, especially abstractive text summarization transformer T5 models with Encoder and Decoder structural design. As there are some challenges or issues which include redundancy, temporal aspect, co-reference, and sentence reordering, linguistic and statistical aspects related to text summarization approaches. Deep learning is very much efficient enough to generate quality based text summaries. But as still there is scope for improvement in quality of text summary further by applying and implementing hybrid approach to text summarization in future from research perspective.

REFERENCES

1. Zhang, M., Zhou, G., Yu, W., Huang, N., & Liu, W. (2022). A comprehensive survey of abstractive text summarization based on deep learning. *Computational Intelligence and Neuroscience, 2022*, 1–21. https://doi.org/10.1155/2022/7132226
2. Shaker, F., Zaid, M. M. A., Al-Khazaali, A. A. T., & Al-Khzaali, Z. A. (2024). Text summarization using different methods for deep learning. *BIO Web of Conferences, 97*, Article 00074. https://doi.org/10.1051/bioconf/20249700074
3. Suleiman, D., & Awajan, A. (2020). Deep learning based abstractive text summarization: Approaches, datasets, evaluation measures, and challenges. *Mathematical Problems in Engineering, 2020*, 1–29. https://doi.org/10.1155/2020/9365340
4. Balaji, N., Deepa Kumari, Bhavatarini, N., Megha, N., Sunil Kumar, P., & Rai, S. (2022). A text summarization using NLP technique (pp. 30–35). https://doi.org/10.1109/discover55800.2022.9974823
5. Shah, N. C. A., & Phadnis, N. N. (2022). Text summarization using extractive and abstractive techniques. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 236–241. https://doi.org/10.32628/cseit228361
6. Aswani, S., Choudhary, K., Shetty, S., & Nur, N. (2024). Automatic text summarization of scientific articles using transformers—A brief review. *Journal of Autonomous Intelligence, 7*(5), 1331. https://doi.org/10.32629/jai.v7i5.1331
7. Rane, N., & Govilkar, S. (2019). Recent trends in deep learning based abstractive text summarization. *International Journal of Recent Technology and Engineering, 8*(3), 3108–3115. https://doi.org/10.35940/IJRTE.C4996.098319
8. Alomari, N., Idris, A., Sabri, A. Q. M., & Alsmadi, I. (2021). Deep reinforcement and transfer learning for abstractive text summarization: A review.

Computer Speech & Language, 71, Article 101276. https://doi.org/10.1016/j.csl.2021.101276
9. Nnadi, G. O., & Bertini, F. (2024). *Survey on abstractive text summarization: Dataset, models, and metrics* (arXiv:2412.17165v1 [cs.AI]). arXiv.
10. El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2020). *Automatic text summarization. (The source details appear incomplete—journal/conference, volume, issue, and pages are missing.)*
11. Shakil, H., Farooq, A., & Kalita, J. (2024). *Abstractive text summarization: State of the art, challenges, and improvements* (arXiv:2409.02413v1 [cs.CL]). arXiv.
12. Shi, T., Keneshloo, Y., Ramakrishnan, N., & Reddy, C. K. (2021). Neural abstractive text summarization with sequence-to-sequence models. *ACM/IMS Transactions on Data Science, 2*(1), 1–37. https://doi.org/10.1145/3419106
13. Masum, K. M., Abujar, S., Talukder, M. A. I., Rabby, A. K. M. S. A., & Hossain, S. A. (2019). Abstractive method of text summarization with sequence to sequence RNNs. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1–5). https://doi.org/10.1109/ICCCNT45670.2019.8944620
14. Batra, P., Chaudhary, S., Bhatt, K., Varshney, S., & Verma, S. (2020). A review: Abstractive text summarization techniques using NLP. IEEE.
15. Ulker, M., & Ozer, A. B. (2024). Abstractive summarization model for summarizing scientific article. *IEEE Access, 12*, 91252–91262. https://doi.org/10.1109/ACCESS.2024.3420163
16. Syed, F., Gaol, F. L., & Matsuo, T. (2021). A survey of the state-of-the-art models in neural abstractive text summarization. *IEEE Access, 9*, 13248–13265. https://doi.org/10.1109/ACCESS.2021.3052783
17. Rehman, T., Das, S., Sanyal, D. K., & Chattopadhyay, S. (2023). *An analysis of abstractive text summarization using pre-trained models* (arXiv:2303.12796v1 [cs.CL]). arXiv.
18. Aksenov, D., Moreno-Schneider, J., Bourgonje, P., Schwarzenberg, R., Hennig, L., & Rehm, G. (2020). *Abstractive text summarization based on language model conditioning and locality modeling* (arXiv:2003.13027v1 [cs.CL]). arXiv.
19. Tsirmpas, D., Gkionis, I., Papadopoulos, G. T., & Mademlis, I. (2024). Neural natural language processing for long texts: A survey on classification and summarization. *Engineering Applications of Artificial Intelligence*. https://doi.org/10.1016/j.engappai.2024.108231
20. Muia, C. M., Oirere, A., & Ndung, R. (2024). A comparative study of transformer-based models for text summarization of news articles. *International Journal of Advanced Trends in Computer Science and Engineering, 13*(2), 37–43. https://doi.org/10.30534/ijatcse/2024/011322024
21. Lubis, R., Safitri, H. R., Irvan, Lubis, M., Hamzah, M. L., Al-Khowarizmi, A., & Nugroho, O. (2023). Enhancing text summarization with a T5 model and Bayesian optimization. *Revue d'Intelligence Artificielle, 37*(5), 1213–1219. https://doi.org/10.18280/ria.370513
22. Guo, M., Ainslie, J., Uthus, D., Ontañón, S., Ni, J., Sung, Y.-H., & Yang, Y. (2022). *LongT5: Efficient text-to-text transformer for long sequences* (arXiv:2112.07916v2 [cs.CL]). arXiv.
23. Wang, M., Xie, P., Du, Y., & Hu, X. (2023). T5-based model for abstractive summarization: A semi-supervised learning approach with consistency loss functions. *Applied Sciences, 13*(12), 7111. https://doi.org/10.3390/app13127111
24. Mercan, O. B., Cavsak, S. N., Deliahmetoglu, A., & Tanberk, S. (2023). *Abstractive text summarization for resumes with cutting edge NLP transformers and LSTM* (arXiv:2306.13315 [cs.CL]). arXiv. https://doi.org/10.48550/arXiv.2306.13315
25. Patil, S., Chavan, L., Mukane, J., & Vora, D. (2022). *State-of-the-art approach to e-learning with cutting edge NLP transformers: Implementing text summarization, question and distractor generation, question answering*. *International Journal of Advanced Computer Science and Applications*. https://www.ijacsa.thesai.org
26. Asmitha, M., Kavitha, C. R., & Radha, D. (2024). Summarizing news: Unleashing the power of BART, GPT-2, T5, and Pegasus models in text summarization (pp. 1–6). https://doi.org/10.1109/CONIT61985.2024.10626617
27. Parabrahma Chari, S., Dinesh Kumar, B., Shonaakshay, C., Thrillokh, B., & Ravi Teja, B. (2024). Automatic text summarization using BART. *International Journal of Creative Research Thoughts*. https://www.ijcrt.org/papers/IJCRT2402277.pdf
28. Bhukya, V. K., & Bhukya, U. (2024). *Abstractive text summarisation using T5 transformer architecture with analysis*. *Research Square*. https://doi.org/10.21203/rs.3.rs-4986903/v1
29. Keswani, G., Bisen, W., Padwad, H., Wankhedkar, Y., Pandey, S., & Soni, A. (2021). Survey and analysis of different transformer models for abstractive text summarization. *International Journal of Mechanical Engineering*, 2881–2882.
30. Karnik, M. P., & Kodavade, D. V. (2023). Abstractive summarization with efficient transformer based approach. *International Journal on Recent and Innovation Trends in Computing and Communication, 11*(4), 291–298. https://doi.org/10.17762/ijritcc.v11i4.6454
31. Jin, H., Zhang, Y., Meng, D., & Wang, J. (2024). *A comprehensive survey on process-oriented automatic text summarization with exploration of LLM-based methods* (arXiv:2403.02901v1 [cs.AI]). arXiv.

32. Tas, O., & Kiyani, F. (2017). A survey automatic text summarization. *PressAcademia Procedia, 5*(1), 205–213. https://doi.org/10.17261/pressacademia.2017.591
33. Tonchar, K., Raut, S., Peshwe, M., Rathod, S., Pagrut, R., & Quazi, K. A. G. A. (2024). Text-summarization using transformer based models. *International Journal of Novel Research and Development, 9*(4), C198–C199. https://www.ijnrd.org
34. Wang, G., & Wu, W. (2023). *Surveying the landscape of text summarization with deep learning: A comprehensive review* (arXiv:2310.09411v1 [cs.CL]). arXiv.
35. Wahab, M. H. H., Ali, N. H., Hamid, N. A. W. A., Subramaniam, S. K., Latip, R., & Othman, M. (2023). A review on optimization-based automatic text summarization approach. *IEEE Access, 12*, 4892–4909. https://doi.org/10.1109/ACCESS.2023.3348075
36. Kumar, Y., Kaur, K., & Kaur, S. (2021). Study of automatic text summarization approaches in different languages. *Artificial Intelligence Review, 54*(8), 5897–5929. https://doi.org/10.1007/s10462-021-09964-4.