Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Exploring Summarization Performance: A Comparison of Pointer Generator, Pegasus, and GPT-3 Models

Gresha Bhatia, Abha Tewari, Dr. Nupur Giri

Vivekanand Education Society Institute of Technology, Mumbai
Email: gresha.bhatia@ves.ac.in

The world is rapidly advancing technologically and the way we communicate is changing with it.We are now able to send messages through text, voice, or video chat, which means that the amount of content being generated every day is growing fast. The problem with all this information is that it can be overwhelming for people to stay on top of everything they need to,especially those who have to read a lot of texts and get the gist of what they need to know. Manual text summarizing is a time-consuming and inherently tedious activity. There is a need for a system that could provide crisp, precise information on a real time basis . This would help the decision makers to comprehend the information quickly and take decisions at a faster rate. This also helps to read important information without having to spend too much time on it. By doing so, you save your time and energy, as well as reduce stress levels. This research paper identifies the need for automating the text summarization . It further elaborates on its key concepts and gives a comparison of the various text summarization models. Moving into a proposed model in itself was a challenge. This research paper dwells on the proposed methodology and performs various evaluation metrics.
ImpactStatement -
The process of text summarization involves using natural language processing to condense information into a shorter, more concise version. The goal is to condense the original document while preserving its essential information This paper conducts a comparison of different text summarization methods, including extractive and abstractive techniques. It also categorizes summarization systems and examines the use of statistical and linguistic approaches for summarization.
Keywords: Natural language processing, Summarization techniques, News generation, quality evaluation.

1. Introduction

The field of Natural language processing (NLP) brings together computer science, artificial intelligence, and linguistics to explore the ways computers understand, process and produce human language. With the increasing use of the World Wide Web, there is a growing need for systems that can automatically retrieve, categorize and summarize documents. This is needed so as to help users navigate information overload. Document summarization forms one of the potential solutions to this problem. Text summarization involves condensing a document's content into a concise form that meets the user's needs. With the vast volume of digital data available online, reading everything is impractical, making information condensation essential. Summarization enables users to efficiently extract relevant information from large datasets. [3] Text summarization has various applications. For example, researchers may use it to generate summaries.This can help them to decide whether a full document warrants further reading or to condense information gathered from the internet. News organizations can utilize multi-document summarization to gather information from multiple sources and provide a cohesive summary.

This paper provides an all-inclusive survey of different summarization techniques, including their superiority and limitations. Section II defines text summarization, while Section III explores related work and past literature. Section IV delves into the various text summarization methods while Section V represents the models used for summarization . A statistical approach for text summarization is elaborated in section VI. Applications of text summarization are explored in section VII. Proposed methodology is indicated in section VIII. Results of the study and its respective evaluation measures are further expressed in section IX.The findings are further summarized through the conclusion in section X.

2. Text Summarization

Text summarization is the process of creating a shorter version of a longer text while preserving its essential information.The goal of text summarization is to help users quickly grasp the main points of a document without having to read the entire text. Text summarization is widely applied in areas like news, document and social media summarization.Text summarization primarily uses two approaches namely extractive summarization and abstractive summarization. Extractive summarization involves selecting specific sentences from the original text and combining them to create a summary. This approach is simple and efficient but may result in a summary that is not fluent and lack coherence. On the other hand, abstractive summarization involves generating new sentences that capture the essential information of the original text. This approach requires more advanced natural language generation techniques and can produce more fluent and coherent summaries. However, it is more challenging and computationally expensive.

3. Literature Review

There have been various approaches proposed for text summarization, including rule-based systems, statistical methods, and deep learning models. In recent years, deep learning-based models have shown great promise in achieving state-of-the-art performance on text summarization tasks.One popular model is the Pointer Generator model, introduced by See et al. in 2017. This model achieves a greater performance on a number of benchmark datasets. It also forms a hybrid model that combines an attention mechanism with a pointer network. However, one limitation of this model is that it is computationally expensive, which can limit its scalability for large-scale datasets. Additionally, the model may struggle with generating novel words and phrases that are not present in the source text. Another model that has gained attention in recent years is Google's Pegasus, a transformer-based model introduced in 2019. Pegasus is a pre-trained model that fine-tunes on the summarization task and has been shown to outperform previous state-of-the-art models on multiple benchmark datasets. However, one limitation of Pegasus is that it requires a large amount of computational resources for training and inference. The OpenAI GPT-3 model is another widely used model for text summarization tasks. GPT-3 is a large transformer-based language model that has been shown to perform well on several natural language processing tasks, including text summarization, language translation, and text completion. However,one limitation of this model is that it is not publicly available. Additionally, the computational resources required for training and inference can be significant, making it less accessible to researchers with limited resources. In addition to these models, researchers have also explored various other techniques for text summarization, including extractive and abstractive summarization, reinforcement learning-based approaches, and unsupervised learning approaches. While these approaches have shown promise, they also have their own limitations and trade-offs in terms of efficiency, accuracy, and scalability.Overall, while there have been significant advancements in text summarization using deep learning-based models, there are still limitations that need to be addressed in terms of computational resources and scalability for large-scale datasets.Future research may focus on developing more efficient and scalable models for text summarization, as well as exploring new approaches and techniques to improve the quality of generated summaries.

4. Types of Summarization Techniques

Automated text summarization has been categorized into many different types of summaries based on their usefulness or purposes. Summarization systems can be classified into several categories, based on the approach considered, type of details analyzed, the type of content explored, the limitations identified, the number of input documents undertaken as well as the language used for representation. The difference between each of them is as expressed below:-

1. Based on approach considered :-

In this, the summarization can either be extractive or abstractive. Summarization by extraction involves taking sentences directly from the root document and adding them to the summary, while summarization by abstraction involves generating new sentences that are semantically related to the original text. Abstractive summarization provides a more generalized summary, but it is more difficult to compute than extractive summarization.

2. Based on the type of detail identified:-

In this method, the summarization can be either informative or indicative. Indicative summarization provides only the core idea of the input document and is usually short, encouraging the reader to read the full document. Informative summarization, on the other hand, provides a concise representation of the original document and can serve as a substitute for it.

3. Based on the type of content represented:-

Based on type of content, summarization can be either generic or query-based. Generic summarization is not specific to any particular user or subject and provides information at the same level of importance. Query-based summarization is specific to the user's needs and provides answers to specific questions.

4.Based on the limitations identified:-

Based on limitation, summarization can be genre-specific or domain-independent. Genre-specific systems are limited to certain types of input, while domain-independent systems can accept a wide range of text and are not dependent on the subject of the document.

5.Based on the number of input documents considered:-

Based on this, the summarization can be single document or multi-document. Single document summarization involves summarizing a single document, while multi-document summarization involves summarizing multiple documents on the same topic.

6. Based on the Language used for representation:-

Based on language, summarization can be monolingual or multilingual. Mono-lingual systems only accept documents in a specific language and produce a summary in the respective language. As the name suggests, multi-lingual systems consider documents in multiple languages and generate summaries in the language specified.

5. Text Summarization Models

Three text summarization models were compared as described below:

1. Pointer Generator Networks:

Pointer Generator Networks were found to be a class of sequence-to-sequence model that gained significant attention in the field of text summarization. The Pointer Generator Network was introduced in 2017 as an alternative approach to abstractive summarization, which can generate new sentences to summarize the content of an input document. Instead of generating new sentences, the Pointer Generator Network used a combination of extractive and abstractive technique. This enabled the model to select and combine parts of the input document and create a summary.[13] In the Pointer Generator Network, the input document was first encoded into a sequence of hidden states. This was then decoded to generate the summary. The decoder is equipped with a pointer mechanism that allows it to select and copy parts of the input document as part of the summary. The pointer mechanism further enabled the decoder to generate new words that were not present in the input document. This Network used a combination of attention mechanisms, pointer networks, as well as a coverage mechanism that improved the quality and diversity of the generated summaries. One of the major advantages of the Pointer Generator Network wass its ability to handle long input documents, which could be a challenge for other summarization models. The Pointer Generator Network performed considerably well on a couple of benchmark datasets, demonstrating its effectiveness in summarizing a wide range of text types and genres. Overall, the Pointer Generator Network formed a powerful tool for text summarization that combined extractive and abstractive techniques to produce high-quality summaries that are faithful to the content of the input document.

2. Google PEGASUS:

Google Pegasus formed one of the text summarization models that was introduced in 2020, as an advanced version of the Transformer architecture. It is used in language processing tasks, which has been fine-tuned for summarization tasks. Pegasus utilizes a training stage that has been pre trained on a huge corpus of text data. Fine-tuning on this dataset is further done to improve its performance on those tasks.[1] [14] One of the key features of Pegasus is its ability to perform abstractive summarization, which involves generating novel sentences that capture the essence of the input text, as opposed to simply selecting and rearranging existing sentences. Pegasus has demonstrated state-of-the-art performance on a variety of summarization datasets, outperforming other popular summarization models like BERTSUM and TextRank. Overall, Pegasus has established itself as a highly effective and versatile text summarization model, with potential applications in a variety of domains, including journalism, research and data analysis.

3. OpenAI GPT-03 (Da Vinci-003)

OpenAI GPT-3 (Da Vinci-003) forms one of the transformer-based neural network based advanced language models for text summarization tasks. It works on around 175 billion parameters and is trained on a huge dataset of web text. This allows it to generate more concise, high-quality summaries that capture the essence of the input document. One of the strengths of GPT-3 for text summarization is its ability to generate abstractive summaries, which involve synthesizing new text rather than just selecting and copying text from the input document. GPT-3 also supports multiple languages, making it a versatile tool for summarizing text in a variety of languages. While GPT-3 is a powerful tool for text summarization, it does have some limitations. Its large size and complexity make it difficult to train and deploy in some settings, and its high computational requirements may make it prohibitively expensive for some use cases. Nonetheless, GPT-3 represents a major advance in the field of natural language processing and has the potential to transform how we summarize and understand text.[2]

6. Statistical Approaches

Statistical approaches in text summarization involve the use of statistical features of sentences, such as the title, location, term frequency, and assigned weights for keywords. These features are used to calculate a score for each sentence, and the highest-scoring sentence is selected to be included in the summary. Below mentioned statistical features find place in extractive text summarization. There are several methods for deciding the importance of a sentence, including:[1]

6.1 The Title Method:

It is based on the premise that mostly the sentences appearing in the document's title are utmost important. There are probable chances of them to be considered as a part of the summary. Based on the number of common words it shares with the title, the score of a sentence is further determined . However, if the document lacks a title, this method may not be effective.

6.2 Location Method:

This method assigns weights to text based on its location in the document. For example, leading sentences, the last few sentences, or the conclusion of a document are assumed to be more important as compared with other sentences in the document. Due to this, there are probable chances of it getting considered as a part of the summary. This method is based on the intuition that headings, bold-formatted text, and text in the beginning or end of the text contain important information.

6.3 tf-idf Method:

The tf-idf Method utilizes the term frequency-inverse document frequency as a numerical statistic to reflect the importance of a word in a document. This weighting factor is generally used for text mining and information retrieval . It is also helpful to filter stop words in text summarization and tagging or categorizing these summaries. Based on the number of occurrences of a word appearing in the document, it is considered that the tf-idf value increases proportionally.

6.4 Cue Word Method:

This method assigns weights to text depending on its importance. Positive weights are assigned to words like ”verified,” ”significant,” ”best,” and ”this paper,” while negative weights are assigned to words like ”hardly” and ”impossible.” Significant sentences can further be determined via the cue words. These words are said to provide a rhetorical context and a source of abstraction between a set of phrases.

.

7. Applications of Automated Text Summarization

This research paper focuses on providing a fair understanding of NLP and ATS research. It aims to contribute to the field by creating new resources, datasets, methods, and tools that cater to the requirements of both research and industry. As NLP gets richer, automated text summarization has become more accessible for usual document summaries and analysis of sentiments. ATS promotes a multidisciplinary approach to research across various fields, including machine learning, natural language processing, cognitive science, and psychology. Another crucial aspect of ATS research is its application, which will be discussed in a subsequent section. Recently, ATS has found wide applications in the domain of information extraction and retrieval, forming questions and answers, textual mining and respective analytics. It also enhances the capabilities of search engines with various applications, including news summaries, email summarization, and domain-specific summarization. The following section presents the applications of ATS.

1. Summarization of Novels or Books: ATS is primarily utilized to condense long texts like books, literary works, or novels into shorter versions, as summarizing brief documents is not effective. It can be difficult to extract meaning from brief texts, and so long documents are considered more suitable for summarization.[10]

2. Summarizing Social Posts and Tweets:Daily: Billions of messages are created on social media platforms like Facebook and Twitter. ATS can be used to obtain useful and crucial text summarization. This information source is valuable with the aid of ATS.[8]

3. Sentiment Analysis (SA): Evaluation of the attitude of people, their emotions, along with their perceptions towards events and circumstances is called Sentiment Analysis. SA categorizes sentiments and further explores the product reviewed by customers as ”Positive” or ”Negative”. These reviews are determined by applying the fuzzy logic technique. Market basket analysis done by the basket analysts, summarize the sentiments or views of people for a specific product.[5]

4. News Summarization:World news obtained from news channels such as CNN, CNBC are further summarized through the ATS model network.It explores the main focus of a news article, which can sometimes serve as its headline.[6] 5. Email Summarization: Emails often lack structure and proper syntax, making them difficult to summarize. ATS typically extracts noun phrases and summarizes email messages through linguistic techniques and machine learning algorithms.[11]

6. Summarizing Legal Documents: Legal questions and discourse functions of earlier cases are explored and utilized to summarize a legal judgment document through the keywords, critical phrase matching, and case-based analysis of the ATS models respectively. [7]

7. Summarizing Biomedical Documents: ATS incorporates graph-based summarization to combine the information obtained from genetic clusters and connections. Genetic clustering determines the theme of any given biological document, whereas its relative importance is represented by its respective connectivity data.[9]

8. Summarizing Scientific Papers: Scientific papers are structured texts that encompass multiple researchers’ perspectives on a given topic. Additionally, the key points in scientific papers are often found in tables and diagrams, rather than generic text. A multi document ATS framework employs two methods to produce a comprehensive review of scientific papers. Firstly, it tracks and collects citations, then employs summarization techniques to analyze the content of the original and cited articles.[7] Some applications that have been implemented to increase the area of utilization of text summarization include:-

1. Image Captioning Summarizer: This module uses computer vision techniques to identify objects, people, and scenes in an image and generates a caption that summarizes the visual content. It can be further enhanced with natural language generation techniques to create more informative and fluent summaries. This module can be useful in applications such as image search engines, social media content analysis, and security monitoring.

2. VideoTranscript Summarizer: This module takes a video as input and generates a transcript by transcribing the spoken words in the video. It can then use the same techniques as those used in traditional text summarization to generate a summary of the video content. This module can be useful in applications such as video content indexing, video search

engines, and video surveillance analysis.

3. Audio to Text Summarizer: This module uses speech recognition software to transcribe audio content into text and then applies text summarization techniques to generate a summary. This module can be useful in applications such as podcast analysis, voice search engines, and voice assistants.

4. Multimodal Summarizer: This module combines information from multiple modalities, such as images, text, and audio, to generate a summary that captures the main content from all modalities. This module can be useful in applications such as multimedia news summarization, multimedia content indexing, and multimedia content search engines.

8. Proposed Methodology

Today's world has an overload of information. Data can be procured from internal sources through flat files, organization databases , operational as well as executive dataset. While external sources comprise data exchanged between the organization, vendor specific data, social media data and much more. At the same time , data can take up any form ranging from textual based, images, videos, alphanumerical . Organizations have hygge amount of textual information in the form of .pdfs, docs, request for proposals, request grants, inventory related documents, legal information etc. It becomes humanly difficult to understand, analyze and comprehend every word written in the text along with its contextual meaning, especially for those who need to go through documents day and night. Text summarization forms one of the solutions to this problem. It allows you to take the time to read something important within less time, understand and get concise information quickly.This, thereby saves time, energy, as well as reduces stress levels. Article summarization is a task that aims to reduce the length of an article while retaining its most important information.The proposed architecture comprises the modules as represented in Fig. 1.The block diagram mainly consists of six sections:

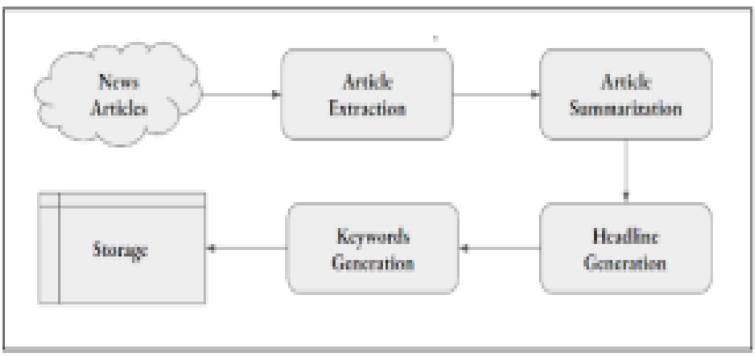


Fig. 1. Block diagram of the proposed system

1. News articles section : In this , all the technology related news articles hosted at websites like TechCrunch, GadgetsNow, EconomicTimes, BBCNews, IndiaToday, are considered..

2. Article Extraction: In this phase, relevant articles are extracted based on RSS by parsing RSS feeds from various sources and using libraries such as 'feedparser' and 'newspaper. The 'feedparser' library helps in parsing the RSS feed, while the 'newspaper' extracts the article from the URL provided in the RSS feed. The extracted articles can then be further processed for text analysis or used for various natural language processing tasks. This process enables automated article extraction and analysis, making it easier to keep up with news and trends in a given field.

3. Article Summarization: This section uses a pre-trained transformer-based model, Pegasus Summarizer,to summarize text from a given article. The library takes in the article as input, applies various techniques like sentence segmentation, tokenization, and attention-based mechanisms to identify important sentences, and generates a summary based on these important sentences. The generated summary is a condensed version of the original article that captures the key points, making it easier to understand and quickly process large amounts of information, Neural networks based model, Long Short-Term Memory

4. Headline Generation: (LSTM) model and natural language processing is used to generate the headline. The model is trained on a large corpus of articles and their corresponding headlines, enabling it to learn to generate high-quality headlines that capture the essence of the article.

5. Keywords Generation: Keyword generation involves calculating the importance of words or phrases within a text corpus based on their frequency and distribution across documents. TF-IDF (Term Frequency-Inverse Document Frequency) is used for the same. Libraries such as 'scikit- learn' and 'nltk' are used to preprocess and tokenize the text data, followed by computing the TF-IDF scores using the 'TfidfVectorizer' class. The top keywords for each document can then be extracted based on the TF-IDFscores, and used for various text analysis tasks.

6. Storage: The Article content, link, author, date, summary and set of keywords are stored in a csv format by giving each summary an unique id. The company's Content Management System can easily retrieve and use these articles in their applications.

The system consists of several steps, each of which con- tributes to the overall effectiveness of the summarization system. The first step involves the extraction of an article using RSS feeds. RSS feeds are a popular way of obtaining news articles from various sources. By extracting articles from RSS feeds, your system can obtain a large amount of content that is relevant to the user's interests. Once an article is extracted, it is passed through a summarization model called PEGASUS. PEGASUS is a state-of-the-art model for text summarization that uses a transformer-based architecture to generate summaries. The model is pre-trained on a large corpus of text and fine-tuned on a summarization task. By using PEGASUS, your system can generate high-quality summaries that capture the most important information from the input article. After summarizing the article, our system generates a headline that summarizes the main point of the article in a concise manner. The headline is usually the first thing that users read, and it plays a crucial role in determining whether users will read the entire article or not. By generating a headline automatically, we can save users time and provide them with a quick overview of the article. In addition to the headline, our system also generates keywords using TF-IDF scores. TF- IDF stands for

Term Frequency-Inverse Document Frequency and is a method for weighting the importance of words in a document. By using TF-IDF, our system can identify the most important words in the article and use them as keywords. These keywords can be used to categorize the article and help users find relevant content. Apart from text-based content, our system also includes modules for summarizing content from images, audio, and video.

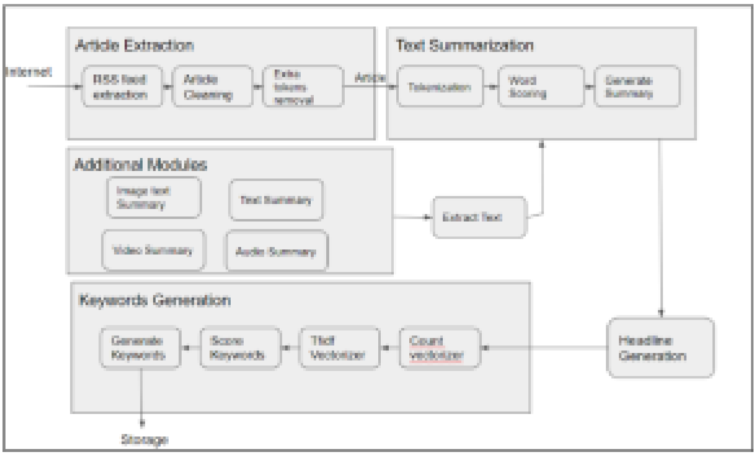


Fig. 2. Modular diagram of the proposed system

For image summarization, we use optical character recognition (OCR) to extract text from images, which is then passed through the summarization model. For audio and video summarization, your system uses automatic speech recognition (ASR) to convert speech to text, which is then summarized using the same model. By adding these modules, we can handle a wide range of content types and provide users with a comprehensive summary of the input article. The steps are very well explained in the Fig. 2

9. Results and Evaluation Measures

Various metrics, such as ROUGE (Recall Oriented Understudy for Gisting Evaluation) and BLEU (Bilingual Evaluation Understudy) are used to quantify the effectiveness of the summarization system. These evaluation metrics are as represented below.

• Recall-Oriented Understudy for Gisting Evaluation (ROUGE) :- It measures the similarity between the generated summary and the reference summary. ROUGE scores measure the overlap between the generated summary and the reference summary, with ROUGE-1, ROUGE-2, and ROUGE-L corresponding to unigrams, bigrams, and longest common subsequences respectively. The Pegasus model achieves high ROUGE scores, indicating that it generates summaries that are highly similar to the reference summaries.

• Bilingual Evaluation Understudy (BLEU):-

It measures the quality of the generated summary in terms of the n-gram overlap with the reference summary. The Pegasus model achieves high BLEU scores as well, indicating that its generated summaries have high precision in terms of n-gram overlap with the reference summaries.

R1/R2/RL	XSum	CNN/DailyMail	Gigaword
BERTShare (Rothe et al., 2019)	38.52/16.12/31.13	39.25/18.09/36.45	38.13/19.81/35.62
MASS (Song et al., 2019)	39.75/17.24/31.95	42.12/19.50/39.01	38.73/19.71/35.96
UniLM (Dong et al., 2019)	-	43.33/20.21/40.51	38.45/19.45/35.75
BART (Lewis et al., 2019)	45.14/22.27/37.25	**44.16**/21.28/40.90	-
T5 (Raffel et al., 2019)	-	43.52/**21.55**/40.69	-
PEGASUS$_{LARGE}$ (C4)	45.20/22.06/36.99	43.90/21.20/40.76	38.75/**19.96/36.14**
PEGASUS$_{LARGE}$ (HugeNews)	**47.21/24.56/39.25**	**44.17/21.47/41.11**	**39.12/19.86/36.24**

Fig. 3. Performance of the PEGASUS model on different benchmark datasets for single-document summarization tasks

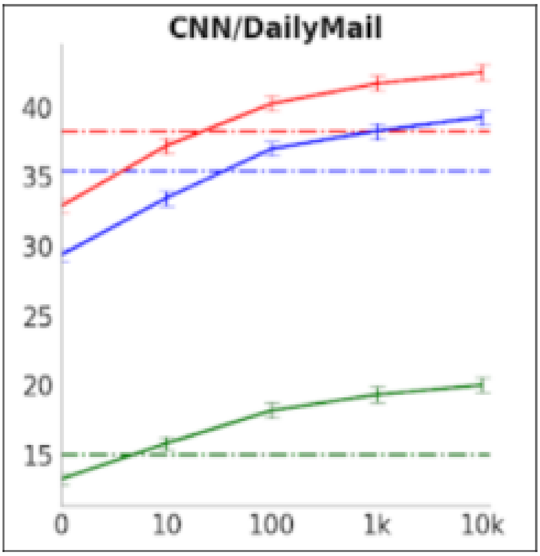
XSum	CNN/DailyMail	NEWSROOM
47.60/24.83/39.64	44.16/21.56/41.30	45.98/34.20/42.18
Multi-News	Gigaword	WikiHow
47.65/18.75/24.95	39.65/20.47/36.76	46.39/22.12/38.41
Reddit TIFU	BIGPATENT	arXiv
27.99/9.81/22.94	52.29/33.08/41.66 ‡	44.21/16.95/25.67
PubMed	AESLC	BillSum
45.97/20.15/28.25	37.68/21.25/36.51	59.67/41.58/47.59

Fig. 4. Performance of the PEGASUS model on different types of summarization tasks, including single-document summarization, multi-document summarization, and summarization with topic constraints

• Coverage score It measures the percentage of input tokens that are covered by the summary. The Pegasus model achieves high coverage scores, indicating that its generated summaries cover a high percentage of the input tokens.

The PEGASUS model was evaluated on different summarization tasks to test its generalization ability and effectiveness. These tasks include single-document summarization,multi-document summarization, and summarization with topic constraints. The model’s performance was also evaluated on out-of-domain data to assess its ability to generate high quality summaries for unseen data sets.The performance of the PEGASUS model on different benchmark dataset is given in fig. 3

The PEGASUS model was evaluated on different summarization tasks to test its generalization ability and effectiveness.These tasks include single-document summarization, multi document summarization, and summarization with topic constraints. The model’s performance was also evaluated on out-of-domain data to assess its ability to generate high quality summaries for unseen datasets given in fig. 4. The CNN/Daily Mail dataset is a popular benchmark for text summarization, and the Pegasus model has achieved state-of-the-art performance on this dataset. Here are some of the performance metrics achieved by the Pegasus model on this dataset. ROUGE-1, ROUGE-2, and ROUGE-L correspond to unigrams, bigrams, and longest common subsequences respectively. Experimental results showed that PEGASUS outperformed other state-of-the-art models on most benchmark datasets, including CNN/Daily Mail news articles, represented by fig 5.



Characteristics	Value obtained
ROUGE-1	44.2
ROUGE-2	21.28
ROUGE-L	41.1
BLEU-1	43.3
BLEU-2	22.1
BLEU-3	13.7
BLEU-4	8.8
Coverage score	64.06

Fig. 5. Performance metrics achieved by the Pegasus model on the CNN/Daily Mail database

We compared the various models we selected for our implementation and the techniques we used to evaluate them: In text summarization, there are two main categories of evaluation methods: intrinsic and extrinsic evaluation. Intrinsic evaluation measures the quality of summaries using only information available in the summary itself. Extrinsic evaluation measures the usefulness of summaries in a specific task, such as information retrieval or question answering. Here are some of the most commonly used evaluation methods for text summarization:

1. Co-selection: Co-selection measures only allow for the use of identical sentences, disregarding the fact that different wording can still convey the same information. It's uncommon for summaries created by different writers to have similar sentences. The precision, recall, and F-measure are used to calculate co-selection.

a. Precision: Precision is the fraction of relevant instances among the retrieved instances.

b. Recall: Recall is the fraction of relevant instances retrieved from a document.

c. F-measure: It is computed by combining recall and precision by taking their harmonic mean.

2. Content-based: Drawbacks of co-selection methods are handled by content-based methods.

Characteristics	Existing Systems	Proposed Systems
Handling multimedia inputs	×	✓
Capturing important information	×	✓
Summarization method	E	E + A
Easy to interpret	×	✓
Coherence	×	✓
Clear Narrative structure	×	✓
Computationally expensive	✓	×
Comprehensive and accurate summaries	×	✓

Fig. 6. Performance metrics achieved by the Pegasus model on the CNN/Daily Mail dataset

a. Cosine Similarity: Cosine Similarity can be measured based on the vector space model.

b. Unit Overlap: Unit Overlap can be calculated based on sets of words or lemmas.

c. Longest Common Subsequence (LCS): the LCS formula is defined based on sequences of words or lemmas.

d. ROUGE: ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is used to evaluate the quality of a summary by comparing it to human-generated ideal summaries.

e. LSA-based method: The LSA-based method involves creating an m x n matrix from m terms and n sentences in the document, given in equation (9). Singular Value Decomposition (SVD) is then applied to the matrix to uncover the document's latent semantic structure.

3. Text Coherence or Quality Evaluation:

a. Grammaticality: The text should be free of non-textual markers, punctuation errors, and incorrect words.

b. Non-redundancy: The text should not contain repetitive information.

c. Referential clarity: Nouns and pronouns should be clearly referred to in the summary.

d. Coherence and structure: The summary should have a well-organized structure and the sentences should flow logically.

5. Automatic Text Summarization Evaluation Programs:

The Document Understanding Conferences (DUC) were held annually from 2001 to 2007 and evaluated different aspects of summarization, including generic summarization of single and multiple documents, query-based summary of multiple documents, topic-based single-document summarization, and multi-document summarization.

Even though training time of Google PEGASUS and GPT-03 are long, pretrained models are available.[12][4].

The proposed system offers a significant improvement over current summarization systems can be seen in fig. 6 by providing a more comprehensive and efficient approach that can handle all types of input data. This can be especially beneficial in fields where multimedia content is common, such as journalism, research, and entertainment.

10. Conclusion

As computer capabilities in natural language processing increase, they will be capable of learning from online in- formation and applying it in real-life situations. With the added ability of natural language generation, machines will be able to both receive and give instructions more effectively. The growth of technology and the widespread use of the internet has resulted in an abundance of information, leading to information overload. To combat this issue, the development of efficient text summarization systems is necessary. One potential answer is to summarize documents using either extractive or abstractive techniques. Extractive summarization is simple to execute, but abstractive summarization is robust as it creates a semantically related summary, though it is more challenging to produce. This paper covers the various summarization methods and their pros and cons.

References

[1] SIGIR ’01: Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, 2001. Association for Computing Machinery.

[2] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeff Wu, Clemens Winter, Christopher Hesse, Mark Chen,Eric Sigler, Ma teusz Litwin, Scott Gray, Benjamin Chess, Jack Clark,Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever,and Dario Amodei. Language models are few-shot learners. ArXiv,

abs/2005.14165, 2020.
[3] Saeedeh Gholamrezazadeh, Mohsen Amini Salehi, and Bahareh Gholamzadeh. A comprehensive survey on text summarization systems.2009 2nd International Conference on Computer Science and its Applications, pages 1–6, 2009.
[4] Yvette Graham. Re-evaluating automatic summarization with BLUE and 192 shades of ROUGE. In Llu ́ıs Marquez, Chris Callison-Burch, and Jian Su, editors, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 128–137, Lisbon,Portugal, September 2015. Association for Computational Linguistics.
[5] Pankaj Gupta, Ritu Tiwari, and Nirmal Robert. Sentiment analysis and text summarization of online reviews: A survey. In 2016 InternationalConference on Communication and Signal Processing (ICCSP), pages 0241–0245, 2016.
[6] Vishal Gupta and Gurpreet Lehal. A survey of text summarization extractive techniques. Journal of Emergin Technologies in Web Intelligence, 2, 08 2010.
[7] Selvani Deepthi Kavila, Vijayasanthi Puli, G. S. V. Prasada Raju, and Rajesh Bandaru. An automatic legal document summarization and search using a hybrid system. In Suresh Chandra Satapathy, Siba K.Udgata, and Bhabendra Narayan Biswal, editors, Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), pages 229–236, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
[8] N. Kumar and M. Reddy. Factual Instance Tweet Summarization and Opinion Analysis of Sport Competition: Proceedings of ICSCSP 2018, Volume 2, pages 153–162. 02 2019.
[9] Hector D. Men ́ endez, Laura Plaza, and David Camacho. Combining ́graph connectivity and genetic clustering to improve biomedical summarization. In 2014 IEEE Congress on Evolutionary Computation (CEC),pages 2740–2747, 2014.
[10] Rada Mihalcea et al, Explorations in automatic book summarization. pgs 380–389, 01 2007.
[11] Smaranda Muresan, Evelyne Tzoukermann, and Judith L. Klavans.Combining linguistic and machine learning techniques for email summarization. In Proceedings of the ACL 2001 Workshop on Computational Natural Language Learning (ConLL), 2001.
[12] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. Bleu a method for automatic evaluation of machine translation. In Pierre Isabelle, Eugene Charniak, and Dekang Lin, editors, Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pages 311–318, Philadelphia, Pennsylvania, USA, July 2002. Association for Computational Linguistics.
[13] Abigail See, Peter J. Liu, and Christopher D. Manning. Get to the point:Summarization with pointer-generator networks. In Regina Barzilayand Min-Yen Kan, editors, Proceedings of the 55th Annual Meetingof the Association for Computational Linguistics (Volume 1: Long Papers), pages 1073–1083, Vancouver, Canada, July 2017. Association for Computational Linguistics.
[14] Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter J. Liu. Pegasus:pre-training with extracted gap-sentences for abstractive summarization.In Proceedings of the 37th International Conference on Machine Learning, ICML'20. JMLR.org, 2020.