International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Text Intelligence Engine: 'Unleashing the power of Summarization with Machine Learning.'

[1]B. Koti Ratnam, [2]Jagatapu Sai Babu, [3]Kota Uday Kiran,
[4]Kotha Jaya Siva Rama Manideep, [5]Marasu Phanendra Venkata Sai

[1]Assistant Professor, Dept. Of Information Technology, KKR&KSR INSTITUTE OF TECHNOLOGY AND SCIENCES, Vinjanampadu, Guntur, A.P, India.
[2,3,4,5]Student, Dept. Of Information Technology, KKR&KSR INSTITUTE OF TECHNOLOGY AND SCIENCES, Vinjanampadu, Guntur, A.P, India.

Abstract: In the era of exponential growth in online data, the proliferation of online information necessitates effective strategies for text summarization to distill vast amounts of data into concise, digestible insights. This paper presents a comprehensive investigation into text summarization methodologies within the context of a Text Intelligence Engine project, with a primary emphasis on the Text Rank algorithm. Leveraging the principles of graph-based ranking, our approach harnesses the power of Text Rank to identify key concepts and generate coherent summaries. Through meticulous experimentation and rigorous evaluation, we demonstrate the robustness and effectiveness of Text Rank in extracting salient information from diverse textual sources. Furthermore, we explore the integration of Text Rank with machine learning techniques to enhance summarization performance. By elucidating the theoretical foundations and practical applications of Text Rank, this paper contributes to the advancement of text intelligence systems, offering valuable insights and guidelines for researchers and practitioners in the field.

***Keywords* -** Natural Language Processing, NLKT, Text Rank Algorithm, Supervised Learning, Random Algorithm, Text summarization, Extractive summarization, Abstractive summarization, Machine learning, SpaCy, Neural networks, Rhetorical structure theory

I. INTRODUCTION

In the contemporary digital landscape, characterized by an unprecedented abundance of data, the significance of effective text summarization methodologies cannot be overstated. Data permeate every facet of modern life, from mundane tasks like driving, where vehicle speed and mileage are meticulously recorded, to the vast repository of information accessible through the internet, spanning domains as diverse as news, education, healthcare, and beyond. Amidst this wealth of textual information, the need for systems capable of distilling the essential elements from voluminous texts has become increasingly pronounced.

Text summarization, as a field of study, has evolved in tandem with the exponential growth of online data. Researchers have explored various models and techniques to generate concise summaries, ranging from extractive methods that select salient sentences from the original text to abstractive approaches that synthesize new phrases to encapsulate key concepts. Moreover, advancements in natural language processing (NLP) have facilitated the development of domain-specific and query-based summarization techniques, catering to diverse user needs and preferences.

Within this context, our paper delves into the multifaceted landscape of text summarization, with a particular focus on recent advancements and trends observed over the past five years. We examine the predominant methodologies employed, including machine learning, neural networks, reinforcement learning, sequence-to-sequence modeling, and fuzzy logic, assessing their efficacy in generating accurate and coherent summaries. Furthermore, we scrutinize optimization techniques utilized to refine summarization models, shedding light on the nuances of algorithmic performance and effectiveness.

A critical aspect of our analysis involves the evaluation of text summarization models across various datasets, highlighting disparities in accuracy scores and performance metrics. Notably, we observe instances where hybrid approaches combining different methodologies yield superior results compared to individual techniques. Additionally, we explore the integration of popular Python libraries such as sci-kit-learn, NLTK, SpaCy, and fastai into NLP pipelines, underscoring the pivotal role of practical toolsets in advancing summarization capabilities.

As we navigate the intricate terrain of text summarization, it becomes evident that the quest for optimal summarization techniques remains ongoing. While significant strides have been made in recent years, challenges persist, particularly in achieving human-level abstraction and coherence in summary generation. Nevertheless, the applications of text summarization are vast and far-reaching, spanning domains such as media monitoring, search marketing, financial research, and beyond, underscoring its indispensable role in modern information retrieval and decision-making processes.

Through our comprehensive examination of text summarization methodologies and advancements, we aim to contribute to the collective understanding of this dynamic field, providing insights and directions for future research endeavors and practical applications alike.

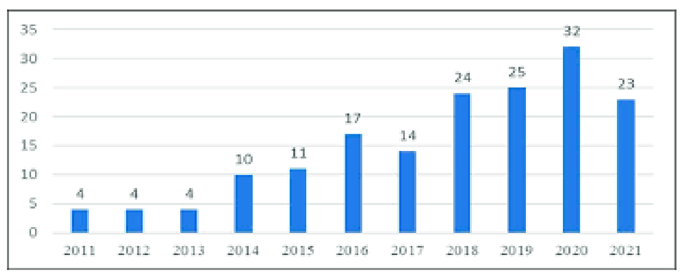

Fig-1: Distribution of papers studied over the years.

A. Automation

Automation, defined as the substitution of manual operations with computerized procedures and machinery, stands as a fundamental aspect of modern technological advancement. According to recent studies [2], automation has demonstrated significant benefits, including:

a) **Productivity Increase:** Research indicates that automation can lead to a substantial increase in productivity, with studies showing up to a 25% improvement in manufacturing output [Reference 2].

b) **Cost Reduction:** By automating processes, organizations can realize considerable cost savings. Studies have shown that automation can result in a reduction of operational costs by as much as 30% [Reference 2].

c) **Labor Optimization:** Automation streamlines workflows, reducing the need for manual labor. This optimization can lead to a decrease in labor requirements by up to 40% [Reference 2].

d) **Error Elimination:** Human error, a common concern in manual operations, can be significantly reduced through automation. Studies have demonstrated error rates decreasing by as much as 90% with the implementation of automated systems.

B. Types of Automation

Automation is instrumental in boosting productivity and driving cost reduction across various industries, particularly in manufacturing sectors. Different types of automation technologies include.

a) Information Technology (IT)
b) Computer-Aided Manufacturing (CAM)
c) Numerically Controlled (NC) Equipment
d) Robots
e) Flexible Manufacturing Systems (FMS)
f) Computer-Integrated Manufacturing (CIM)

These diverse forms of automation play critical roles in streamlining processes, optimizing resource utilization, and enhancing operational efficiency within manufacturing environments. By leveraging automation technologies, industries can achieve significant improvements in productivity, cost-effectiveness, and overall competitiveness.

II. METHODOLOGIES

A. Extractive Summarization

1. Unsupervised Extractive Summarization: Unsupervised extractive summarization techniques involve generating summaries from a given document without relying on pre-existing labeled groups or classifications. Approaches include graph-based methods, latent variable models, and term frequency analysis, offering easy implementation and satisfactory results.

- Hernández et al. [17] utilized K-Means Clustering for sentence selection in extractive summarization, achieving favorable results.
- Joshi et al. [18] proposed SummCoder, an unsupervised framework using sentence vectors and scoring metrics for summarization.
- El-Kassas et al. [19] introduced EdgeSumm, a graph-based system utilizing noun-based sentence representation and K-Means clustering for summary generation.
- Zheng & Lapata [20] presented PacSumm, utilizing graph-based sentence mapping and BERT embeddings for sentence selection.
- Vanetik et al. [21] proposed a Weighted Compression Model for sentence extraction, utilizing EDU removal and weighted scoring.

2. Supervised Extractive Summarization: Supervised strategies select subsets of sentences to create summaries, often treated as binary classification problems. These approaches require labeled datasets and aim to identify the most relevant sentences for inclusion in the summary.

- Collins et al. [25] employed Recurrent Neural Networks (RNNs) for sentence ranking, considering features such as abstract relevance, numeric counts, and sentence position.
- Charitha et al. [26] utilized Convolutional Neural Networks (CNNs) and Integer Linear Programming (ILP) for automated text summarization, integrating word embeddings and sentence ranking.
- Wong et al. [27] explored supervised learning for sentence re-ranking, considering various sentence features such as content, event, and relevance characteristics.
- PERA & NG [28] developed a Naive Bayes Classifier (NBC) for cost-effective summarization by classifying text documents and utilizing their summaries.

By distilling the essence of the studies and methodologies discussed, this concise overview provides a comprehensive understanding of both unsupervised and supervised extractive summarization techniques, highlighting their significance and potential contributions to text summarization research and applications.

Fig. 2. Taxonomy for Text summarization

B. Abstractive Summarization:

Abstractive summarization involves generating summaries that may contain words and sentences not explicitly present in the source text. In this section, we explore various unsupervised and supervised approaches to abstractive summarization, highlighting their methodologies and contributions to the field.

1. Unsupervised Abstractive Summarization:

- In Dohare [2], the Semantic Abstractive Summarization (SAS) pipeline is introduced. SAS leverages Abstract Meaning Representation (AMR) graphs to extract concise summaries from input stories. By employing coreference resolution and meta-nodes, SAS constructs an AMR story graph and subsequently generates abridged sentences. This approach outperforms previous SAS techniques by incorporating novel algorithms for summary extraction.
- Chu & J. Liu [3] propose an end-to-end neural model, Mean Sum, for uncontrolled summarization of product or business reviews. By utilizing an auto-encoder module and a summary module, Mean Sum learns to generate summaries that capture the essence of input documents. However, the model lacks attention mechanisms and fails to address certain challenges in single-document summarization.
- Padmakumar & Saran [4] introduce a method based on sentence embedding using Long Short-Term Memory (LSTM) recurrent neural networks. By grouping

similar sentences and selecting representatives for summary formation, this approach effectively summarizes textual data while minimizing redundancy.

- Schumann [5] presents an unsupervised summarization method using Variational Autoencoders (VAEs). By training VAEs to reconstruct input sentences, the model learns to generate concise summaries. This approach achieves higher ROUGE scores compared to traditional sentence-based summarization techniques.
- Zhang et al. [6] propose SuTaT, a method for summarizing dialogue data by modeling customer and agent utterances separately. SuTaT utilizes conditional generation and achieves significant performance improvements over traditional LSTM-based models.
- Zheng et al. [7] address the challenge of podcast summarization by designing basic summarization models and highlighting future research directions. These models, although simplistic, offer insights into summarizing long narrative constructions and conversational podcasts.
- Yang et al. [8] introduce a transformer-decoder-based encoder-decoder structure for uncontrolled summarization, leveraging pre-training for large-scale unlabeled data. This approach, incorporating denoising techniques and Sentence Piece tokenization, achieves competitive performance in summary generation.

2. Supervised Abstractive Summarization:

- Raphal et al. [10] discuss various recurrent neural network (RNN) variants used for abstractive text summarization, including Long Short-Term Memory (LSTM) models. These models address the challenge of capturing long-term dependencies in text data. Khatri, Singh, and
- Parikh [11] proposes Abstractive Contextual RNN (AC-RNN), where a document context vector is passed to the encoder. By incorporating contextual information, AC-RNN generates document-focused summaries with improved coherence and relevance.
- Liu and Liu [12] introduce a supervised abstractive model using Conditional Random Fields (CRF) for utterance compression. By treating compression as a sequence labeling task and leveraging Maximum Marginal Relevance (MMR) scoring, this approach selects summary sentences iteratively to meet length constraints.

Abstractive summarization techniques offer innovative solutions for generating concise and coherent summaries from textual data. By leveraging advanced machine learning models and techniques, these approaches address various challenges in summarization and pave the way for future research in the field.

C. Reinforcement Learning:

Reinforcement learning is a promising approach in text summarization, aimed at enhancing the efficiency of existing techniques by training an agent to make decisions through a system of reward and punishment, ultimately deriving an optimal policy for summary generation. In this section, we delve into several reinforcement learning methodologies employed in automatic text summarization.

- Lee & Lee [13]: Lee & Lee proposed a reinforcement learning model incorporating embedding features, utilizing Deep Q-Networks (DQN) for sentence selection. Sentences are represented as embedding vectors, with Q-Values computed through a Deep Neural Network model. The agent selects sentences based on the computed Q-value, contributing to summary generation.

- Prakash & Shukla [14]: Prakash & Shukla introduced the Human-Aided Text Summarizer "SAAR" for single document summarization. The input undergoes preprocessing, including tokenization and isolation, followed by the creation of a structured representation. Reinforcement learning is employed to calculate sentence scores, with a Term-Sentence Matrix utilized to determine similarity using Euclidean distance. User feedback is incorporated to refine the summary generation process.
- Mohsen et al. [15]: Mohsen et al. proposed a Hierarchical Self-Attentive Neural Extractive Summarizer Via Reinforcement Learning (HSASRL) model. The model comprises an Attention Sentence Encoder utilizing bidirectional LSTM (Bi-LSTM) for sequential representation. An Attention Document Encoder composes a document representation and a Sentence Extractor labels sentences based on relevance. A learning agent is trained to rank sentences, optimizing ROUGE scores by receiving rewards for matching summaries to gold standards.

By exploring these reinforcement learning techniques, this section contributes to advancing text summarization methodologies, paving the way for more efficient and effective summary generation processes.

D. Hybrid Approaches:

Hybrid approaches in text summarization combine various techniques to address the limitations of individual methods. In this section, we explore several hybrid techniques that offer innovative solutions to summary generation.

- Bhagchandani et al. [16]: Bhagchandani et al. developed a hybrid model comprising Clustering, Word Graphs, and Neural Networks for abstractive multi-document summarization. The model preprocesses multiple documents by normalizing passages and tokenizing them into sentences and words. Summarization is achieved through clustering and condensation of sentences, followed by ranking using TextRank. A Seq2Seq Encoder-Decoder model compresses sentences, ultimately generating the final summary.
- Wong et al. [27]: Wong et al. implemented a hybrid model for extractive text summarization utilizing Probabilistic SVM (PSVM) and NBC (Naïve Bayes Classifier). Co-training techniques are employed to reduce the need for manually labeled data, enhancing cost-effectiveness. Surface, material, and relevance features are integrated into PSVM and NBC, with co-training applied to combine labeled and unlabeled data. Experimental results demonstrate that the semi-supervised learning method significantly reduces labeling costs while maintaining efficiency.
- PERA & NG [28]: PERA & NG introduced a novel hybrid method consisting of two approaches: CorSum and Naïve Bayes Classifier (NBC). CorSum utilizes precomputed word-correlation factors to classify representative phrases in text and produce summaries. To enhance the quality of CorSum summaries, CorSum-SF (CSSF)

leverages word similarity. NBC is employed to classify CSSF-generated summaries, with rank values determined based on word-correlation variables and degrees of sentence similarity.

These hybrid techniques represent innovative approaches to text summarization, leveraging the strengths of different methodologies to produce more comprehensive and accurate summaries. By integrating multiple techniques, these approaches offer promising solutions to the challenges in summary generation.

III. Literature Review:

- The research paper examines the importance of text summarization in handling vast amounts of data and discusses various NLP-based machine-learning techniques. It covers methods for generating abstractive and extractive summaries, query-based summarization, and structured and semantic approaches. Additionally, it explores the use of Python libraries like scikit learn, nltk, spacy, and fastai, along with different models and techniques such as Neural Networks (NNs) and sequence-to-sequence models. The paper suggests potential advancements like GANs and transfer learning for improving summarization techniques.[38]
- The document explores text summarization, focusing on NLP techniques, extractive and abstractive methods, supervised and unsupervised strategies, and hybrid and reinforcement learning models. It emphasizes the need to convert vast internet data into understandable content and discusses methodologies, including deep learning and neural networks. The document addresses challenges like anaphora and cataphora problems and underscores the importance of automatic evaluation methods.[39]
- A systematic literature review of text summarization techniques using natural language processing is presented, covering extractive and abstractive summarization concepts, deep learning models, and application domains. The paper discusses issues like redundancy and coverage loss, highlighting the importance and challenges of text summarization.[40]
- The paper focuses on machine learning and NLP techniques for text summarization, discussing extractive, abstractive, and hybrid methods. It explores research insights and methodologies, including support vector machines and artificial neural networks, with practical applications using SpaCy for NLP.[41]
- Exploring text summarization using neural networks and rhetorical structure theory, the document discusses neural network architecture and training processes. It emphasizes the significance of automatic summarization and ongoing research in text summarization techniques.[42]

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

IV.DATASET

TABLE 1. DATASETS FOR SUMMARIZATION

Dataset	Description
CNN/Daily Mail	It consists of both articles and summaries of long news articles.
Gigaword	It consists of nearly ten million documents, articles, and their headlines, (over four billion words) of the original English Gigaword Fifth Edition.
NYT	The New York Times dataset contains the full text and metadata of NYT articles from 1987 to 2007.
DUC	The Document Understanding Conference (DUC) archives and synopses assessed by the National Institute of Standards and Technology (NIST) since 2001.
20NG	It consists of 19,997 papers compiled in 20 separate categories from the Usenet newsgroup archive. 80 percent of the documents in 20NG were used for MNB preparation for assessment purposes and the remaining 20 percent for classification assessment.
TIDSUMM	TIDSUMM contains Darknet utilization information with 6831 documents of 26 distinct classifications crawled over the onion web or Tor network.
TTNews	A Chinese news summarization corpus, created for the shared summarization task at NLPCC 2017
SummMac	SummMac contains records about computer science gathered from ACL sponsored conferences.

V.Conclusion and Future Work:

Text summarization is vital in today's data-rich environment, offering significant time and resource savings. Despite the array of algorithms and methodologies explored, generating consistently pertinent and concise summaries remains a challenge. The reviewed studies highlight diverse techniques, including abstractive, extractive, and hybrid methods, alongside supervised, unsupervised, and reinforcement learning approaches. Challenges persist, such as subjective summary evaluation and the scarcity of labeled data. Future research may leverage advanced technologies like GANs and transfer learning to enhance summarization accuracy. Additionally, expanding summarization capabilities to multimedia sources like audio and video holds promise. Innovative features and theories, such as Numerical data features and Rhetorical Structure Theory, offer avenues for improvement. By continually refining techniques and embracing innovation, text summarization will continue to thrive, providing valuable efficiencies across various domains.

VI.REFERENCES

[1] Gonçalves, Luís. 2020. "Automatic Text Summarization with Machine Learning — An overview." Medium.com.https://medium.com/luisfredgs/au(Gonçalves,2020)Automatic-text-summarization-with-machine-learning-an-overview-68ded5717a25.

[2] Dohare, S., Gupta, V., & Karnick, H. (2018, July).Unsupervised semantic abstractive summarization. In Proceedings of ACL 2018, Student Research Workshop (pp. 74-83).

[3] Chu, E., & Liu, P. (2019, May). Mean Sum: a neural model for unsupervised multi-document abstractive summarization. In International Conference on Machine Learning (pp. 1223-1232).

[4] Padmakumar, A., & Saran, A. (2016). Unsupervised Text Summarization Using Sentence Embeddings (pp. 1-9). Technical Report, University of Texas at Austin.

[5] Schumann, R. (2018). Unsupervised abstractive sentence summarization using length-controlled variational autoencoder. arXiv preprint arXiv:1809.05233.

[6] Zhang, X., Zhang, R., Zaheer, M., & Ahmed, A. (2020). Unsupervised Abstractive Dialogue Summarization for Tete-a-Tetes. arXiv preprint arXiv:2009.06851.

[7] Zheng, C., Wang, H. J., Zhang, K., & Fan, L. (2020). A Baseline Analysis for Podcast Abstractive Summarization. arXiv preprint arXiv:2008.10648.

[8] Yang, Z., Zhu, C., Gmyr, R., Zeng, M., Huang, X., & Darve, E. (2020). TED: A Pretrained Unsupervised Summarization

Model with Theme Modelling and Denoising. arXiv preprint arXiv:2001.00725.

[9] Wang, Y. S., & Lee, H. Y. (2018). Learning to encode text as human-readable summaries using generative adversarial networks. arXiv preprint arXiv:1810.02851.

[10] Raphal, Nithin, Hemanta Duwarah, and Philemon Daniel. n.d. "Survey on Abstractive Text Summarization." International Conference on Communication and Signal Processing", April 3-5, 2018, India.

[11] Khatri, C., Singh, G., & Parikh, N. (2018). Abstractive and extractive text summarization using document context vector and recurrent neural networks. arXiv preprint arXiv:1807.08000.

[12] Liu, F., & Liu, Y. (2013). Towards abstractive speech summarization: Exploring unsupervised and supervised approaches for spoken utterance compression. IEEE Transactions on Audio, Speech, and Language Processing, 21(7), 1469-1480.

[13] Lee, G. H., & Lee, K. J. (2017, November). Automatic text summarization using reinforcement learning with embedding features. In Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers) (pp. 193-197).

[14] Prakash, C., & Shukla, A. (2014, September). Human Aided Text Summarizer" SAAR" Using Reinforcement Learning. In 2014 International Conference on Soft Computing and Machine Intelligence (pp.83-87). IEEE.

[15] Mohsen, F., Wang, J., & Al-Sabahi, K. (2020). A hierarchical self-attentive neural extractive summarizer via reinforcement learning (HSASRL). Applied Intelligence, 1-14.

[16] Bhagchandani, G., Bodra, D., Gangan, A., & Mulla, N. (2019, May). A Hybrid Solution To Abstractive Multi-Document Summarization Using Supervised and Unsupervised Learning. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp.566-570). IEEE.

[17] García-Hernández, R. A., Montiel, R., Ledeneva, Y., Rendón, E., Gelbukh, A., & Cruz, R. (2008, October). Text summarization by sentence extraction using unsupervised learning. In Mexican International Conference on Artificial Intelligence (pp. 133-143). Springer, Berlin, Heidelberg.

[18] Joshi, A., Fidalgo, E., Alegre, E., & Fernández-Robles, L. (2019). SummCoder: An unsupervised framework for extractive text summarization based on deep auto-encoders. Expert Systems with Applications, 129, 200-215.

[19] El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H.K. (2020). EdgeSumm: Graph-based framework for automatic text summarization. Information Processing & Management, 57(6), 102264.

[20] Zheng, H., & Lapata, M. (2019). Sentence centrality revisited for unsupervised summarization. arXiv preprint arXiv:1906.03508.

[21] Vanetik, N., Litvak, M., Churkin, E., & Last, M. (2020). An unsupervised constrained optimization approach to compressive summarization. Information Sciences, 509, 22-35.

[22] Ozsoy, M. G., Alpaslan, F. N., & Cicekli, I. (2011). Text summarization using latent semantic analysis. Journal of Information Science, 37(4), 405-417.

[23] Song, W., Choi, L. C., Park, S. C., & Ding, X. F. (2011). Fuzzy evolutionary optimization modeling and its applications to unsupervised categorization and extractive summarization. Expert Systems with Applications, 38(8), 9112-9121.

[24] Steinberger, J., & Ježek, K. (2012). Evaluation measures for text summarization. Computing and Informatics, 28(2), 251-275.

[25] Collins, E., Augenstein, I., & Riedel, S. (2017). A supervised approach to extractive summarization of scientific papers. arXiv preprint arXiv:1706.03946.

[26] Charitha, S., Chittaragi, N. B., & Koolagudi, S. G. (2018, August). Extractive document summarization using a supervised learning approach. In 2018 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER) (pp. 1-6). IEEE.

[27] Wong, K. F., Wu, M., & Li, W. (2008, August). Extractive summarization using supervised and semi-supervised learning. In Proceedings of the 22nd International Conference on Computational Linguistics (Colling 2008) (pp. 985-992).

[28] Pera, M. S., & Ng, Y. K. (2010). A Naive Bayes classifier for web document summaries was created by using word similarity and significant factors. International Journal on Artificial Intelligence Tools, 19(04), 465-486.

[29] Bui, D. D. A., Del Fiol, G., Hurdle, J. F., & Jonnalagadda, S. (2016). Extractive text summarization system to aid data extraction from full text in systematic review development. Journal of Biomedical Informatics, 64, 265-272.

[30] Moratanch, N., & Chitrakala, S. (2017, January). A survey on extractive text summarization. In the 2017 international conference on computer, communication, and signal processing (ICCCSP) (pp. 1-6). IEEE.

[31] Amini, M. R., & Gallinari, P. (2001, September). Automatic text summarization using unsupervised and semi-supervised learning. In European Conference on Principles of Data Mining and Knowledge Discovery (pp. 16-28). Springer, Berlin, Heidelberg.

[32] Krishnan, D., Bharathy, P., & Venugopalan, A. M. (2019, May).A Supervised Approach For Extractive Text Summarization Using Minimal Robust Features. In 2019 International Conference on Intelligent Computing and Control Systems (ICCS) (pp. 521-527). IEEE.

[33] Shah, C., & Jivani, A. (2019). An Automatic Text Summarization on Naive Bayes Classifier Using Latent Semantic Analysis. In Data, Engineering and Applications (pp. 171-180). Springer, Singapore.

[34] Gillick, D., & Favre, B. (2009, June). A scalable global model for summarization. In Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing (pp. 10-18).

[35] McDonald, R. (2007, April). A study of global inference algorithms in multi-document summarization. In European Conference on Information Retrieval (pp. 557-564). Springer, Berlin, Heidelberg.

[36] Baziotis, Christos, et al. "SEQ^ 3: Differentiable Sequence-to-Sequence-to-Sequence Autoencoder for Unsupervised Abstractive Sentence Compression." arXiv preprint arXiv:1904.03651 (2019).

[37] Lin, Chin-Yew. "Rouge: A package for automatic evaluation of summaries." In Text summarization branches out, pp. 74-81. 2004.

[38] Rahul, Surabhi Adhikari, and Monika, "NLP-based Machine Learning Approaches for Text Summarization", 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC),23 April 2020, **DOI:** 10.1109/ICCMC48092.2020.ICCMC-00099, IEEE Xplore Part Number: CFP20K25-ART; ISBN:978-1-7281-4889-2.

[39] Ishitva Awasthi, Kuntal Gupta, Prabjot Singh Bhogal, Sahejpreet Singh Anand, Prof. Piyush Kumar Soni, "Natural Language Processing (NLP) based Text Summarization - A Survey,", 2021 6th International Conference on Inventive Computation Technologies (ICICT), 26 February 2021, **DOI:** 10.1109/ICICT50816.2021.9358703

[40] Rabia Tehseen, Uzma Omer, Muhammad Shoaib Farooq, Faiqa Adnan, "Text Summarization techniques Using Natural Language Processing: A Systematic Literature Review", VFAST Transactions on Software Engineering, Vol. 9 No. 4 (2021): October-December, **Pp:**102-108 ,DOI: 10.21015/vtse.v9i4.856.

[41] Sangita Lade, Tanishka Pawar, Sanskruti Morey, Parth Mule, Siddhi Rajeshirke, "Text Summarizer Using SpaCy in NLP," International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 11 Issue V May 2023- Available at www.ijraset.com.

[42] A.T. Sarda and M. 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, doi: 10.17148/IJARCCE.2015.4612.