

Text Summarization Techniques Using Natural Language Processing: A Systematic Literature Review

Rabia Tehseen¹, Uzma Omer^{1,2}, Muhammad Shoaib Farooq^{1*}, Faiqa Adnan¹

¹University of Management and Technology, Lahore, Pakistan

²University of Education, Lahore, Pakistan

*Corresponding author email address: shoaib.farooq@umt.edu.pk

ABSTRACT

In recent years, data has been growing rapidly in almost every domain. Due to this excessiveness of data, there is a need for an automatic text summarizer that summarizes long and numerical data especially textual data without losing its content. Text summarization has been under research for decades and researchers used different summarization methods by using natural language processing and combining various algorithms. This paper presents a systematic literature review by showing a survey of text summarization methods and explains the accuracy of these methods used for text summarization. The paper first introduced some concepts of extractive and abstractive text summarization and also define how deep learning models can be used for the improvement of text summarization. This paper aims to identify the current utilization of text summarization in different application domains. Different methodologies are discussed for text summarization. To carry out this SLR, twenty-four published articles have been chosen carefully for this domain. Moreover, it discusses issues and challenges which are investigated in different application domains using text summarization methods. Lastly, the existing work of different researchers has been carried out for further discussion.

KEYWORDS

Natural language processing (NLP), Text summarization, Extractive Method (EXT), Abstractive Method (ABS), Deep learning.

JOURNAL INFO

HISTORY: Received: October 11, 2021

Accepted: December 15, 2021

Published: December 31, 2021

INTRODUCTION

Text summarization is one of the applications of natural language processing which is not only challenging job but also gaining attention more nowadays. Text summarization is an interaction of making a correct summary of the original document by maintaining its original content. That helps the quick and simple data recovery from the document [1], text summarization that joins the source archive into an improved structure, protecting the general idea and complete information. Without text summarization, people have to read the complete document to get to the main point which is a total wastage of human time [2]. Reading all the documents, articles, and various types of data to physically sum up is amazingly tedious and very tiring for people [3]. Text summarization helps in saving important time for the reader because using a text summarizer people don't have to read the whole text document. In the 1950s, text summarization first started, and from that time forward summarization has been improving. Extraction strategies are replicating the data which is more significant by the system to the summarization [4]. The strategies Exploit is the utilization of natural language processing and factual techniques are utilized for creating summaries. What's more, the old-style moved toward text summarization offered has set up the reason for the order of text summarization procedures. The importance of text summarization is progressively being used in the business areas, in the regions of media transmission, data recovery, data mining, and in word handling with a high probability idea of success [5]. The major focus of this article is to find out how text

summarization can be utilized for numerous application tasks e.g. surveys of motion pictures, similar to email summary, news headline, prints of understudy notes, and sum up data for businessman and government authorities, sum up the medical information for specialists, [6] [7]. The purpose of this research is to present a systematic literature review in the field of text summarization using natural language processing. The structure of this paper is as follows: Section II explains the background and discusses abstractive and extractive summarization, Section III explains research methodology and further discusses research objectives, research questions, search strategy, screening of relevant articles, inclusion and exclusion criteria, and by adopting data extraction strategy, discussed quality assessment score and discussed research questions in detail. In Section IV comparison of different methodologies is discussed. Issues and challenges to text summarization have been explained in section V. Finally, the review has been concluded in section VI.

LITERATURE REVIEW

Text summarization is emerging more in today's world. Nowadays, 21st century is known as the Digital age due to trending technologies in information technology. Automatic text summarization solves the problem by reducing the size of the document preserving its original content to the document [8]. In this systematic review, different techniques and methods are proposed for text summarization based on different types of applications. Some are NLP, supervised ML techniques, NNs, KNN [9].



Different algorithms are used like the word vector embedding, k-nearest neighbour algorithm, human learning algorithm, etc. There is a comparison between different methods on the dataset and described the accuracy of these methods. There are different type of summarization techniques used but mostly used techniques are extractive summarization and abstractive summarization which is discussed in below section.

EXTRACTIVE SUMMARIZATION

Extractive Summarization generates a summary based on important words and sentences from the original document preserving their original content and sentences and words that have the highest ranking are selected and simply generate summary [9]. By using statistical features like title, term frequency, location and assigning scores to the words which are more relevant extractive summarization extract important words and sentences in document [10]. Extractive summarization is so simple to accomplish but sometimes there occurs an error such as miscommunication and ambiguity in text summary as presented in Figure 1.

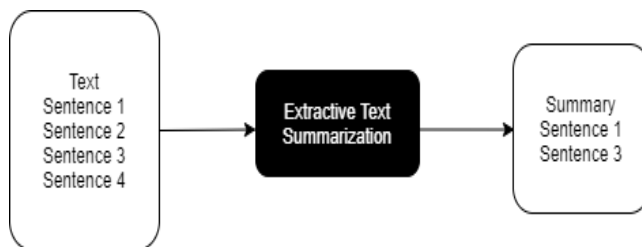


Figure 1. Extractive Summarization

ABSTRACTIVE SUMMARIZATION

This summarization technique first understands the documents and generates the summary of that documents by focusing on the main idea of the document. It is just like that human summarizes the document. Abstractive Text summarization is usually based on a linguistic approach. This approach tries to find the connection between words and focus on the main idea of the document and generate a summary [2]. Usually, this approach is hard to implement. Abstractive text summarization is a simple strategy to achieve that generates an appropriate and correct summary with less uncertainty as shown in Figure 2. It uses advanced heuristic algorithms for summarization [3]. It also reduces the redundancy in summary. It needs data processing which includes: 1) Remove unwanted sentences and words. 2) Tokenize articles into summaries by focusing on the main idea of the original document.

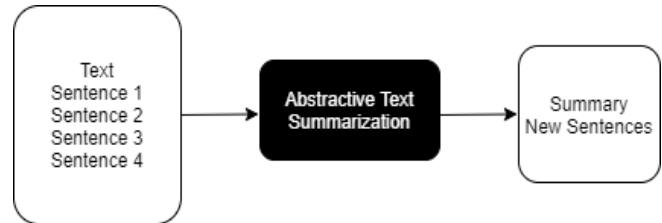


Figure 2. Abstractive Summarization

RESEARCH METHODOLOGY

A systematic literature review (SLR) is chosen as the research methodology for this paper. The main aim of this analysis is to offer the prevailing work of various researchers and build a model that summarizes the text from different application domains. We have followed the methodology proposed by [16] to make our research more impactful in terms of study selection and results. After finalizing the research questions a search protocol has been identified. The research methodology for this systematic review is illustrated in figure 3.

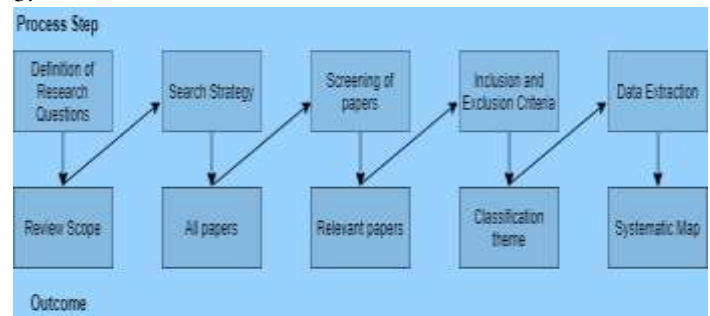


Figure 3. An SLR process model

RESEARCH OBJECTIVES

The main goal of this study is as follows:

- RO1: The major concentration is to identify various summarization techniques in the area of text summarization using natural language processing.
- RO2: Descriptive representation of how various datasets have been used over the years by using several summarization methods.
- RO3: The main focus is to recognize several gaps and challenges faced by different text summarization methods during summarizing text documents.

RESEARCH QUESTIONS (RQS)

To carry out this SLR successfully, the most important research questions have been identified. We have designed different questions which are related to our research that is Text Summarization. This SLR addresses some of the following research questions with corresponding motivation in table 1.

Table 1. RQ and major motivation

| | Research Question | Motivation |
|------------|---|--|
| RQ1 | What are different types of text summarization techniques for various applications domains? | To understand various text summarization techniques in the SLR. |
| RQ2 | What are the major datasets used in text summarization? | In an SLR, we have used several datasets and discussed them in detail. |
| RQ3 | What are the issues and challenges in text summarization? | To identify various research issues and implementation challenges in detail. |
| RQ4 | How deep learning model can be used in text summarization? | To identify how deep learning model can be used for text summarization |

SEARCH STRATEGY

The second phase of SLR is to plan the search strategy and to gather published research articles from our chosen field that is text summarization. We conducted a search based on specific keywords like text summarization techniques using NLP. Internet research has been performed to gather articles from different digital libraries. We gathered different articles from different sources. The articles which we chose for leading our research are collected from the following digital libraries which are: IEEE, Springer, Science Direct, and Google Scholar as presented in Table 2.

table 2. terms and keywords used in search

| Sources | Search String | Context |
|---|--|--------------------|
| IEEE Xplore, Google Scholar, Science Direct, Springer | ("Text Summarization based machine learning OR deep learning") AND ("Automatic Text Summarization") OR "Text Summarization techniques using NLP" | Text Summarization |

STUDY COLLECTION

Not all the papers in our search are relevant so they need to be filtered according to actual relevancy. When the same paper appears in more than one source then it was considered only once according to our search order. In the first stage of screening, papers were nominated based on their titles and omitted those papers that were inappropriate to the research area. For example, our search string sometimes returns those articles which are not relevant to our research area or have a different meaning. In this phase of screening, we read the abstracts of the selected paper during the first phase. We have gathered published articles from different digital libraries for conducting the SLR. About 40% of papers are selected from IEEE, 50% papers are selected from google scholar, and 5% papers are selected from springer and ACM. In this work, different search strategies are selected and different types of papers are chosen from different digital libraries as presented in Figure 4.

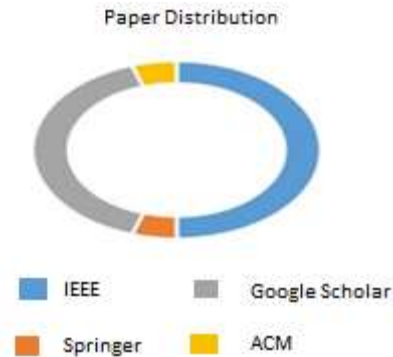


Figure 4. Distribution of relevant papers

INCLUSION CRITERIA

In an SLR, we focus on high-quality papers that define different text summarization approaches using different techniques and methods. The following inclusion criteria are defined:

- A research article on Text Summarization in different languages is found.
- Research articles presenting different methods on different types of datasets.
- Research articles presenting different Text Summarization techniques like ML, NLP, and DL.

EXCLUSION CRITERIA

Some of the papers that appear in more than one source were discarded and only include those papers which are related to our article and search string. Different papers were found which is not related to our search string. The following exclusion criteria are defined:

- Research articles that are not related to Text Summarization.
- Research articles that focus on Text Classification methods and approaches.
- Articles not representing new and emerging ideas
- Some of the papers which are not in the English language are also excluded.

DATA EXTRACTION

The data extraction approach is designed to deliver a set of promising answers to each research question. We have followed the data extraction strategy proposed by [17] to answer our question.

RQ1: The answer to this question is given by discussing different methods of text summarization

RQ2: Discuss the datasets used over the years by using different methods.

RQ3: The defined research question identifies major issues and challenges that occur in text summarization.

QUALITY ASSESSMENT SCORE

As the scoring strategy is defined, each score was given to each selected article presented in Table 3. 40% of articles are considered to be highest-ranking articles, 30% of articles obtain average ranking, and just 10% of articles are considered low ranking articles.

TABLE 3: quality assessment score and classification

| Ref. No. | P. Type | I-Score | E-Score | Total Score |
|----------|---------|---------|---------|-------------|
| [1] | Journal | 7.0 | 2 | 9 |
| [3] | Journal | 5.5 | 2 | 7.5 |
| [7] | Conf. | 8.0 | 2 | 10 |
| [15] | Journal | 7.5 | 2 | 9.5 |
| [16] | Journal | 4.5 | 2 | 6.5 |
| [20] | Journal | 5.5 | 2 | 7.5 |
| [24] | Conf. | 5.0 | 2 | 7 |
| [9] | Conf. | 8.0 | 2 | 10 |
| [12] | Journal | 6.5 | 2 | 8.5 |
| [21] | Conf. | 5.0 | 2 | 7 |
| [17] | Journal | 5.5 | 2 | 7.5 |
| [2] | Conf. | 7.0 | 2 | 9 |
| [19] | Journal | 8.0 | 2 | 10 |
| [6] | Conf. | 6.5 | 2 | 8.5 |

ANALYSIS AND DISCUSSION

We have studied and gathered different published papers for conducting this systematic literature review and among these papers and designed some of the research questions. The explanation of these questions is as follows:

RQ1: What are different types of text summarization techniques for different application domains: Based on various applications several types of summarization methods can be used and classified based on these categories as presented in figure 5. In addition to abstractive and extractive type, there is a numerous type of summaries which is needed according to the type of application. Some of them are discussed as:

- Based on the type of detail summary can be either indicative or informative. An indicative summary gives a quick view of a long document, it only gives the key idea of the original document. Suppose when you buy a novel then at the back end of the novel it delivers the main idea of the novel to the reader. Whereas an informative summary gives a concise view of the original document as it does not provide a quick overview of the original text.
- Base on the type of content summary can be either generic or query-based. A generic summary can be used by any user and it provides only the author's point of view not from the user's perspective. The summary it

provides is at the same level of importance. A query-based summary is a question-answer type, a user has to define the subject of the text in the form of a query and the system extracts only that information which the user has searched. Only specific information can be retrieved according to user interest.

- Based on input document summary can be either a single document summary or a multi-document summary. Single document summary can only accept input text of the single document as it is easy to produce a summary of a single document. The multi-document summary can produce a summary of numerous documents of the same subject as sometimes it is hard to implement.
- Based on the type of language summary can be either a monolingual summary or multilingual summary. A monolingual summary can accept input only with a specific type of language and produce a summary based on that specific type of language as it cannot handle many languages input text and needed to work only with one language. A multilingual summary can accept documents in form of many languages as it is not limited to only one language. It is difficult to implement. [18, 27]

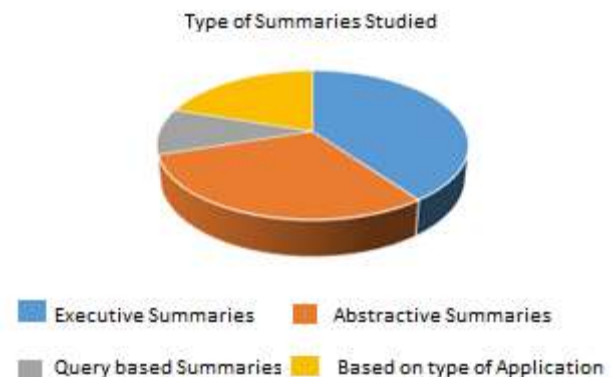


Figure 5. Type of summarization studied

RQ2: What are the major datasets used in text summarization: In an SLR, published papers from 2014 to 2020 are studied, and from these papers, we have seen that different datasets are tested on different models to train by using text summarization approaches. In the past few years, numerous datasets have been made for the analysis of text summarization. Without data, it will be difficult to authenticate them and modify system parameters according to different datasets used. Some of the datasets which are used:

- DUC (the document understanding conference) dataset contains 500 news articles along with their summaries written by human authors. The DUC dataset cannot be used for training models with such large numbers of parameters. So, this dataset can be used only along with other datasets to train different models.

- Another dataset is a collection of 10 million news articles that is Gigaword which is used for summarization. It is considered to be the largest dataset.
- Another corpus is NEWSROOM which comprises 1.3 million news articles and is considered the most recent large-scale dataset introduced in text summarization. This corpus combines extractive and abstractive strategies to generate summaries. The quality of summaries, in this case, is high.
- The CAST corpus is another dataset that is a collection of various documents where each sentence is stated as essential or non-essential. This sort of dataset could help in developing sentence selection and sentence reduction algorithms.
- CNN/Daily Mail dataset is used for training and evaluation purposes in text summarization. It contains bullet points that define the article. Multi sentences summaries are generated by concatenating bullet points[19,20,28].
- There are multiple datasets from different sources like tweets from the user, text documents, etc. these datasets can also be used in text summarization. Different kinds of the dataset used for different purposes in text summarization presented in figure 6.

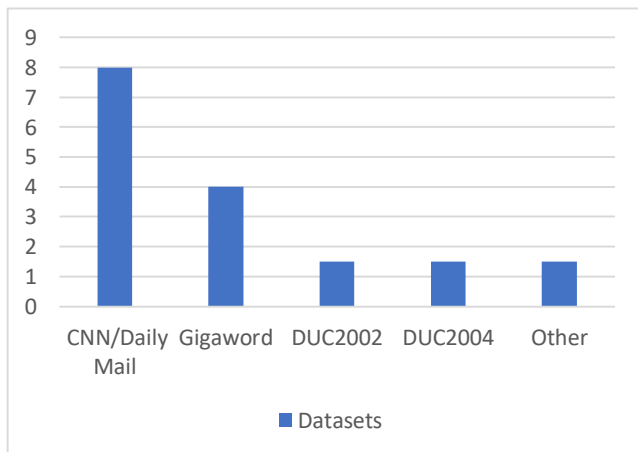


Figure 6. Datasets used over the years

RQ3: What are the issues and challenges in text summarization: Different research issues and implementation challenges occur in text summarization. Some of them are:

- In the case of multi-document summarization, several issues occur during the evaluation of summaries such as redundancy of data, sentence order, or grammatical errors which makes it very difficult to achieve a high-quality summary.
- The quality of summary may vary from system to system or from person to person some may feel that a particular set of sentences may be important for the summary generation and others may feel that other sets of sentences are required for a summary.

- Another problem is irrelevant content in a summary. The core purpose of summarization is to extract meaningful and appropriate content [29].
- Another problem is loss of coverage, in-text summarization it is impossible to cover all the topics for summarization. If we try to cover all possible features then there is a probability of redundancy.
 - To obtain a high-quality summary, quality keywords are required for summary generation [21].

RQ4: How deep learning model can be used for text summarization: Text summarization is normally achieved by the natural language processing techniques by using different algorithms like page rank algorithm [11]. But these algorithms just fulfill the objectives of text summarization they cannot make new sentences or paragraphs which are not in that document just like humans. They can also have grammatical errors. By using deep learning it can generate an efficient and fast model of text summarization. The deep learning model helps us to generate complete summaries of the document which are also grammatically correct [12]. Deep Learning is used with a fuzzy logic algorithm which improved the efficiency of our model. The training phase is used to generate text summarization by given input documents by using a Deep learning algorithm along with fuzzy logic classifier [30].

TRAINING PHASE

The training Phase uses a deep learning algorithm to generate a Text summary. Those features which are extracted from multiple text documents are considered as the important ones for the summarization process [13,31]. In this phase, this approach defines the feature shown in Table 4.

Table 4. Training features

| Assigning scores to words and sentences. | Title similarity. | Similar points. |
|--|----------------------------|---------------------|
| Sentence to Centroid similarity Feature | Number of numerals feature | POS Tagger Feature. |

FEATURE MATRIX PHASE

The above-discussed features for multi-document text summarization are considered to be the most important ones. The text document which is used for summarization is subjected for the feature extraction and a set of features are extracted from the document [14]. Apart from those above features, an additional attribute called a class label for each sentence is also added with the feature matrix. Here, the Class labels along with the class values of each sentence are assigned [15].

COMPARISON OF SUMMARIZATION METHODS

In this section, researchers used several methods and algorithms to train models and to generate text summarization of long articles. The comparison of these methods on a particular dataset as shown by [22], [23], [25] and the papers analyzed by researchers over the years have been presented in Table 5.

TABLE 5. COMPARISON OF METHODS

| Year of Publication | Methods | Dataset | Remarks |
|---------------------|--|--|--|
| 2014 | K-clustering algorithm | Dataset used was from DUC2001 and DUC2002. | The results compared with ROGUE1 and ROGUE2 scores show improvement than another summarizer. The graph-based algorithm was used to improve the result. |
| 2016 | Human Learning Optimization Algorithm. | Text documents as input are used. | This model tried to find the balance between sentence repetition and focus on the main idea and produce a concise summary. |
| 2017 | K-nearest algorithm. | Paragraph as an input used. | Similarity scores are found which can be used in medicine, law. |
| 2017 | Word vector embedding, NN methods used. | 100 news articles from CNN news Corpus with its abstractive summary is used. | This suggested model beats other online text summarizers when compared with ROGUE scores. |
| 2018 | Query-based extractive summarization method using TF-IDF, fuzzy logic. | DUC2007 corpus is used as an input. | A precise 250-word summary is created and also matched with the ROGUE score. |
| 2020 | Extractive text summarization in the Urdu language. | Abstractive Urdu summary corpus. | Most Pakistani people read Urdu language novels on social media. |

ISSUES AND CHALLENGES

In this section, several challenges and limitations have been seen during summarizing the documents which are given as [24-29]

- Redundancy plays a major role in text summarization. As sometimes there is a possibility that one can find redundant elements in summary. A summary is more accurate as long as it contains non-redundant elements. If we measure comparison between elements in a document then there is a chance to minimize redundancy in summary.
- Another problem is irrelevant content in a summary. The core focus of the summarization system is to extract appropriate content from documents to give a quick view. If we consider all the features from the text document to generate a summary then it can lead to irrelevancy. So, it is important to know which features to extract to generate a high-quality summary.
- Another problem is the loss of coverage. It is impossible to cover all topics during summarizing the text document because when we try to cover all topics during summarization then there is a high chance of redundancy and irrelevancy in summary. Coverage of topics is an important feature in generic summarization but it always does not produce a good summary. But this is not compulsory in the case of query-based summarization [30]. The current summarization techniques produced a summary that does not pay much attention to covering all the topics in the document. This problem arises mostly in multi-document summarization when there are many topics to cover than in a single document. Some approaches

produce a summary by covering all topics but there is a high chance of redundancy [31].

A good summary should be understandable and consistent. Consistency and readability mean sentences, in summary, should be related to each other so that it is easy for the reader to understand.

CONCLUSION

In this SLR, several text summarization techniques are introduced as an information-driven semi-regulated way to deal with extractive/abstractive and several other types of summarization methods based on the application domain. The study was conducted using well-defined and organized criteria and selected twenty-four published articles to carry out this SLR effectively. We propose a straightforward methodology for creating a summary without the need for domain information. It has been seen that due to the enormous amount of data in this era text summarization plays a very important role in saving the consumer time. We have discussed various summarization methods and algorithms. By combining these methods and algorithms they produced a good summary.

Text summarization has been demonstrated to be helpful for natural language processing tasks. However, the summary produced by these methods will not always guarantee that the summary is always accurate. Occasionally it is inappropriate to the original document. We have also seen that how some of the existing methods suffer from several kinds of challenges. Furthermore, this issue is ongoing and various people have done their work on text summarization. There is not any particular model that creates an efficient summary. So, in upcoming years there is a chance

that summarization systems discussed will be reformed for a more accurate summary.

CREDIT AUTHOR STATEMENT:

Rabia Tehseen: Conceptualization, Methodology, Software
Uzma Omer: Data curation, Writing- Original draft preparation.
Faiqa Adnan: Visualization, Investigation.
Muhammad Shoaib Farooq: Supervision.
Uzma Omer: Software, Validation,
Rabia Tehseen: Writing- Reviewing and Editing.

COMPLIANCE WITH ETHICAL STANDARDS

It is declared that all authors don't have any conflict of interest. Furthermore, informed consent was obtained from all individual participants included in the study.

REFERENCES

- [1] "Dalwadi, Bijal Patel, Nikita, and Suthar Sanket, 'A Review Paper on Text Summarization for Indian Languages' IJSRD - International Journal for Scientific Research Development, Vol. 5, Issue 07, 2017.
- [2] Chandra Khatri, Sumanvoleti, Sathish Veeraraghavan, Nish Parikh, Atiq Islam, Shifa Mahmood, Neeraj Garg, and Vivek Singh, "Algorithmic Content Generation for Products".
- [3] "Proceedings of IEEE International Conference on Big Data, Santa Clara, pp.2945-2947, CA 2015."
- [4] Deepali K. Gaikwad and C. Namrata Mahender, "A Review Paper on Text Summarization". International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016.
- [5] Batra, S. Chaudhary, K. Bhatt, S. Varshney and S. Verma, "A Review: Abstractive Text Summarization Techniques using NLP," 2020 International Conference on Advances in Computing, Communication Materials (ICACCM), Dehradun, India, 2020, pp. 23-28, doi: 10.1109/ICACCM50413.2020.9213079.
- [6] Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assefi, Saeid Safaei, Elizabeth D. Trippe, Juan B. Gutierrez, and Krys Kochut, "Text Summarization Techniques: A Brief Survey". In Proceedings of arXiv, USA, July 2017.
- [7] Thomaidou, I. Lourentzou, P. Katsivelis-Perakis, and M. Vazirgiannis, "Automated Snippet Generation for Online Advertising", Proceedings of ACM International Conference on Information and Knowledge Management (CIKM'13), San Francisco, pp.1841- 1844, USA, 2013.
- [8] Huong Thanh Le and Tien Manh Le, "An approach to Abstractive Text Summarization", In proceeding of International Conference of Soft Computing and Pattern Recognition (SoCPaR), Hanoi, Vietnam, Dec 2013.
- [9] "Divya, Kasimahanthi, Kambala Sneha, Baisetti Sowmya, and G Sankara Rao."
- [10] "Text Summarization Using Deep Learning' 07, no. 05 (2020): 5."
- [11] Sutskever, Ilya Vinyals, Oriol and Le, Quoc, "Sequence to Sequence Learning with Neural Networks", Advances in Neural Information Processing Systems, 2014.
- [12] "Ramesh Nallapati, Bowen Zhou, Cicero Nogueira dos Santos, Caglar Gulcehre, and Bing Xiang, "Abstractive Text Summarization."
- [13] "The SIGNLL Conference on Computational Natural Language Learning (CoNLL), 2016."
- [14] "Afzal and K. Mehmood, 'Spam filtering of bi-lingual tweets using machine learning,' in 2016 18th International Conference."
- [15] A. Fattah, "A hybrid machine learning model for multi-document summarization," Appl. Intell., vol. 40, no. 4, pp. 592–600, Jun. 2014, doi: 10.1007/s10489-013-0490-0.
- [16] Omer. U.. Farooq. M. S.. & Abid. A. (2021). Introductory programming course: review and future implications. *PeerJ Computer Science*, 7, e647.
- [17] "Ouhbi, Sofia, et al. 'Requirements engineering education: a systematic mapping study.' Requirements Engineering 20.2 (2015): 119-138.
- [18] Munot, Nikita, and Sharvari S. Govilkar. "Comparative study of text summarization methods." *International Journal of Computer Applications* 102.12 (2014).
- [19] Koupaei, Mahnaz, and William Yang Wang. "Wikihow: A large scale text summarization dataset." *arXiv preprint arXiv:1810.09305* (2018).
- [20] Suleiman, Dima, and Arafat Awajan. "Deep learning based abstractive text summarization: Approaches, datasets, evaluation measures, and challenges." *Mathematical Problems in Engineering* 2020 (2020).
- [21] "Jo, 'K nearest neighbor for text summarization using feature similarity,' in 2017 International Conference on Communication."
- [22] Ordóñez, Y. Zhang, and S. L. Johnsson, "Scalable machine learning computing a data summarization matrix with a parallel array DBMS," *Distrib. Parallel Databases*, vol. 37, no. 3, pp. 329–350, Sep. 2019, doi: 10.1007/s10619-018-7229-1.
- [23] Adhikari, Surabhi. "Nlp based machine learning approaches for text summarization." *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2020.
- [24] "Verma, Pradeepika, and Anshul Verma. 'A review on text summarization techniques.' Journal of Scientific Research 64.1 (2020)."
- [25] Tehseen, R., Farooq, M. S., & Abid, A. (2020). Earthquake prediction using expert systems: a systematic mapping study. *Sustainability*, 12(6), 2420.
- [26] Farooq, M. S., Riaz, S., Abid, A., Umer, T., & Zikria, Y. B. (2020). Role of IoT technology in agriculture: A systematic literature review. *Electronics*, 9(2), 319.
- [27] Farooq, M. S., Khan, S. A., Abid, K., Ahmad, F., Naeem, M. A., Shafiq, M., & Abid, A. (2015). Taxonomy and design considerations for comments in programming languages: a quality perspective. *Journal of Quality and Technology Management*, 10(2), 167-182.
- [28] Obaid, I., Farooq, M. S., & Abid, A. (2020). Gamification for recruitment and job training: model, taxonomy, and challenges. *IEEE Access*, 8, 65164-65178.
- [29] Farooq, M. S., Riaz, S., Abid, A., Abid, K., & Naeem, M. A. (2019). A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming. *IEEE Access*, 7, 156237-156271.
- [30] Abid, A., Hussain, N., Abid, K., Ahmad, F., Farooq, M. S., Farooq, U., ... & Sabir, N. (2016). A survey on search results diversification techniques. *Neural Computing and Applications*, 27(5), 1207-1229.
- [31] Naeem, A., Farooq, M. S., Khelifi, A., & Abid, A. (2020). Malignant melanoma classification using deep learning: datasets, performance measurements, challenges and opportunities. *IEEE Access*, 8, 110575-110597.