Extractive Automatic Text Summarization Techniques for Afaan Oromoo – Afroasiatic Language in Ethiopia

RAMATA MOSISSA GICHILA*

Department of Information Science,
Jimma University, Mettu, Mettu, Oromia,
Ethiopia E-mail: mosissaebissa@gmail.com

DANIEL ALEMNEH+

Department of Information Science,
University of Pretoria, Pretoria, South Africa
E-mail: Daniel.Alemneh@up.ac.za

Text summary has become a vital and more popular domain to preserve and highlight the core purpose of textual information as the amount of online information and resource texts has grown. Text summarization is the task of extracting key information from a text document. Text summarizing research in Afaan Oromoo is still rare and hasn't been thoroughly assessed. This study's primary goal was to evaluate the performance and method of extractive models on automatic extractive text summarization for Afaan Oromoo. Automatic Text summarization approaches can be classified as extractive or abstractive. Automatic abstractive text summarization was not included in this study. Existing automatic extractive text summarizing algorithms take key sentences from the source manuscript and provide a summary without changing the data. This paper examined and assessed some studies on the Afaan Oromoo Language Text Summarization system with a focus on methods and performance. In addition, the automatic extractive text summarization domain's challenges in Afaan Oromoo's were also discussed. This paper used a systematic literature review method to examine the most recent literature in automatic extractive text summarization as it relates to the Afaan Oromoo language. We used a search engine, Google Scholar, ResearchGate, CiteseerX, peer-reviewed papers, and Academia to gather papers. After the papers were selected, the performance and automatic text summarization methods were evaluated.

Keywords: *automatic text summarization, extractive summarization, abstractive summarization, afaan oromoo language*

1. Introduction

1.1. Background of Text summarization

In the past, humans were used to summarizing the text in a different language in their own way, but in recent days because of the increasing volume of data, it is difficult for human beings to navigate the huge amount of data. To control this problem, text summarization is needed. Text summarization is extremely active in the field of machine learning and generates clearly expressed forms of large documents keeping most of the original information. It is carried out on single or multiple documents of a similar kind. Text summarization has attracted the attention of researchers from 1950 to today. Luhn was the first researcher of computer science and information science who made an automatic text summarization system. Luhn, (1958) specified the advantage of words depends on their frequency measures. He claimed that words that appear frequently in a text give a good notion of the document's content, even if there are terms that appear frequently but are not content bearing.

Text summarization is the practice of distilling the most significant information from a source (or sources) to produce a condensed version for specific users and tasks (Lin & Chen, 2010). It is a tool that can assist users in sorting through a large amount of data and can be carried out by human experts on a certain topic, but it is time-consuming and costly (Regassa, 2017). In today's fast-growing information era, text summary becomes a crucial and timely instrument for assisting and taking text information (Jezek & Steinberger, 2008).

Condensing the original text into a shorter version while keeping the original message of the content is known as text summarization. Manually summarizing massive amounts of text is extremely challenging for humans (Gupta & Lehal, 2010).

Automatic textual content summarization has been considerably utilized in numerous fields like science, medicine, law, engineering, etc. One example is researchers have targeted producing summaries of doctors' prescriptions, and that has proved very beneficial to patients. Similarly, lengthy information articles have been summarized and readers can consume multiple records on numerous subjects inside a brief span.

Different text summarizations depend on the form of summary we need to be produced from sources. There are two categories of techniques: extractive and abstractive. Extractive summarization is a method that identifies important text and reproduces the summary from original documents. The summary produced from this approach is completely selected from the sentences or sequence of words in the original document (Alguliev & Aliguliyev, 2009) In addition to the complete sentences, extractives can have paragraphs and paraphrases. The problem with the extractive approach is that it produces incoherent summaries from original text or lacks a sense of balance and consistency. In this approach, the extraction of sentences can be out of the context meaning and anaphoric references can also be broken (Rejhan et al., 2009).

When we come to abstractive summarization, it can compose novel sentences which are even unseen in the original sources or original document. This approach is usually built from the existing content but uses advanced methods. However, abstractive approaches require deep natural language processing such as inference; semantic representation, and natural language generation (Alguliev & Aliguliyev, 2009). When the abstraction approach is used for text summarization in deep learning problems, it can solve the grammar inconsistencies of the other approach or extractive method.

Afaan Oromoo is a prominent African language that is widely spoken and utilized in much of Ethiopia, as well as some sections of neighboring Kenya and Somalia (Abera, 1988). It is spoken by the Oromo, Ethiopia's largest ethnic group, accounting for 34.5 percent of the country's population. A few individuals of other ethnicities, who come into contact with the Oromos, like the Omotic-speaking Bambassi and the Nilo-Saharan-speaking Kwama in northwestern Oromia, speak it as a second language. Currently, Afaan Oromoo is the official language of the Oromia Regional State.

1.2. Problem Definition

The history of research work on text summarization started in the 1950s using the extractive system approach (Luhn, 1958). He proposed that words existing many times in a text provide a good idea about the content of the document though there are words that appear very frequently but are not content bearing. As a result, he cut off these words by determining a fixed threshold. The idea of the researcher was recognized and used in many automatic information processing systems. The system uses properties like term filtering and word frequency. Luhn carefully observed that the location of a sentence in a text gives some hint about the importance of the sentence. Thus, he suggested word frequency, cue phrases, title and heading words, and sentence location as extraction features.

Digital information in Afaan Oromoo is growing since Afaan Oromoo became the official language of the Oromia regional state. Currently, there are several newspapers, articles, printed media, and other news releases in Afaan Oromo. There are sources of newspapers like Bariisaa, Kallacha OromiyaA, and radio broadcasts in Afaan Oromoo by Ethiopian Radio, Radio Fana, OBN, and VOA Afaan Oromoo. Also, magazines, judiciary documents, and office reports constitute some portion of the documents produced in the language.

Currently, some study in an automatic artificial text summarization has been conducted for Ethiopian languages, mainly for Amharic text in many areas using various methodologies, and a few contributions have also been made for the Afaan Oromoo language on plagiarism detection, sentiment analysis, disambiguation, and extractive automatic text summarizer using traditional analysis (traditional algorithm). Few studies have compared extractive

and abstractive document summarization (Carenini et al., 2006). Findings indicate that extraction and abstraction worked roughly as well as in other languages. The current effort is a contribution to the development of Ethiopian language natural language processing applications, and it expands the area of text summarizing study by looking into its applicability to the Afaan Oromoo language. Therefore, the objective of this paper is to review the literature on the performance, methods, and challenges of extractive automatic text summarizers for Afaan Oromoo.

1.3. Research questions

In the end, this research project will attempt to answer the following study questions:

What is the performance of automatic extractive text summarization for Afaan Oromo?

What are the challenges in Afaan Oromoo automatic extractive text summarization?

What are the methods used in automatic extractive text summarization?

2. Methodology

A systematic literature review was conducted in automatic text summarization for Afaan Oromoo Language. We used a search engine, Google Scholar, ResearchGate, CiteseerX, peer-reviewed papers, and Academia to gather papers. Only studies published after 2013 were considered in the review of papers on Afaan Oromoo text summarization. After the papers were selected, the performance and automatic text summarization methods were evaluated.

3. Literature Review

3.1. Text Summarization

Text summarization research may be traced back to the 1950s when the first extraction system was created (Luhn,1958). He claimed that words that appear frequently in a text give a good notion of the document's content, even if there are terms that appear frequently but are not content bearing. As a result, they attempted to eliminate these terms by establishing a defined threshold. Many automatic information processing systems recognize and use Luhn's concept. It is a technical article summarizing system with a domain-specific single document summarization system. Term filtering and word frequency are among the features used by the algorithm (low-frequency terms are removed). Sentences are weighted according to the significant terms in them, and sentence segmentation and extraction are conducted. He meticulously defined the human extraction principles, noting that the placement of a sentence in a text provides some insight into its relevance. As an extraction feature, he offered word frequency, cue phrases, title and heading terms, and sentence location. Edmundson's system (Edmundson, 1969), like Luhn's, is a single document that is domain-specific (that deals with technical articles). Different text summarizations depend on the form of summary we need to produce from sources. There are two categories of techniques: extractive and abstractive.

3.1.1. Extractive Text Summarization

The extractive summarization method selects important sentences from the original text document and connects them into a shorter form without changing or altering the main text

(Shivangi & Rachana, 2018). According to Kylmenko et al. (2020), extractive summarization methods create summaries by concatenating numerous sentences (text units) from the text to be summarized in the exact order in which they appear. The fundamental goal of these systems is to figure out which sentences are essential enough to include in the summary. For many years, extractive methods have been the primary focus of text summarization researchers. Many recent approaches treat extractive summarization as a sequence labeling task, with each label indicating whether a sentence should be included or excluded from the summary.

3.1.2. Abstractive Text Summarization

According to Pai (2014) natural language processing, semantic representation and modification, text interpretation and production are all topics covered by abstraction summarization. The goal of the abstractive method is to extract essential sentences as a type of computed document summary. This can be accomplished in several ways. Structured and semantic-based approaches to abstractive text summarization are examples. Structured approaches use a variety of schemas to encode important aspects of documents, such as a tree, ontology, lead and body phrases, and template and rule-based schemas, whereas semantic-based approaches are more concerned with the text's semantics and thus rely on the document's information representation to summarize the text. The multimodal semantic method, information item technique, and semantic graph-based method are all semantic-based methodologies (Suleiman & Awajan, 2020).

3.2. Single Document Summarization

Single document text summarization is a summary from a single source document, using single document text summarization (Shivangi & Rachana, 2018). This sort of text summarization technique takes only one document as input, then employs several techniques to extract relevant sentences from the source document, after which a summary is constructed from the retrieved sentences. Summary generation is more comprehensible, syntactically, or semantically correct, and, most importantly, in a reduced form. To minimize repetitions when summarizing various papers, one must identify and locate topic overlaps. It is also required to decide what to do with the rest, deal with potential document contradictions, and, if necessary, arrange events from diverse sources along a single timeline.

3.3. Multi-Document Summarization

The goal of multi-document text summarizing is to create a summary from multiple source documents. The goal of multi-document text summarizing is to extract useful information from each source document and then create a summary that meets the needs of humans (Modi & Oza, 2019). To create the new summary, extractive text summarization involves selecting terms and sentences from the base text. Procedures entail rating the importance of phrases to select only those that are most relevant to the source's implication (Asawa et al., 2020).

3.4. Extractive Summarization Methods

The goal of extractive summarizers is to pull out the most important information by highlighting the most important sentences in the document while also keeping the summary's repetition to a minimum.

3.4.1. Term Frequency-Inverse Document Frequency (TFIDF) method

According to Zhang et al. (2005), the term frequency-inverse document frequency (TF-

IDF) of scores of words is used to denote documents. In this instance, the term frequency refers to the average number of occurrences (per document) across the cluster. The IDF value is calculated using the complete corpus. The summarizer accepts texts that have previously been clustered as input. Each cluster is referred to as a theme. Words with the highest term frequency and inverse document frequency (TF-IDF) scores in that cluster represent the theme.

The conventional weighted term-frequency and inverse sentence frequency paradigm are used to build the bag-of-words model at the sentence level, where sentence frequency is the number of sentences in the document that contain that term. These sentence vectors are then ranked according to their closeness to the query, with the highest-scoring sentences being selected for inclusion in the summary. This is a straightforward application of the information retrieval paradigm to summary.

Nonstop words that occur most frequently in the document(s) can be used as query terms to build a generic summary. These words provide generic summaries since they describe the document's theme. For sentences, term frequency is usually 0 or 1, because the same

content word does not present frequently in a sentence. If users create query words in the same way they do for information retrieval, query-based summary production will become generic.

3.4.2. Cluster-Based Method

According to Shiva Kumar and Soumya (2015), documents are often constructed in such a way that they address several themes one after the other in a logical order. They are usually divided into sections, either expressly or implicitly. This categorization should be applied to summaries as well, which should address various "themes" that arise in the texts. Some summarizers use clustering to add this feature. Document clustering becomes almost required to build a coherent summary when the document collection for which the summary is being prepared is of completely distinct themes.

The graph-theoretic approach combines related sentences using both sentence ranking and clustering, which are both employed in graph models. Singular nonmatrix factorization is used to cluster sentences in the text. Finally, while clustering and ranking sentences in a document, the weighted graph model technique utilized in this approach examines the discourse relationship between sentences. Hariharan and Srinivasan, (2009) studied a way of summarizing news stories using a Graph-Based approach 2009. An adjacency matrix, which is the cornerstone of Graph-Based methods, is used to represent the measure of similarity between phrases in this method. In this work, two different strategies are studied. The authors' first suggestion was the cumulative sum approach. The second strategy studied was the degree of centrality, which was a previously existing method. A novel way for evaluating the adjacency matrix was proposed in this paper using the aforementioned two techniques, which add two metrics: Effectiveness 1 and Effectiveness 2. These are useful for comparing system summaries to human summaries. Comprehensive studies have shown that this method is superior to traditional methods and that there is still room for advancement in this field of text summarization.

3.4.3. Latent Semantic Analysis Method

According to Ozsoy et al. (2011), Latent Semantic Analysis (LSA) is a statistical-algebraic method for extracting latent semantic structures in words and sentences. It is an unsupervised method that does not necessitate any training or prior knowledge. LSA gathers information from the context of the input document, such as whether words are used together, and which common terms appear in different phrases. The presence of a large number of common terms between sentences implies that they are semantically connected. The meaning of a sentence is determined by the words in it, and the meanings of words are determined by the sentences in which they appear. The interrelationships between sentences and words are discovered using Singular Value Decomposition, an algebraic approach. In addition to being able to model links between words and phrases, single value decomposition can also reduce noise, which helps to increase accuracy.

3.4.4. Machine Learning Approach

According to Sirohi et al. (2021), these methods transform the unsupervised summarizing task into a supervised classifying task that works on the text. The sentence from the inputted document is categorized as "instant" (summary) or "non-instanced (non-summary)" using an algorithm trained from examples and a document with training sets (i.e a set of documents and their numerous summaries generated by humans). The goal of the machine-learning-based summarization method is to score the sentence Query-based

3.4.5. Query-based based extractive text summarization

According to Rajendran and Shravan (2020), the sentences in a given document are graded based on the frequency counts of the query-based text summarization system (words or phrases). The sentences that contain query phrases receive a better score than those that only contain single query terms. The highest-scoring sentences are then included in the output summary, together with their structural context. Text fragments can be taken from various parts or subsections. The resultant summary is the result of combining these

extracts. The number of extracted sentences and the extent to which their context is displayed is determined by the summary frame size, which is set to the maximum screen size that can be viewed without scrolling. When a sentence is chosen for inclusion in the summary by the sentence extraction algorithm, some of the heads in that context are also chosen.

4. Related Work in Afaan Oromoo Language

Dinegde and Tachbelie (2014) tried constructing an Open Oromo Text Summarizer (OOTS). The reference summary was created and used to assess the system output's performance (system summary). An intrinsic technique was used to accomplish the evaluation. They built three approaches for summarizing Afaan Oromoo news material and examined their performance both objectively and subjectively in this study. S1 is a summarizer that uses term frequency and position methods without using the Afaan Oromoo stemmer and other lexicons (synonyms and abbreviations), S2 is a summarizer that uses a combination of term frequency and position methods with the Afaan Oromoo stem specific language language-specific lexicons (synonyms and abbreviations), and S3 is a summarizer that uses improved position method and term frequency as well as (synonyms and abbreviations). Its summarizing mechanism is based on sentence extraction. It was utilized for the phases of processing, sentence ranking, and summary generation. Two evaluation types were used: subjective and objective. For a summary derived using three distinct methodologies, subjective and objective evaluations were undertaken. Both evaluations' findings are consistent; in every example, the S3 summary outperformed the other techniques.

The average in formativeness of the summary is (34.37%, 37.5 %, and 62.5 %) for the three methods (S1, S2, and S3), respectively; the average language quality is (59.37 %, 60 %, and 65 %), and the average coherence and structure are (21.87 %, 28.12 %, and 75 %). The objective assessment result, on the other hand, demonstrates that the average f-measure score for (S1, S2, and S3) is (34 %, 47 %, and 81 %). Dinegde and Tachbelie (2014) concluded that the results of the study's summarizers, like those of previous extraction-based summarizers, lack consistency. Based on the result they recommended abstract summarization more advanced way of avoiding such an issue.

Kannaiya Raja et al. (2019) tried to create Afaan Oromoo Text Summarization by Frequency and Sentence Position Methods. To determine the most significant sentence for extracting a summary, the summarizer combines term frequency and sentence position algorithms with language-specific lexicons. The extraction technique for single news text is the technique proposed for this study. The most significant sentences from the document are extracted and shown to the reader using the extraction process. They aimed to create an algorithm that can summarize a document in the Afaan Oromoo language based on its performance, both objectively and subjectively. The technique proposed for this study was the extraction technique for single news text. Using the extraction technique most important sentences from the document are extracted and displayed to the reader. Term frequency and sentence position algorithm language-specific lexicons were employed in this paper to provide weights to the sentences to be extracted for the summary. An intrinsic method was used to accomplish the evaluation. It included subjective (qualitative) as well as objective (quantitative) evaluation techniques. The four human subjects (expert journalists) are involved in both measurements. The linguistic quality, informativeness, and coherence of the automatically generated summaries were assessed using subjective judgment.

The recall and precision criteria were used to assess its performance. It compares the extracts to the reference summary given an input text, a human's (reference) summary, and a summarizer's extract. Tokenizing, stop-word elimination, stemming, and parsing are all part of the preprocessing steps (breaking the input document into a collection of sentences). Following the formatting and stemming of an input document, the document is broken down into a series of sentences, which are then rated based on two key characteristics. Sentence position and term frequency (TF). This paper missed the finding or its performance and recommendation for the prospects.

Based on the user's inquiry, Jilo et al. (2021), developed a document summary for the Afaan Oromo language. The TF-IDF word weight methodology was employed in the development of the query-based architecture. For morphological analysis, development tools such as HornMorpho are used, whereas, for text processing, Natural Language Processing Toolkit is used. The system has experimented with different extraction rates of 10%, 20%, and 30%. For objective analysis, recall, precision, and f-measure were evaluated; subjective analysis was evaluated by language consultants. At a summary extraction, the proposed system registered f-measures of 90%, 91%, and 93%, according to the results of the evaluation. The performance of the system improved to 91.3% F-measure by applying a morphological analysis tool, even though the additional study is still required to improve the Afaan Oromoo text summary.

Gemechi (2021) attempted to make at Afaan Oromo News Text Summarization Using Sentence Scoring Method. The researcher conducted experiments on ten selected topics from a total of 30 gathered topics using the extractive approach. The natural language tool kit (NLTK) created the system using the Python programming language. The developed approach computes the sentence's score by summing the scores of each word and then computing the sentence's score. The method creates the summary by extracting top-scoring sentences at three different extraction rates: 20%, 30%, and 40%. Automatic summary performed 74%, 78%, and 86% in terms of informativeness at 20%, 30%, and 40% extraction rates. The system performed 86.1 % of the objective evaluation with the three metrics recall, precision, and F-score generated. As a result, a summarizer with better performance in this work. However, evaluating the three metrics of precision, recall, and f-measure during a summary review objectively is difficult. The author recommends new features and future research prospects.

Table 1. Comparison of research performance in Afaan Oromoo

S/N	Author/s	Extractive Method/s	P	R	F
1	(Dinegde & Tachbelie, 2014)	preprocessing, sentence ranking, and summary generation	0.54	0.54	0.54
2	(Kannaiya Raja et al., 2019)	Frequency and Sentence Position Methods	missed	missed	missed
3	Jilo et al. (2021)	TF-IDF term weight Method	missed	missed	0.93
4	(Gemechis,2021)	Sentence scoring Method	0.861	0.861	0.861

5. Challenges and Future Research Directions for Extractive Text Summarization

It is difficult to evaluate summaries (either automatically or manually). The fundamental issue with assessment is the inability to create a standard against which the findings of the systems can be compared. Furthermore, determining a proper summary is difficult due to the possibility that the system will provide a superior summary that differs from any human summary that is used as a rough approximation to correct output. According to Lalithamani et al. (2014), the problem contents elections has not been solved. Because people are so different, subjective authors may choose completely different sentences. Paraphrasing is the process of combining two or more sentences expressed in different languages to represent the same concept. There is a method for evaluating summaries automatically using paraphrases (para Eva).

Extractive summarizing (selecting and replicating lengthy sentences from professional publications) is used by the majority of text summarization s systems. Even though humans may cut and paste important material from a text, they frequently rewrite sentences or combine similar facts into a single statement. The low levels of inter-annotator agreement reported during manual evaluations suggest that the future of this field of research is heavily reliant on the ability to develop efficient methods for automatically evaluating systems.

6. Conclusion

This review has revealed a variety of mechanisms for extractive text summarizing. The

process of extractive summarization is highly coherent, less redundant, and cohesive (summary and information-rich). The goal is to provide a thorough examination and comparison of various methodologies and strategies for extractive text summarization.

Even though summarization research has been going on for a long time, there is still a long way to go. In large-scale applications, simple sentence removal has shown satisfactory results. Some trends in automatic summary system evaluation have been highlighted. However, in the context of time and space difficulty, the work has not fully focused on the various issues of extractive text summarization. Extractive text summarization work carried out in Afaan oromoo language is very limited when we compare it with other languages spoken by more than 40 million similar to Afaan Oromoo.

References

Abera N. (1988). Long vowels in Afan Oromo: A generic approach, Master's thesis, School of graduate studies, Addis Ababa University, Ethiopia.

Alguliev, R., & Aliguliyev, R. (2009). Evolutionary Algorithm For Extractive Text Summarization. *Intelligent Information Management*, *01*(02), 128–138. Https://Doi.Org/10.4236/Iim.2009.12019

Asawa, Y., Balaji, V., & Dey, I. I. (2020). *Modern Multi-Document Text Summarization Techniques*. *1*, 654–670. Https://Doi.Org/10.35940/Ijrte.A1945.059120

Dinegde, G. D., & Tachbelie, M. Y. (2014). Afan Oromo News Text Summarizer. *International Journal Of Computer Applications*, *103*(4), 975–8887.

Edmundson, H. P. (1969). New Methods In Automatic Extracting. *Journal Of The Acm (Jacm)*, *16*(2), 264–285. Https://Doi.Org/10.1145/321510.321519

Gupta, V., & Lehal, G. S. (2010). A Survey Of Text Summarization Extractive Techniques. *Journal Of Emerging Technologies In Web Intelligence*, *2*(3), 258–268. Https://Doi.Org/10.4304/Jetwi.2.3.258-268

Hariharan, S., & Srinivasan, R. (2009). Studies On Graph-Based Approaches For Single and Multi-Document Summarizations. *International Journal Of Computer Theory And Engineering*, *1*(5), 519–526. Https://Doi.Org/10.7763/Ijcte.2009.V1.84

Jezek, K., & Steinberger, J. (2008). Automatic Text Summarization (The State Of The Art 2007 And New Challenges). *Proceedings Of Znalosti*, 1–12.

Jilo, J. A., Alemu, A. T., Abera, F. Z., & Rashid, F. (2021). An Integrated Development Of A Query-Based Document Summarization For Afaan Oromo Using Morphological Analysis. *Indian Journal Of Science And Technology*, *14*(38), 2946–2952. Https://Doi.Org/10.17485/Ijst/V14i38.1182

Kannaiya Raja, N., Bakala, N., & Suresh, S. (2019). Nlp: Text Summarization By Frequency And Sentence Position Methods. *International Journal Of Recent Technology And Engineering*, *8*(3), 3869–3872. Https://Doi.Org/10.35940/Ijrte.C5088.098319

Lalithamani, N., Sukumaran, R., Alagammai, K., Sowmya, K. K., Divyalakshmi, V., & Shanmugapriya, S. (2014). A Mixed-Initiative Approach For Summarizing Discussions Coupled With Sentimental Analysis. *Acm International Conference Proceeding Series*, *10-11-Octo*. Https://Doi.Org/10.1145/2660859.2660910

Lin, S. H., & Chen, B. (2010). A Risk Minimization Framework For Extractive Speech Summarization. *Acl 2010 - 48th Annual Meeting Of The Association For Computational Linguistics, Proceedings Of The Conference*, *April*, 79–87.

Modi, S., & Oza, R. (2019). Review On Abstractive Text Summarization Techniques (Atst) For Single And Multi Documents. *2018 International Conference On Computing, Power And Communication Technologies, Gucon 2018*, 1173–1176. Https://Doi.Org/10.1109/Gucon.2018.8674894

Ndyalivana & Shibeshi, 2019 Automatic Text Summarization Using an Advanced Stemmer Algorithm: A Case Study of the Xhosa Language

Ozsoy, M. G., Alpaslan, F. N., & Cicekli, I. (2011). Text Summarization Using Latent Semantic Analysis. *Journal Of Information Science*, *37*(4), 405–417. Https://Doi.Org/10.1177/0165551511408848

Rajendran, S., & Shravan, V. (2020). Query-Based Text Summarization Using Averaged Query. *International Journal Of Psychosocial Rehabilitation*, *July*. Https://Www.Researchgate.Net/Publication/342638073%0aquery-Based

Regassa, M. G. (2017). *Topic-Based Tigrigna Text Summarization Using Wordnet A Thesis Submitted To The Department Of Computer Addis Ababa University College Of Natural Science*.

Luhn H. P. (1958), "The Automatic Creation of Literature Abstracts", pp. 159-165.

Shiva Kumar, K. M., & Soumya, R. (2015). Text Summarization Using Clustering Technique And Svm Technique. *International Journal Of Applied Engineering Research*, *10*(10), 25511–25519.

Sirohi, N. K., Bansal, D. M., & Rajan, D. S. N. R. (2021). Text Summarization Approaches Using Machine Learning & Lstm. *Revista Gestão Inovação E Tecnologias*, *11*(4), 5010–5026. Https://Doi.Org/10.47059/Revistageintec.V11i4.2526

Suleiman, D., & Awajan, A. (2020). *Deep Learning-Based Abstractive Text Summarization: Approaches, Datasets, Evaluation Measures, And Challenges*. *2020*.

Zhang, Y., Zincir-Heywood, N., & Milios, E. (2005). Narrative Text Classification For Automatic Key Phrase Extraction In Web Document Corpora. *Proceedings Of The International Workshop On Web Information And Data Management*

Wide, 51–58. Https://Doi.Org/10.1145/1097047.1097059