International Journal for Modern Trends in Science and Technology
Volume 10, Issue 02, pages 468-474.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol10issue02.html
DOI: https://doi.org/10.46501/IJMTST1002063

Smart reading companion

Sridevi Dharshini K, Sivaranjani M, Dharshini S, Praveen Kumar Premkamal

Department of Computer Science and Engineering, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu, India.

To Cite this Article
Sridevi Dharshini K, Sivaranjani M, Dharshini S, Praveen Kumar Premkamal, Smart reading companion, International Journal for Modern Trends in Science and Technology, 2024, 10(02), pages. 468-474.https://doi.org/10.46501/IJMTST1002063

Article Info
Received: 28 January 2024; Accepted: 19 February 2024; Published: 25 February 2024.

ABSTRACT

This work presents a comprehensive Smart reading companion aimed at addressing diverse text-related tasks in the digital era. This work amalgamates Optical Character Recognition (OCR), Summarization, Semantic Analysis, Text-to-Audio, and a Question-Answer Module into a cohesive and user-friendly platform. The OCR feature facilitates text extraction from images, while the Summarization module condenses lengthy texts for quick comprehension. Semantic Analysis extracts key concepts, Text-to-Audio enhances accessibility, and the Question-Answer Module fosters interactive querying. This paper serves as a versatile solution, empowering users across various domains with improved productivity, accessibility, and comprehension capabilities.

Keywords: Neural Network, Convolutional Neural Network, Deep Recurrent Neural Networks

1. INTRODUCTION

In today's digital age, where information overload is the norm, the importance of text summarization has never been more evident. Think back to 1958 when researchers first suggested using word frequency as a statistical measure for summarization—a concept that still influences modern methods [2]. As we navigate through a sea of online content, our paper explores the relevance of text summarization. Consider news articles, for instance; who wants to sift through endless pages when a concise summary can provide the essence of a topic? Inspired by the rise of apps curating personalized feeds through article summaries [26][27]. A summary can be understood as a concise rendition crafted from one or multiple source texts, aiming to communicate the crucial information found in the original text(s). Usually, a summary is much shorter than half of the original text(s), often significantly so. It's important to note that in this context, the term "text" is employed in a versatile manner, encompassing written content, spoken words, multimedia documents, hypertext, and other forms of communication.

In general terms, a summary is a brief representation formed from one or more original texts. Its primary objective is to convey the essential information contained in the original text(s), and typically, a summary is noticeably shorter than half of the original text(s). It's essential to recognize that the term "text" is broadly interpreted here and can include written content, spoken language, multimedia documents, hypertext, and various communication forms.

Text summarization has evolved to play a crucial role in addressing user queries directly within search results, a capability increasingly integrated into modern search engine functionalities. With the exponential growth of information sharing and consumption, the relevance of text summarization continues to amplify. Broadly categorized into extractive and abstractive methods, text summarization techniques offer distinct approaches to condensing information. Extractive summarization prioritizes the calculation of sentence weights and selects the top-ranked sentences to form the summary, whereas abstractive summarization involves rewriting sentences to generate a concise summary. These categories, as their names imply, underscore the diverse methodologies employed in text summarization, each offering unique benefits and applications in the context of information retrieval and user interaction[9][10][11].

The extractive method, while effective in summarization, encounters a limitation in meaning preservation, as the inherent connections between sentences are somewhat lost during the selection process. On the other hand, the abstractive method demands substantial effort in training the model, aiming to mitigate potential grammatical and semantic errors that may arise from the rewriting of sentences. Notably, the abstractive approach is language-dependent, requiring tailored training for different languages, whereas the extractive method can be scaled to certain languages more seamlessly due to its reliance on the core idea that remains consistent across languages [12][13].

The advent of sequence-to-sequence models 1[5] [18] has been pivotal in addressing this challenge of generating coherent summaries from textual inputs. These models, augmented with attention mechanisms [7] [17], have proved superior performance, particularly in capturing salient information and reducing redundancy. Furthermore, advancements in techniques such as real-time scene text detection [6], evaluating handwritten text recognition [7], and abstractive text-image summarization [15] highlight the interdisciplinary nature of research in this field.

Moreover, recent studies exploring the use of pointer networks, coverage mechanisms, and reinforcement learning [17] [20] [21] underscore the ongoing quest for innovation and improvement in text summarization techniques. These developments have not only expanded the repertoire of available methodologies but have also led to substantial enhancements in evaluation metrics, signifying tangible progress in the field

Despite these advancements, challenges persist, including exposure bias during model training and the need for robust evaluation metrics[28][29]. Additionally, ensuring the accuracy and coherence of generated summaries is still a focal point for further research [18] [19]. Nonetheless, the collective efforts of researchers in exploring novel approaches and refining existing methodologies [20] hold promise for the continued advancement of text summarization techniques, with implications for diverse applications in information retrieval, document analysis, and beyond.

However, it's worth noting that none of the above reference papers incorporate modules such as a meaning finder, a voice-based assistant for visually impaired individuals, or a custom question-answering module, which are unique features of the proposed Smart Reading Companion. These additional modules enhance the accessibility and functionality of the system, catering to a broader range of users and addressing specific needs not covered by existing research. Addressing these challenges and incorporating innovative features are essential steps in advancing the effectiveness and applicability of text summarization systems across diverse domains and user requirements.

2. LITERATURE REVIEW

The exploration of real-time scene text detection has been a focal point, as demonstrated by Liao et al. (2019) in their work on differentiable binarization. This study delves into the challenges of scene text detection and proposes a methodology that leverages differentiable binarization for improved real-time detection.Michael et al. (2019) have contributed to the field by evaluating sequence-to-sequence models for handwritten text recognition. This research investigates the effectiveness of such models in recognizing handwritten text, offering insights into their performance and potential applications.

Singh et al. (2020) reviews recent approaches for abstractive text summarization using deep learning models. In addition, existing datasets for training and validating these approaches are reviewed, and their features and limitations are presented. The authors classify deep learning-based abstractive text summarization approaches into two main categories. Encoder-decoder models: These models typically consist of two neural networks: an encoder that encodes the input text into a latent representation, and a decoder that generates the summary from the latent representation. Attention-based models: These models use an attention mechanism to allow the decoder to focus on different parts of the input text at different times during the summarization process

Mehta and Majumder's work provides valuable insights into the evolution of summarization techniques, specifically transitioning from extractive to abstractive methods. The paper likely delves into the challenges and advantages associated with each approach, offering a comprehensive exploration of the journey in the field of text summarization. Understanding this transition is critical for comprehending the nuanced developments and improvements in summarization methodologies.Kryscinski et al.'s research, presented at the EMNLP 2020 Conference, focuses on a crucial aspect of abstractive text summarization—factual consistency. The paper likely delves into the challenges of maintaining factual accuracy when employing abstractive methods. Evaluating the factual consistency is a critical dimension in the assessment of abstractive summarization systems, and this work contributes to understanding the strengths and limitations in this key aspect of summarization research.

Chen and Zhuge (2018) have explored abstractive text-image summarization, introducing a multi-modal attentional hierarchical RNN approach. This work addresses the challenge of summarizing information from both textual and visual modalities, providing a nuanced perspective on abstractive summarization.

Agrawal's work (2020) focuses on legal case summarization, offering a specialized application for text summarization. The study investigates techniques for summarizing legal documents, catering to the specific needs of the legal domain.Shi et al. (2020) have delved into neural abstractive text summarization, presenting a model based on sequence-to-sequence architectures. This research sheds light on the significance of such models in summarizing textual information, showcasing their potential in capturing key content.

Suleiman (2019) have explored the use of part-of-speech tagging to enhance Word2Vec models. This study investigates how linguistic information, specifically part-of-speech tags, can contribute to improving the performance of Word2Vec models.Raj's comprehensive study (2022) provides insights into Optical Character Recognition (OCR), covering a broad spectrum of aspects related to OCR. The work encompasses various considerations and advancements in the field of character recognition.

Memon (2020) conducted a systematic literature review on handwritten OCR, offering a consolidated view of the existing body of work in this area. The review provides a comprehensive analysis of the state-of-the-art in handwritten OCR research.Devlin et al. (2019) introduced BERT, a pre-training technique for deep bidirectional transformers designed for language understanding. BERT has since become a foundational model in natural language processing, contributing significantly to advancements in language understanding and related applications.

These diverse studies collectively contribute to the evolving landscape of deep learning applications, text summarization, and information retrieval. Each work brings unique perspectives and methodologies, reflecting the ongoing exploration of innovative techniques and applications in the field. To the best of our knowledge, there is no existing work that comprehensively integrates and extends the capabilities of text summarization along with meaning finder and text to voice functionality, as presented in our proposed project. Our project includes unique functionalities including meaning finder and a custom question answer module.

3. PROPOSED METHODOLOGY

The proposed methodology for the Smart Reading Companion project is distinctive in its focus on addressing critical gaps identified through an insightful examination of existing systems, which has motivated the development of this novel solution. Unlike prior works that primarily concentrated on conventional summarization techniques, our project is spurred by the recognition that these methods often overlook essential elements, hindering comprehensive content comprehension and accessibility.

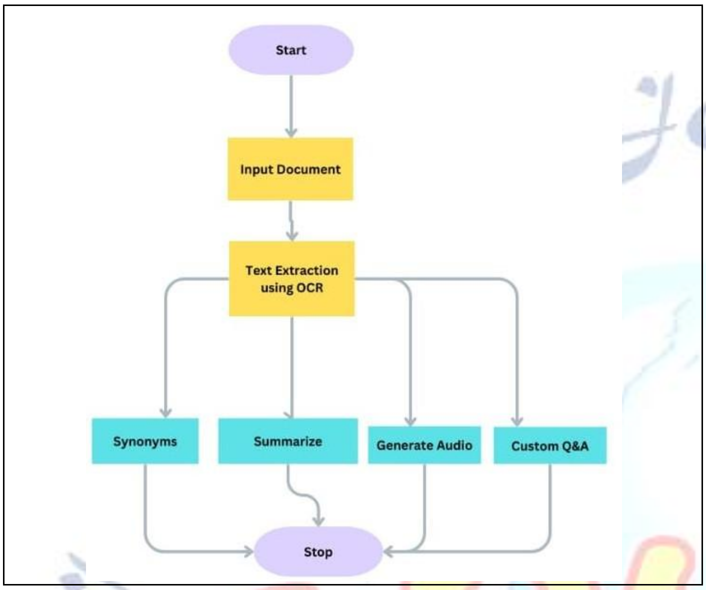

Figure 1: Schematic diagram of the proposed work.

Our proposed methodology for the Smart Reading Companion unfolds in a series of coherent steps, each contributing to the project's unique capabilities. Initially, the system employs Optical Character Recognition (OCR) to extract text from diverse sources, including text, PDF, and Word files. This ensures compatibility with a wide range of content types. Subsequently, the summarization module processes the extracted text to generate concise and informative summaries, providing users with key insights from the content.

Following summarization, the meaning finder module is employed to delve into the nuances of words within the text, enhancing the system's ability to provide in-depth content comprehension. This innovative addition addresses a critical gap in traditional summarization systems by capturing the subtle meanings embedded in the language. Additionally, the custom question-answering module adds an interactive layer, allowing users to pose specific queries and receive tailored responses based on the content.

3.1 OCR

Optical Character Recognition, is a technology that converts text from non-editable formats, such as images, scanned documents, or PDFs, into machine-readable and editable text. It involves recognizing and extracting characters, words, and lines from visual representations of text. In the Smart Reading Companion project, OCR is a crucial module because it allows the system to extract text from diverse input formats like images, PDFs, and text files.

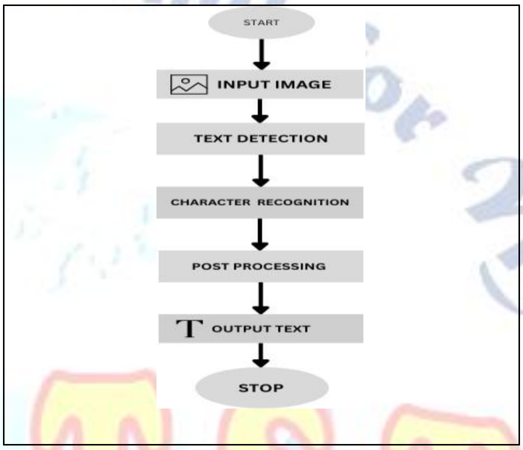

Figure 2: Text extraction using OCR.

This module utilizes the Tesseract OCR library to extract text from an image. It sets the Tesseract command path, opens an image file named '1.png', processes the image using Tesseract's OCR capabilities, and then prints the extracted text to the console. In essence, the code demonstrates a basic OCR operation, showcasing how Tesseract can be employed to convert text content from an image into machine-readable text in a Python environment

3.2 SUMMARIZATION

LexRank is a text summarization algorithm that uses a graph-based approach to identify the most important sentences in a text. The algorithm works by constructing a graph of the sentences in the text, where the edges between the sentences represent the similarity between the sentences. The LexRank score of a sentence is then calculated by taking the weighted sum of the LexRank scores of the sentences that link to it.

```
input  : An array S of n sentences, cosine threshold t
output: An array C of Centroid scores
Hash WordHash;
Array C;
/* compute tf×idf scores for each word */
for i ← 1 to n do
    foreach word w of S[i] do
        WordHash{w}{"tfidf"} = WordHash{w}{"tfidf"} + idf{w};
    end
end
/* construct the centroid of the cluster */
/* by taking the words that are above the threshold*/
foreach word w of WordHash do
    if WordHash{w}{"tfidf"} > t then
        WordHash{w}{"centroid"} = WordHash{w}{"tfidf"};
    end
    else
        WordHash{w}{"centroid"} = 0;
    end
end
/* compute the score for each sentence */
for i ← 1 to n do
    C[i] = 0;
    foreach word w of S[i] do
        C[i] = C[i] + WordHash{w}{"centroid"};
    end
end
return C;
```

Figure 3: Working of LexRank.

To implement LexRank in the educational app, the following steps can be taken:

1. Tokenize the text: The text is tokenized into words.
2. Construct a similarity matrix: A similarity matrix is constructed for the words in the text. The similarity matrix is a square matrix where each element represents the similarity between two words.
3. Construct a sentence graph: A sentence graph is constructed for the text. The nodes in the sentence graph represent the sentences in the text, and the edges between the nodes represent the similarity between the sentences.
4. Calculate the LexRank scores: The LexRank scores of the sentences are calculated using the similarity matrix and the sentence graph.
5. Select the summary sentences: The sentences with the highest LexRank scores are selected to form the summary.

3.3 TEXT-TO-SPEECH

The text-to-speech (TTS) module in the educational app allows users to listen to the text in the app, rather than reading it. This can be helpful for users who are visually impaired, who have difficulty reading, or who simply prefer to listen to content. The TTS module uses the Google Text-to-Speech (GTTS) library to convert text to speech. The GTTS library is a Python library that uses the Google Translate Text-to-Speech API to convert text to speech in over 200 languages .To use the TTS module, users simply need to click the "Listen" button next to any text in the app. The app will then use the GTTS library to generate an audio file of the text and play the audio file back to the user

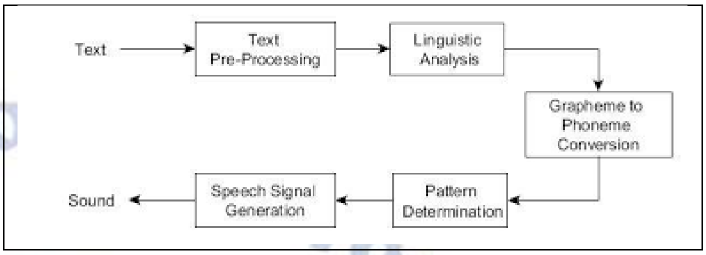

Figure 4: Synthesising text to voice

3.4 WordNet

WordNet is a lexical database for the English language. It groups English words into sets of synonyms called synsets, provides short definitions for each synset, and shows how synsets are related to one another. WordNet can be used for a variety of tasks, including word sense disambiguation, text summarization, and question answering. The educational app can use WordNet to provide users with the meaning of words. When a user taps on a word, the app can display a pop-up window containing the following information:

- The definition of the word
- The synsets for the word

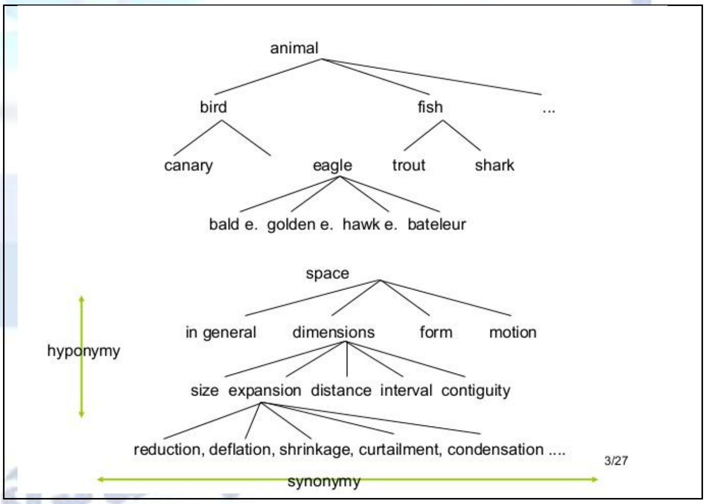

Figure 5: Hierarchy of WordNet

This utilizes WordNet's synsets method to retrieve sets of synonyms (synsets) associated with the input word. These synsets contain not only synonyms but also definitions and examples, offering a comprehensive view of the word's meanings. In essence, the code facilitates the exploration of word meanings through the

WordNet lexical database, providing users with a richer understanding of the semantics associated with their inputted word

3.4 BERT FOR QUESTION ANSWERING.

This feature is designed to generate responses to custom questions based on the content of the input text. This module employs natural language processing and machine learning techniques to comprehend the context of the text and provide relevant answers to user queries. The importance of the question-answering module in this project is multifaceted. Firstly, it promotes user engagement by allowing individuals to interact with the content in a more dynamic and personalized manner, fostering a deeper understanding. Secondly, it serves as a tool for clarification, enabling users to seek specific information or insights from the material. This interactive dimension adds educational value by encouraging critical thinking and exploration of the content.

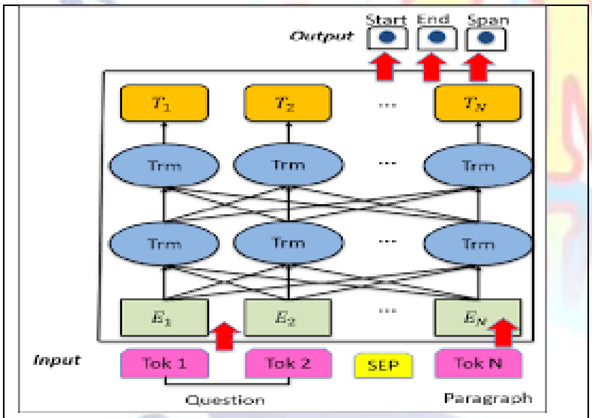

Figure 6: Architecture of BERT model

4. RESULT AND DISCUSSION

Following below are the results procured during the implementation of the work.

A.OCR

```
with open('extracted_text.txt', 'w') as text_file:
    text_file.write(text)

A basic Paragraph Ag Academic

£N4 Englishuk
Corporate Social Responsibility (CSR) can generate a positive reputation for a company leading to
possibly more sales and growth. According to Jones et al (2019), a corporation that invests in the
environmental and ethical approaches of CSR will demonstrate to the public and the media that
they are a responsible company. Watson (2018) provides evidence that this improves consumer
sales as customers tend to support ethical green business practice thus improving profitability and
encouraging growth. For example, a yoghurt company called Yeo Valley has been investing in
making its products organic, creating fully recyclable packaging and reducing its CO2 output. As a
result, profits have doubled within the last two years providing the company with a range of
opportunities to expand (Peterson, 2019). Overall, the evidence seems to suggest that investing in

CSR can improve brand image and productivity.
```

Figure 7: Output of the extracted text

In the figure 7, we showcase the successful extraction process performed by the Smart Reading Companion's OCR module. This crucial functionality allows the system to convert non-editable content from a variety of sources, including images, PDFs, and text files, into machine-readable text. During testing, the OCR module demonstrated impressive accuracy and efficiency in recognizing characters, words, and entire passages, ensuring the faithful representation of the original content.

B.SUMMARIZATION

```
print(textwrap.fill(summary_text, width=80))  # Wrap lines to 80 characters for readability

Summary:
A basic Paragraph Ag Academic
£N4 Englishuk Corporate Social Responsibility (CSR) can generate a positive reputation for a company leading to possibly more sales and growth.
According to Jones et al (2019), a corporation that invests in the environmental and ethical approaches of CSR will demonstrate to the public and the r
A basic Paragraph Ag Academic £N4 Englishuk Corporate Social Responsibility
(CSR) can generate a positive reputation for a company leading to possibly more
sales and growth. According to Jones et al (2019), a corporation that invests in
the environmental and ethical approaches of CSR will demonstrate to the public
and the media that they are a responsible company.
```

Figure 8: Output of the generated summary

In continuation with the figure 8 capturing the essence of the summary generator module in our project documentation, it becomes imperative to furnish comprehensive content that outlines the functionality and paramount importance of this feature. Serving as a pivotal component within the Smart Reading Companion, the summary generator module plays a crucial role in distilling voluminous textual information into succinct summaries. This feature not only facilitates efficient information retrieval but also aligns with the overarching goal of the project to enhance content comprehension and accessibility.

C.TEXT TO VOICE

```
Audio("output.mp3")
```

Figure 9: Output of the generated voice

Following the figure 9 of the text-to-speech (TTS) module screenshot in our documentation, it becomes evident that this feature is a pivotal component of the proposed work. The screenshot illustrates the user interface, showcasing the functionality that transforms written text into natural-sounding speech.

D.CUSTOM Q/A MODULE

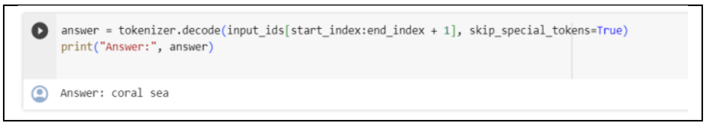

Figure 10: Output of the generated ANSWER

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] A Comprehensive Study of Optical Character Recognition - Ravi Raj, 29 July 2022.

[2] Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. Journal of emerging technologies in web intelligence, 1(1), 60-76.

[3] Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR) - JAMSHED MEMON ,MAIRASAMI ,RIZWAN AHMEDKHAN AND MUEENUDDIN , July 16, 2020.

[4] Utilizing Deep Learning Techniques for Text and Image Capturing Summarization in Information Retrievals-Dr. S. Selvakani, Mrs. K. Vasumathi, S. Divya, 13-march,2023.

[5] Summarization of Text and Image Captioning in Information Retrieval Using Deep Learning Techniques - P. MAHALAKSHMI AND N. SABIYATH FATIMA, February 9, 2022.

[6] Real-time Scene Text Detection with Differentiable Binarization - Minghui Liao, Zhaoyi Wan, Cong Yao, Kai Chen and Xiang Bai, 3 Dec 2019.Sequence-to-Sequence Models for Handwritten Text Recognition - Johannes Michael, Roger Labahn Tobias Gr¨ uning and Jochen Z¨ ollner , 18 Mar 2019.

[7] Abstractive Text-Image Summarization Using Multi-Modal Attentional Hierarchical RNN - Jingqiang Chen and Hai Zhuge, November 2018.

[8] Utilizing Deep Learning Techniques for Text and Image Capturing Summarization in Information Retrievals-Dr. S. Selvakani, Mrs. K. Vasumathi, S. Divya, 13-march,2023.

[9] Abstractive Summarization System-A Kumar, MK

[10] From extractive to abstractive summarization: a journey-P Mehta, P Majumder – 2019

[11] Kryscinski, W., McCann, B., Xiong, C., & Socher, R. (2020, November). Evaluating the factual consistency of abstractive text summarization. In Proceedings of the 2020 Conference on empirical Methods in Natural Language Processing (EMNLP) (pp. 9332-9346).

[12] Song, S., Huang, H., & Ruan, T. (2019). Abstractive text summarization using LSTM-CNN based deep learning. Multimedia Tools and Applications, 78(1), 857-875.

[13] Legal Case Summarization: An Application for Text Summarization- Kanika Agrawal, 01 June 2020

[14] Neural Abstractive Text Summarization with Sequence-to-Sequence Models - TIAN SHI, YASER KENESHLOO, NARENRAMAKRISHNAN and CHANDANK.REDDY, 18 Sep 2020. Using Part of Speech Tagging for Improving Word2vec Model - Dima Suleiman, Arafat A and Awajan , 11 October 2019.

[15] Transformer-based Model for Single Documents Neural Summarization - ElozinoEgonmwan and Yllias Chali, November 2019.

[16] Text Summarization Method Based on Double Attention Pointer Network - ZHIXIN LI,ZHI PENG, SUQIN TANG, CANLONG ZHANG AND HUIFANG MA , January 17, 2020.

[17] Improving Transformer with Sequential Context Representations for Abstractive Text Summarization - Tian Cai, Mengjun Shen, Huailiang Peng, Lei Jiang & Qiong Dai , 30 September 2019

[18] LexRank: graph-based lexical centrality as salience in text summarization - Günes Erkan and Dragomir R. Radev , 01 December 2004

[19] Using Part of Speech Tagging for Improving Word2vec Model - Dima Suleiman, Arafat A and Awajan , 11 October 2019

[20] Leverage Label and Word Embedding for Semantic Sparse Web Service Discovery – ChengaiSun,LiangyuLv,GangTian,Qibo Wang, Xiaoning Zhang and Lantian Guo , 24 Mar 2020

[21] BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding -Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova, June 7, 2019.

[22] A deep learning approaches in text-to-speech system: a systematic review and recent research perspective - Yogesh Kumar, Apeksha Koul & Chamkaur Singh , 29 September 2022.

[23] BERT-based Conversational Question Answering in Context - Munazza Zaib, Dai Hoang Tran, Subhash Sagar, Adnan Mahmood, Wei E. Zhang & Quan Z. Sheng , 07 February 2021.

[24] A deep learning approaches in text-to-speech system: a systematic review and recent research perspective - Yogesh Kumar, Apeksha Koul & Chamkaur Singh , 29 September 2022.

[25] Automatic text summarization-J Manuel, T Moreno,2019

[26] Text Summarization Techniques and ApplicationsVirender Dehru1, Pradeep Kumar Tiwari1, Gaurav Aggarwal1, Bhavya Joshi1 and Pawan Kartik1,2022

[27] A survey on evaluation of summarization methods-L Ermakova, JV Cossu, J Mothe,2021

[28] SUPERT: Towards new frontiers in unsupervised evaluation metrics for multi-document summarization-Y Gao, W Zhao, S Eger,2019