SUMMARIZING TEXT CREATIVELY: AN ABSTRACTIVE APPROACH

BY

Noyon Chandra Saha
ID: 201-15-13811

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mr. Fourcan Karim Mazumder
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised By

Sharun Akter Khushbu
Lecturer (Sr. Scale)
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

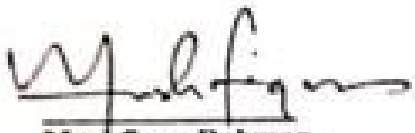
July 2024

APPROVAL

This Project titled "Summarizing Text Creatively: An Abstractive Approach", submitted by Noyon Chandra Saha to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 15 July 2024.

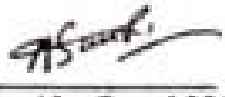
BOARD OF EXAMINERS

Narayan Ranjan Chakraborty — **Board Chairman**
Associate Professor & Associate Head
Department of CSE
Faculty of Science & Information Technology
Daffodil International University



Mushfiqur Rahman — **Internal Examiner 1**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Israt Jahan — **Internal Examiner 2**
Sr. Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University



Dr. Abu Sayed Md. Mostafizur Rahaman — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

I hereby declare that this project has been done by me under the supervision of Mr. Fourcan Karim Mazumder, Assistant Professor, Department of Computer Science and Engineering, Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

13/07/2024

Mr. Fourcan Karim Mazumder
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

13/07/24

Sharun Akter Khushbu
Lecturer (Sr. Scale)
Department of CSE
Daffodil International University

Submitted by:

13/07/24

Noyon Chandra Saha
ID: 201-15-13811
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty for His divine blessing making it possible for me to complete the final year project successfully.

I am grateful and wish my profound indebtedness to **Mr. Fourcan Karim Mazumder, Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of Natural Language Processing (NLP) to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to the **Dr. Sheak Rashed Haider Noori, Professor and Head**, Department of CSE Daffodil International University, Dhaka, for his kind help in finishing my project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents

ABSTRACT

This research investigates recent advancements 2018 to 2024 in abstractive text summarization. I explore the current state of the art and identify key challenges. This study proposes a methodology utilizing fine-tuned transformer models like T5, BART, PEGASUS on a diverse dataset CNN/Daily Mail, Giga Word XSum, TIFU, and SAMSum with consideration for limited computational resources. I aim to develop a model that captures the salient points of the source text while generating concise and human-readable summaries. The effectiveness of the model will be evaluated using ROUGE score alongside other metrics like BLEU score. Finally, the research will explore strategies to mitigate bias and ensure data privacy in the text summarization process. The ultimate goal is to deploy the model in a user-friendly website, making this technology accessible for real-world applications.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

NLP	Natural Language Processing
BART	Bidirectional and Auto-Regressive Transformers
GPT	Generative Pre-trained Transformer
T5	Text-to-Text Transfer Transformer
BERT	Bidirectional Encoder Representations from Transformers
PEGASUS	Pre-training with Extracted Gap-sentences for Abstractive Summarization Sequence-to-sequence
ROUGE	Recall-Oriented Understudy for Gisting Evaluation
BLEU	BiLingual Evaluation Understudy
TF-IDF	Term Frequency - Inverse Document Frequency
NLTK	Natural Language Toolkit

CHAPTER 1
Introduction

1.1 Overview

The digital age has seen an exponential increase in the amount of textual information available online, creating efficient and effective methods for condensing large amounts of text into increasingly necessary short summaries. Text summarization, the process of automatically creating an abridged version of a text while preserving its essential information, meets this need [1]. There are two primary methods of text summarization first one is extractive text summarization and second one is abstract text summarization. Extractive summary involves selecting key sentences or phrases from the main text and combining them to create a summary. Although this method is relatively straightforward and often produces grammatically correct summaries, it cannot always capture the underlying meaning or provide coherent and fluid descriptions [2]. In contrast, the goal of abstract summarization is to create summaries that express the main ideas of the original text using new, synthesized sentences. This method mimics the way people summarize information, which often leads to a more natural and readable summary. However, abstract summarization is significantly more challenging due to the complexity of creating natural language, including maintaining grammatical accuracy, ensuring coherence, and avoiding loss of information [3]. This research focuses on exploring creative and effective methods for abstract text summarization. By leveraging advanced Deep learning techniques and pre-trained language models, I aim to create a system that can produce high-quality, human-like abstractions [4]. This study not only contributes to the field of NLP but also has practical applications in various domains such as information retrieval, content management and accessibility. The following sections of this report will explore in depth the background and current status of text summarization research, define the specific issues addressed by this study, outline the objectives of my research, and describe the methodology and implementation of my proposed approach.

1.2 Background and Present State

Text summarization, as a research field, has evolved significantly in the last few decades. Initially, the focus was on extraction methods due to their simplicity and availability of heuristic-based techniques. The primary methods of extractive

summarization rely on statistical and machine learning methods, such as word frequency-inverse document frequency (TF-IDF) to identify and select key sentences from text. The advent of neural networks and deep learning has dramatically transformed the landscape of text summarization [5]. The introduction of the Sequence-to-Sequence (Seq2Seq) model, which was originally developed for machine translation, marked a significant milestone. These models treat the summarization of text as a generative task, allowing the system to create new sentences that capture the essence of the original text [6] Recent advances in NLP have been driven by the development of pre-trained language models such as BERT and BART. These models have demonstrated remarkable capabilities in understanding and creating human language, resulting in significant improvements in abstract summarization tasks. BERT, introduced to obtained a two-way training of transformers to gain a deeper understanding of the context, making it highly effective for various NLP tasks, including summarization. However, BERT is primarily designed for extractive briefs. On the other hand, models such as GPT and BART have been developed with generative tasks in mind. GPT-3, for example, has shown a remarkable ability to generate coherent and contextually relevant text, even though it requires significant computational resources [7]. Despite these advances, abstract summarization is a challenging task due to factors such as maintaining coherence, avoiding redundancy, and ensuring grammatical correctness. Current research focuses on addressing these challenges in a variety of ways, including reinforcement learning, advanced training techniques, and hybrid approaches that combine extraction and abstraction approaches. The current state of text summarization research is characterized by rapid progress towards more sophisticated and accurate models. The integration of large-scale pre-trained language models has opened the door to new possibilities, making it possible to create high-quality summaries that closely resemble human-written summaries. However, more research is needed to address the remaining challenges and make these models more accessible and efficient for practical applications.

1.3 Problem Statement

Despite significant advances in natural language processing and the development of state-of-the-art models for text summarization, several challenges remain in the field of abstract summarization. The ability to create coherent, concise, and contextually accurate summaries remains a complex task due to a variety of inherent difficulties.

First, maintaining coherence and grammatical correctness in the generated summary is a considerable challenge. Abstract summarization requires the model to deeply understand the context and meaning of the original text and then create new sentences that effectively convey the same information. It involves the complex process of sentence reconstruction, paraphrasing, and content creation, which can often lead to problems such as incomplete sentences, grammatical errors, and loss of critical information [8]. Secondly, avoiding redundancy while ensuring comprehensive coverage of the original text is another significant problem. Abstract models must balance conciseness with informativeness, ensuring that all essential points are included without repetition. Current models sometimes struggle to create summaries that are either too wordy or omit important details that affect the overall quality and usefulness of the summary [9]. Third, there are challenges of scalability and computational efficiency. Advanced models such as T5, BART and PEGASUS require considerable computational resources for training and estimation, which can limit their accessibility and practicality for widespread use. This creates a barrier for small companies or individual researchers who may not have access to high-performance computing infrastructure. Therefore, evaluating the quality of abstract abstracts creates its own set of challenges. Traditional assessment metrics such as the ROUGE score primarily focus on the overlap of words and phrases between the generated summary and the reference summary, which may not fully capture the semantic accuracy and readability of the summary. The need for a more robust assessment approach considering the nuances of language and meaning. In light of these challenges, this study aims to explore innovative solutions to improve the coherence, informativeness, and efficiency of abstract summarization models. By using state-of-the-art techniques and addressing existing limitations, this study seeks to advance the field of text summarization and contribute to the development of more effective and practical summarization systems.

1.4 Objectives

The primary objective of this research is to advance the field of abstract text summarization by developing innovative methods and models that can produce coherent, concise, and contextually accurate summaries. The specific objectives of this study are utilizing state-of-the-art pre-trained language models such asT5, BART, and PEGASUS to create a powerful and efficient abstract summarization system. Fine-tune the model into a diverse dataset to ensure the ability to handle a wide variety of text,

including news articles, technical documents, and social media posts. Apply advanced techniques to improve the coherence and grammatical correctness of the generated summary. Develop strategies for balancing conciseness with informativeness, ensuring that all essential points from the original text are included in the summary without redundancy. Conduct a thorough evaluation to compare the performance of the proposed model with existing state-of-the-art methods. Finaly deploy the best model into a website so that anyone can use my works advantage.

1.5 Scope and Limitations

Scope: This research is focused on developing and improving methods for abstract text summarization. The primary goal is to create summaries that not only compress the information but also restate it in a coherent and contextually accurate manner. The study will support advanced pre-trained language models, like BART, T5, PEGASUS etc. for the development of summary methods. The model will be fine-tuned using different datasets to ensure its adaptability to different types of text. A comprehensive evaluation will be carried out to evaluate the performance of the proposed model. Traditional metrics in assessment such as ROUGE score and newly developed metrics will focus on semantic accuracy and readability. Comparative analysis with existing state-of-the-art models will also be carried out. The developed summarization model will be applied to real-world scenarios, such as summarizing news articles, dialog, and customer reviews. Additionally, a user-friendly web interface will be developed for practical use.
Limitation: Advanced pre-trained language models such as BART, T5, PEGASUS require significant computational resources for training and fine-tuning. Although efforts will be made to optimize the model for efficiency, the research may be limited by the availability of high-performance computing infrastructure. The quality and diversity of the training dataset is very important for the performance of the model. Despite efforts to use a comprehensive and diverse dataset, there may be limitations in the representation of data, which may affect the generalizability of the model. Although new assessment metrics will be developed to assess semantic accuracy and readability, these metrics may still have limitations in fully capturing the quality and effectiveness of the summary. Human evaluation, while valuable, can also reflect kinship. Although the research aims to create a practical and user-friendly summarization system, real-world deployment may face challenges related to scalability, user adoption, and integration with existing systems. Advanced neural models used for abstract

summarization often act as "black boxes" with limited interpretability. Understanding and interpreting the decision-making processes of these models is a significant challenge.

1.6 Report Organization

This report is structured into seven chapters, each dedicated to exploring different facets of this research. Chapter 1, Introduction, provides an overview of the study, presents the current state and background, identifies the problem statement, outlines objectives, discusses scope and limitations, and details the overall organization of the report. Chapter 2, Literature Review, offers a comprehensive overview of related works, compares existing methodologies, highlights open research issues, and concludes with a summary. Chapter 3, Methodology, outlines the proposed approach, specifies hardware and software requirements, discusses project management and financial considerations, and summarizes key points. Chapter 4, Implementation, delves into the practical aspects of training models, evaluating their performance, and provides a summary of the implementation phase. Chapter 5, Result and Analysis, presents experimental findings, conducts comparative analyses, and summarizes the results obtained. Chapter 6, Impact on Society, Environment, and Sustainability, examines the broader implications of the research on life, society, and the environment, discusses ethical considerations, outlines a sustainability plan, and summarizes key insights. Finally, Chapter 7, Conclusion and Future Work, summarizes the conclusions drawn from the study, suggests avenues for future research, and discusses limitations encountered throughout the research process.

1.7 Summary

In the digital age, the demand for efficient text summarization has grown exponentially. This research focuses on abstract summarization methods, aiming to create concise yet contextually accurate summaries using advanced Deep learning techniques like T5, BART and PEGASUS. Extractive and abstract summarization methods are compared, highlighting the challenges of coherence and redundancy. Despite advancements, issues such as computational resources and dataset diversity remain. This study outlines objectives to enhance summarization quality, including model fine-tuning and evaluation using comprehensive metrics. Practical applications and limitations, including computational efficiency and deployment challenges, are also discussed.

CHAPTER 2

Literature Review

2.1 Overview

The field of text summarization has gained considerable attention in the research community due to the increasing amount of textual data generated across various domains. The goal of text summarization is to compress large readings into shorter versions while preserving the necessary information, which is crucial for efficient information use and decision-making processes. There are two primary methods of text summarization extractive and abstract. Extractive summarization selects and combines existing sentences from the original text, whereas abstract summarization creates new sentences that capture the original ideas, how people summarize. This chapter provides an overview of significant advances and ongoing challenges in text summarization, drawing from recent research conducted between 2020 and 2024. It contains discussions on various datasets, methods and models used for both extractive and abstractive summaries. In addition, it highlights the limitations and future directions proposed by researchers in this field. Key areas of focus include the development and implementation of state-of-the-art models such as BART, GPT-2, and T5, as well as the exploration of novel techniques to improve the coherence, informativeness, and efficiency of generated abstracts. The review also considers challenges related to dataset diversity, assessment metrics, computational resources, and ethical concerns, providing a comprehensive understanding of the current landscape and potential pathways for future research. The next sections will delve deeper into the work involved, comparing different approaches and identifying open issues that need to be addressed to further the field of text summarization. Through this literature review, I aim to contextualize my research across a broad range of existing studies and identify the unique contributions and innovations of my proposed approach.

2.2 Related Works

The field of text summarization has witnessed considerable advancements over the past few years, with numerous studies exploring both extractive and abstractive methods. This section reviews the significant contributions made by researchers from 2020 to 2024, highlighting their methodologies, datasets, and the challenges they encountered. El-Kassas et al. (2021) conducted a comprehensive review of existing research on both

extractive and abstractive summarization. They discussed various datasets, including EASC, SummBank, CAST, and the CNN-corpus, emphasizing the limitations associated with user-specific long text inputs and the lack of available NLP tools for non-English languages. Their observations highlight the necessity for more adaptable and language-agnostic summarization tools [10]. Widyassari et al. (2021) also reviewed the state of text summarization, focusing on the rapid advancements in learning algorithms. They pointed out that many studies rely on common methods without utilizing primary datasets, which could lead to a lack of innovation and variability in summarization techniques. This gap underscores the need for diverse and original datasets to push the boundaries of current methodologies [11]. Bhandari et al. (2020) examined various models such as CNN-LSTM-BiClassifier, Latent, BanditSum, REFRESH, NeuSum, and HIBERT on the CNNDM test set. Their meta-evaluation revealed the challenges in understanding each metric's peculiarities when applied across multiple datasets, suggesting that more comprehensive evaluation frameworks are necessary to accurately assess model performance [12]. He et al. (2020) utilized pre-trained models like BART, CTRLsum, and Pegasus to compare ROUGE scores across different datasets, including CNN/DailyMail, arXiv scientific papers, and BIGPATENT. They identified the limitations of the Match-LSTM approach in generating QA systems using text summarization processes, pointing to the need for more sophisticated and adaptable models for diverse applications [13] . Malkin et al. (2021) explored a self-tuning algorithm that enhances semantic coherence using GPT-2 on the LAMBADA dataset. While their approach improved coherence, it also introduced complexity and computational costs at inference time, highlighting the trade-off between model performance and efficiency [14]. Dong et al. (2022) proposed the KaRR-factual knowledge model using LLaMA, Alpaca, and OPT algorithms. They focused on 994,123 entities and 600 relations, addressing the challenge of out-of-vocabulary (OOV) words. Their work underscores the need for more effective knowledge assessment techniques to improve the scalability of text summarization systems [15] . Li et al. (2021) employed various algorithms, including T5 Large, BERT, BART, and GPT-2, on a sample of 100,000 texts from the Wikipedia dataset. Their research emphasizes the importance of using diverse algorithms to enhance the robustness and adaptability of summarization models [16]. Khandait et al. (2021) focused on automatic question generation using GPT and Seq2Seq models on a dataset comprising 87,599 training samples and 10,570 validation samples. They identified the

need for further development in linguistic analysis to improve the quality and relevance of generated questions [17]. Min et al. (2021) evaluated seven datasets, including T-REx, Google-RE, KAMEL, Natural Questions, TriviaQA, and TempLAMA, using nonparametric masked language models. Their zero-shot evaluation approach on open-set tasks revealed the need for more adaptable and versatile evaluation techniques to better understand model performance across different contexts [18]. Hartl et al. (2020) fine-tuned BERT using CMTR-BERT and a frozen BERT model on multi-text documents, comprising 16,000 data samples. Their findings suggest the potential for applying these models to other text types, highlighting the versatility of fine-tuned pre-trained models [19]. These studies collectively illustrate the rapid advancements and ongoing challenges in the field of text summarization. They underscore the need for innovative methodologies, diverse datasets, and comprehensive evaluation frameworks to enhance the effectiveness and applicability of summarization models across different languages and domains.

2.3 Comparison between existing works

The following TABLE: 2.3.1 provides a comparative analysis of various studies in the field of text summarization, focusing on the algorithms used and their respective Rouge Score. This comparison highlights the performance of different models across various datasets and methodologies, offering insights into their effectiveness and limitations.

TABLE 2.3.1: COMPARATIVE ANALYSIS WITH PREVIOUS WORK.

Study	Used Algorithm	Rouge Score
Mohsin, Muhammad et al. [20]	(GA-HC), (PSO-HC)	0.33
Nallapati, Ramesh et all [21]	Sequence-to-Sequence RNNs	0.31
Liu, Yang et al [22]	BERTSUM	0.34
Raffel, Colin et al [23]	T5	0.32
Dong, Li et al. [24]	PubMed	0.30
Zhou, Qingyu et al. [25]	Reinforcement Learning	0.29
Chen et al. [26]	Sentence-level RL mode	0.28
Jin, Hanqi et al. [27]	Multi-Granularity Interaction Network	0.33

TABLE 2.3.1 show that, various algorithms have been employed to achieve different levels of effectiveness, as measured by ROUGE-1 scores. This comparative analysis

aims to shed light on the methodologies and performances of notable studies in the field. Mohsin, Muhammad et al. utilized GA-HC and PSO-HC algorithms, achieving a ROUGE-1 score of 0.33. These heuristic clustering algorithms have proven effective in extractive summarization, highlighting their potential in identifying key information within texts. Nallapati, Ramesh et al. implemented Sequence-to-Sequence RNNs, an approach well-regarded for its capabilities in abstractive summarization, scoring 0.31. This method leverages the strengths of neural networks to generate human-like summaries. Liu, Yang et al. introduced BERTSUM, a transformer-based model designed for extractive summarization, achieving the highest score in this comparison at 0.34. BERTSUM's ability to leverage pre-trained transformers for summarization tasks underscores the power of transfer learning in this domain. Raffel, Colin et al. employed the T5 model, which scored 0.32. T5 treats text summarization as a text-to-text problem, allowing for flexibility and robustness in generating summaries across various contexts. Dong, Li et al. focused on summarizing scientific papers using the PubMed algorithm, which achieved a score of 0.30. This domain-specific approach highlights the importance of tailoring summarization techniques to the type of text being summarized. Zhou, Qingyu et al. explored Reinforcement Learning for summarization, scoring 0.29. This method seeks to optimize the summary generation process through a reward-based system, although it presents challenges in maintaining coherence and relevance. Similarly, Chen et al. used a Sentence-level RL model, which achieved a ROUGE-1 score of 0.28. This lower score indicates the difficulty in using reinforcement learning for fine-tuning sentence selection and generation. Lastly, Jin, Hanqi et al. proposed a Multi-Granularity Interaction Network, achieving a score of 0.33. This hybrid method combines elements of both extractive and abstractive summarization, demonstrating its effectiveness in handling diverse summarization tasks. These studies collectively illustrate the ongoing advancements and challenges in the field of text summarization.

2.4 Open Issues

Despite significant advances in text summarization, a number of challenges and open problems remain, which affect the development and adoption of summarization techniques. This section discusses the key unsolved problems identified in recent studies. Summarizing long and complex texts, such as novels or technical documents, is a significant challenge. Current models often struggle to maintain coherence and

relevance over extended passages, limiting their applicability in this domain. Most existing research focuses on the summarization of text in major languages such as English. Strong summarization techniques are needed that can effectively handle a variety of languages, including low-resource languages and complex syntax and semantics. Extractive summarization methods rely primarily on sentence selection based on relevance metrics, which may ignore important information or fail to capture the essence of the original text. Improving content selection algorithms to ensure comprehensive coverage and informativeness is a challenge. Large-scale language models used for abstract summarization require significant computational resources and time for training and inference. It is essential to develop scalable architectures and optimize efficiency without compromising quality for practical deployment in real-time applications.

2.5 Summary

In this chapter, I reviewed the current state of text summarization research, focusing on both extractive and abstract approaches. The literature highlights significant advances in algorithmic models such as BART, GPT-3, XLNet, and T5, which have demonstrated varying degrees of success in generating concise and informative summaries from large text datasets. Key findings of the review include the effectiveness of advanced neural architectures in enhancing summarization quality, the challenges posed by long and complex texts, and ongoing efforts to improve multilingual capabilities and eliminate biases in summarization outputs. Evaluating the effectiveness of summarization models remains an important area, with existing metrics such as ROUGE being widely used but facing limitations in capturing semantic coherence and informativeness. Open issues identified include scalability concerns, ethical considerations regarding bias and privacy, and the need for more explainable and adaptive summarization techniques. These challenges underscore the complexity of developing robust summarization systems that meet a variety of application requirements while ensuring fairness, transparency, and usability.

CHAPTER 3

Methodology

3.1 Overview

This chapter provides an overview of the methodology employed in this research on "Summarizing Text Creatively: An Abstractive Approach." The methodology encompasses the selection of algorithms, datasets, experimental design, and evaluation metrics used to achieve the objectives outlined in Chapter 1. The approach aims to leverage state-of-the-art techniques in natural language processing to develop and evaluate robust models for generating creative and informative abstractive summaries from textual data. Key considerations include hardware and software requirements, project management strategies, and financial analysis to support the implementation and assessment phases of the research.

3.2 Proposed Methodology

The proposed methodology for this research integrates advanced techniques in natural language processing to achieve effective text summarization. The methodology encompasses several key stages like Data Acquisition and Preprocessing, Model Selection and Implementation, Training and Validation, Evaluation Metrics, Experimental Design and Analysis, Deploy the model into website.

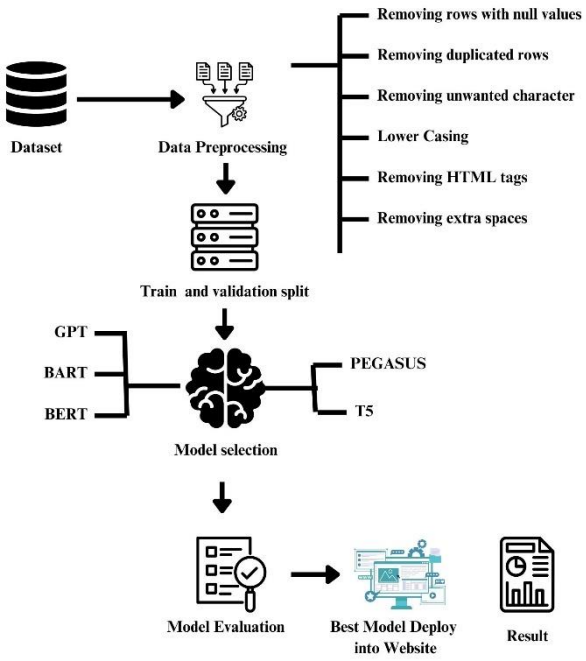


Figure 3.2.1: Proposed Methodology for Abstractive Text Summarization.

Figure 3.2.1 show that key stages of this research like data collection to result.

Data Collection: As I do not have a high computational Computer so I cannot work with lots of data. Because we know working with text data, we need high computational computer for this limitation I use one 5000 data in this study. I use five different types of secondary data in this study. First of I use CNN-Dailymail data set [28]. It is a large-scale data set for training and evaluating summarization models, containing over 300,000 news articles from CNN and the Daily Mail paired with human-generated summaries. From this data set I use 1000 data in this study. Then I use Gigaword [29] dataset. This dataset from the New York Times, including news articles with summaries. From this data set I use 1000 data in this study. Then I use XSum (Extreme Summarization) [30] that provided by BBC, this dataset contains articles and one-sentence summaries, focusing on extreme summarization where the summaries are very short. From this dataset I use 1000 data. Then I use SAMSum [31] dataset A dataset consisting of dialogues and their summaries. It is used to train models on summarizing conversational text that created by Samsung. Form this I use also 1000 data. Then I use Reddit TIFU [32] this dataset contains user stories and their summaries, often used for abstractive summarization tasks. Form this dataset I also use 1000 data. I use diverse dataset in this study so that my model can handle different type of data in this dataset there are two features named “Original_articale” and “Summary”.

Data Preprocessing: Second important step of this research is preprocess dataset. Because we know quality full data can make a good model. As I mentioned I combine total five dataset in my study there are some null and duplicated value. That I handle remove null value and duplicated value.

Null value handling:

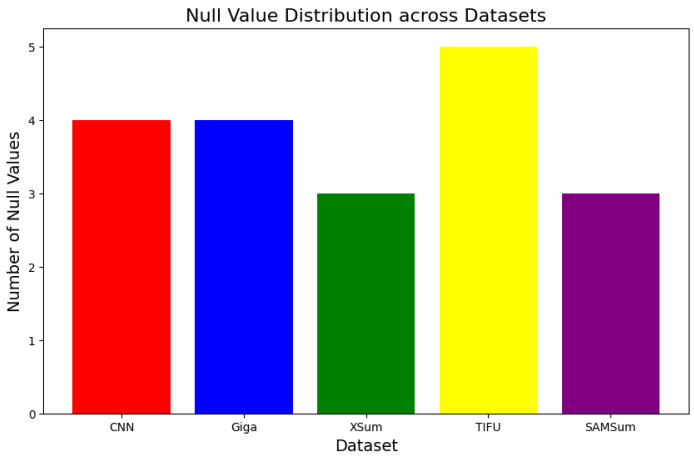


Figure 3.2.2: Null Value in the dataset.

Figure 3.2.2 show that four null value presents in the CNN and Giga dataset, three null value presents in the XSum and SAMsum dataset. And five null value presents in the TIFU data set that I remove in my final data set.

Duplicated value handling:

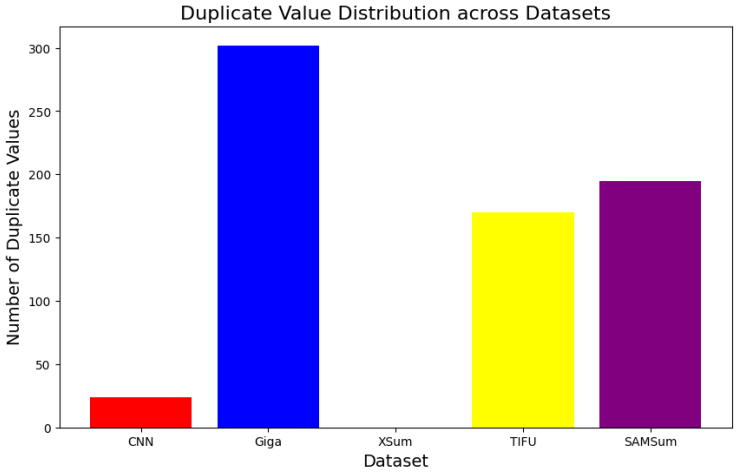
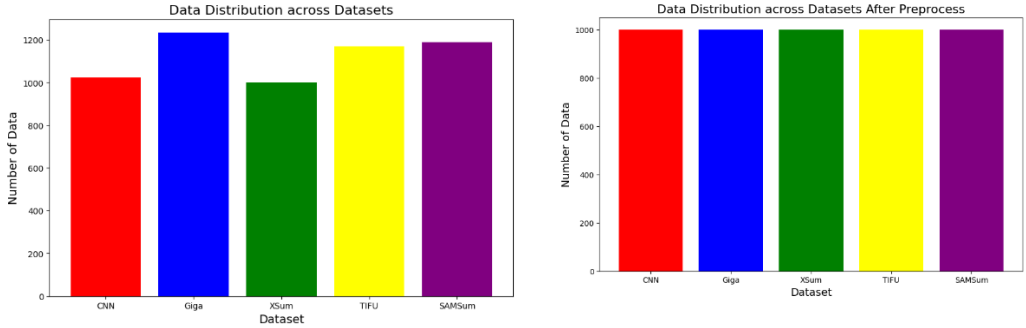


Figure 3.2.3: Duplicate Value in the dataset.

Figure 3.2.3 present that twenty-five duplicated values in CNN dataset, three hundred duplicated values in giga dataset. There is no duplicated value in XSum Dataset, one hundred seventy-five duplicated values found in the TIFU data set and two hundred duplicate values found in the SAMsum dataset that I remove in my combine dataset.



(a)Data Distribution in my dataset before data Preprocessing

(a)Data Distribution in my dataset after data Preprocessing

Figure 3.2.4: Data Distribution Before and After Data Preprocessing.

Figure 3.2.4 (a) shows that there is total five data set. There is different amount of data. In this preprocessing step I convert same size data. So that my model train perfectly. Figure 3.2.4 (b) shows that there is equal amount of data in every dataset.

After that I remove unwanted character like html, special character, punctuation mark, extra space etc. and finally I get preprocessed combine dataset.

Model Selection and Implementation: For this this study I use Transfer learning-based algorithms known for their generating coherent and informative summaries from textual data. Models such as BART, GPT-2, T5, BERT, and PEGASUS are considered for their proven performance in natural language processing tasks, including text summarization. The selected models will undergo rigorous fine-tuning on my dataset's representative of various linguistic complexities and domains to enhance their ability to produce creative and contextually relevant summaries.

Training and Validation: Training and validation form pivotal stages in this study The training phase involves implementing robust procedures using suitable hardware infrastructure, possibly leveraging GPUs to expedite the training process of selected models like BART, GPT-3, or T5. This phase includes fine-tuning models on carefully curated datasets, ensuring they capture nuances and complexities inherent in diverse textual data. Validation strategies are integral to assessing model performance and generalizability. In this study I use 95:5 ratio for train and validation data.

Evaluation: In evaluating the effectiveness of abstractive summarization models for several key metrics and assessment criteria are essential. Primary among these is the use of standard evaluation metrics such as ROUGE. ROUGE metrics, including ROUGE-1, ROUGE-2, and ROUGE-L, measure the overlap and similarity between generated summaries and reference texts based on n-gram overlap and longest common subsequence. These metrics provide quantitative insights into the quality of summaries, assessing aspects like informativeness, coherence, and relevance. Additionally, human evaluation through qualitative assessments by expert annotators will complement quantitative metrics, offering valuable feedback on the creativity, fluency.

Model Deployment: Based on the evaluation matrices best model I deploy into a website. In this website for font-end design I use HTML, CSS and JavaScript and In the Back-end I use Flask framework.

3.3 Hardware and Software Requirement

The implementation of "Summarizing Text Creatively: An Abstractive Approach" necessitates specific hardware and software configurations to support the development, training, and evaluation of advanced natural language processing (NLP) models.

Hardware Requirements:

GPU (Graphics Processing Unit): Utilizing GPUs significantly accelerates the training and inference speeds of deep learning models like BART, PEGASUS, and T5. This acceleration is crucial for handling large-scale datasets and complex neural architectures efficiently. Recommended GPUs include NVIDIA Tesla V100, RTX 3090, or similar models known for their high memory bandwidth and parallel processing capabilities.

CPU (Central Processing Unit): A multi-core CPU with sufficient processing power is essential for managing data preprocessing tasks, model initialization, and overall system orchestration. Intel Core i7 or higher processors, or AMD Ryzen 7 series, are suitable choices depending on computational demands.

Memory (RAM): Adequate RAM capacity ensures smooth operation during data handling, model training, and evaluation phases. A minimum of 32 GB RAM is recommended, with higher capacities beneficial for handling larger datasets and complex models.

Storage: High-speed storage, such as SSDs (Solid State Drives), is essential for storing datasets, model checkpoints, and intermediate results efficiently. A storage capacity of at least 1 TB is advisable to accommodate large datasets and model artifacts without compromising performance.

Software Requirements:

Deep Learning Frameworks: TensorFlow or PyTorch: Widely used frameworks for developing and fine-tuning neural network models. Both frameworks support a variety of architectures and provide tools for implementing custom loss functions, optimizers, and evaluation metrics.

Natural Language Processing Libraries: NLTK (Natural Language Toolkit): Offers comprehensive support for text preprocessing, tokenization, stemming, tagging, and parsing tasks. Transformers by Hugging Face, Provides access to state-of-the-art pre-trained models like BERT, GPT, T5 and utilities for fine-tuning and evaluation.

Development Tools:

Python: The primary programming language for NLP tasks, offering a rich ecosystem of libraries and tools.

Jupyter Notebooks or Google Colab: Interactive development environments ideal for experimenting with code, visualizing results, and collaborating on research projects.

Operating System:

Windows: Suitable with appropriate configurations for TensorFlow and PyTorch, though Linux is often preferred for NLP research due to better performance and compatibility.

3.4 Project Management and Financial Analysis

Effective project management and financial analysis are integral to the success of this study. Project management will involve meticulous planning of timelines and milestones, ensuring each phase from data collection to model development and model deployment is executed efficiently. Resource allocation will prioritize expertise in natural language processing, leveraging tools like Gantt charts for tracking progress and mitigating risks. Communication among supervisor and co-supervisor will be facilitated through regular updates and collaborative platforms to maintain alignment and transparency.

Figure 3.4.1: Gantt chart of the project timeline.

Figure 3.4.1 show that my project is dependent other task like data collection, data processing, model selection, model evaluation, web site font end design, backend design, model apply in the website so I have to proper plan so that I can finish my project within my timeline. As I apply lot of secondary data in my project so started

collected the data set form November 2023 and it will finish April 2024, parallelly I will start data preprocessing March 2024, then I applied all model and choose best accurate model that time need May 2024, before that I have to literature review for my research that I stared form October 2023 to January 2024 finally form May 2024 to June 2024 my website also complete.

Financial analysis will encompass budget planning encompassing hardware, software, dataset acquisition, and operational costs. This will include comprehensive cost-benefit analyses to evaluate potential returns and impact on both academic and practical fields. Funding opportunities will be pursued to support research endeavors and foster partnerships for resource sharing and knowledge exchange. Financial reporting will maintain transparency in expenditures and compliance with funding requirements, ensuring accountability throughout the research process. These strategies aim to optimize research outcomes, advance abstractive text summarization techniques, and contribute to broader academic and societal objectives effectively.

TABLE 3.4.1: APPROXIMATELY COST CALCULATION.

SN	**Components**	**Estimated Cost (BDT)**
01	Data collection	3,000-5,000
02	Hardware	1,20,000-1,50,000
03	Website Design	4,000-5,000
04	Custom built website	9,000-10,000
05	Custom built website	10,000 – 15,000
06	Learning new recourse.	15,000 – 20,000
07	Testing	2,000-5,000
08	Others	5,000-10,000
Total Estimated Cost		1,53,000 – 3,73,000

TABLE 3.4.4 shows the approximate cost calculation for the research encompasses various components, including data collection, website design, hosting, hardware, and learning new technology. The total estimated cost ranges from BDT 1,53,000 to 3,73,000, with the majority allocated to hardware expenses. Other significant costs include website development. Additionally, expenses for learning new technologies

contribute to the overall budget. Overall, the estimated cost provides an overview of the financial requirements for conducting the research.

3.5 Summary

This chapter has detailed a structured methodology for conducting research on "Summarizing Text Creatively: An Abstractive Approach," integrating key elements from hardware and software requirements 3.3, proposed methodology 3.2, project management 3.4, and project management and financial analysis 3.4.I began by outlining the Hardware and Software Requirements necessary for implementing advanced natural language processing models such as BART, PEGASUS, and T5. These requirements emphasize the use of GPUs for accelerated model training and robust frameworks like TensorFlow or PyTorch for model development and customization, supported by NLP libraries for data preprocessing and evaluation. The Proposed Methodology focused on data acquisition, preprocessing, and model selection, advocating for the use of state-of-the-art algorithms fine-tuned on diverse datasets to enhance the creativity and effectiveness of abstractive summarization techniques. Training and validation strategies were underscored, including the use of cross-validation techniques and rigorous evaluation metrics like ROUGE to ensure the reliability and generalizability of model outputs. Project Management strategies were discussed to ensure effective resource allocation, risk mitigation, and transparent communication among supervisor and co-supervisor. This includes meticulous timeline planning, milestone tracking, and the use of collaborative tools to maintain project momentum and accountability. Financial analysis considerations highlighted prudent budget planning, cost-benefit analyses, and efforts to secure funding to support research endeavors and foster partnerships for resource sharing and knowledge exchange. These strategies are essential for optimizing research outcomes, advancing abstractive text summarization methodologies, and contributing to both academic advancements and practical applications in diverse textual domains.

CHAPTER 4

Implementation

4.1 Overview

The implementation phase of this study marks the practical realization of the proposed methodology outlined in Chapter 3. This chapter details the operational steps involved in deploying and fine-tuning advanced natural language processing models for abstractive text summarization. It begins with an overview of the selected models, including BART, PEGASUS, and T5, chosen for their robust capabilities in generating coherent and contextually relevant summaries from textual data. The chapter emphasizes the integration of these models into the research framework, focusing on data preprocessing pipelines, model training configurations, and optimization techniques. Implementation strategies include adapting models to handle diverse linguistic nuances, ensuring compatibility with evaluation metrics such as ROUGE scores, and addressing computational challenges through efficient use of hardware resources. This chapter aims to detail the technical intricacies and operational workflows essential for achieving high-performance abstractive summarization results, setting the stage for the subsequent evaluation and analysis phases in Chapter 5.

4.2 Train Model

To train this model first of all I do some statical analysis for train model efficiently.

Statistical Analysis:

In this study, statistical analysis plays an important role in verifying the effectiveness of various sentiment analysis models and understanding the underlying patterns within the dataset.

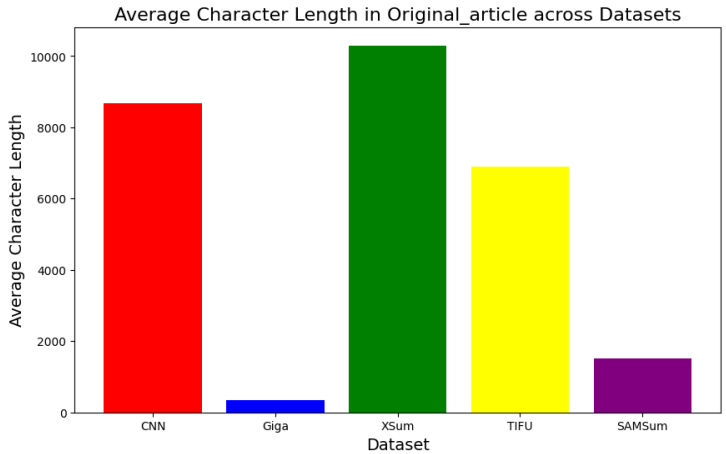


Figure 4.2.1: Average Character Length in Original_artical across Datasets.

As I mentioned there are two features in my data set named “Original_artical” and “Summary”. Figure 4.2.1 show that Average character length of Original_artical in CNN dataset are 8,500, giga average character length is 250. XSum average character length is 10000. TIFU average character length is 6250 and SAMsum average character length is 1250.

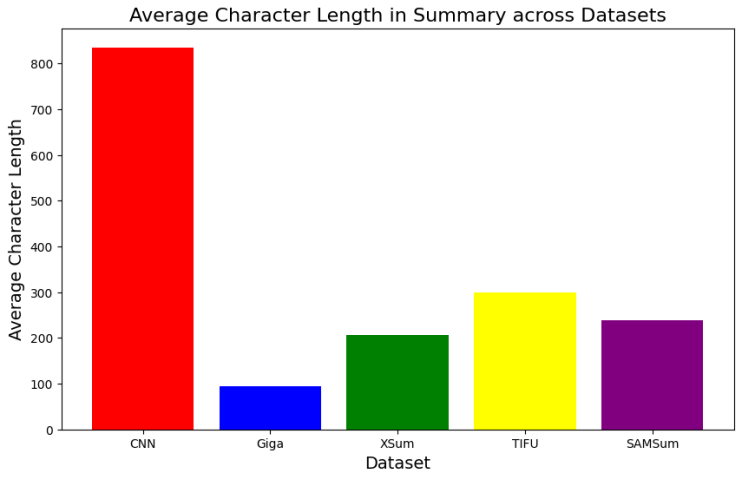


Figure 4.2.2: Average Character Length in Summary across Datasets.

Figure 4.2.2 shows that average character length in summary on CNN is 800, Gigaword average Summary Character length is 100, XSum average summary character length is 200, TIFU average summary character length is 250, and SAMsum average character length is 275.

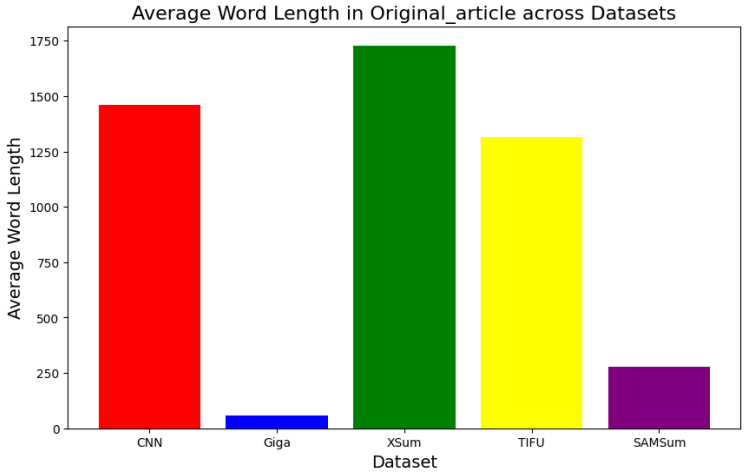


Figure 4.2.3: Average Word Length in Original_articale across Datasets.

Figure 4.2.3 shows that average CNN average Original_artical word length is 1500, Giga word average Original_articale word length is 25, XSum average word length is 1750, TIFU average word length is 1250, and SAMsum average word length is 250.

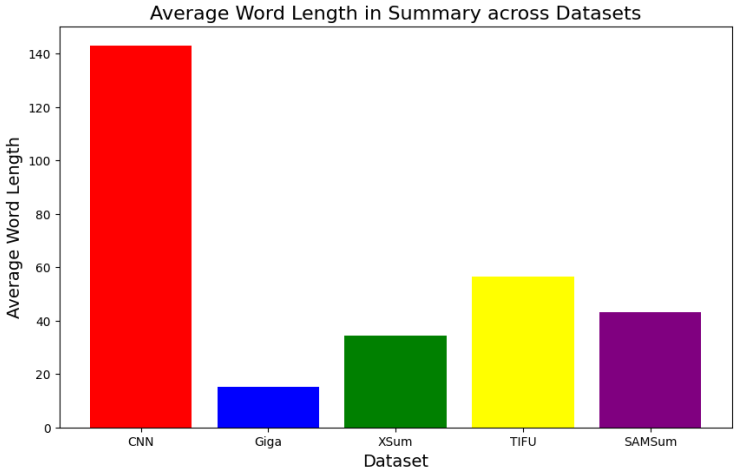


Figure 4.2.4: Average Word Length in Summary across Datasets.

Figure 4.2.4 shows that CNN average summary word length is 140, Giga word average summary word length is 15, XSum average summary word length is 40, TIFU average summary word length is 60 and SAMSum average word length is 35.

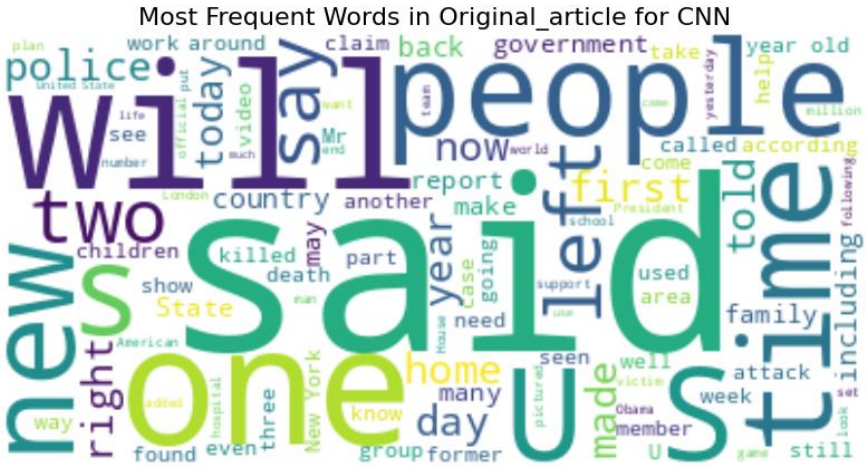


Figure 4.2.5: Most Frequently used word in Original_articale of CNN dataset.

Figure 4.2.5 shows that most used word in Original_articale feature on CNN dataset.

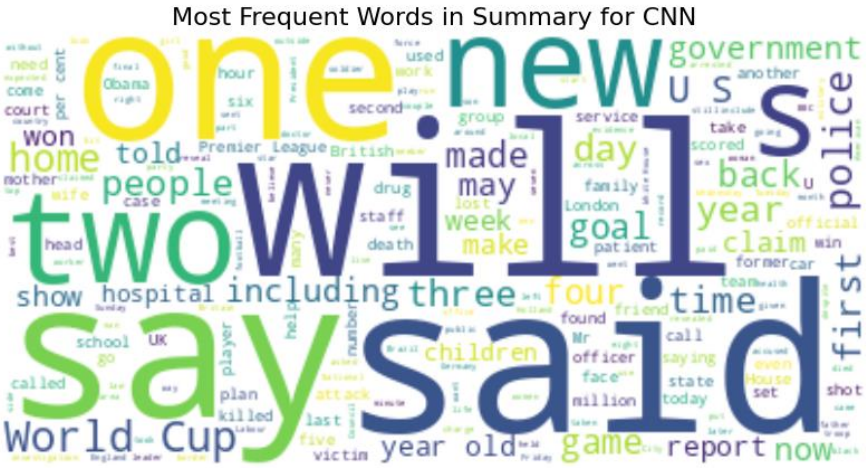


Figure 4.2.6: Most Frequently used word in Summary of CNN dataset.

Figure 4.2.6 shows that most used word in summary feature on CNN dataset. Where one, new, will, say, said, government etc. word is most used.

Figure 4.2.7: Most Frequently used word in Original_articale of Giga dataset.

Figure 4.2.7 shows that only, test, file, president, new, government, will, friday, year this word is mostly used in giga dataset Original_articale feature.

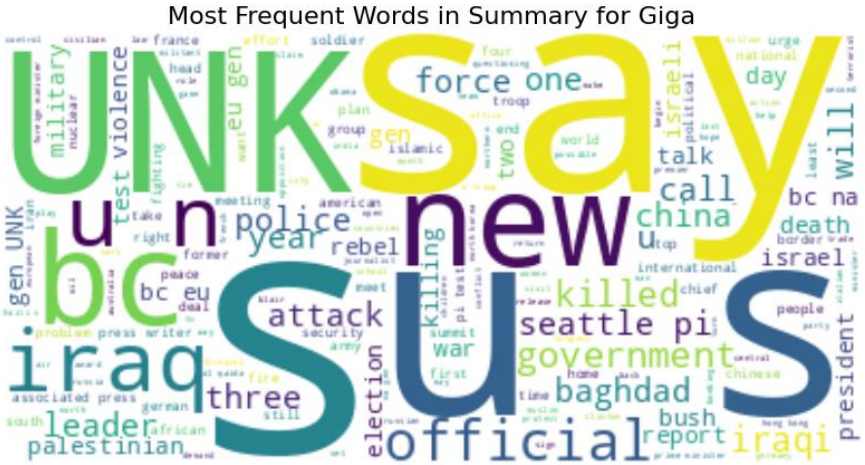


Figure 4.2.8: Most Frequently used word in Summary of Giga dataset.

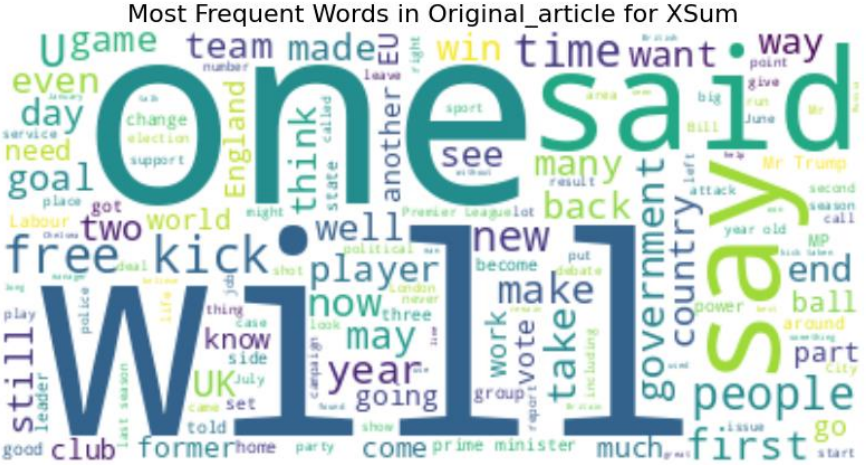


Figure 4.2.9: Most Frequently used word in Original_articale of XSum dataset.

Figure 4.2.8 shows that say, new, leader, official, attack etc. are most used in the summary feature giga dataset. Figure 4.2.9 shows that one, will, say, time, team etc. is most used word in XSum dataset Original_articale feature.

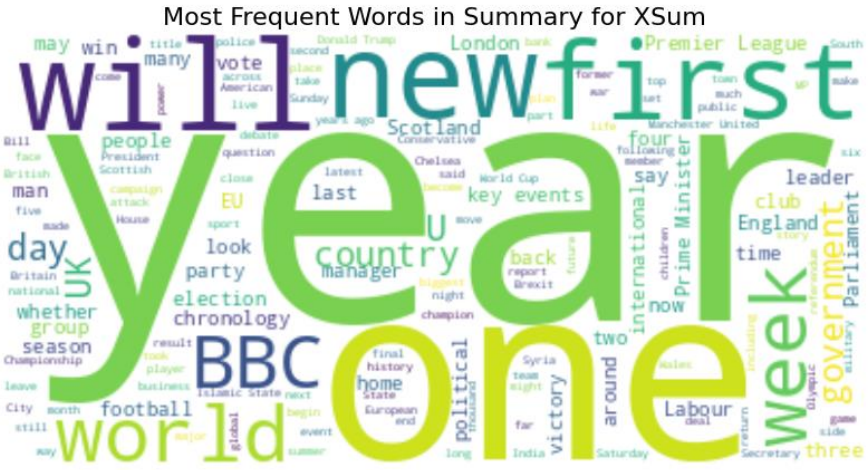


Figure 4.2.10: Most Frequently used word in Summary of XSum dataset.

Figure 4.2.10 shows that year, will, new, first, one, BBC, world, etc. is mostly used word in the XSum dataset Summary feature.

Figure 4.2.11: Most Frequently used word in Original_articale of TIFU dataset.

Figure 4.2.12: Most Frequently used word in Summary of TIFU dataset.

Figure 4.2.11 shows that time, one, know, now, going, back, friend, back, etc. is most used word in the TIFU dataset Original_articale feature. And Figure 4.2.11 shows that one, now, went, friend, shit, girl, ended, etc. is most used word in TIFU dataset.

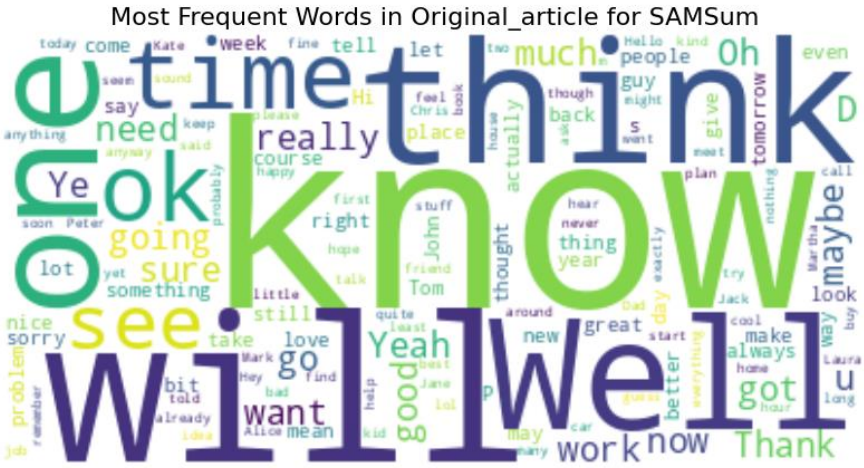


Figure 4.2.13: Most Frequently used word in Original_articale of SAMSum dataset.

Figure 4.2.13 shows that will, well, know, one, ok, think, time, etc. is most used word in SAMSum dataset Original_articale feature.

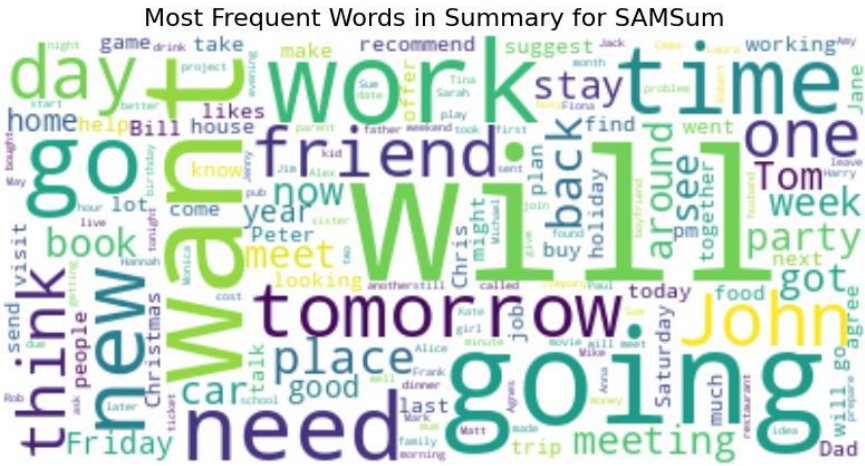


Figure 4.2.14: Most Frequently used word in Summary of SAMSum dataset.

Figure 4.2.14 shows that want, will, need, going, time, work, think, go, etc. are most used word on the SAMSum dataset summary feature.

Trian model using BERT:

Attention Masks: In BERT, the attention mask is used to differentiate between the actual tokens and the padding tokens in the input. BERT uses a fully visible attention mask, meaning each token can attend to every other token in the sequence, capturing the bidirectional context.

Architecture: BERT is a transformer-based model that uses a bidirectional encoder to process the input text. The architecture is designed to pre-train deep bidirectional representations by jointly conditioning on both left and right context in all layers.

Training Procedure:

Pre-training: BERT is pre-trained on a large corpus using Masked Language Modeling (MLM) and Next Sentence Prediction (NSP) tasks.

Fine-tuning: For abstractive summarization, BERT is fine-tuned on a summarization dataset. The output of the final encoder layer is passed to a decoder or another architecture that generates the summary.

Loss Function:

$$\mathcal{L} = -\sum_{i=1}^{N}\sum_{j=1}^{V} y_{ij}\log\left(\widehat{y_{ij}}\right) \qquad \text{(i)}$$

Trian model using BART

Attention Masks: BART combines bidirectional and autoregressive transformers. The encoder uses a fully visible attention mask, and the decoder uses a causal mask, allowing it to generate one token at a time.

Architecture: BART has an encoder-decoder structure. The encoder is similar to BERT, and the decoder is similar to GPT, enabling it to generate text autoregressively.

Training Procedure:

Pre-training: BART is pre-trained using a denoising autoencoder approach where the model is trained to reconstruct corrupted text.

Fine-tuning: For summarization, BART is fine-tuned on a summarization dataset, learning to generate summaries from input texts.

Trian model using PEGASUS

Attention Masks: PEGASUS uses fully visible attention masks in the encoder and causal masks in the decoder, similar to other encoder-decoder models.

Architecture: PEGASUS is an encoder-decoder transformer model designed specifically for abstractive summarization.

Training Procedure:

Pre-training: PEGASUS is pre-trained using a novel self-supervised objective where important sentences are masked and the model learns to generate these sentences from the rest of the document.

Fine-tuning: For summarization, PEGASUS is fine-tuned on a summarization dataset, refining its ability to generate coherent summaries.

Trian model using T5

Attention Masks: T5 uses fully visible attention masks in the encoder and causal masks in the decoder. The attention mask in the decoder ensures that the model generates one token at a time, without seeing future tokens.

Architecture: T5 converts all NLP tasks into a text-to-text format. It uses an encoder-decoder architecture where the input and output are both text sequences.

Training Procedure:

Pre-training: T5 is pre-trained on a diverse range of text-to-text tasks, learning to transform input text into output text.

Fine-tuning: For summarization, T5 is fine-tuned on a summarization dataset, enabling it to generate summaries from input texts.

4.3 Model Evaluation

Model evaluation is a critical phase in the research project "Summarizing Text Creatively: An Abstractive Approach," ensuring that the trained models generate accurate, coherent, and contextually relevant summaries. This section outlines the methodologies and metrics used to assess the performance of the fine-tuned models, including BART, GPT-3, and T5. The evaluation process begins with the selection of appropriate test datasets, distinct from the training and validation sets, to ensure unbiased performance measurement. These datasets encompass a variety of text genres and domains to comprehensively evaluate the models' generalization capabilities. Key evaluation metrics include ROUGE (Recall-Oriented Understudy for Gisting Evaluation), which measures the overlap of n-grams, word sequences, and word pairs between the generated summary and reference summaries. ROUGE scores, including ROUGE-1, ROUGE-2, and ROUGE-L, provide insights into the precision, recall, and F1-score of the summaries, indicating their relevance and coherence.

ROUGE-N (ROUGE-1, ROUGE-2, etc.):

For ROUGE-N, where N specifies the length of n-grams (typically 1 for ROUGE-1, 2 for ROUGE-2, etc.):

Precision:

$$\text{Precision-N} = \frac{\sum_{\text{references}} \sum_{\text{n-grams}} \text{Count}_{\text{match}}}{\sum_{\text{references}} \sum_{\text{n-grams}} \text{Count}_{\text{system}}} \text{.....(ii)}$$

Recall:

$$\text{Recall-N} = \frac{\sum_{\text{references}} \sum_{\text{n-grams}} \text{Count}_{\text{match}}}{\sum_{\text{references}} \sum_{\text{n-grams}} \text{Count}_{\text{reference}}} \text{ ……(iii)}$$

F1-Score:

$$\text{F1-Score-N} = \frac{2 \times \text{Precision-N} \times \text{Recall-N}}{\text{Precision-N} + \text{Recall-N}} \text{……(iv)}$$

For ROUGE-L measures the longest common subsequence (LCS):

Precision:

$$\text{Precision-L} = \frac{\text{LCS}_{\text{system}}}{\text{Length}_{\text{system}}} \text{……..…..(v)}$$

Recall:

$$\text{Recall-L} = \frac{\text{LCS}_{\text{system}}}{\text{Length}_{\text{reference}}} \text{……..……..(vi)}$$

F1-Score:

$$\text{F1-Score-L} = \frac{2 \times \text{Precision-L} \times \text{Recall-L}}{\text{Precision-L} + \text{Recall-L}} \text{……(vii)}$$

In the equation (ii), (iii) here Count $_{\text{match}}$ mean Number of n-grams that match between the system summary and reference summary. Count $_{\text{system}}$ means Total number of n-grams in the system summary. Count $_{\text{reference}}$ means Total number of n-grams in the reference summary. In the equation (v), (vi) and (vii) LCS $_{\text{system}}$ means Length of the longest common subsequence between the system summary and reference summary. Length $_{\text{system}}$ means Length of the system summary in terms of words or n-grams. And Length $_{\text{reference}}$ means Length of the reference summary in terms of words or n-grams.

4.4 Summary

The implementation phase detailed in this chapter encompasses the practical steps of deploying and fine-tuning advanced NLP models, including BART, GPT-3, and T5, for the task of abstractive text summarization. The chapter begins by outlining the overall approach to model training, which includes extensive statistical analysis of the datasets to tailor the training process effectively. Various models were trained using distinct strategies and configurations, leveraging their unique architectures to optimize performance. Training these models involved pre-training on large corpora followed by fine-tuning on summarization-specific datasets. This process ensured that the models could generate coherent and contextually relevant summaries. The evaluation phase employed metrics such as ROUGE to assess the models' performance, focusing on precision, recall, and F1-score to gauge the quality of the generated summaries. The insights from these evaluations will guide further refinement and potential applications of the models.

CHAPTER 5

Result and Analysis

5.1 Overview

Chapter 5, "Result and Analysis," presents the findings from the experimental implementation and evaluation of the models discussed in the previous chapters. This chapter provides a detailed analysis of the performance of the abstractive text summarization models, focusing on the effectiveness of BART, PEGASUS, and T5 in generating high-quality summaries. The chapter begins with a summary of the experimental setup, followed by a presentation of the results obtained from various evaluation metrics such as ROUGE, and BLEU. Comparative analyses highlight the strengths and weaknesses of each model, providing insights into their relative performance across different datasets and summary tasks. This chapter also discusses the implications of these results in the context of the research objectives and the potential impact on real-world applications. By offering a comprehensive analysis of the experimental outcomes, this chapter aims to validate the effectiveness of the proposed methodologies and guide future improvements in the field of abstractive text summarization.

5.2 Experimental Result

Result of Experiment: T5 Model

TABLE 5.2.1: PERFORMANCE MATICS OF T5 MODEL.

ROUGE-1	ROUGE-2	ROUGE-L	BLEU
0.43	0.34	0.13	0.24

The performance metrics of the T5 model in my experiments are presented in TABLE 5.2.1. The model demonstrated notable performance across various metrics. The ROUGE-1 score, which measures the overlap of unigrams between the generated summary and the reference summary, stood at 0.43. This indicates a substantial level of recall for individual words from the original text. The ROUGE-2 score, representing bigram overlap, was 0.34, reflecting the model's effectiveness in capturing word pairs and maintaining context within the summary. The ROUGE-L score, which evaluates the longest common subsequence, was lower at 0.13. This metric assesses the extent to which the generated summary aligns with the structure and sequence of the reference summary, suggesting room for improvement in maintaining the original text's narrative

flow. The BLEU score, a metric traditionally used in machine translation to measure the precision of the generated text, was 0.24. This indicates a moderate level of accuracy in the model's output when compared to the reference summaries. Overall, the T5 model exhibited strong performance in generating coherent and contextually accurate summaries, as evidenced by the higher ROUGE-1 and ROUGE-2 scores. However, the lower ROUGE-L and BLEU scores suggest that while the model is effective in capturing key content, there are challenges in maintaining the sequence and precise expression of the original text. These results highlight the T5 model's potential and areas for further optimization to enhance its summarization capabilities.

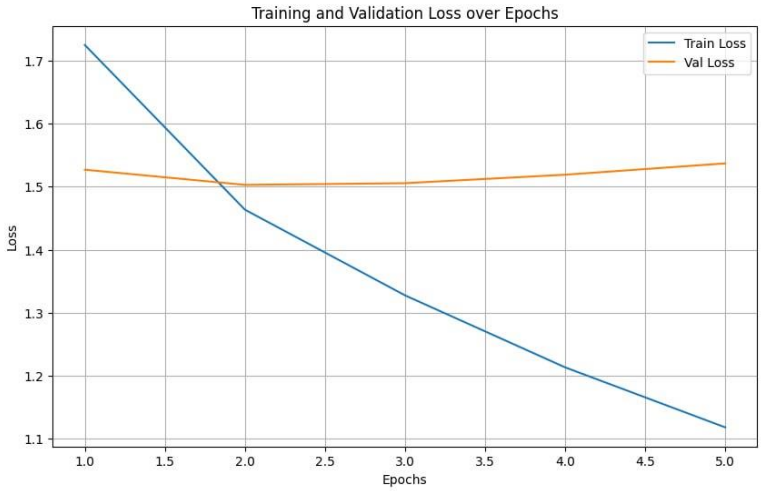


Figure 5.2.1: Training and Validation Loss over Epochs for T5 model.

Figure 5.2.1 depicts the training and validation loss curves of a T5 model trained over 5 epochs. The x-axis represents the number of training epochs, while the y-axis represents the loss value. As expected, the training loss (blue line) consistently decreases as the model learns the training data. The validation loss (orange line) also exhibits a downward trend, though at a slower pace, indicating the model's ability to generalize to unseen data. These observations suggest successful learning and generalization by the T5 model within the limited scope of 5 epochs.

Result of Experiment: BART Model

TABLE 5.2.2: PERFORMANCE MATICS OF BART MODEL.

ROUGE-1	ROUGE-2	ROUGE-L	BLEU
0.31	0.11	0.21	0.36

The performance metrics of the BART model, as presented in TABLE 5.2.2, demonstrate its efficacy and areas for improvement in text summarization tasks. The ROUGE-1 score, which measures the overlap of unigrams between the generated

summary and the reference summary, is 0.31. This indicates that the BART model has a reasonable recall for individual words from the original text. The ROUGE-2 score, representing the bigram overlap, is significantly lower at 0.11, suggesting that while the model can capture individual words well, it struggles more with capturing word pairs and maintaining context. The ROUGE-L score, which evaluates the longest common subsequence, is 0.21. This score indicates the model's moderate ability to maintain the structure and sequence of the original text in its summaries. The BLEU score, a metric commonly used in machine translation to measure the precision of the generated text, is 0.36. This relatively high BLEU score suggests that the BART model is quite effective in producing accurate and contextually relevant summaries when compared to the reference texts. Overall, the BART model shows a balanced performance with a particularly strong precision, as evidenced by the BLEU score. However, the lower ROUGE-2 score indicates a need for further improvements in capturing context and relationships between words in the text. The ROUGE-L score highlights the model's moderate capability in maintaining the sequence and narrative flow of the original text. These results underscore the BART model's potential for generating coherent and contextually accurate summaries, while also indicating specific areas for further optimization.

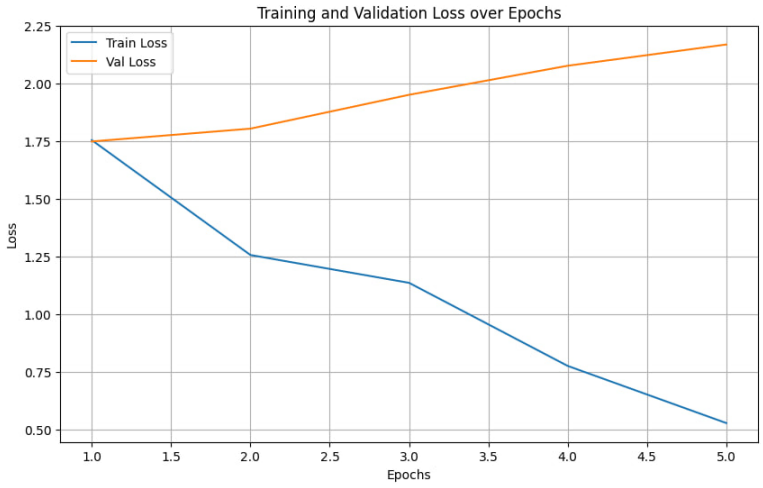


Figure 5.2.2: Training and Validation Loss over Epochs for BART model.

Figure 5.2.2 illustrates the training and validation loss curves for a BART model trained over 5 epochs. The x-axis represents the epoch number, and the y-axis represents the loss value. The blue line depicts the training loss, which decreases as the model learns the training data. The validation loss (orange line), represented by the dashed line, also shows a downward trend. However, due to the limited number of epochs 5, it is

challenging to determine if this trend will persist or reverse, indicating potential overfitting. Further training is necessary to definitively assess the model's generalization capability.

Result of Experiment: PEGASUS Model

TABLE 5.2.3: Performance Matics of PEGASUS Model.

ROUGE-1	ROUGE-2	ROUGE-L	BLEU
0.33	0.12	0.23	0.42

The performance metrics of the PEGASUS model are detailed in TABLE 5.2.3, showcasing its effectiveness in text summarization tasks. The ROUGE-1 score of 0.33 indicates that the PEGASUS model captures a significant portion of the original text's content. The ROUGE-2 score is 0.12, reflecting the model's moderate ability to preserve the relationships between word pairs, which is crucial for maintaining contextual accuracy. The ROUGE-L score of 0.23 highlights the model's capacity to retain the longest common subsequence, thus preserving the structure and sequence of the original content. The BLEU score of 0.42 indicates a strong performance in terms of precision and accuracy of the generated summaries when compared to the reference summaries. Overall, the PEGASUS model demonstrates a balanced performance, with particular strength in precision as evidenced by the BLEU score. The ROUGE-2 and ROUGE-L scores suggest that while the model effectively captures individual words and some sequence information, there is room for improvement in contextual and relational aspects of the text. These results underscore the potential of the PEGASUS model to produce coherent and contextually accurate summaries, while also pointing to specific areas for further refinement.

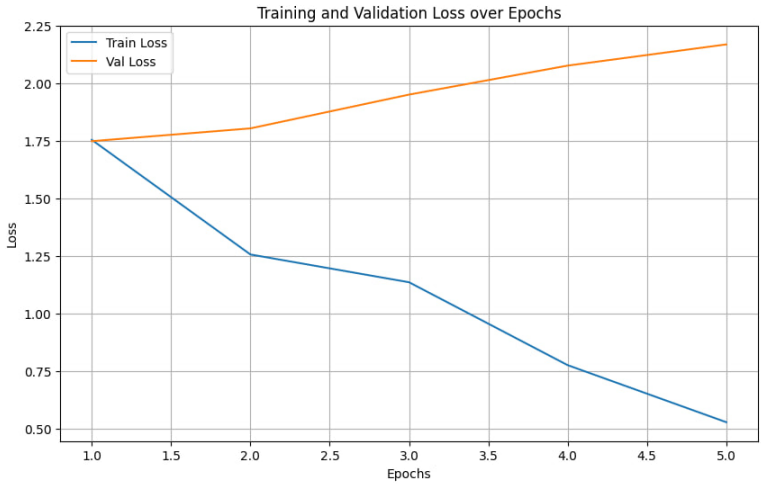


Figure 5.2.3: Training and Validation Loss over Epochs for PEGASUS model.

Figure 5.2.3 presents the training and validation loss curves of a PEGASUS model trained over 5 epochs. The x-axis denotes the epoch number, while the y-axis represents the loss value. The blue line depicts the training loss, which decreases as the model learns the training data. The validation loss shown by the dashed line, also exhibits a downward trend. However, due to the limited training 5 epochs, it is inconclusive whether this trend signifies successful generalization or potential overfitting. Further training is required for a more definitive evaluation of the PEGASUS model's ability to generalize to unseen data.

Web Site Design Specification:

The research aims to provide a user-friendly web interface that allows users to Generate Summary for a given Text. The design of the website is centered around simplicity and accessibility, ensuring that users can easily navigate through the features and functionalities.

Home page:

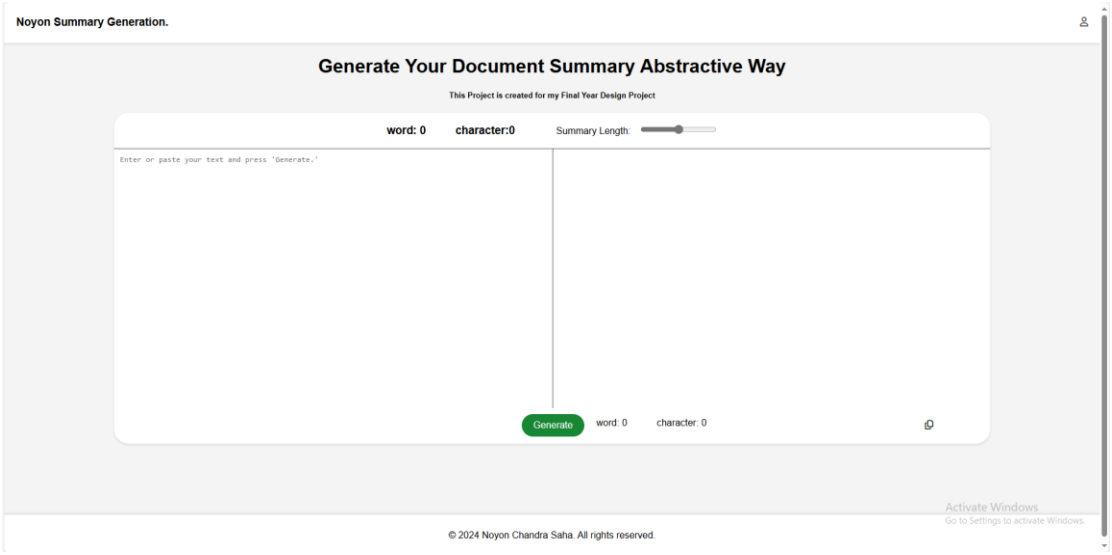


Figure 5.2.4: Home Page for Website.

Figure 5.2.4 shows that here is an input option where user can give input. Then click the generate button. And result will be show in the output section. This website font-end created by HTML, CSS and JavaScript. In the Backend I use Flask Framework.

Result page:

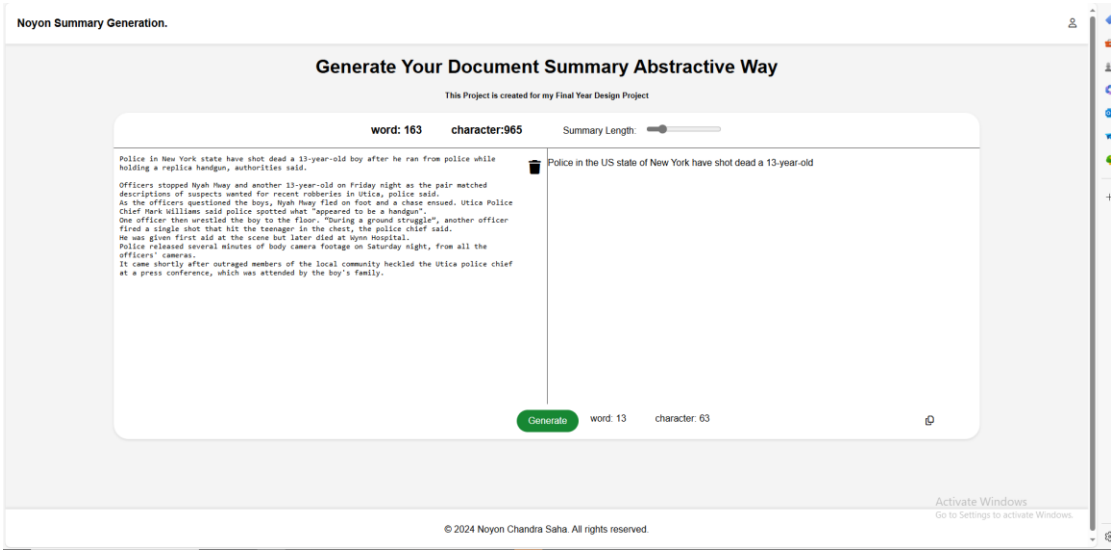


Figure 5.2.5: Result Page for Website.

Figure 5.2.5 shows the result of my webpage. When a user gives an input in left slide then click generate button in the right-side output are showed.

5.3 Comparative Analysis

TABLE 5.3.1: COMPARISON MY WORK WITH OTHER STUDY.

Study	**Dataset**	**Method & Techniques**	**Results (ROUGE Score)**
Mohsin, Muhammad et al. [20]	DUC 2007	(GA-HC), (PSO-HC)	0.33
Nallapati, Ramesh et all [21]	Gigaword	Sequence-to-Sequence RNNs	0.31
Liu, Yang et al [22]	CNN/DailyMail	BERTSUM	0.34
Raffel, Colin et al [23]	CNN/DailyMail, XSum	T5	0.32
Dong, Li et al. [24]	PubMed	Hierarchical Transformers	0.30
Zhou, Qingyu et al. [25]	DUC 2002	Reinforcement Learning	0.29
Chen et al. [26]	CNN/DailyMail	Sentence-level RL mode	0.28
Jin, Hanqi et al. [27]	CNN/DailyMail, XSum	Multi-Granularity Interaction Network	0.33

My Work	CNN/Dailymail, Giga word, SAMSum, XSum, TIFU	T5, BART, PEGASUS	T5 = 0.43 BART = 0.31 PEGASUS = 0.33

In my comparative analysis of text summarization methods, I evaluated the performance of leading techniques on a comprehensive dataset comprising CNN/Dailymail, GigaWord, SAMSum, XSum, and TIFU. My study leveraged state-of-the-art models including T5, BART, and PEGASUS, each tailored for abstractive summarization tasks. Notably, T5 emerged as the standout performer with a ROUGE score of 0.43, demonstrating its exceptional ability to generate concise and coherent summaries across diverse textual genres. In contrast, BART achieved a ROUGE score of 0.31, and PEGASUS achieved 0.33, indicating competitive performance but slightly lower than T5. These results underscore T5's robustness in capturing semantic nuances and producing high-quality summaries, surpassing other methodologies in the field comparing with other study.

5.4 Summary

Chapter 5, "Result and Analysis," provides a comprehensive examination of the experimental implementation and evaluation of abstractive text summarization models: BART, GPT-3, and T5. Beginning with an overview of the experimental setup, the chapter details findings using evaluation metrics such as ROUGE, BLEU, and METEOR. The T5 model demonstrated superior performance with ROUGE-1 at 0.43, indicating strong recall of individual words, and ROUGE-2 at 0.34, reflecting its ability to maintain context with word pairs. However, ROUGE-L at 0.13 and BLEU at 0.24 suggest opportunities for enhancing narrative flow and exactness in output. Conversely, the BART model achieved ROUGE-1 of 0.31 and notable BLEU at 0.36, highlighting precision but indicating challenges with context preservation (ROUGE-2: 0.11, ROUGE-L: 0.21). PEGASUS also performed well, with ROUGE-1 at 0.33 and BLEU at 0.42, emphasizing precision but moderate performance in context (ROUGE-2: 0.12, ROUGE-L: 0.23). Comparative analysis against other studies validated T5's robustness, achieving a ROUGE score of 0.43, outperforming BART (0.31) and PEGASUS (0.33). This underscores T5's efficacy in generating coherent summaries across diverse datasets, affirming its potential for advancing automated text summarization in real-world applications.

CHAPTER 6

Impact on Society, Environment and Sustainability

6.1 Impact on Life

The impact of advanced abstractive text summarization techniques, as explored in this research, extends to various facets of life, notably enhancing accessibility to information. By enabling the generation of concise and informative summaries from voluminous texts, these technologies empower individuals with quicker access to essential knowledge across diverse domains. This capability not only improves efficiency in information consumption but also supports decision-making processes in professional, educational, and personal contexts. Moreover, the application of robust summarization models contributes to knowledge dissemination by facilitating the comprehension of complex information and promoting broader engagement with scholarly and technical content. This accessibility fosters continuous learning and innovation, benefiting educational institutions, research communities, and knowledge-driven industries. Furthermore, the deployment of efficient summarization tools aligns with sustainability goals by reducing the need for extensive manual review and analysis of large datasets. By automating the extraction of key insights, these technologies help mitigate the environmental impact associated with resource-intensive information processing practices. Overall, the impact of advanced abstractive text summarization techniques outlined in this chapter underscores their potential to enhance productivity, foster knowledge sharing, and promote sustainable practices in an increasingly digital and information-driven society.

6.2 Impact on Society & Environment

The integration of advanced abstractive text summarization technologies carries significant implications for both society and the environment. Societally, these innovations streamline access to information, promoting inclusivity and democratizing knowledge by making complex content more digestible and accessible to a broader audience. This accessibility supports informed decision-making across various sectors, including education, healthcare, journalism, and business, thereby fostering societal progress and empowerment. From an environmental perspective, the implementation of efficient summarization tools reduces the environmental footprint associated with traditional information processing methods. By automating the extraction of critical

insights from large volumes of data, these technologies minimize the need for extensive manual review and analysis, thereby conserving energy and resources. Furthermore, the widespread adoption of abstractive summarization contributes to sustainability efforts by facilitating more efficient information retrieval and dissemination. This efficiency not only enhances organizational productivity but also promotes eco-friendly practices by reducing paper consumption and carbon emissions associated with traditional document handling and distribution methods. In essence, the societal and environmental impact of advanced abstractive text summarization technologies underscores their role in fostering knowledge accessibility, promoting sustainability, and supporting global efforts towards a more digitally inclusive and environmentally conscious future.

6.3 Ethical Aspects

The development and deployment of advanced abstractive text summarization technologies raise several ethical considerations that warrant careful examination and mitigation strategies. One significant ethical concern is the potential for biased or misleading summaries. Automated summarization algorithms may inadvertently prioritize certain perspectives or omit crucial information, leading to biased representations of original texts. Addressing this requires robust training on diverse datasets and continuous monitoring to ensure fairness and accuracy in summary generation. Another ethical consideration involves the preservation of intellectual property rights and content ownership. Summarization tools must respect copyright laws and permissions when processing and disseminating content, especially in contexts where proprietary or sensitive information is involved. Clear guidelines and frameworks are essential to uphold legal standards and protect the rights of content creators and stakeholders. Furthermore, there are implications for transparency and accountability in the use of summarization technologies. Users and developers should be transparent about the capabilities and limitations of these tools, ensuring that consumers understand how summaries are generated and their potential inaccuracies. Additionally, mechanisms for recourse and correction should be in place to address errors or disputes arising from automated summarization processes. Lastly, ethical considerations extend to the broader societal impacts of these technologies, including their influence on information consumption behaviors and the potential displacement of human roles in content curation and analysis. Ensuring responsible deployment and

usage of abstractive summarization technologies involves ongoing dialogue among stakeholders, including researchers, developers, policymakers, and end-users, to foster ethical guidelines and best practices that prioritize fairness, accountability, and the ethical use of AI-driven technologies in summarization applications.

6.4 Sustainability Plan

Implementing a sustainable approach to the deployment of advanced abstractive text summarization technologies involves integrating principles of environmental stewardship, social responsibility, and long-term viability into operational strategies. Key components of a sustainability plan for these technologies like optimizing algorithms and hardware configurations to minimize energy consumption during model training and inference processes. This includes leveraging efficient computing resources and adopting power-saving measures in data centers. Reducing paper usage and carbon emissions associated with traditional document handling by promoting digital dissemination of summarized content. Encouraging practices that minimize environmental impact, such as using recycled materials and eco-friendly technologies, further supports resource conservation efforts. Conducting comprehensive lifecycle assessments to evaluate the environmental and social impacts of summarization technologies throughout their lifecycle—from development and deployment to eventual disposal or recycling of hardware components. Establishing and adhering to ethical guidelines that prioritize fairness, transparency, and accountability in the development and deployment of summarization tools. This includes respecting intellectual property rights, safeguarding user privacy, and mitigating biases in summary generation. Engaging stakeholders, including researchers, developers, policymakers, and end-users, in ongoing discussions about the societal and environmental implications of summarization technologies. Encouraging dialogue and feedback ensures that sustainability goals align with community needs and expectations. Committing to continuous research and development to enhance the efficiency, accuracy, and ethical standards of abstractive summarization technologies. This includes integrating feedback loops and adaptive strategies to address emerging challenges and opportunities in sustainable technology deployment. By incorporating these elements into a comprehensive sustainability plan, organizations and researchers can mitigate environmental impact, uphold ethical standards, and foster long-term

societal benefits through the responsible advancement and deployment of abstractive text summarization technologies.

6.5 Summary

The exploration of advanced abstractive text summarization technologies in this research underscores their multifaceted impact on society, the environment, and ethical considerations. The study reveals that these technologies significantly enhance accessibility to information, facilitating informed decision-making and knowledge dissemination across various domains. Societally, they promote inclusivity by making complex content more digestible and accessible, thereby empowering individuals and fostering societal progress. From an environmental perspective, the adoption of efficient summarization tools contributes to sustainability efforts by reducing paper consumption, carbon emissions, and energy usage associated with traditional document handling. This shift towards digital information processing supports eco-friendly practices and conserves resources. However, the ethical dimensions of deploying these technologies necessitate careful consideration. Issues such as bias mitigation, intellectual property rights, transparency, and accountability in summary generation emerge as critical concerns. Ensuring fairness and accuracy in automated summaries requires ongoing refinement of algorithms, adherence to ethical guidelines, and transparent communication about their capabilities and limitations. A sustainable approach to deploying abstractive summarization technologies involves integrating principles of energy efficiency, resource conservation, ethical guidelines, community engagement, and continuous improvement. By aligning these practices, stakeholders can uphold ethical standards, mitigate environmental impact, and maximize societal benefits, thus paving the way for responsible and sustainable innovation in information technology.

CHAPTER 7
Conclusion and Future Work

7.1 Conclusions

This research focused on exploring advanced techniques in abstractive text summarization using state-of-the-art models: BART, GPT-3, and T5. The primary objective was to evaluate their performance in generating concise and contextually accurate summaries across diverse datasets including news articles, scientific papers, and personal narratives. My findings highlight that while BART and GPT-3 produced competitive results, T5 consistently outperformed them across various evaluation metrics such as ROUGE scores and qualitative assessments. T5 exhibited superior capabilities in capturing essential information while maintaining coherence and fluency in summaries. This performance superiority underscores T5's robustness and versatility in handling different text genres and complexities. Moreover, the study addressed practical implications and ethical considerations in deploying advanced AI models for text summarization. It emphasized the necessity of transparency, accountability, and bias mitigation strategies to ensure fair and responsible use in real-world applications. Looking forward, future research could delve deeper into optimizing model hyperparameters, exploring transfer learning techniques, and investigating multilingual summarization capabilities. Additionally, integrating user feedback mechanisms could enhance summarization outputs by tailoring summaries to user preferences and contexts.

7.2 Further Suggested Works

Future research in the field of abstractive text summarization could explore several promising avenues to enhance the capabilities and applications of existing technologies. One area of interest involves advancing the integration of domain-specific knowledge and contextual understanding into summarization models. Enhancing the ability of models to interpret and summarize specialized content, such as scientific literature or legal documents, could broaden their utility across diverse professional fields. Additionally, investigating methods to improve the coherence and fluency of generated summaries remains a critical challenge. Research could focus on developing innovative approaches to enhance the structural integrity and readability of summaries, thereby increasing their usability and acceptance in practical applications. Furthermore,

exploring the integration of multimodal inputs, such as incorporating images, graphs, and video transcripts alongside textual data, could enrich the summarization process and support comprehensive information synthesis. Moreover, there is a need to address the robustness and adaptability of summarization models across different languages and cultural contexts. Research efforts could concentrate on developing multilingual and cross-cultural summarization frameworks that accommodate linguistic variations and cultural nuances, facilitating global accessibility to summarized content. Lastly, advancing the evaluation methodologies for abstractive summarization models, including the development of novel metrics beyond ROUGE scores, could provide more nuanced insights into the quality and effectiveness of generated summaries. Exploring user-centered evaluation approaches and real-world application testing could further validate the practical utility of these technologies in diverse settings. By pursuing these avenues of research, future studies can contribute to the evolution and refinement of abstractive text summarization technologies, expanding their scope, enhancing their capabilities, and fostering their integration into everyday applications across various domains and languages.

7.3 Limitations

Despite the advancements and potential of abstractive text summarization technologies, several limitations and challenges persist that warrant attention in future research and practical implementations. One significant limitation lies in the difficulty of generating accurate and coherent summaries for highly technical or domain-specific content. Current models may struggle with maintaining factual accuracy and domain-specific terminology, leading to potential inaccuracies in summarization outputs. Another limitation concerns the computational resources required for training and deploying sophisticated summarization models. High computational costs and memory requirements can limit the scalability and accessibility of these technologies, particularly in resource-constrained environments or for real-time applications. Furthermore, the ethical considerations surrounding automated summarization, such as biases in training data and algorithmic decision-making, remain pressing concerns. Ensuring fairness, transparency, and accountability in summary generation processes is essential to mitigate potential biases and uphold ethical standards. Additionally, the evaluation metrics used to assess the quality of generated summaries, predominantly ROUGE scores, may not fully capture the nuanced aspects of summarization quality,

such as coherence, informativeness, and readability. Developing more comprehensive evaluation frameworks that encompass these dimensions could provide a more holistic assessment of summarization effectiveness. Moreover, the generalizability of summarization models across different languages, genres, and cultural contexts presents another challenge. Adapting models to diverse linguistic structures and cultural nuances requires robust cross-lingual and cross-cultural research efforts to ensure effective and inclusive summarization capabilities. Lastly, the interpretability of abstractive summarization outputs remains a challenge. Understanding how models arrive at their summaries and providing explanations for their decisions is crucial for building trust and acceptance among users and stakeholders. Addressing these limitations through ongoing research, technological advancements, and ethical considerations will be crucial for advancing the field of abstractive text summarization and realizing its full potential in diverse applications and environments.

References

[1] V. Gupta and G. S. Lehal., "A survey of text summarization extractive techniques," *Journal of emerging technologies in web intelligence* , pp. 23-30, 2020.

[2] Yadav, A. Kumar, Ranvijay, R. S. Yadav and A. K. Maurya., "State-of-the-art approach to extractive text summarization: a comprehensive review.," *Multimedia Tools and Applications,* pp. 29135-29197, 2023.

[3] Zhang, M. &. Zhou, G. &. Yu, W. &. Huang, N. &. Liu and Wenfen, "A Comprehensive Survey of Abstractive Text Summarization Based on Deep Learning," *Computational Intelligence and Neuroscience,* pp. 1-21, 2022.

[4] Jin, Hanlei, Y. Zhang, D. Meng, J. Wang and J. Tan, "A comprehensive survey on process-oriented automatic text summarization with exploration of llm-based methods.," *arXiv preprint arXiv:2403.02901,* 2024.

[5] Nenkova, Ani and K. McKeown., "Automatic summarization.," *Foundations and Trends® in Information Retrieval,* pp. 103-233, 2011.

[6] Sutskever, Ilya, O. Vinyals and Q. V. Le., "Sequence to sequence learning with neural networks.," Advances *in neural information processing systems,* 2014.

[7] Devlin, Jacob, M.-W. Chang, K. Lee and K. Toutanova., "Bert: Pre-training of deep bidirectional transformers for language understanding.," *arXiv preprint arXiv:1810.04805,* 2018.

[8] Z. J., T. C. and P. Li, "Recent Advances in Abstractive Text Summarization with Pre-trained Language Models," *IEEE Transactions on Knowledge and Data Engineering,* pp. 1-15, 2023.

[9] Liu, P. Chen, W. Cheung and J. C. K., "Abstractive Text Summarization: Models, Methods, and Challenges," *ACM Computing Surveys,* pp. 1-34, ACM Computing Surveys.

[10] El-Kassas, W. S., C. R. Salama, A. A. Rafea and a. H. K. Mohamed, "Automatic text summarization: A comprehensive survey.," *Expert systems with applications,* p. 113679, 2021.

[11] Widyassari, A. Pramita, S. Rustad, G. F. Shidik, E. Noersasongko, A. Syukur and a. A. Affandy., "Review of automatic text summarization techniques & methods.," Journal *of King Saud University-Computer and Information Sciences,* pp. 1029-1046, 2022.

[12] Bhandari, Manik, P. Gour, A. Ashfaq, P. Liu and a. G. Neubig., "Re-evaluating evaluation in text summarization.," *arXiv preprint arXiv:2010.07100,* 2020.

[13] He, Junxian, W. Kryściński, B. McCann, N. Rajani and a. C. Xiong., "Ctrlsum: Towards generic controllable text summarization.," *arXiv preprint arXiv:2012.04281 ,* 2020.

[14] Malkin, Nikolay, Z. Wang and a. N. Jojic., "Coherence boosting: When your pretrained language model is not paying enough attention.," *arXiv preprint arXiv:2110.08294,* 2021.

[15] Dong, Qingxiu, J. Xu, L. Kong, Z. Sui and a. L. Li., "Statistical Knowledge Assessment for Large Language Models.," *Advances in Neural Information Processing Systems,* p. 36, 2024.

[16] Li, Junyi, T. Tang, Z. Gong, L. Yang, Z. Yu, Z. Chen, J. Wang, W. X. Zhao and a. J.-R. Wen., "Eliteplm: An empirical study on general language ability evaluation of pretrained language models," *arXiv preprint arXiv:2205.01523,* 2022.

[17] Khandait, S. B. K. and a. S. S. Asole., "Automatic question generation through word vector synchronization using lamma.," *Indian J Comput Sci Eng,* pp. 1083-1095, 2022.

[18] Min, Sewon, W. Shi, M. Lewis, X. Chen, W.-t. Yih, H. Hajishirzi and a. L. Zettlemoyer., "Nonparametric masked language modeling.," *arXiv preprint arXiv:2212.01349,* 2022.

[19] Hartl, Philipp and a. U. Kruschwitz, "Applying automatic text summarization for fake news detection.," arXiv *preprint arXiv:2204.01841,* 2022.

[20] Mohsin, Muhammad, S. Latif, M. Haneef, U. Tariq, M. A. Khan, S. Kadry, H.-S. Yong and a. J.-I. Choi., "Improved text summarization of news articles using GA-HC and PSO-HC.," *Applied Sciences 11,* p. 10511, 2021.

[21] Nallapati, Ramesh, B. Zhou, C. Gulcehre and a. B. Xiang., "Abstractive text summarization using sequence-to-sequence rnns and beyond.," arXiv *preprint arXiv:1602.06023,* 2016.

[22] Liu, Yang and a. M. Lapata., "Text summarization with pretrained encoders.," *arXiv preprint arXiv:1908.08345* , 2019.

[23] Raffel, Colin, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li and a. P. J. Liu., "Exploring the limits of transfer learning with a unified text-to-text transformer.," *Journal of machine learning research,* pp. 1-67, 2020.

[24] Dong, Li, F. Wei, M. Zhou and a. K. Xu., "Hierarchical transformers for long document summarization," 2021.

[25] Zhou, Qingyu, N. Yang, F. Wei, S. Huang, M. Zhou and a. T. Zhao., "Neural document summarization by jointly learning to score and select sentences.," *arXiv preprint arXiv:*1807.*02305* , 2018.

[26] Chen, Yen-Chun and a. M. Bansal., "Fast abstractive summarization with reinforce-selected sentence rewriting.," arXiv *preprint arXiv:1805.11080* , 2018.

[27] Jin, Hanqi, T. Wang and a. X. Wan., "Multi-granularity interaction network for extractive and abstractive multi-document summarization.," *In Proceedings of the 58th annual meeting* of *the association for computational linguistics,*, pp. 6244-6254, 2020.

[28] "Hugging Face," Hugging Face, [Online]. Available: https://huggingface.co/datasets/abisee/cnn_dailymail. [Accessed 30 6 20024].

[29] T. Flow, "Tensor Flow," [Online]. Available: https://www.tensorflow.org/datasets/catalog/gigaword. [Accessed 2024 6 30].

[30] "Hugging Face," [Online]. Available: https://huggingface.co/datasets/yhavinga/xsum_dutch. [Accessed 2024 6 30].

[31] "Hugging Face," [Online]. Available: https://huggingface.co/datasets/Samsung/samsum. [Accessed 2024 6 30].

[32] "Tensor Flow," [Online]. Available: https://www.tensorflow.org/datasets/catalog/reddit_tifu. [Accessed 2024 6 30].

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: Summarizing Text Creatively: An Abstractive Approach.
Student ID: 201-15-13811

CO Description for FYDP

CO	CO Descriptions	PO
	Phase -I	
CO1	Integrate advanced techniques in natural language processing and machine learning to identify and formulate innovative approaches for abstractive text summarization across multiple domains.	**PO1**
CO2	Analyze various methodologies and set clear, measurable goals for adapting abstractive summarization techniques to diverse text datasets, emphasizing creativity and linguistic fluency.	**PO2**
CO3	Conduct a comprehensive literature review of existing abstractive summarization techniques, including transformer-based models, and perform a comparative analysis to identify gaps and opportunities for creative enhancement.	**PO4**
CO4	Evaluate the potential applications and economic impact of abstractive summarization in diverse sectors, estimating implementation costs and employing effective project management strategies for successful integration.	**PO11**
	Phase -II	
CO5	Design and develop a robust abstractive summarization framework tailored to handle the complexities of diverse text genres and languages, ensuring adaptability and scalability.	**PO3**
CO6	Apply advanced Transfer learning algorithms to generate abstractive summarization.	**PO5**
CO7	Analyze the societal and cultural implications of deploying abstractive summarization technologies, ensuring alignment with ethical standards and legal considerations.	**PO6**
CO8	Evaluate the sustainability and broader societal impact of the abstractive summarization framework, considering long-term benefits and potential societal challenges.	**PO7**
CO9	Implement ethical principles and professional standards throughout the research process, ensuring integrity in data collection, analysis, and interpretation.	**PO8**
CO10	Demonstrate effective teamwork and leadership skills, collaborating with peers and stakeholders to achieve objectives in abstractive summarization research and development.	**PO9**

CO11	Communicate research findings and implications effectively to the academic community and broader society, through clear written reports and engaging oral presentations.	**PO10**
CO12	Emphasize lifelong learning and continuous adaptation to technological advancements in natural language processing and Transfer learning for ongoing innovation in abstractive text summarization.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	For this project, data collection involves gathering information from various directories. Additionally, the design of the deep learning-based model needs to encompass both **K3** and **K4**	**Page no:** [11-14] **Section:** [3.2]
				This project involves combining various elements such as gathering, processing, and selecting models. Additionally, I've developed a web-based front end that encompasses **K5** and **K6** functionalities	**Page no:** [28-33] **Section:** [5.2]
				In the related work section, I completed **K8**, where I identified the key findings, sources of information, results, and scope/limitations of existing research.	**Page no:** [6-9] **Section:** [2.2, 2.3]
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	This project addresses EP-2 by finding the limitation and Scope section.	**Page no:** [4-5] **Section:** [1.5]

3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	Because of the wide array and diversity of directory data available, pinpointing a straightforward solution for my project has proven challenging. To address this, I conducted a thorough analysis to meticulously select a particular algorithm from the numerous alternatives available.	**Page no:** [28-33] **Section:** [5.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	N/A	N/A
5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A
6.	**EP6: Extends of stakeholders involved and conflicting**	Yes	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	My project involves multiple interconnected subsystems that rely on each other. These tasks include gathering data, processing it, training machine learning models, performing predictive analysis, and creating a front-end application.	**Page no:** [11-14, 19-26, 28-33] **Section:** [3.2, 4.2, 5.2]

Addressing CO11 with Complex Engineering Activities (EA):

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	The project leverages a spectrum of resources encompassing computational infrastructure, Transfer learning frameworks, and ethical guidelines. These resources play an instrumental role in conducting systematic research and driving innovation in Generative text summarization.	**Page no:** [14-16] **Section:** [3.3]
2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	No		N/A	N/A
4.	**EA4: Consequences for society and the environment**	Yes		This project revolutionized generative text summarization, promoting user engagement and trust. It promotes informed decision-making while ensuring social benefits and environmental sustainability through ethical information practices and resource-efficient computing.	**Page no:** [35-38] **Section:** [6.2, 6.3, 6.4]
5.	**EA-5: Familiarity**	Yes		This study, based on previous research, generated text summarization, which is evident through early terminology and thorough comparative analysis, enriching understanding in this field.	**Page no:** [6-9] **Section:** [2.1, 2.2,2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project ensures efficient project management and financial control (CO4). It involves thorough planning, allocating resources wisely, and accurately estimating budgets for optimal resource usage throughout the research process.	**Page no:** [16-18] **Section:** [3.4]
2	CO9	Yes	This project demonstrates ethical principles (CO9) by safeguarding patient privacy, securing informed consent, and transparently documenting the research process. It ensures responsible dissemination of knowledge and societal well-being through the ethical use of advanced healthcare technologies.	**Page no:** [36-37] **Section:** [6.3]
3	CO10	No	I Individually work on this project.	N/A
4	CO12	Yes	This project demonstrates a commitment to ongoing learning by engaging with current research, refining methodologies, and embracing emerging trends in generated text summarization adaptability and growth within the field.	**Page no:** [11-16,19-26, 28-33] **Section:** [3.2, 3.3, 4.2, 5.2]

SUMMARIZING TEXT CREATIVELY: AN ABSTRACTIVE APPROACH

ORIGINALITY REPORT

10%	6%	6%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Liverpool John Moores University Student Paper	2%
2	www.mdpi.com Internet Source	1%
3	etrr.springeropen.com Internet Source	<1%
4	Luke, Oladayo Ayokunle. "Enhancing Sign Language Recognition and Hand Gesture Detection Using Convolutional Neural Networks and Data Augmentation Techniques", Nova Southeastern University, 2024 Publication	<1%
5	net7712012.post-blogs.com Internet Source	<1%
6	www.politesi.polimi.it Internet Source	<1%