**ASSIGNMENT # 1**

**SUBMITTED BY: SUMAN NAZIR**

**REGISTRATION NO: FA22-BCS-116**

**Date: 03 October 2024**

# **A Novel Approach to Highlight Text Summarization Using BART**

**Abstract**

Abstractive text summarization is used to generate brief but meaningful summaries that focus on the main points of the original text. However, this task can be difficult, especially when dealing with different types of data. This problem is even more prominent when summaries need to be both brief and informative. In this research, we worked on fine-tuning the **facebook/bart-large-xsum** model using a dataset specifically designed for highlight summaries, called **knkarthick/highlightsum**. Our aim was to train the model to create summaries that focus on the crux of the dataset. Other models often create summaries that are too long. This dataset focuses on summarizing articles into short highlight sentences, presenting a unique challenge for generating high-quality, informative summaries. The objective of the research was to fine-tune the model for highlight-based summarization, making the summaries shorter, clearer, and more informative. We evaluated our improvements using ROUGE, and we have made our code open-source so that others can copy and build on this work. This research is significant for the NLP field because it shows how transformer-based models can be adapted for more specialized summarization tasks. This makes it easier to generate summaries for large amounts of information, such as news articles, legal documents, or technical reports.

**Keywords**

1. Abstractive text summarization
2. Natural Language Processing
3. Fine-tuning
4. Highlight-based summaries
5. ROUGE evaluation
6. Automated content generation