**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. This problem is even more prominent when summaries need to be both concise and informative. In this study, we fine-tuned the **facebook/bart-large-xsum model**, a pre-trained sequence-to-sequence transformer, for the task of abstractive text summarization using the **knkarthick/highlightsum** dataset. The aim was to train the model to create summaries that focus on the key points, much like a set of highlights. This dataset focuses on summarizing articles into concise highlight sentences, presenting a unique challenge for generating high-quality, informative summaries. The objective of the research was to enhance the model's ability to produce coherent and accurate summaries, suitable for summarizing textual content across various domains. We judge our improvements using ROUGE, and we have made our code open-source so that others can replicate and build on this work. Additionally, the research provides a complete, executable codebase to facilitate reproducibility and further exploration by the research community. 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 scores
6. Automated content generation