Problem Statement : **Deep Learning for Text Summarization**

**Team Name :** The Artificial Minds (Group No: 19)

**Members Associated :**

Anushka Ghosh - 21052055

Ipsita Majhi - 2128115

Sujoy Dutta - 22052688

Aheeron Changmai - 2228004

Problem Description: The objective of this project is to implement a Deep Reinforcement Learning based system for text summarization and question answering. Given a source text document or chat history, the system should perform the following tasks:

Text Summarization: Generate a concise and coherent summary of the input text, capturing the essential information.

Web or Mobile Application: Develop a user-friendly application.

Approach Used : Transformers

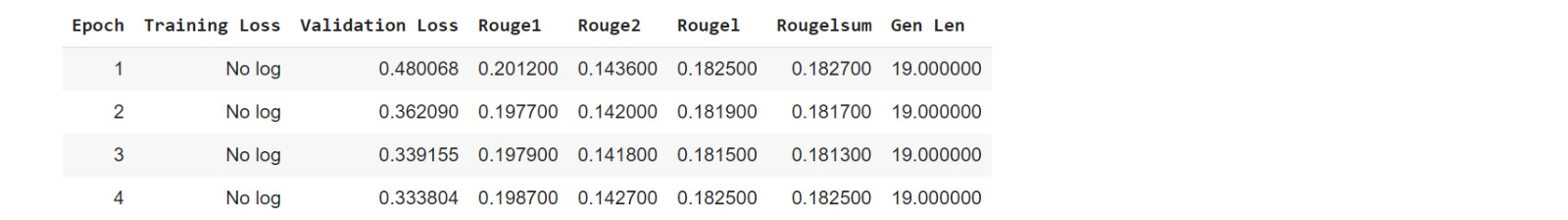
Models Used ; Hugging Face

[gopalkalpande/bbc-news-summary](REPORT_TXT_SUMMARIZATION.docx)

Content This dataset for extractive text summarization has four hundred and seventeen political news articles of BBC from 2004 to 2005 in the News Articles folder. For each articles, five summaries are provided in the Summaries folder. The first clause of the text of articles is the respective title

Evaluation Metrics :

1. ROUGE (Recall-Oriented Understudy for Gisting Evaluation) : ROUGE (Recall-Oriented Understudy for Gisting Evaluation) is a set of metrics used for the automatic evaluation of machine-generated text, especially in the context of text summarization. It measures the quality of a summary by comparing it to one or more reference summaries. ROUGE focuses on the recall of content in the generated summary concerning the reference summaries.



Metric Type:

This column specifies the ROUGE metric being reported, such as ROUGE-1 (unigram overlap), ROUGE-2 (bigram overlap), ROUGE-L (longest common subsequence), ROUGE-W (weighted word overlap), ROUGELsum (sentence overlap).

Precision, Recall, F1 Score:

For each ROUGE metric, there are typically three columns representing precision, recall, and F1 score. These values are calculated based on the overlapping n-grams (or other units) between the system-generated summary and the reference summary.

1. User-interactivity and User-friendliness.