

Folder Structure Overview:

1. Dataset_Sample_Test_Cases (CNN Daily Mail):

This folder contains sample datasets along with reference summaries, which are used for testing and evaluating the performance of text summarization models.

2. Models and Testing (Google Colab):

This folder includes the Google Colab notebooks where various text summarization models and algorithms have been tested. The models include:

- **Extractive Summarization:** TextRank and LexRank
- **Abstractive Summarization:** T5 and BART The purpose of this folder is to experiment with different techniques for both extractive and abstractive summarization.

3. requirements:

This file lists all the dependencies required to run the project. It is typically a requirements.txt file, which can be used to install the necessary packages and libraries.

4. Readme for App (langchain_streamlit.py):

This file provides instructions for running the application, offering a step-by-step guide to set up and launch the application. It details the setup, dependencies, and configuration steps.

5. langchain_streamlit_app:

This is the main codebase for the application, likely developed with Streamlit. It integrates LangChain (a framework for building applications with large language models) to facilitate text summarization tasks. This script handles user interactions and integrates the models for summarization.

6. langchain_textsum(datasettetesting):

This file contains the core application logic tailored for use in Google Colab. It focuses on testing the summarization models using different datasets, making it easy to run experiments and evaluate performance directly in a Colab environment.