# Text Summarization Project Documentation

# Introduction

# This project aims to implement a text summarization tool. Text summarization is the process of creating a concise and coherent summary of a longer text while preserving its main points. The summarization is done using natural language processing techniques and a pre-trained language model.

# Project Files

# The project consists of the following files:

# text\_summarization.py: This Python file contains the code for the text summarization algorithm. It uses the spacy library for natural language processing.

# index.html: This is the HTML file for the user interface where users can input the text they want to summarize.

# summary.html: This HTML file displays the original text and its summary after processing.

# Code Overview

# text\_summarization.py

# Libraries Used

# spacy: This library is used for natural language processing.

# STOP\_WORDS from spacy.lang.en.stop\_words: It provides a set of common English stop words.

# punctuation from string: It provides a string containing all ASCII punctuation characters.

# nlargest from heapq: It is used to find the largest elements in a list.

# Functions

# summarizer(document): This function takes a document as input and returns the summary, the original document in spaCy format, the length of the original text, and the length of the summary.

# Code Flow

# The script begins by importing necessary libraries and defining the document to be summarized.

# The summarizer function is defined. Inside this function, the following steps are performed:

# Stopwords and punctuation are imported and stored in variables.

# The spaCy English language model is loaded.

# The document is processed using spaCy.

# Tokens, word frequencies, and sentence tokens are extracted.

# Word frequencies are calculated, and each word's frequency is normalized.

# Sentence scores are computed based on word frequencies.

# A selection length is calculated for the summary.

# The summary sentences are selected and joined into a final summary.

# The function returns the summary, the original document in spaCy format, the length of the original text, and the length of the summary.

# index.html

# This HTML file contains a form where users can input the text they want to summarize. It includes a text area for input and a submit button.

# summary.html

# This HTML file displays the original text and its summary side by side. It includes placeholders {{original\_text}} and {{summary}} which are filled in with the actual text during runtime.

# **app.py**: This Python file serves as the backend of the text summarization tool. It connects the user interface with the text summarization algorithm implemented in text\_summarization.py

# How to Use

# Open the index.html file in a web browser.

# Enter the text you want to summarize in the text area.

# Click the submit button.

# The text\_summarization.py script processes the input and returns the summary.

# The summary.html page displays the original text and its summary. STUCTURE OF THE FILE DIC

# TEXT\_SUMMARIZATION\_WITH SPACY/

# │

# ├── \_\_pychache/

# │ └── templates/

# │ ├── index.html

# │ └── summary.html

# │

# ├── static/

# │ └── styles.css

# │

# │ ├── app.py

# | |──tempCodeRunnerFile.py

# │ └── text\_summarization.py

# │

# └── requirements.txt

HOW TO RUN THE PROGRAM

# NAVIGATE TO APP.PY FILE AFTER FINISHING ALL THE CODE REQURIED TO YOUR PROJECT

# USE python app.py command in terimal and hit the enter button

# Then we can see a link like this

# 

# ­­­

# You can see that a link was shown in the terminal go near by the link then you will get a message as follow to link link on it,then you can a page like this

# 

# Enter the text and hit the submit button the you will navigate into a page summary.html then their you can find the original text and summary on that text

# 

# Final page output of our project

# Conclusion

# This project provides a simple yet effective text summarization tool. It utilizes natural language processing techniques to generate concise summaries from longer texts. The user interface allows for easy input and visualization of the original text and its summary. This tool can be further extended and customized to suit specific use cases and requirements.

# GITHUB LINK :- https://github.com/KIET-NEST-PROJECTS-KBH/E-4F-419

# LINKEDLIN LINK:- https://www.linkedin.com/posts/pavan-kumar-b9634521a\_textsummarization-nlp-machinelearning-activity-7118313957089521664-ZaHf?utm\_source=share&utm\_medium=member\_desktop