## Introduction

Text summarization is a crucial task in natural language processing that involves condensing a lengthy document into a shorter version while preserving its essential information. Latent Semantic Analysis (LSA) is a powerful technique used for text summarization that leverages the underlying semantic structure of the text.

## Problem Statement

The problem we are addressing is the overwhelming amount of text data available, which makes it challenging to extract relevant information quickly. Traditional methods of reading through all the content are time-consuming and impractical, especially when dealing with large documents or multiple sources. This is where text summarization comes in as a valuable tool.

## Objective

1. Deploy a text summarization system using Flask.
2. Display the summarized text.
3. Implement TF-IDF, SVD, sentence scoring and summary extraction.

## Methodology

1. Requirement Identification
2. Literature Review
3. SMMRY

SMMRY.com is an online tool designed to provide concise summaries of longer texts. It uses an algorithm to identify the most critical parts of a document, highlighting key points and reducing the length of the original text significantly. The primary aim is to save users time by distilling essential information from articles, papers, and other lengthy documents.

1. SCRIBBR

Scribbr.com is an online platform known for its academic services, including proofreading, plagiarism checking, and citation generation. Recently, Scribbr has introduced text summarization tools to assist students and researchers in quickly digesting large volumes of academic content.

1. RESOOMER

Resoomer offers both extractive and partially abstractive summarization, aiming to identify and highlight the main ideas and key points from a given text. The tool supports multiple languages and is accessible via web browsers and browser extensions.

1. Requirement Analysis
2. Functional Requirements
3. Pre Processing
4. Text Summarization
5. Post Processing
6. User Interface
7. Non-Functional Requirements
8. Speed
9. Usability
10. Reliability

The use-case diagram for depicting the use-case model of the system is as follows:

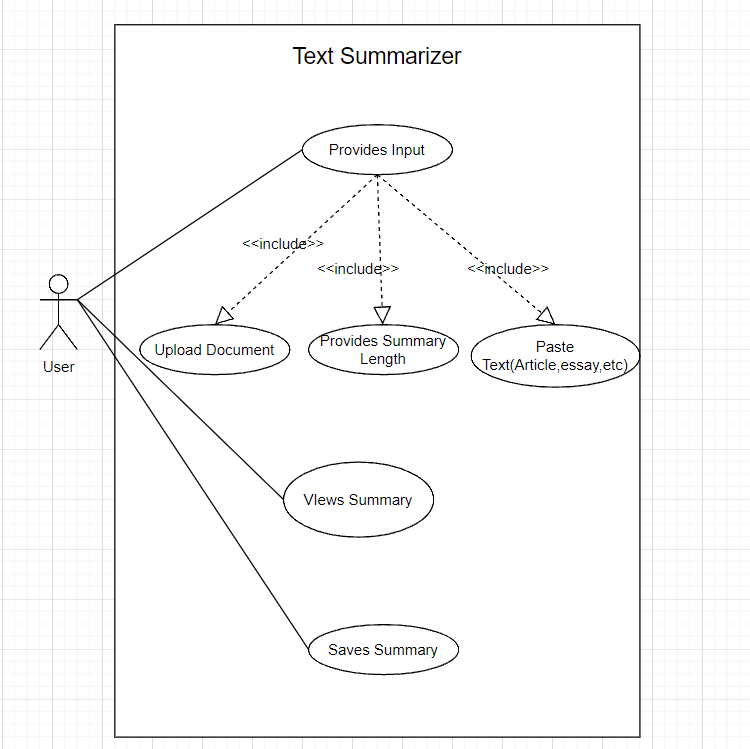


Fig: Use Case Diagram

Use Case Description Table:

|  |  |  |
| --- | --- | --- |
| S.N | Use-Case Identifier | UC005-Views Summary |
| 1 | Primary Actor | User |
| 2 | Secondary Actor |  |
| 3 | Description | The user can view the summary generated by the system. |
| 4 | Pre-Condition | The user has to provide input and wait for the system to generate the summary. |
| 5 | Post-Condition | The summary is generated with no of sentences specified by the user. |
| 6 | Failure Scenario | The input language is other than English. |

Feasibility Study

1. Technical Feasibility

The project can be accomplished by leveraging the knowledge and skills acquired from relevant coursework.

1. Operational Feasibility

The project will be managed and maintained by team members, requiring a moderate time commitment.

1. Economic Feasibility

All the resources required to execute this project are freely and easily available with no additional expenses.

1. Schedule Feasibility

The tentative schedule of the project is outlined using a Gantt Chart.

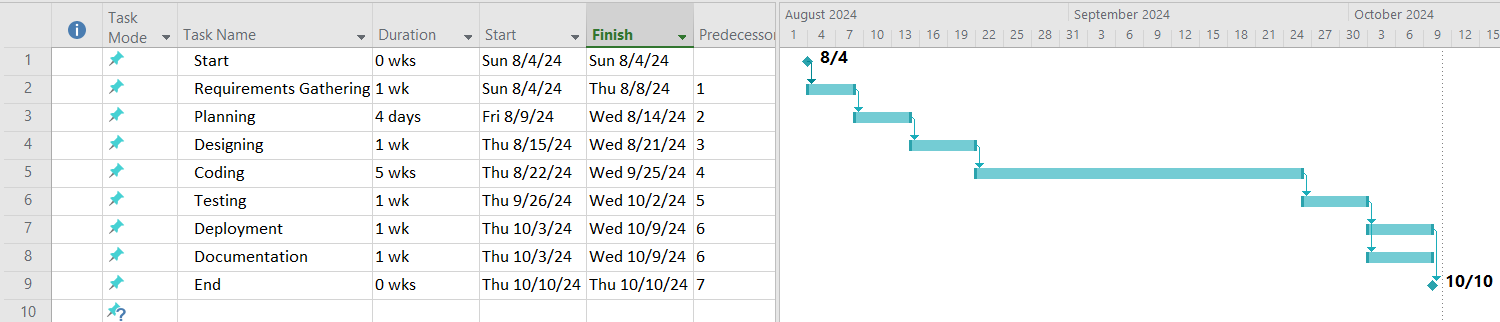


Figure: Work Breakdown Structure

Flowchart

**Expected** **Outcomes**

Figure: Wireframe Diagram

**References**

1. Ozsoy, Makbule & Alpaslan, Ferda & Cicekli, Ilyas. (2011). Text summarization using Latent Semantic Analysis. J. Information Science. 37. 405-417. 10.1177/0165551511408848.
2. Geetha J K and Deepamala N, "Kannada text summarization using Latent Semantic Analysis," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, 2015, pp. 1508-1512, doi: 10.1109/ICACCI.2015.7275826.
3. Q. Liu, J. Wang, D. Zhang, Y. Yang and N. Wang, "Text Features Extraction based on TF-IDF Associating Semantic," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China, 2018, pp. 2338-2343, doi: 10.1109/CompComm.2018.8780663.
4. G. Erkan and D. R. Radev, "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization," Journal of Artificial Intelligence Research, vol. 22, pp. 457-479, 2004. DOI: 10.1613/jair.1523.