

TRANSUMDOCS

Abstract

In today's rapidly evolving digital landscape, the ability to efficiently manage and comprehend vast amounts of information has become essential across various domains, including academia, business, and personal use. To address these challenges, this project proposes the development of an innovative web application designed to simplify the process of summarizing and translating content from diverse input sources, such as PDF documents, extensive text entries, and images containing intricate or dense text. The application focuses on two primary functionalities: summarization and translation. By leveraging advanced Natural Language Processing (NLP) techniques, the application will distill large volumes of text into concise and coherent summaries that retain the essential information while significantly reducing the length and complexity of the original content. Additionally, Optical Character Recognition (OCR) technology will be integrated to convert image-based content into readable text, ensuring users can effectively process and understand information from various formats, including scanned documents and photos. To further enhance the application's utility, the summarized content will be translated into multiple languages using the Google Translate API or similar services, ensuring high-quality, contextually accurate translations that make the information accessible to a global audience. The proposed solution integrates these cutting-edge technologies into a seamless, user-friendly platform where users can upload documents, process and summarize content, and receive translations—all within a single interface. By combining state-of-the-art NLP, OCR capabilities, and robust translation services, the application aims to streamline the information processing workflow, broaden the accessibility of textual data across different languages and formats, and contribute to a more inclusive and accessible digital communication environment for users worldwide.

Group 5:

Tesa Mariam Biju
Afsana Rassak
Karthika R
Abhinand Shaji