Introduction to Robotic Process Automation

TEXT SUMMARIZATION BOT

REG NO. :220701029
NAME :ARUN MOHAN S
GUIDE NAME :JINU SOPHIA
DESIGNATION :B.E COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

Abstract

This project showcases an advanced **Robotic Process Automation (RPA)** workflow designed to automate the extraction, summarization, and sharing of student-provided notes in PDF format. Utilizing UiPath's **Extract PDF** activity, the system extracts text from student-submitted PDFs. The extracted text is then processed through UiPath's **Generative AI** capabilities to produce a concise summary. This summarized content is stored locally for future access and is subsequently distributed to recipients via email using the **SMTP protocol**.

By automating the entire process—PDF text extraction, AI-driven summarization, local storage, and email distribution—this solution eliminates manual intervention, streamlining the management and dissemination of educational content. The project demonstrates the synergy of RPA and AI, improving efficiency and ensuring timely sharing of knowledge.

Need for the Proposed System

Increasing Volume of Education Content

- Automates the extraction and summarization of student notes, addressing the challenge of handling large volumes of content efficiently.

Reducing Manual Effort

- Eliminates time-consuming manual summarization through the use of **RPA** and **AI**, improving accuracy and consistency.

Improving Accessibility To Information

- Provides quick, concise summaries, enabling users to quickly grasp essential content and make informed decisions.

Scalability for Diverse Applications

- Adapts to various use cases—such as lecture notes and research papers—without the need for extensive customization.

Advantages of the Proposed System

- Automates student note handling with AI, reducing manual effort.
- Handles and summarizes notes in various languages
- Delivers seamless, real-time content summarization and sharing.
- Ensures meaningful and precise language processing.
- Provides continuous, dependable support with automation.

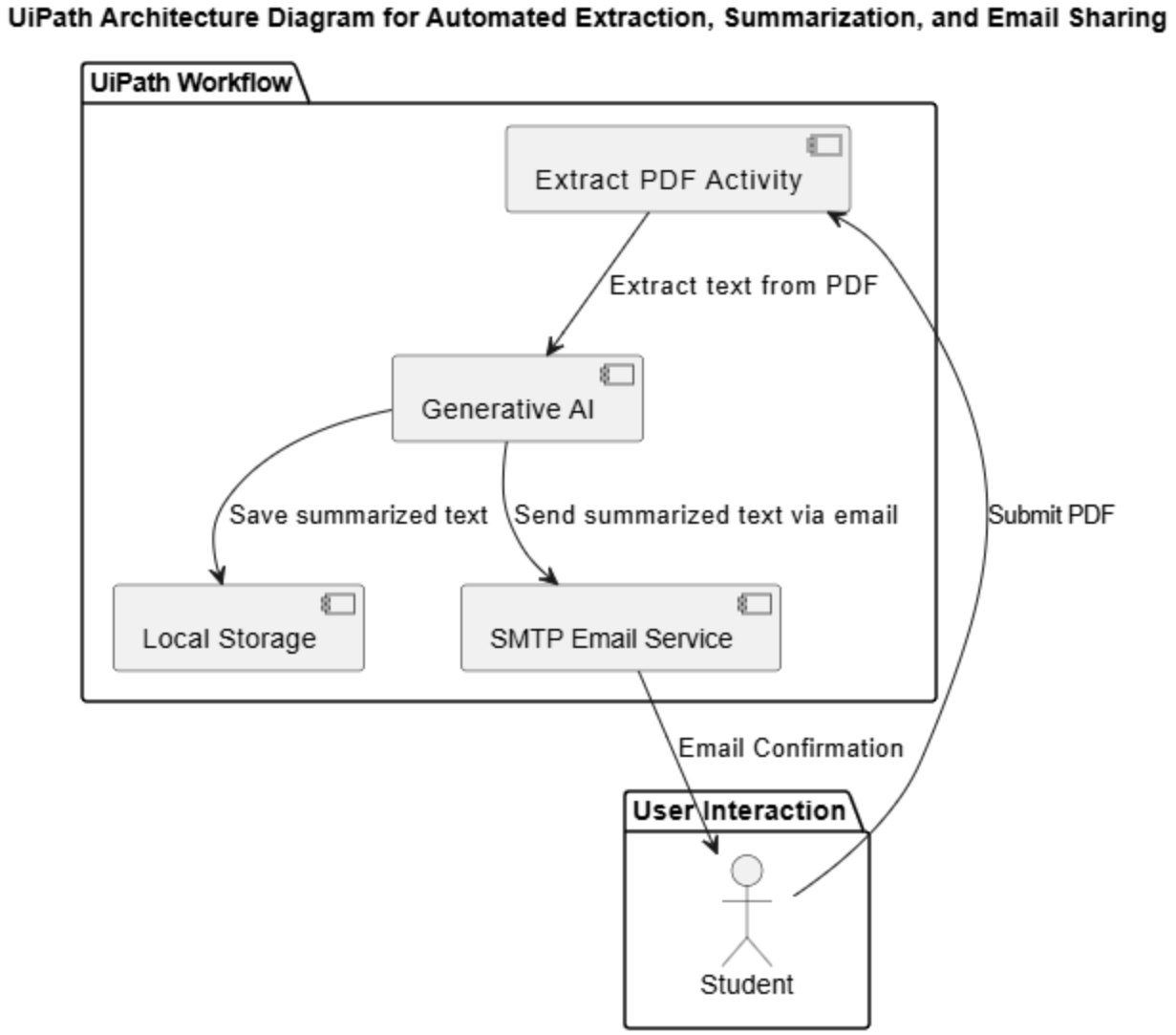
Literature Survey

- Paper 1
- Advantages
- Disadvantages
- Paper 2
- Advantages
- Disadvantages

Main Objective

The main objective of this project is to develop an automated system using UiPath that extracts, summarizes, and shares educational content, simplifying content management for students and administrators. Leveraging advanced AI and automation, the system aims to provide accurate and efficient text processing, including multilingual support, ensuring accessibility and adaptability across diverse use cases. By automating these processes, it facilitates seamless communication, enhances learning, and provides a reliable and scalable solution for managing educational materials effectively.

Architecture



System Requirements

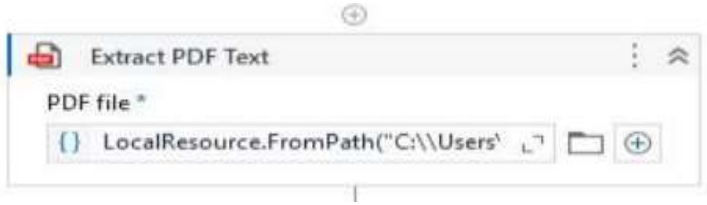
- **Hardware**
  - **Processo**r: Intel Core i5 (Min), Core i7 (Rec)
  - **RAM**: 8 GB (Min), 16 GB (Rec)
  - **Storage:** 250 GB SSD (Min), 500 GB SSD (Rec)
  - **GPU (**Optional): NVIDIA RTX 3060 or higher (for NMT)
  - **Network**: Stable broadband (10 Mbps+)

- **Software**
  - **OS**: Windows 10/11, Ubuntu 20.04+, macOS
  - **APIs**: Google Translate API, OpenAI GPT, Microsoft Translator

Functional Description

Name of the Module 1: Pdf Text Extraction Module

- This module is responsible for extracting text content from PDF files submitted by students. Using UiPath's Extract PDF activity, the system processes studentprovided PDF documents, regardless of format or structure, to retrieve textual data. It ensures that the extracted complex layouts, and prepares it for further processing.



Name of the Module 2: Email Sending Module

- This module facilitates the automated sharing of summarized content with designated recipients via email. Using SMTP protocols, the system formats and sends the summarized text to specified email addresses. It ensures seamless communication by automating the email-sending process, including error handling and confirmation of delivery.

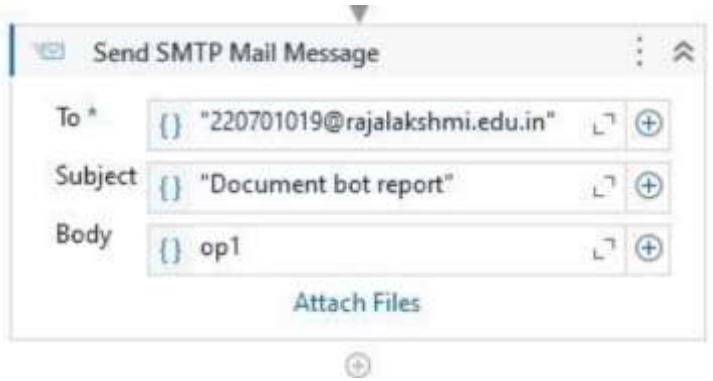
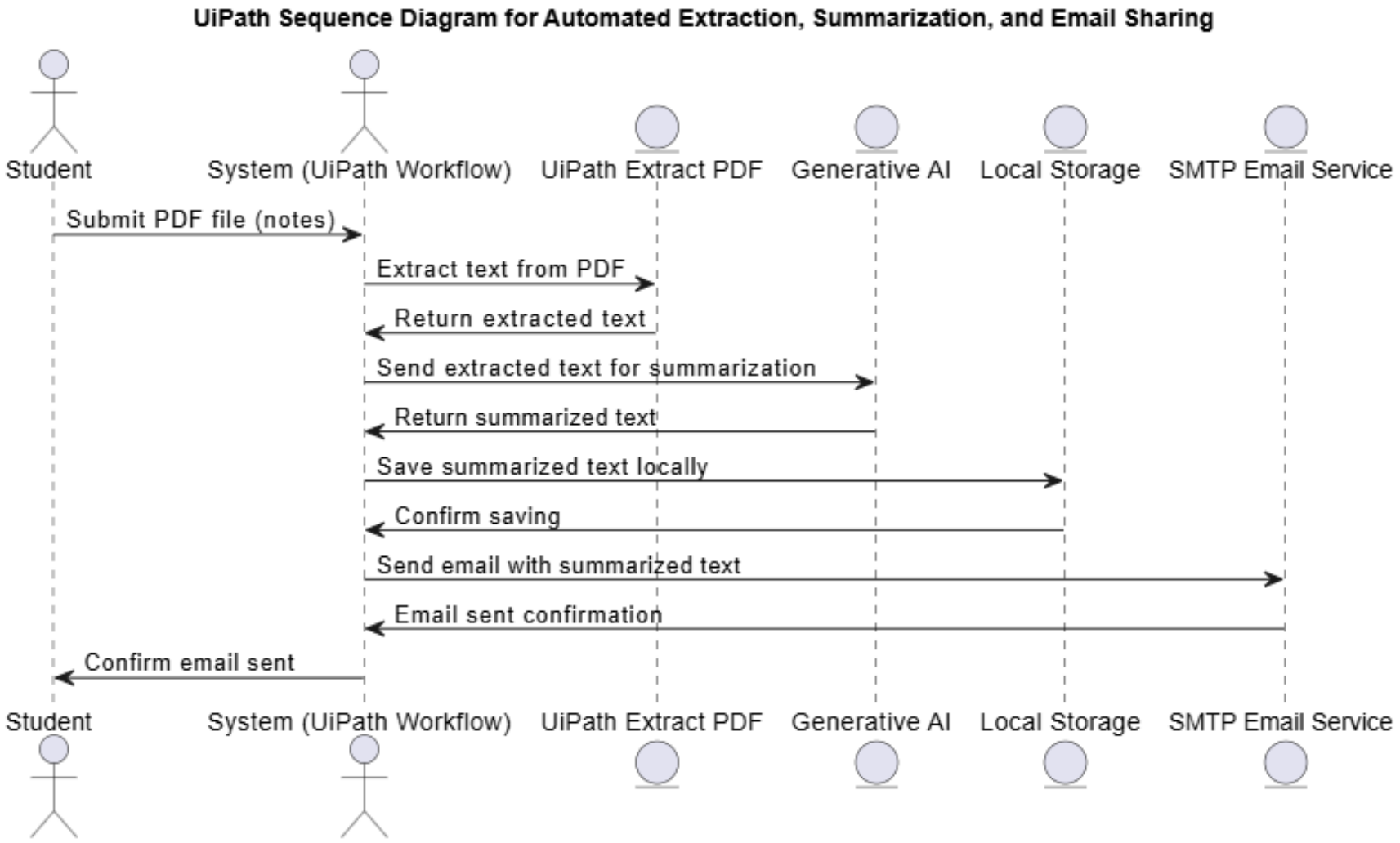
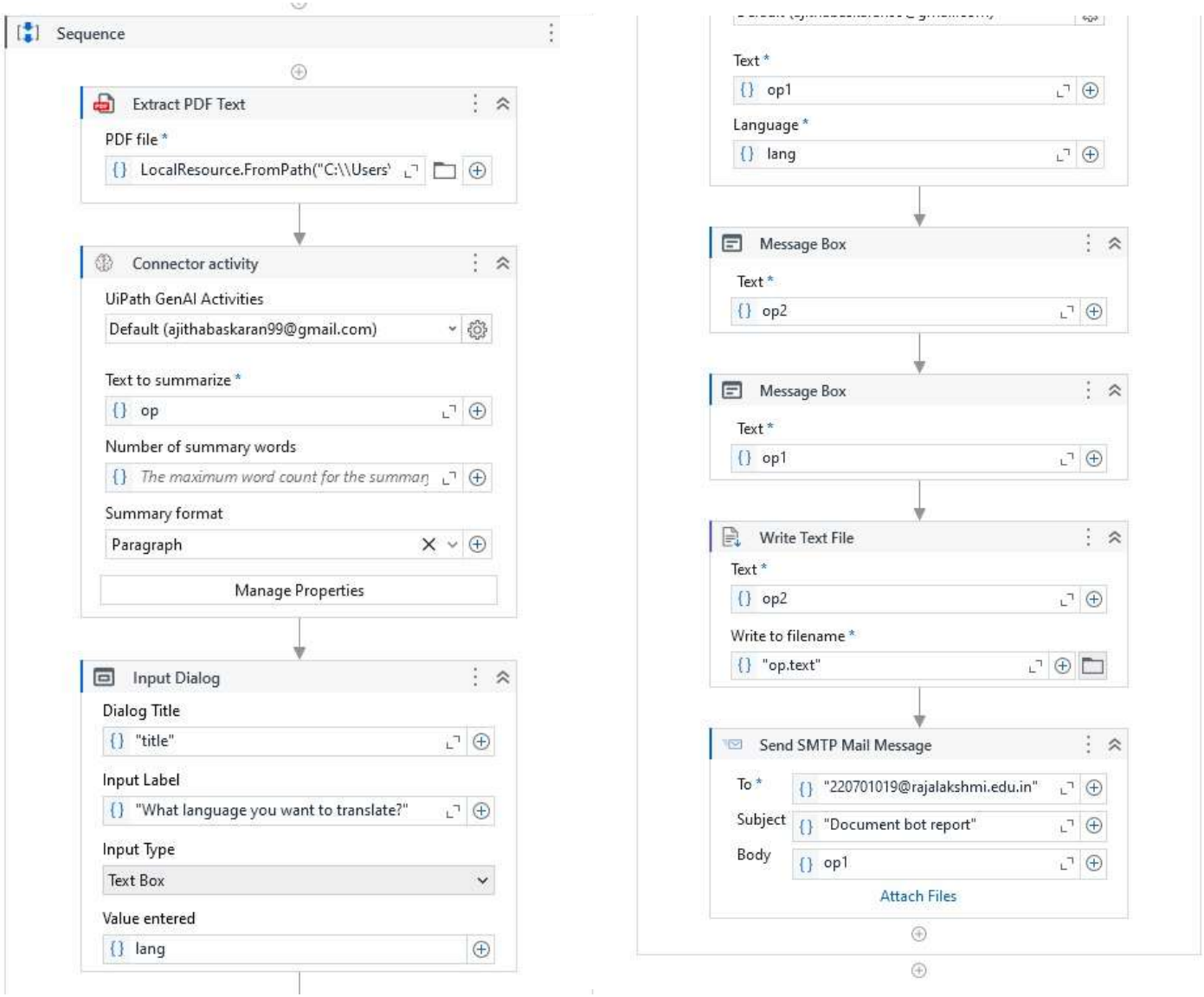


Table Design



Process Design



Implementation

- Implementation of Module 1
  - Message of summarized text:

- Implementation of Module 2
  - Sending the summarized text through email:

Testing

- Description
- Screen shots

Conclusions

This project successfully demonstrates the automation of extracting, summarizing, and sharing educational content from PDF files submitted by students. By leveraging UiPath's capabilities, such as the Extract PDF activity and integration with Generative AI, the workflow provides an efficient, reliable, and scalable solution for managing large volumes of student notes. The automated process eliminates the need for manual intervention, reducing errors and improving overall productivity.

The integration of advanced modules like text summarization, local storage, and automated email sharing ensures that summarized content is concise, accessible, and effectively distributed to the intended recipients. Additionally, the implementation of error-handling mechanisms and admin controls enhances the system's robustness, making it suitable for real-world applications in educational institutions or organizations dealing with extensive document management.

References

https://medium.com/luisfredgs/automatic-text-summarization-with-machine-learning-an-overview-68ded5717a25

https://en.wikipedia.org/wiki/Automatic_summarization

https://www.kdnuggets.com/2019/11/getting-started-automated-text-summarization.html

Queries

Demonstration

Thank You