EMAIL INVOICE DOWNLOADER BOT

A PROJECT REPORT

Submitted by

NALIN KARTHIK K (220701178)

in partial fulfillment for the course

OAI1903 - INTRODUCTION TO ROBOTIC PROCESS AUTOMATION

for the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE
RAJALAKSHMI NAGAR
THANDALAM CHENNAI – 602105 NOVEMBER 2024

BONAFIDE CERTIFICATE

Certified that this project report **"EMAIL INVOICE DOWNLOADER BOT"** is the bonafide work of **"NALIN KARTHIK K (22070178)"** who carried out the project work for the subject OAI1903 - Introduction to Robotic Process Automation under my supervision.

SIGNATURE
Mrs. G. M. Sasikala, M.E.,
SUPERVISOR,
Assistant Professor, Department of
Computer Science and Engineering,
Rajalakshmi Engineering College,
Rajalakshmi Nagar,
Thandalam,
Chennai- 602105.

Submitted to Project and Viva Voce Examination for the subject

OAI1903 - Introduction to Robotic Process Automation held on ____.

Internal Examiner   **External Examiner**

ABSTRACT

This project leverages Robotic Process Automation (RPA) with UiPath to automate the retrieval and storage of invoice attachments from email. The system streamlines the invoice management process by scanning incoming emails for specific keywords (e.g., "Invoice") and automatically downloading the relevant attachments. Once identified, attachments are saved locally in a specified file path, reducing manual effort and enhancing operational efficiency. This automation is designed to handle multiple emails and diverse attachment formats, ensuring a robust, scalable solution adaptable to various business requirements. The project incorporates error handling and logging mechanisms to maintain reliability and allow easy tracking of processed emails and downloaded invoices. By implementing this solution, organizations can save time, minimize errors in invoice handling, and enable better workflow management for accounts and finance departments.

The project follows a structured workflow:

1. **Email Retrieval**: The system connects to an email inbox (e.g., Outlook or IMAP) to access incoming messages.
2. **Keyword Filtering**: Each email is checked for the keyword, and only matching emails are processed.
3. **Attachment Download**: Relevant attachments are downloaded and stored in a pre-defined local path.
4. **Error Handling and Logging**: The process includes robust error-handling mechanisms and logs each step, tracking successful downloads and noting any issues.

This automated approach significantly reduces manual effort in invoice management, minimizes the risk of errors, and provides a consistent and efficient solution for storing financial documents. It is easily adaptable to different business environments, ensuring scalability for various organizational needs, and enables the finance team to focus on higher-value tasks rather than repetitive administrative duties.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	ACCRONYM
RPA	Robotic Process Automation
SMTP	Simple Mail Transfer Protocol
IMAP	Internet Message Access Protocol
OCR	Optical Character Recognition
API	Application Programming Interface

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

In today's digital landscape, automating routine business processes is crucial for efficiency and productivity. One of the most repetitive yet essential tasks within organizations is managing financial documents, specifically invoices. Traditionally, invoice management involves manually monitoring email inboxes, identifying relevant documents, downloading attachments, and organizing them in a structured manner. However, this approach is time-consuming, prone to error, and requires significant manual intervention. To address these challenges, this project implements a Robotic Process Automation (RPA) solution using UiPath to streamline the process of downloading, storing, and managing invoice attachments from email systems.

The main objective of this project is to automate the collection and storage of invoices from email attachments based on specific keywords, reducing the time and resources required for these tasks. The UiPath workflow is designed to connect to an email inbox, filter emails that contain keywords such as "Invoice" in the subject or body, and save the attachments from matching emails in a designated local folder. This process is efficient and accurate, allowing organizations to manage invoices more effectively while minimizing the risk of missing or misplacing important financial documents.

This automated invoice retrieval solution offers significant benefits to organizations. First, it greatly reduces manual work, freeing up resources for higher-value tasks. Second, the automation ensures consistency and reduces errors that might arise from manual handling. Third, with the flexibility of UiPath, this solution can be tailored to meet the specific needs of different business environments, making it scalable and adaptable.

By leveraging RPA through UiPath, this project demonstrates how automation can optimize invoice management, resulting in increased operational efficiency, reduced errors, and improved document accessibility. In the long term, this solution not only saves time but also enhances data accuracy, allowing finance and accounting departments to better support organizational goals.

1.2 OBJECTIVE

The primary objective of this project is to develop an efficient and automated solution for downloading and organizing invoice attachments from email using Robotic Process Automation (RPA) with UiPath. By implementing this automation, the project aims to reduce manual workload, improve accuracy, and streamline the invoice management process for organizations, particularly for finance and accounting departments.

This project's objectives are as follows:

1. **Automate Invoice Retrieval from Email**:
   - Develop a UiPath workflow to automatically connect to an email inbox and monitor incoming emails, identifying those containing invoice attachments based on a specific keyword (e.g., "Invoice") in the subject line or body text.
   - Ensure compatibility with popular email protocols (e.g., IMAP, Outlook) to make the solution versatile and adaptable for various email systems.
2. **Accurate Filtering and Keyword Matching**:
   - Implement a keyword filtering mechanism to selectively process only relevant emails. This ensures that only invoices are downloaded, minimizing processing time and reducing the risk of irrelevant file downloads.
3. **Download and Organize Attachments**:
   - Develop a method within the workflow to download attachments from filtered emails and save them in a pre-specified local directory. This organization promotes easy access and structured storage for financial records.
   - Implement the ability to handle various attachment formats (e.g., PDF, image files) to accommodate diverse invoice types.
4. **Incorporate Error Handling and Logging**:
   - Integrate error-handling mechanisms to manage potential issues, such as network interruptions or unsupported file formats. This ensures the process remains robust and resilient.
   - Set up a logging system that records each step, capturing successful downloads and logging any errors, to facilitate audit trails and provide insights into system performance.
5. **Improve Efficiency and Reduce Manual Effort**:
   - Streamline the invoice processing workflow, reducing manual effort in locating and downloading invoices and decreasing errors associated with manual handling.
   - Enable finance teams to focus on strategic and analytical tasks rather than repetitive administrative duties, adding value to the organization.

By achieving these objectives, this project aims to offer a scalable, reliable, and user-friendly RPA solution that enhances efficiency, accuracy, and accessibility in invoice management. This automation will support organizations in optimizing their document handling processes, reducing costs associated with manual tasks, and enabling smoother financial operations.

1.3 EXISTING SYSTEM

In most organizations, invoice processing relies on manual or semi-automated methods, where finance or administrative personnel monitor email inboxes to identify and retrieve invoices. Typically, the process includes manually scanning emails, identifying messages containing invoices based on keywords in the subject or content, downloading attachments, and storing them in a designated local or cloud-based folder. This approach requires significant time, effort, and attention to detail, especially in high-volume email environments, which can result in increased labor costs and errors in document handling.

Challenges in the Existing System:

1. **Time-Consuming and Labor-Intensive**:
   - Manually monitoring email inboxes and identifying invoice-related messages is slow and repetitive. Employees must open each email to verify its relevance and download attachments, which consumes time and can delay downstream processes, especially when dealing with large volumes of emails.
2. **Prone to Human Error**:
   - Manual processes increase the likelihood of errors, such as misclassifying emails, overlooking attachments, or failing to save files in the correct location. Human error can lead to misplaced invoices, delays in payment processing, and financial discrepancies.
3. **Inefficiency in Data Handling and Organization**:
   - With manual methods, it is challenging to maintain a consistent organization of files. Employees may save attachments in different formats or folders, creating inconsistencies and making document retrieval more difficult. This can hinder the efficiency of finance teams and impede access to critical documents when needed.
4. **Limited Tracking and Transparency**:
   - Manual systems lack real-time tracking, making it difficult to monitor which invoices have been processed and which are pending. This can create bottlenecks, reduce transparency, and limit oversight into invoice processing status, especially when multiple team members are involved.
5. **Scalability Issues**:
   - As organizations grow, the volume of incoming invoices increases, making it unsustainable to rely on manual handling. High email volume can quickly overwhelm employees, increasing the risk of delayed or missed payments and reducing overall efficiency.

Need for Automation:

To address these challenges, automation offers a solution that significantly improves speed, accuracy, and organization. A robotic process automation (RPA) tool like UiPath can streamline this workflow by automatically scanning emails for specific keywords, retrieving relevant attachments, and storing them in a designated directory. With automated error handling and logging, the system also provides better transparency and traceability in document management.

1.4 PROPOSED SYSTEM

The proposed system harnesses the power of Robotic Process Automation (RPA) through UiPath to create a streamlined, automated solution for downloading and organizing invoice attachments from emails. This automation addresses the inefficiencies and error-prone nature of traditional manual invoice processing, providing finance and accounting teams with a faster, more reliable way to manage these crucial documents. By implementing UiPath, the system can seamlessly retrieve invoices from emails, store them in an organized local directory, and significantly reduce the need for repetitive manual work.

In this automated workflow, UiPath is configured to connect to the designated email inbox, such as Outlook or an IMAP-compatible provider, and automatically scan incoming messages for relevant content. The system actively monitors the inbox at regular intervals or on demand, ensuring that no email is overlooked. Filtering mechanisms use keywords, such as "Invoice" in the subject or body, to target only emails containing invoices, effectively reducing the processing load by ignoring irrelevant messages.

When the system detects an email containing an invoice, it automatically retrieves the attachment and saves it in a specified local directory. The directory structure is predefined for easy accessibility and standardized storage, allowing for efficient document organization and quick retrieval. This automated attachment handling can accommodate a variety of file types, including PDFs and images, ensuring the flexibility needed for different invoice formats.

CHAPTER 2 LITERATURE REVIEW

2.1 Survey on Robotic Process Automation (RPA) in Education

This survey explores the adoption, effectiveness, and challenges of Robotic Process Automation (RPA) in educational settings. As education institutions increasingly embrace digital transformation, RPA has emerged as a key tool for automating repetitive administrative tasks, thereby improving efficiency and reducing operational costs. The survey examines the current uses of RPA in areas like admissions processing, grading, student information management, and report generation. Respondents are asked about the specific RPA solutions implemented, the benefits observed (such as time saved and error reduction), and the challenges faced in deploying RPA, such as integration with legacy systems or limited technical expertise among staff. Additionally, the survey seeks to understand how educators and administrators perceive the role of RPA in enhancing student engagement and academic outcomes. Results from this survey will provide insights into how educational institutions can optimize their processes, allowing faculty and staff to focus more on student development and less on administrative burdens.

2.2 Survey on Template-Based Document Automation

Template-based document automation is transforming the way organizations manage large volumes of standardized documents, such as contracts, invoices, certificates, and reports. This survey investigates the use, effectiveness, and impact of template-based document automation in various sectors, including education, finance, healthcare, and government. The survey addresses the types of documents most frequently automated, the software tools utilized, and the benefits, including time savings, improved accuracy, and consistency across documents. Participants are also asked about any obstacles they face, such as customization limitations, setup complexity, or integration challenges with other systems. Additionally, the survey seeks feedback on the impact of document automation on employee satisfaction, as it reduces repetitive tasks, and on organizational efficiency, as it accelerates workflows. Insights from this survey will help identify best practices in template-based automation, contributing to more efficient and scalable document management across industries.

2.3 Survey on Error Reduction in Document Workflows

This survey examines the strategies and tools used to reduce errors in document-intensive workflows across organizations. Document errors, often due to manual data entry, misinterpretation, or formatting issues, can result in significant delays, compliance risks, and additional costs. The survey focuses on identifying common sources of errors within document workflows and the technologies deployed to address them, such as RPA, template-based automation, optical character recognition (OCR), and AI-driven data validation. Respondents are asked to provide insights into the effectiveness of these solutions in minimizing mistakes, their impact on quality assurance processes, and their role in maintaining compliance standards. Additionally, the survey explores the perceived improvements in employee productivity and customer satisfaction resulting from reduced

errors in document handling. Findings from this survey will highlight effective approaches to error reduction, ultimately helping organizations establish more reliable and streamlined workflows.

2.4 Summary of the Intersection of RPA and Document Automation

The intersection of Robotic Process Automation (RPA) and document automation represents a transformative approach to managing complex, high-volume document processes in various sectors. By integrating RPA with document automation tools, organizations can create workflows that not only handle repetitive tasks but also streamline document generation, processing, and storage. RPA enables the automation of data extraction, input, and document routing, while template-based document automation provides consistency and structure to document creation. Together, these technologies minimize human error, reduce manual intervention, and accelerate workflows, benefiting areas such as finance (e.g., invoicing), legal (e.g., contract management), and healthcare (e.g., patient records). Moreover, the combined use of RPA and document automation enhances compliance with regulatory standards, as automated workflows are more consistent and easier to audit. The result is an efficient, scalable, and reliable approach to document management, allowing organizations to focus resources on more strategic initiatives while maintaining accuracy and accountability in their document workflows.

CHAPTER 3 SYSTEM DESIGN

3.1 SYSTEM FLOW DIAGRAM

A flowchart is a type of diagram that represents an algorithm, workflow or process. The flowchart shows the steps as boxes of various kinds, and their order by connecting the boxes with arrows. This diagrammatic representation illustrates a solution model to a given problem. The system flow diagram for this project is in Fig. 3.1.
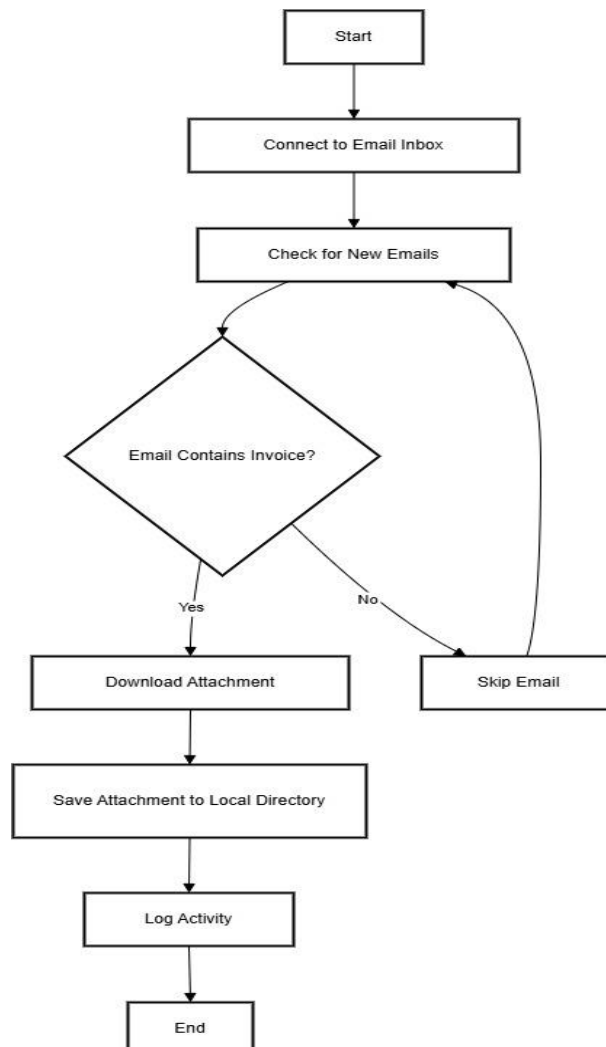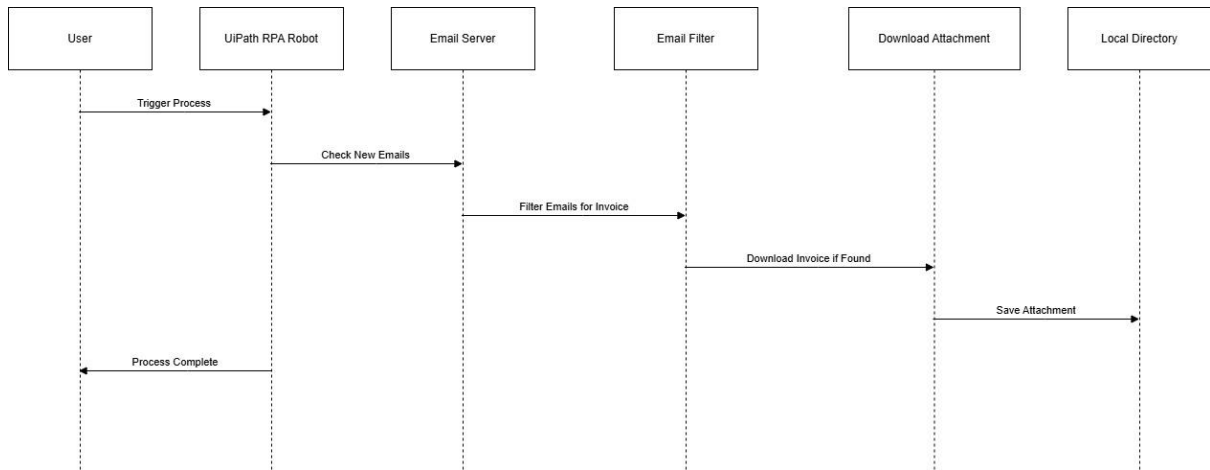
3.2 ARCHITECTURE DIAGRAM

An architecture diagram is a graphical representation of a set of concepts, that are part of an architecture, including their principles, elements and components. The architecture diagram for this project is in Fig. 3.2.
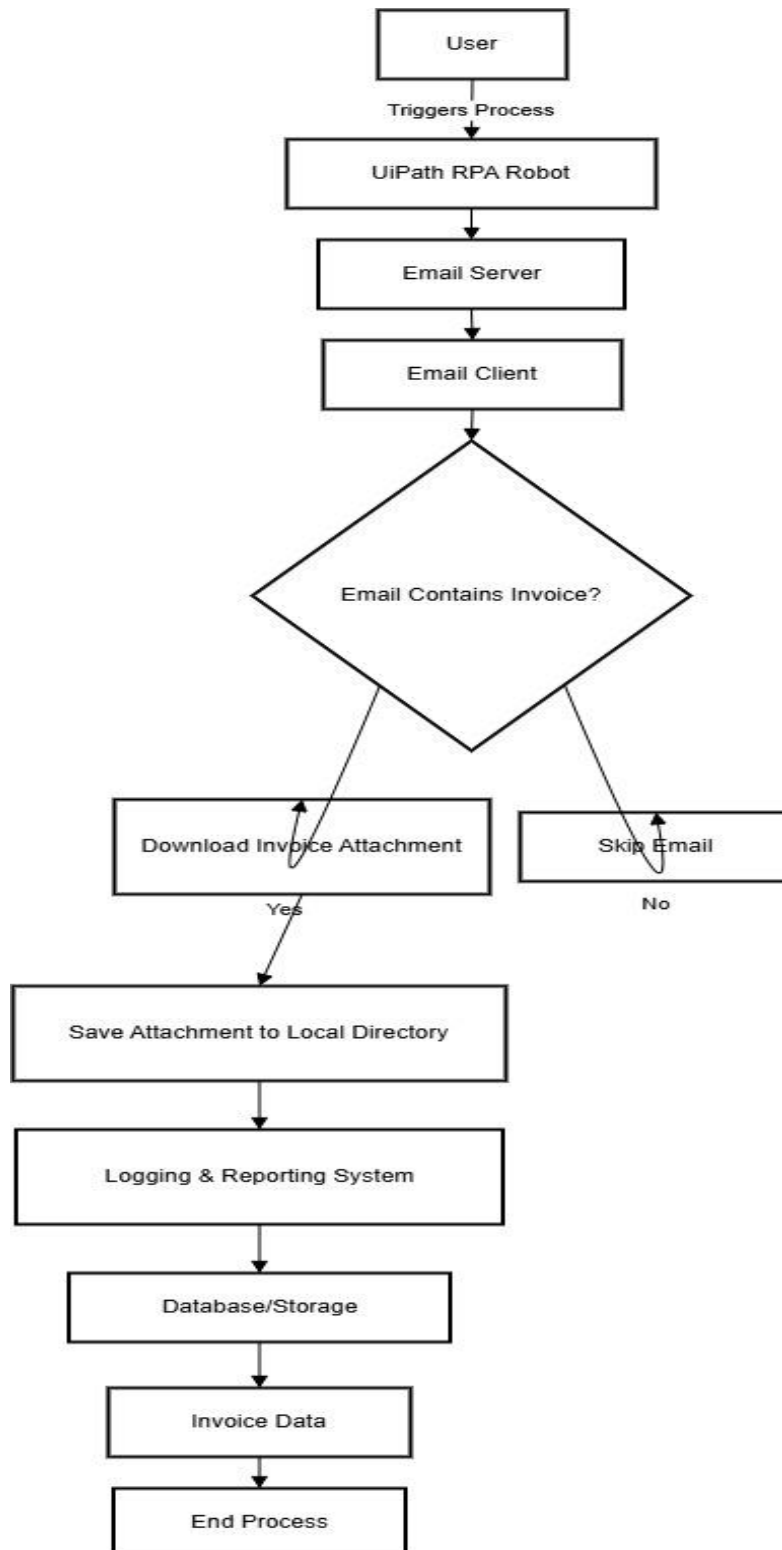
3.3 SEQUENCE DIAGRAM

A sequence diagram is a type of interaction diagram because it describe and s how in what order a group of objects works together. The sequence diagram for this project is in Fig. 3.3.

User	UiPath RPA Robot	Email Server	Email Filter	Download Attachment	Local Directory

Trigger Process

Check New Emails

Filter Emails for Invoice

Download Invoice if Found

Save Attachment

Process Complete

CHAPTER 4 PROJECT DESCRIPTION

The Robotic Process Automation (RPA) system is designed to automate the process of downloading and organizing invoice attachments from emails. The primary goal of the system is to reduce manual effort, minimize human error, and streamline the workflow of invoice processing. By integrating UiPath RPA technology with an email server, this system can automatically retrieve emails, analyze the content to identify invoices, and then store the attachments in a predefined local directory. This automated workflow not only saves time but also ensures consistency, accuracy, and compliance with document management standards.

4.1 MODULES

4.1.1 INPUT HANDLING AND INITIALIZATION

The input handling phase begins when the RPA bot connects to the email server (such as IMAP or Outlook) configured with the correct credentials. The bot then retrieves new emails from the inbox at regular intervals. Each email is fetched along with its metadata, such as the subject, sender, and attachment. Initialization involves setting up the system parameters such as:

- The email account to be monitored.
- The folder or path where attachments will be saved.
- Keywords (such as "Invoice") to identify relevant emails.
- File formats and attachment types supported by the system.

The system ensures that the right email inbox is connected, all necessary configurations are in place, and the bot is ready to start processing incoming emails.

4.1.2 CONTENT ANALYSIS

Once an email is retrieved, the system moves on to content analysis, where the bot inspects the subject and body of the email for predefined keywords or patterns that indicate an invoice. This analysis is essential for filtering out irrelevant emails, ensuring that only emails containing invoices are processed further. The content analysis phase includes:

- Checking for the keyword "Invoice" or other relevant phrases within the subject or body.
- Verifying whether the email contains an attachment, and if so, whether the attachment type matches supported formats (e.g., PDF, Excel).
- Extracting key details from the attachment's metadata, if needed (e.g., invoice number, date).
- Applying rules to flag any unusual or non-conforming emails that need to be reviewed manually.

This ensures that only relevant emails with invoices are passed onto the next step in the process.

4.1.3 RESULT MANAGEMENT

After downloading the attachment, the RPA bot proceeds to manage the results. This includes:

- Saving the attachment to a predefined local directory. The directory can be structured by dates, client names, or invoice numbers for easy retrieval.
- Storing the downloaded invoice data in a centralized location, such as a database or file system, for further analysis or audit purposes.
- Keeping a log of all processed invoices, including metadata such as email subject, sender, and processing timestamp.
- Handling errors, such as failed downloads or unsupported attachment formats, by notifying the administrator or logging the error for troubleshooting.

This result management phase ensures that every downloaded invoice is stored securely and can be retrieved efficiently when needed.

4.1.4 COMPLETION AND REPORTING

Once the system completes the processing of an invoice, it generates a report summarizing the actions taken. The report typically includes:

- A success/failure log for each email processed, indicating whether the attachment was downloaded and stored successfully.
- Information on any errors encountered (e.g., unsupported file type, missing attachments) along with a description of the issue.
- A summary of how many invoices were processed and stored, helping administrators track performance and efficiency.
- Notification to the user or administrator indicating the completion of the process.

This final reporting phase provides transparency, allows users to track the progress of the task, and ensures that any issues can be quickly addressed.
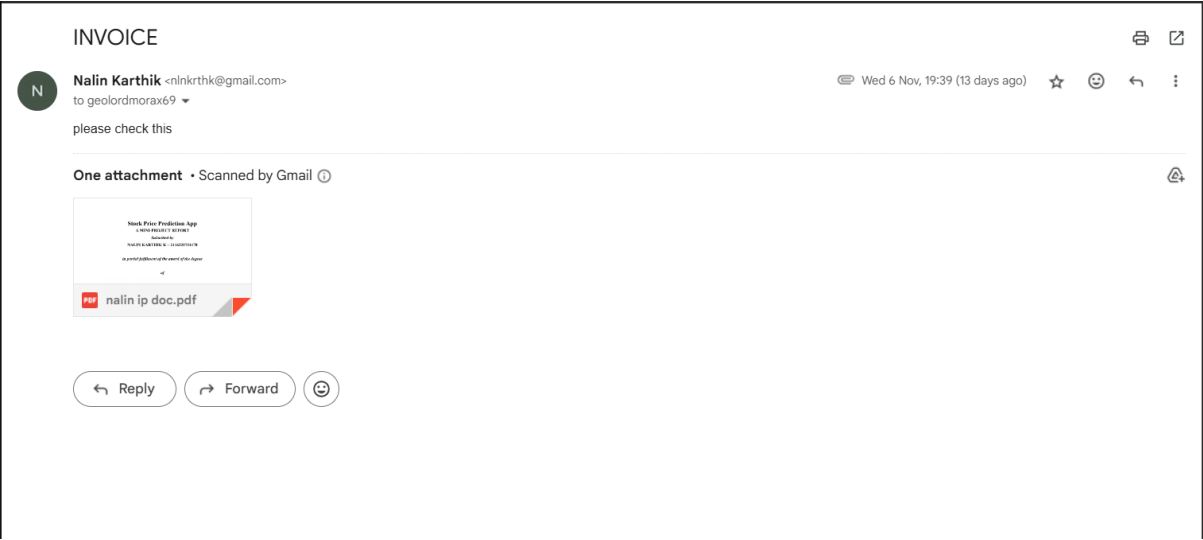
CHAPTER 5 OUTPUT SCREENSHOTS

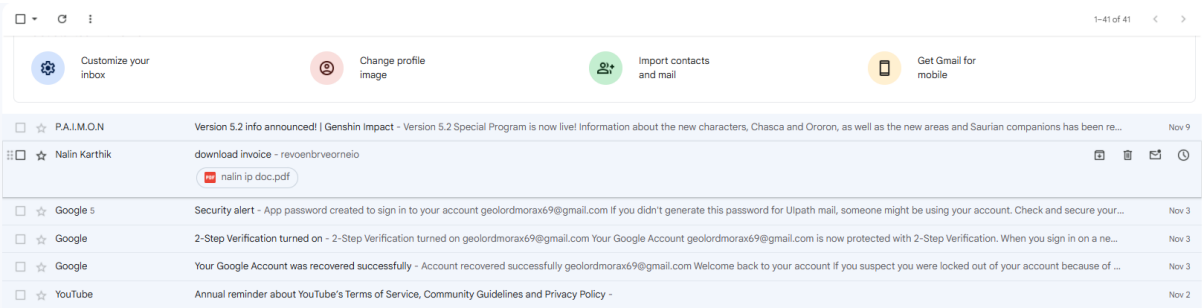
Fig 5.1 sending email with an invoice
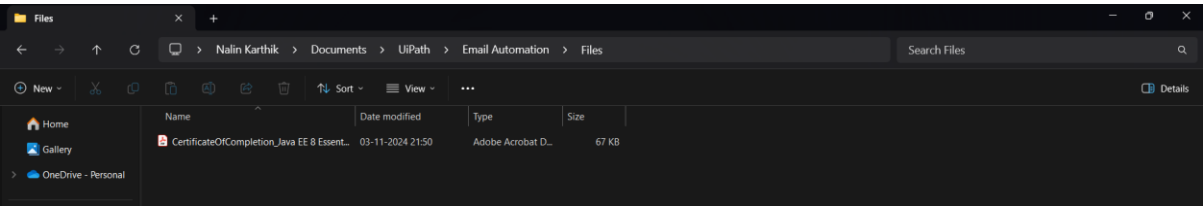

Fig 5.2 Email is received

Fig 5.3 invoice is stored in the path specified

CHAPTER 6

CONCLUSION

The implementation of this Robotic Process Automation (RPA) system for automatically downloading and organizing invoice attachments from emails marks a significant improvement in the efficiency and accuracy of document processing workflows. By leveraging UiPath's RPA capabilities, the system automates the entire process, from connecting to the email server and filtering relevant emails to downloading and storing attachments in a local directory.

The automated system reduces the manual effort required for invoice management, significantly lowering the risk of human error and streamlining the workflow. Furthermore, the structured storage of invoices ensures that they are easily accessible for future retrieval and auditing. The logging and reporting features provide transparency and allow for easy tracking of the system's performance, ensuring that any errors are promptly addressed.

Overall, this RPA-based solution not only saves time and resources but also enhances accuracy, compliance, and productivity. As organizations continue to embrace automation, this system serves as a valuable tool for improving document management and optimizing administrative tasks, allowing employees to focus on more strategic initiatives.
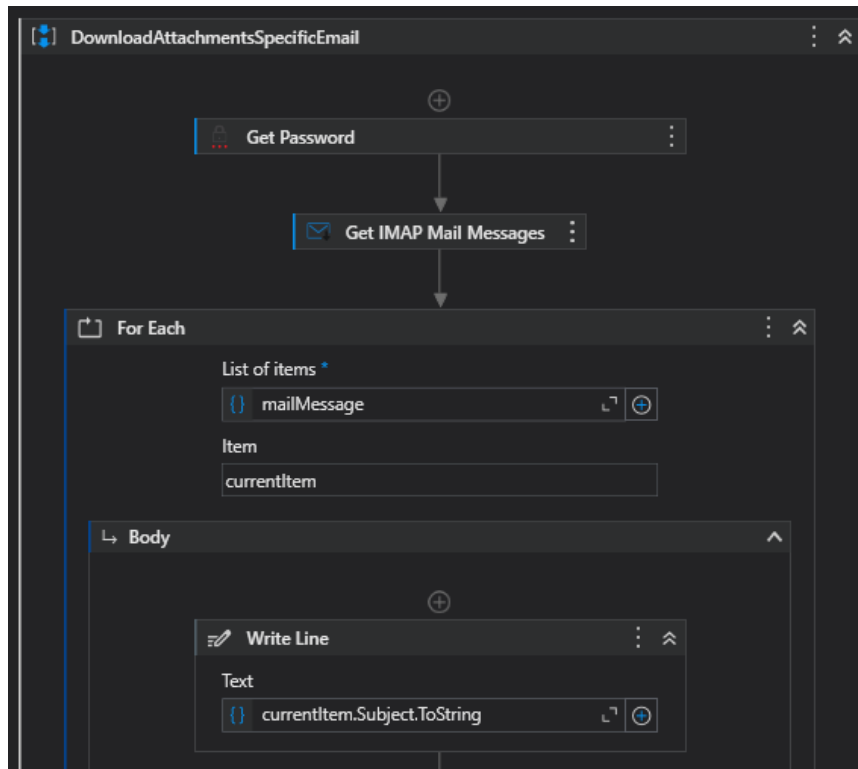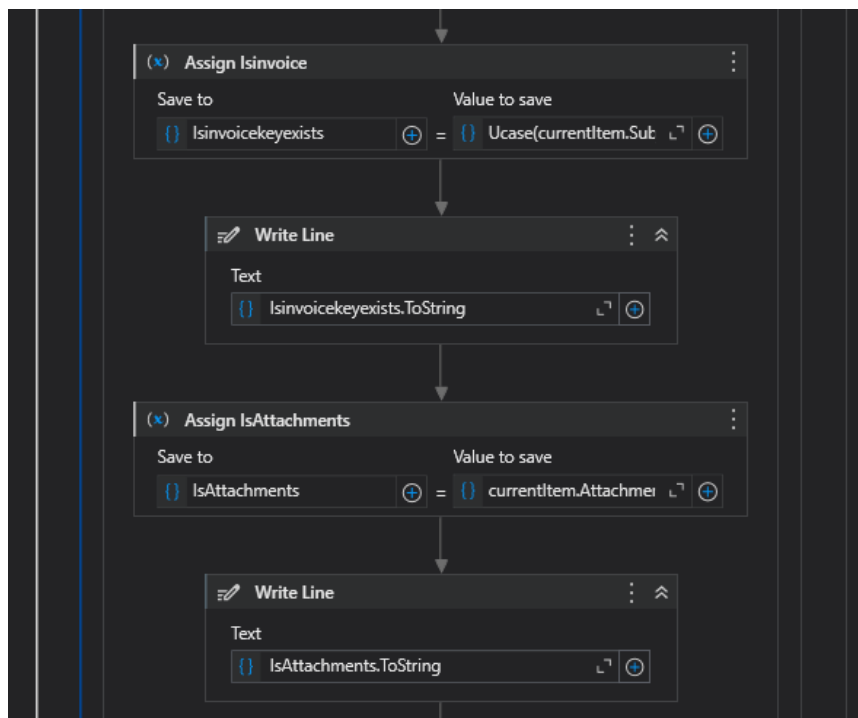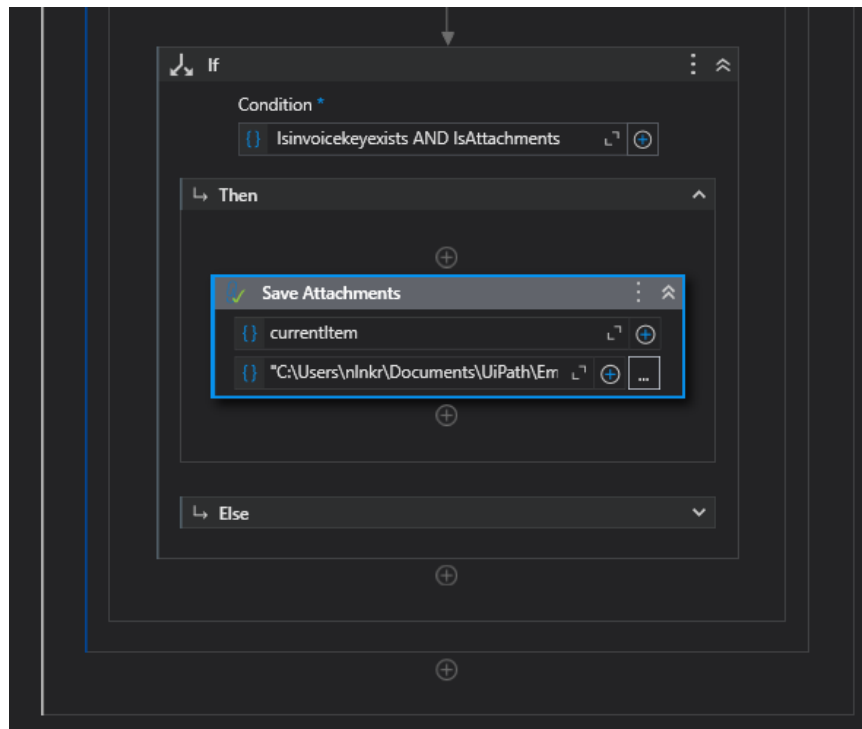
5.APPENDIX PROCESS WORK FLOW

Fig 5.1

Fig 5.2

Fig 5.3

REFERENCES

1. Kuppusamy, Palanivel & Joseph K, Suresh. (2020). Robotic Process Automation to Smart Education. 3775.

2. Patil, Dr & Mane, Vinod & Patil, Dr. (2019). Social Innovation in Education System by using Robotic Process Automation (Rpa). International Journal of Innovative Technology and Exploring Engineering. 8. 3757-3760. 10.35940/ijitee.K2148.0981119.

3. Elkhatat, A.M., Elsaid, K. & Almeer, S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *Int J Educ Integr* **19**, 17 (2023). https://doi.org/10.1007/s40979-023-00140-5

[4] H. Alamleh, A. A. S. AlQahtani and A. ElSaid, "Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning," 2023 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 2023, pp. 154-158, doi: 10.1109/SIEDS58326.2023.10137767.

[5] Tomáš Foltýnek, Norman Meuschke, and Bela Gipp. 2019. Academic Plagiarism Detection: A Systematic Literature Review. ACM Comput. Surv. 52, 6, Article 112 (November 2020), 42 pages. https://doi.org/10.1145/3345317

[6] H. A. Chowdhury, D. K. Bhattacharyya, "Plagiarism: Taxonomy, Tools and Detection Techniques", 19th National Convention on Knowledge, Library and Information Networking, 2018.