BOT FOR LATEST NEWS UPDATES AND FACTS

A PROJECT REPORT

Submitted by

HARSHNI P (220701089)

in partial fulfillment for the course

OAI1903 - INTRODUCTION TO ROBOTIC PROCESS AUTOMATION

for the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND DESIGN

RAJALAKSHMI ENGINEERING COLLEGE
RAJALAKSHMI NAGAR THANDALAM CHENNAI – 602 105

NOVEMBER 2023

RAJALAKSHMI ENGINEERING COLLEGE CHENNAI - 602105

BONAFIDE CERTIFICATE

Certified that this project report **"Bot For Latest News Updates And Facts"** is the bonafide work of **"HARSHNI P (220701089)"** who carried out the project work for the subject OAI1903- Introduction to Robotic Process Automation under my supervision**.**

Mrs. J. Jinu Sophia, M.E. (Ph.D.),

Assistant Professor (SG),

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 _____.

ABSTRACT

“This UiPath automation bot streamlines the process of obtaining and sharing the latest news updates by combining web scraping and email automation. Users input a specific topic and a Gmail address, prompting the bot to search for the most recent and relevant news articles on the topic from trusted sources. The bot extracts article titles and URLs and compiles them into an email that is sent directly to the provided Gmail address. This automated solution ensures users receive real-time news updates without the need for manual searching, enhancing convenience and productivity. By leveraging UiPath's advanced web automation capabilities and secure email functions, the bot effectively simplifies how users stay informed. Potential enhancements include adding news summaries for a quick overview, scheduled updates for regular information delivery, and source filtering for tailored results. This tool is particularly valuable for busy professionals and individuals who wish to remain current with minimal effort, making it a seamless, efficient solution for personalized news consumption”.

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavour to put forth this report. Our sincere thanks to our Chairman **Mr. S. Meganathan, B.E, F.I.E.,** our Vice Chairman **Mr. Abhay Shankar Meganathan, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) Thangam Meganathan, Ph.D.,** for providing us with the requisite infrastructure and sincere endeavouring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. Murugesan, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P.Kumar, M.E., Ph.D.,** Professor and Head of the Department of Computer Science and Design for her guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guides, **Mrs. J. Jinu Sophia, M.E., (Ph.D.),** Assistant Professor (SG), Department of Computer Science and Engineering for their valuable guidance throughout the course of the project. We are very glad to thank our Project Coordinators, **Dr. N.Durai Murugan, M.E., Ph.D.,** Associate Professor, and**Mr. B.Bhuvaneswaran, M.E.,** Assistant Professor (SG), Department of Computer Science and Engineering for their useful tips during our review to build our project.

Harshni P (220701089)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	ACCRONYM
RPA	Robotic Process Automation
AI	Artificial Intelligence
API	Application Programming Interface
CV	Computer Vision
OCR	Optical Character Recognition

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Staying updated with the latest news is essential in today's fast-paced world, yet manually searching for relevant articles can be time-consuming and repetitive. To address this challenge, automation tools can simplify and expedite the process, ensuring that users receive timely information with minimal effort. The use of automation technology in gathering and sharing news updates allows for an efficient, seamless flow of information directly to users' preferred platforms.

This UiPath automation bot is designed to optimize the way users access the latest news by harnessing the power of web scraping and email automation. Users provide a search topic and their Gmail ID as input, enabling the bot to search for recent and relevant articles on the specified topic across trusted online sources. By extracting the URLs of these articles and sending them directly to the provided email, the bot significantly reduces the time and effort needed to stay informed.

With this tool, users benefit from instant updates delivered straight to their inbox, making it ideal for professionals, researchers, and anyone who values staying informed. The bot can be further enhanced to include summaries, scheduled updates, and customizable source preferences, ensuring a personalized and comprehensive news consumption experience.

1.2 OBJECTIVE

The objective of this UiPath automation bot is to provide users with a simplified, efficient method for staying updated on specific news topics of interest. By combining web scraping with email automation, the bot searches for the latest, most relevant articles on a given topic, extracts their URLs, and sends this information directly to the user's Gmail inbox. This solution eliminates the need for manual searches, ensuring timely delivery of news updates tailored to each user's preferences. The bot aims to enhance convenience, save time, and deliver valuable information effortlessly, making it ideal for individuals and professionals seeking streamlined news access.

1.3 EXISTING SYSTEM

In existing systems, staying updated with the latest news typically requires manual searching through news websites, search engines, or dedicated news apps. Users must input queries, filter through numerous articles to find relevant content, and manage how they save or share the information. While some news aggregator apps or RSS feeds provide automated updates, they often lack customization and the ability to directly deliver curated results to users' email inboxes. Moreover, integrating this information into personalized workflows may require additional tools or manual steps, making the process less efficient for users who need timely, topic-specific updates on a regular basis.

1.4 PROPOSED SYSTEM

The proposed system is a UiPath automation bot that revolutionizes how users stay informed by automating the process of obtaining and sharing the latest news. Users simply provide a search topic and a Gmail ID as input, and the bot handles the rest. It employs web scraping to search for recent, relevant news articles on the given topic from trusted sources and extracts their URLs.

The bot then compiles these URLs into a concise email and sends it directly to the user's Gmail account. This approach eliminates the need for manual searching, improves efficiency, and ensures users receive timely, customized updates effortlessly.

CHAPTER 2

LITERATURE REVIEW

2.1 Web Scraping Technologies

Web scraping is a widely used technique for extracting data from websites. Studies have highlighted various tools and frameworks, such as Beautiful Soup, Scrapy, and UiPath itself, which offer robust solutions for collecting web data. According to existing research, the integration of web scraping in automation projects has proven effective for accessing real-time information. Automation platforms like UiPath, with built-in activities and seamless integration capabilities, have made these tasks more accessible and efficient for non-programmers as well.

2.2 Email Automation and Notification Systems

Email automation is another area with extensive research, focusing on how systems can automate the delivery of tailored content. Various platforms, including UiPath, allow for configuring SMTP protocols and leveraging APIs like Gmail's for secure and automated email dispatch. Literature points out that automated email systems increase productivity and streamline communication, especially when tailored to user preferences. Combining web scraping with email automation creates a powerful tool for timely dissemination of customized information.

2.3 News Aggregation and Customization

Research on news aggregation indicates that while many apps and websites exist to keep users informed, customization and user-specific updates are often limited. RSS feeds and aggregator services like Feedly or Google News offer a degree of automation but lack personalized integration into other platforms.

Studies suggest that tailored solutions that align content retrieval with direct delivery channels, such as email, offer an advantage in terms of user convenience and real-time responsiveness.

2.4 UiPath in Automation Projects

UiPath has been extensively studied as a versatile robotic process automation (RPA) tool. Its drag-and-drop interface and extensive library of pre-built activities simplify automation tasks like data extraction and email handling. Case studies and technical analyses emphasize UiPath's success in automating repetitive tasks, reducing human error, and enhancing workflow efficiency in various industries. The bot proposed in this project builds on this established use of UiPath by combining its web scraping capabilities with email automation to create a seamless user experience for receiving news updates.

2.5 Conclusion

The literature shows that while technologies for web scraping, email automation, and news aggregation exist independently, the combination of these in a single automated bot using UiPath is relatively novel. This project aims to bridge the gap by providing a comprehensive solution that ensures timely, relevant information delivery tailored to user preferences.

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.

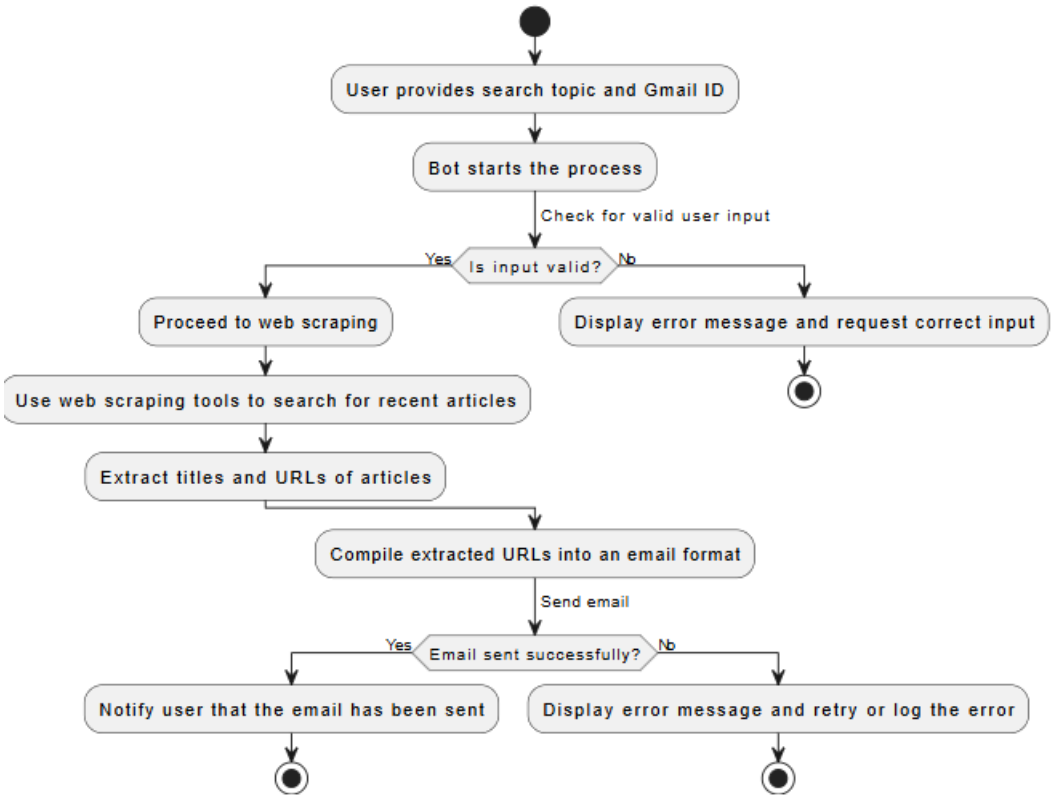


Fig 3.1 System Flow Diagram

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.

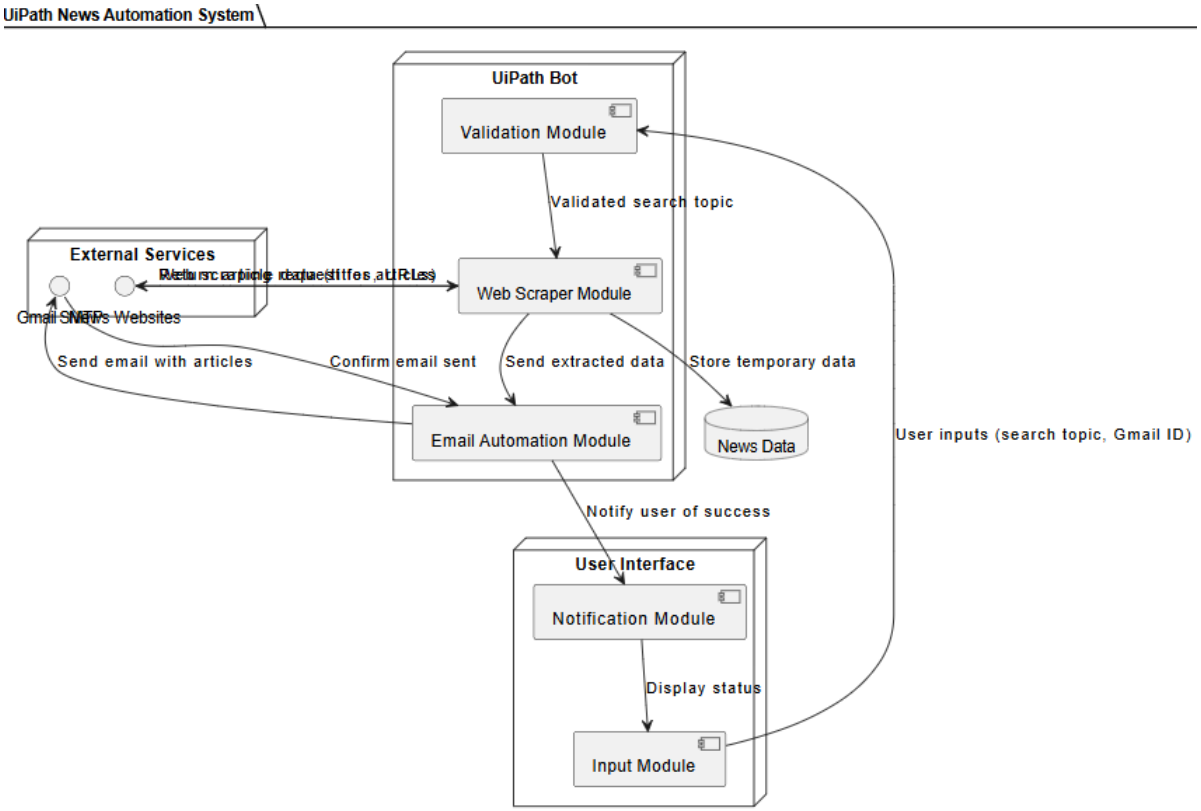


Fig 3.2 Architecture Diagram

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.

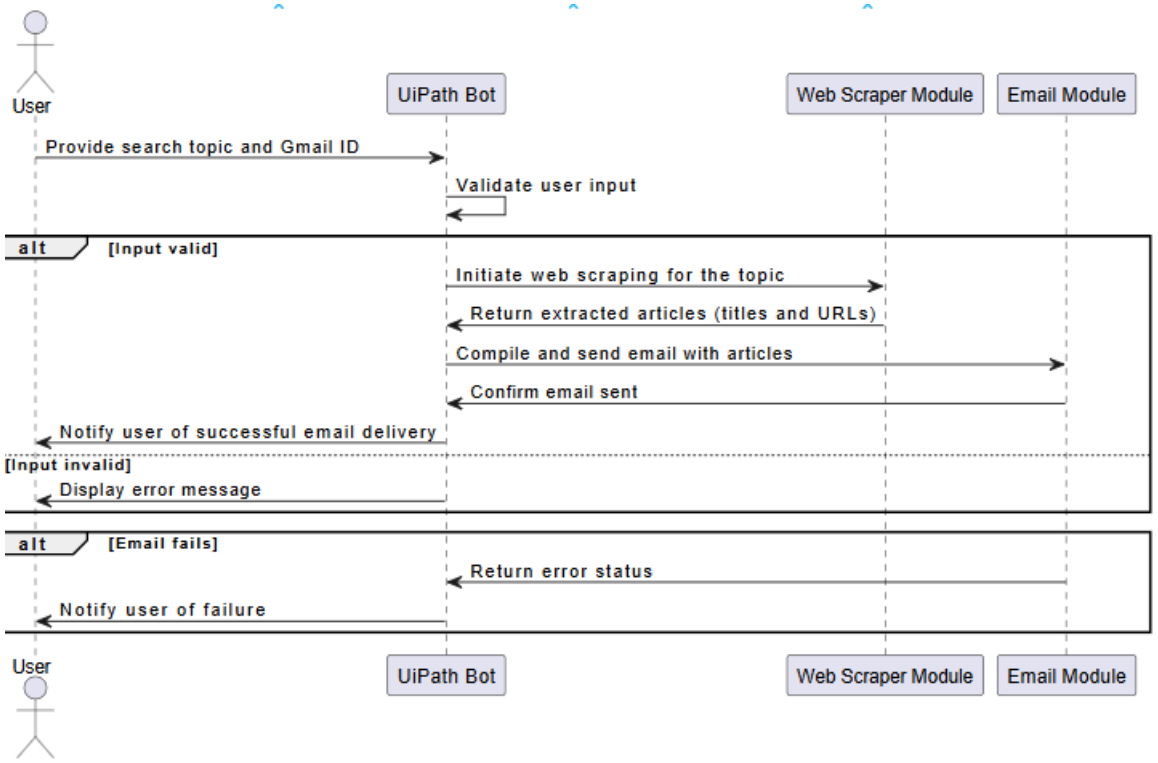


Fig 3.3 Sequence Diagram

CHAPTER 4

PROJECT DESCRIPTION

This UiPath automation bot simplifies obtaining and sharing news updates. Users input a search topic and Gmail ID, prompting the bot to extract relevant articles using web scraping and email them directly. The project automates the process for real-time, customized news delivery, enhancing efficiency and convenience for staying informed.

4.1. MODULES:

4.1.1 Input Module

- **Function**: Captures user input, including the search topic and the Gmail ID where the news updates should be sent.
- **Features**: Provides an interface for users to enter data and initiates the automation process upon submission.

4.1.2 Validation Module

- **Function**: Ensures that the user-provided input is valid and in the correct format (e.g., checks if the Gmail ID is formatted properly and if the search topic is not empty).
- **Features**: Returns feedback or prompts the user for correction if the input is invalid.

4.1.3 Web Scraper Module

- **Function**: Automates the process of searching and extracting news articles relevant to the provided topic from various online sources.
- **Features**: Utilizes UiPath's web automation tools to navigate web pages, extract article titles and URLs, and handle potential web page variations.
- **Challenges Addressed**: Handles dynamic web content and ensures compliance with ethical scraping practices.

4.1.4 Data Processing Module

- **Function**: Processes the scraped data by compiling and organizing it into a readable format for the email body.
- **Features**: May include basic text formatting and data cleaning to ensure the extracted URLs and titles are presented neatly.

4.1.5 Email Automation Module

- **Function**: Sends the processed news data directly to the user's provided Gmail ID.
- **Features**: Integrates with Gmail's SMTP server or API to securely send emails. Includes error handling to manage potential email delivery failures and retries if necessary.

4.1.6 Notification Module

- **Function**: Provides feedback to the user after the process is completed, notifying them whether the email was sent successfully or if there were any issues.
- **Features**: Can include desktop notifications or in-app messages indicating the status of the automation.

4.1.7 Logging and Error Handling Module

- **Function**: Logs each step of the process to help with troubleshooting and to keep a record of activities.
- **Features**: Captures errors or failures during web scraping, data processing, or email sending and logs them for review. Offers retry mechanisms for error-prone tasks.

4.1.8 Scheduler Module (Optional Enhancement)

- **Function**: Allows users to schedule automated news updates at regular intervals (e.g., daily or weekly).
- **Features**: Configures automated triggers within UiPath for periodic task execution.

CHAPTER 5

OUTPUT SCREENSHOTS

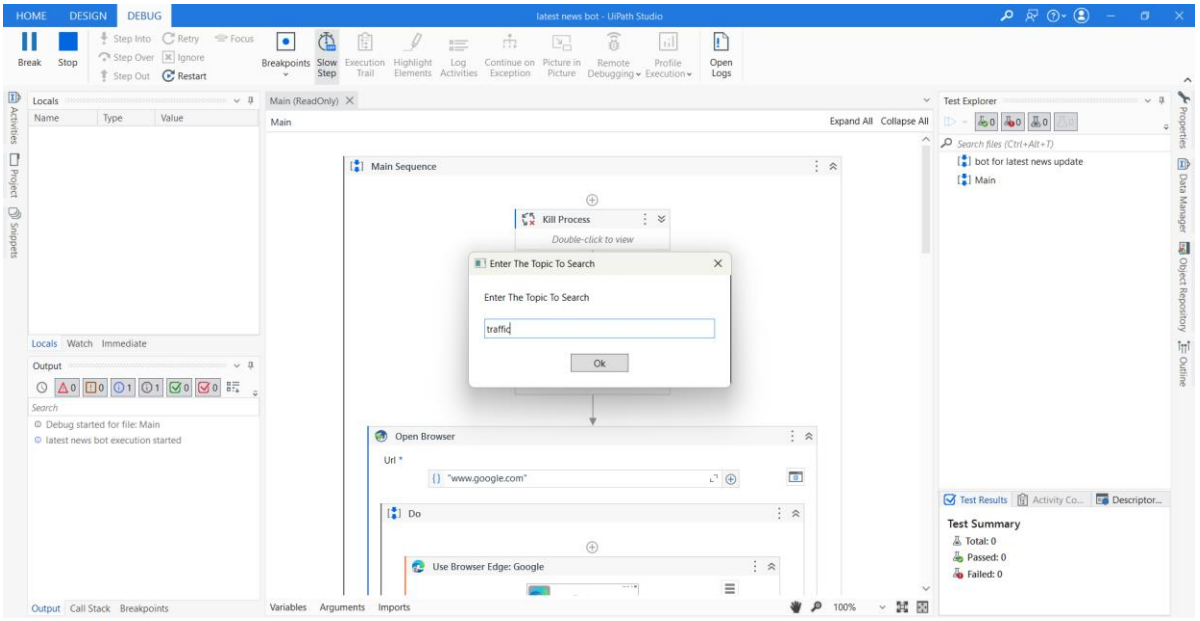


Fig 5.1 – Input Dialog

The bot gets the topic to searched from the user as shown in Fig 5.1.

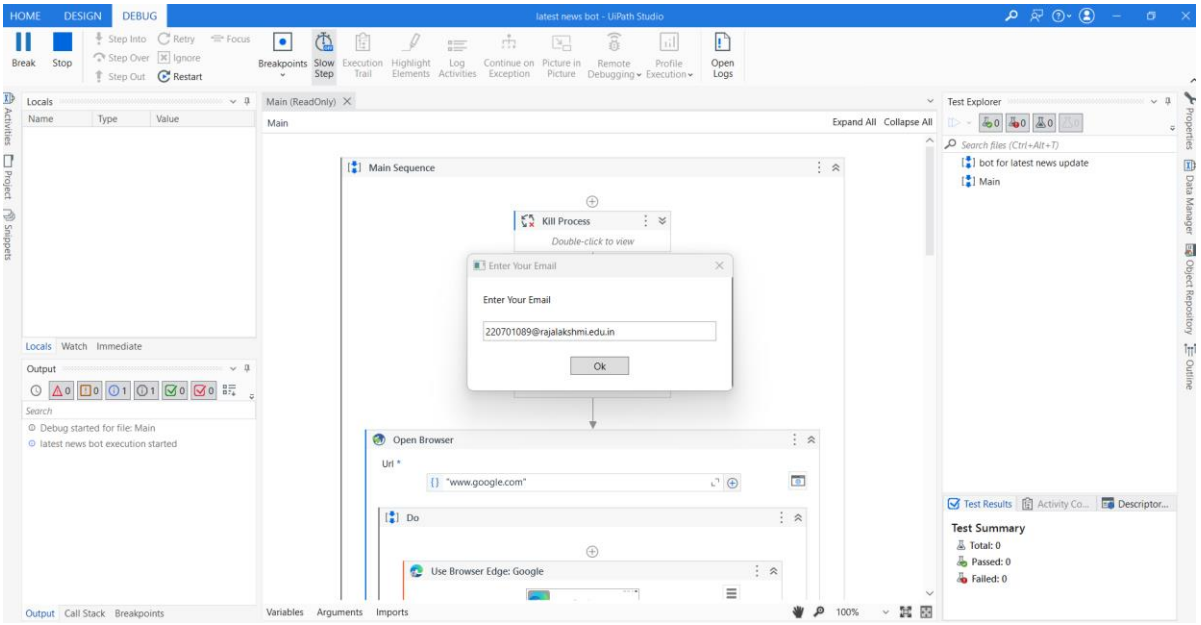


Fig 5.2 – Input Dialog

The bot get specific mail id to send the url of topic searched as shown in Fig 5.2.

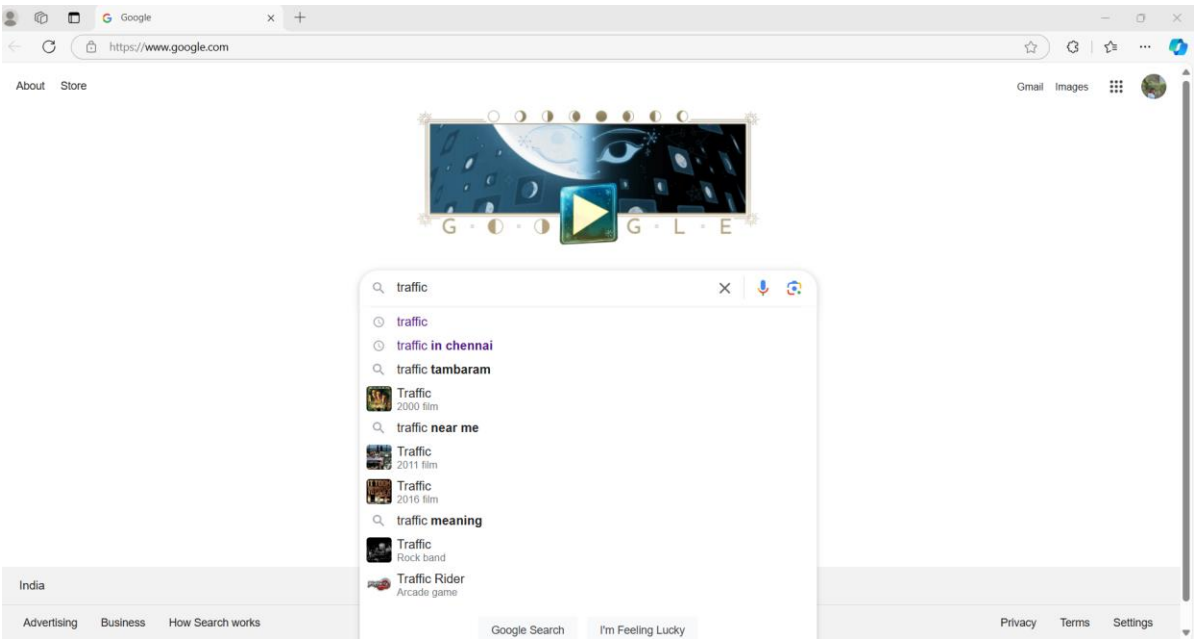


Fig 5.3 The topic is searched in edge

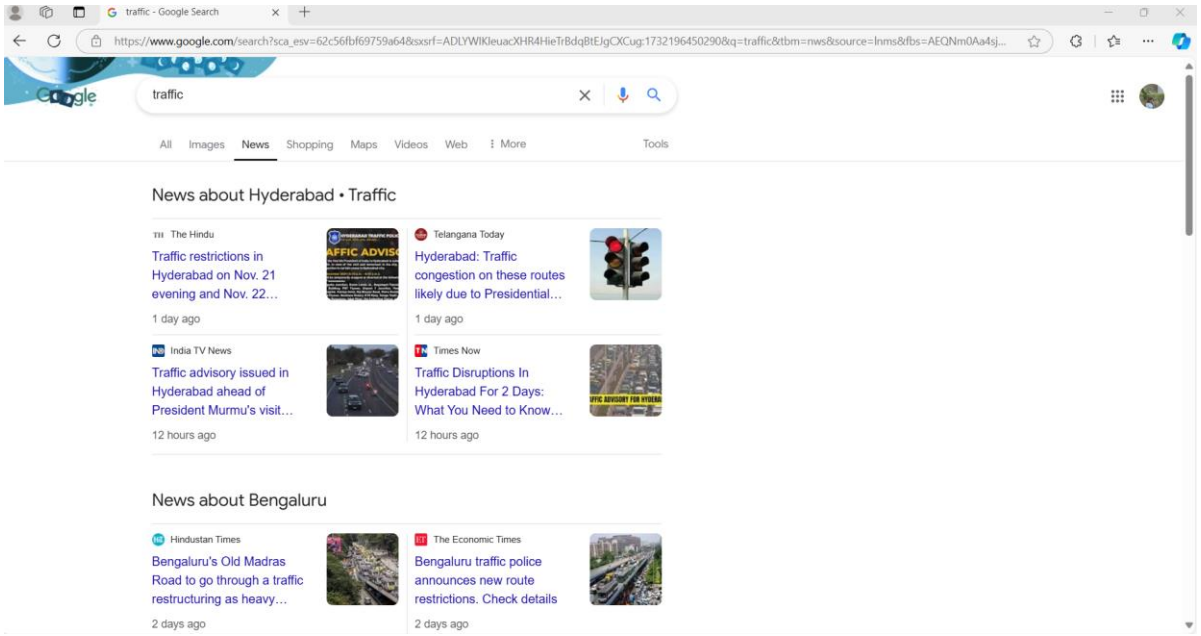


Fig 5.4 – Then move to news information

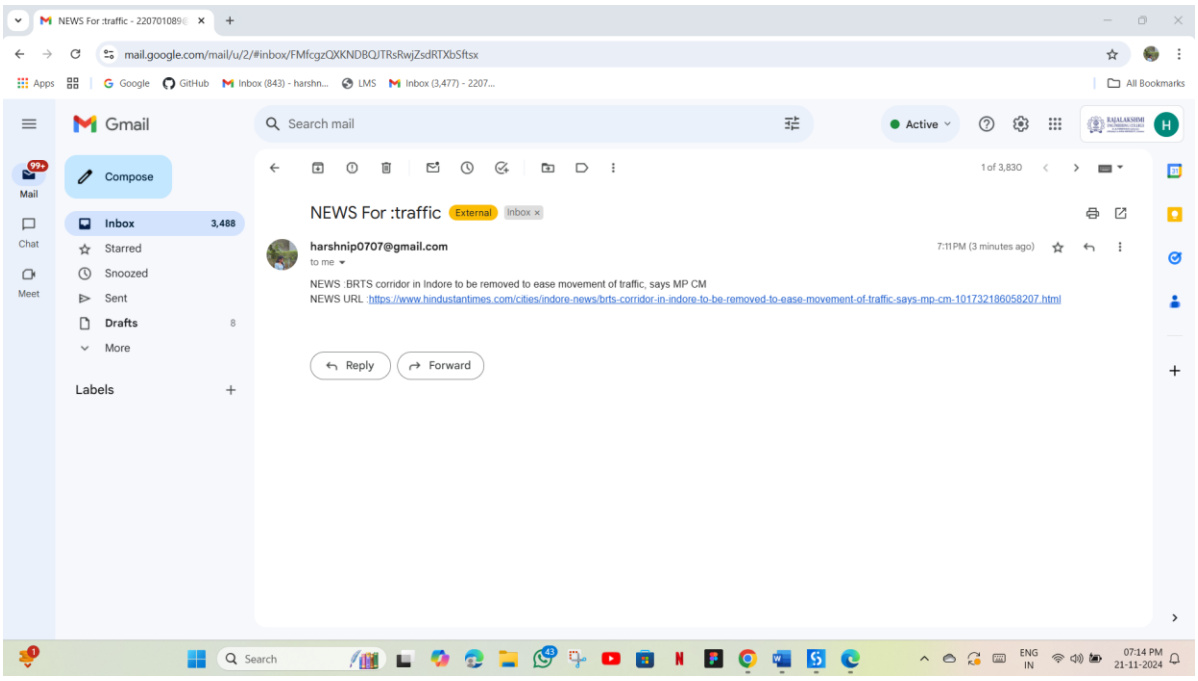


Fig 5.5 – Send to the mail

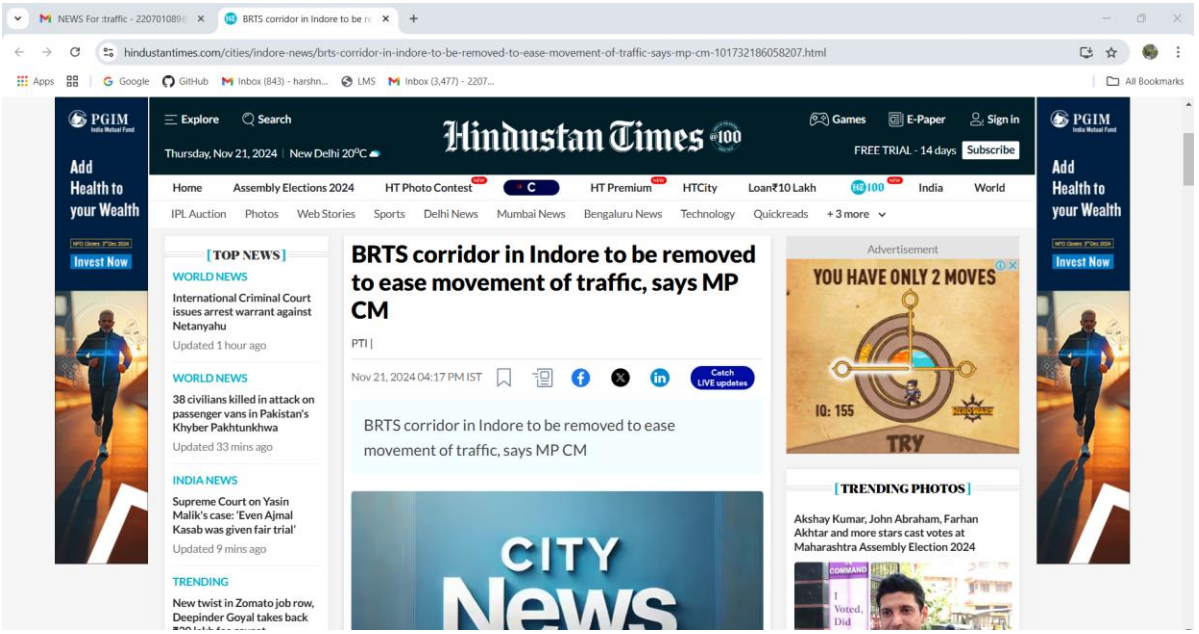


Fig 5.6 Then url opens to the latest news website

CHAPTER 6

CONCLUSION

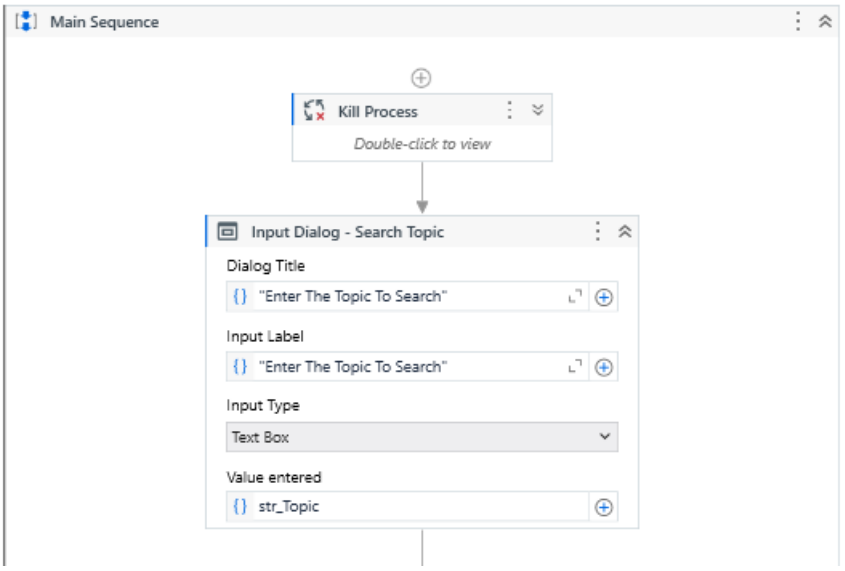
The UiPath automation bot designed for extracting and sharing the latest news provides an effective solution for streamlining information access. By combining web scraping with email automation, this project allows users to stay informed effortlessly and in real-time. The automated process reduces manual work, ensuring users receive timely, relevant updates directly in their inboxes with minimal interaction.

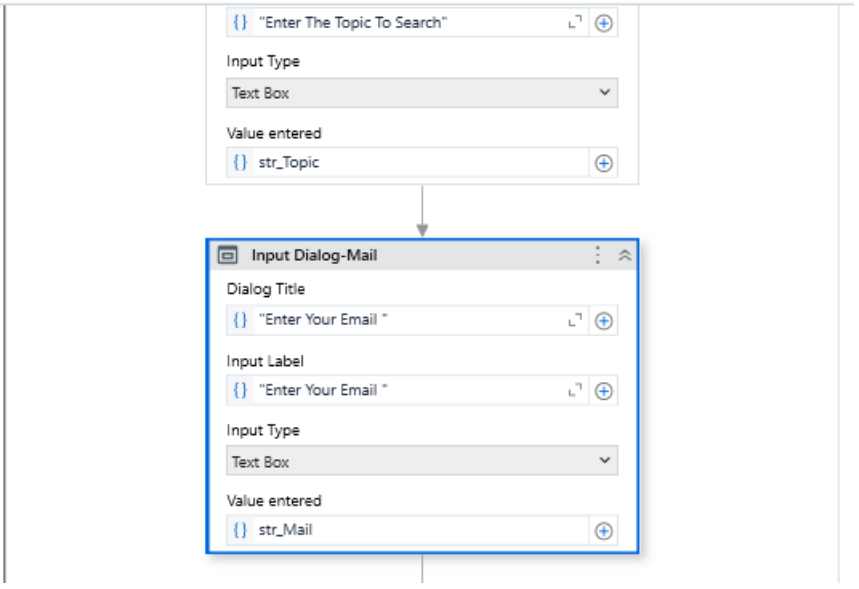
This system's modular architecture, including input validation, data extraction, and email automation, ensures flexibility, scalability, and ease of maintenance. The project can be further enhanced with features like periodic scheduling, source customization, and article summaries, making it even more user-friendly and tailored to individual preferences. These extensions would increase the bot's utility for a wide range of users, from professionals to news enthusiasts.

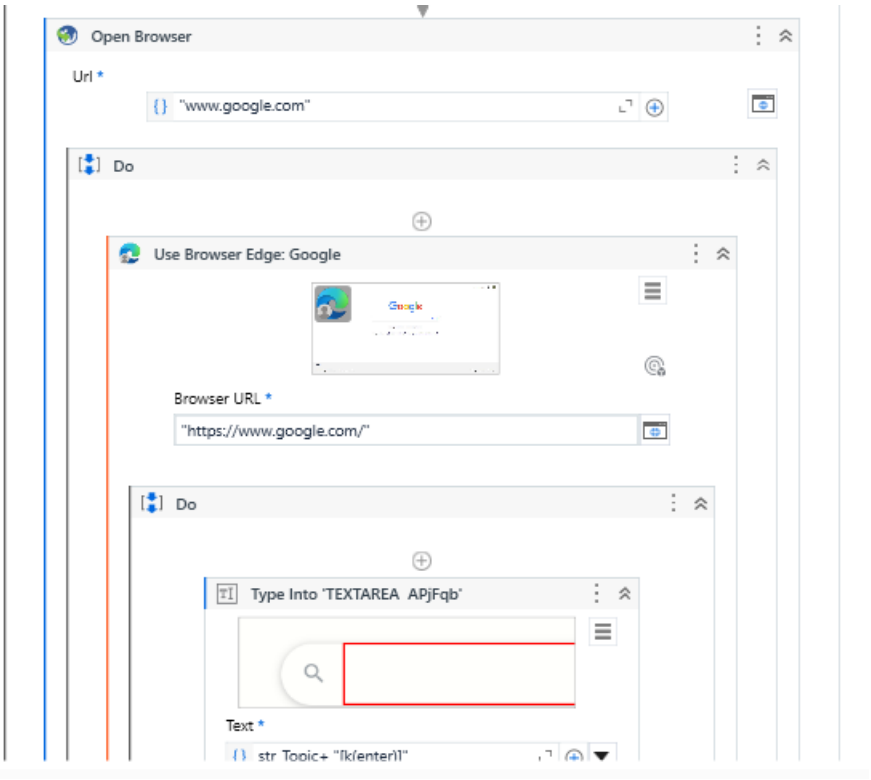
In conclusion, this project demonstrates the power of integrating automation with daily information retrieval needs. By eliminating repetitive tasks and enhancing user convenience, the UiPath bot positions itself as a valuable tool for personalized, efficient news consumption. With potential for further improvements, it sets a foundation for more sophisticated automation solutions in the realm of information sharing.

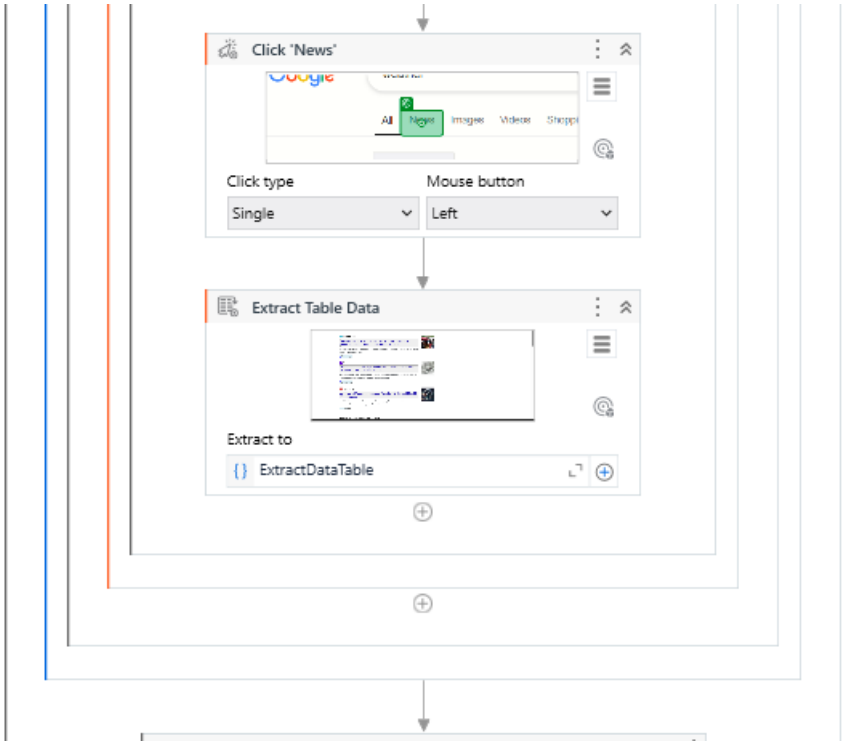
APPENDIX

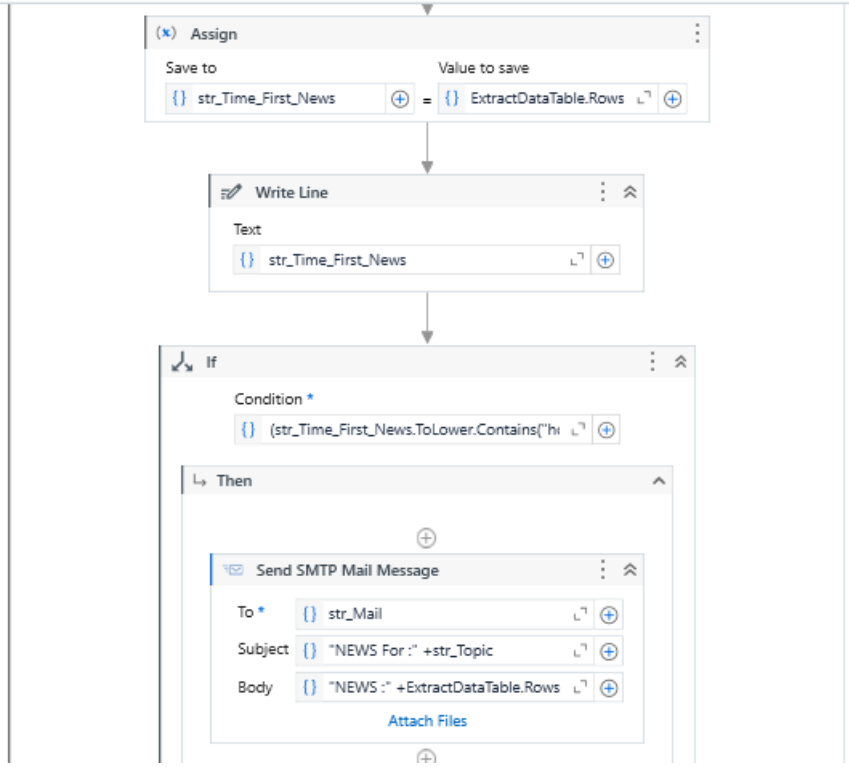
PROCESS WORKFLOW

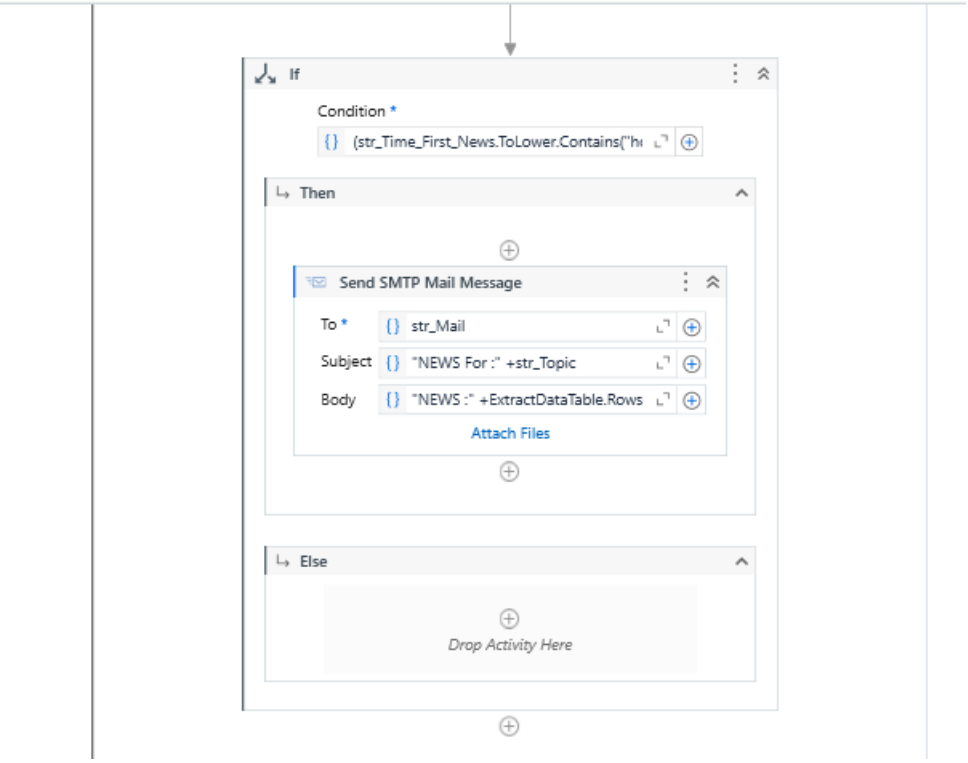


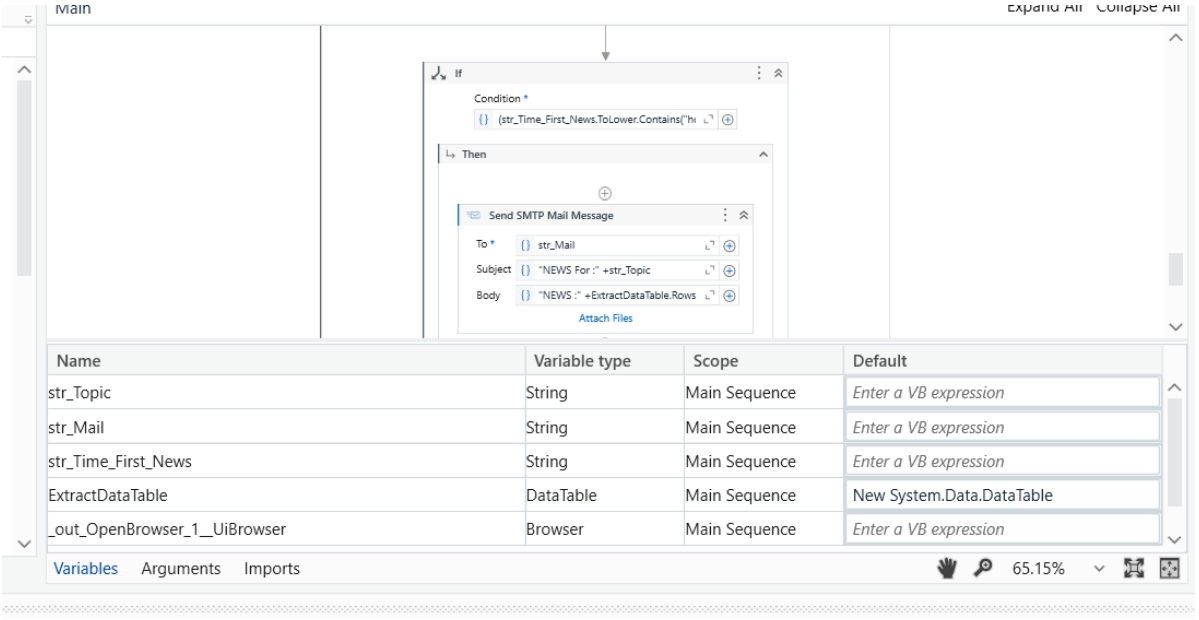



Open Browser
Url *
"www.google.com"
Do
Use Browser Edge: Google
Browser URL *
"https://www.google.com/"
Do
Type Into 'TEXTAREA APjFqb'
Text *


Click 'News'
Click type
Mouse button
Single
Left
Extract Table Data
Extract to
ExtractDataTable


Assign
Save to
Value to save
str_Time_First_News
=
ExtractDataTable.Rows
Write Line
Text
str_Time_First_News
If
Condition *
(str_Time_First_News.ToLower.Contains("h
Then
Send SMTP Mail Message
To *
str_Mail
Subject
"NEWS For :" +str_Topic
Body
"NEWS :" +ExtractDataTable.Rows
Attach Files


If
Condition *
(str_Time_First_News.ToLower.Contains("h
Then
Send SMTP Mail Message
To *
str_Mail
Subject
"NEWS For :" +str_Topic
Body
"NEWS :" +ExtractDataTable.Rows
Attach Files
Else
Drop Activity Here


Main
Expand All
Collapse All
If
Condition *
(str_Time_First_News.ToLower.Contains("h
Then
Send SMTP Mail Message
To *
str_Mail
Subject
"NEWS For :" +str_Topic
Body
"NEWS :" +ExtractDataTable.Rows
Attach Files
Name
Variable type
Scope
Default
str_Topic
String
Main Sequence
Enter a VB expression
str_Mail
String
Main Sequence
Enter a VB expression
str_Time_First_News
String
Main Sequence
Enter a VB expression
ExtractDataTable
DataTable
Main Sequence
New System.Data.DataTable
_out_OpenBrowser_1__UiBrowser
Browser
Main Sequence
Enter a VB expression
Variables
Arguments
Imports
65.15%

REFERENCES

[1] Kuppusamy, Palanivel & Joseph K, Suresh. (2020). Robotic Process Automation to Latest News. 3775.

[2] Patil, Dr & Mane, Vinod & Patil, Dr. (2019) News Generator 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 AIgenerated 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.