EMAIL CLIENT AUTOMATION WITH RPA

Akshay P N[1], Nisarga Kalagi[2], Deeksha Shetty[3], Ramalingam H M[4]

[1,2,3]UG Scholar, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India
[4]Senior Assistant Professor, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India

ABSTRACT: In today's internet era Electronic mail (Email) is a widely used communication channel commonly used to control customer inquiries such as complaints, feedbacks, reviews, and suggestions. With an increase in the number of customers, there is a significant increase in the emails being received daily, which needs to be segregated, to ensure that a proper reply is sent to all the senders in an organization and to prevent overwhelming and messy email accumulation. Now, the problem with the traditional method is that a manual workforce team cannot sit and segregate every email as the humongous amount of emails get generated daily. Apart from that it is quite a tiresome job and cannot be done by a single employee or a team of employees. So, the businesses industry and other industries can simply automate the email processing task of segregating common emails into specified folders as per the organization's requirements. The proposed system is a smart email client prototype that automates the task of email segregation and routing and by providing fixed reply using Robotic process automation (RPA) that runs on UiPath software.

KEYWORDS: Robotic Process Automation, Electronic mail, Identification.

I. INTRODUCTION

More than half the global population uses email application. The email communication services have become an inexorable part of human lives. In the year 2019, the number of email users reached the 3.9 billion mark. In 2020 it is expected that the number of email users to rise to 4 billion. The predicted user growth rate for the next four years is 3 percent, which is around 100 million users more each year. So, in 2023 the number of email users worldwide should be approximately 4.3 billion [11].

The number of emails sent and received per day, even in a midsize organization, is large and requires huge storage size. With the recent development of technologies that can store and manage big data, these high volume emails can be handled effectively. But managing the mail for taking cognitive decisions on hourly, or even daily, basis requires a lot of manual effort and high overhead cost. In addition to huge volume, the variety in its content (subject matter of discussion) may require human agents to possess knowledge in various domains to be able to manage these emails.

Given the workload that emails generate, there has been a longstanding desire to automate various aspects of email processing. Suppose a support Organization that needs to answer thousands of emails that get bombarded at their inboxes for a response, then that organization could get away from such a situation and have RPA look after them. Emails can be segregated into groups and responses to such emails can be given by the RPA solution whereas the critical ones which are not ascertained into a group can be handled by the respective personnel.

Due to the increase in the number of incoming emails, the user needs more efficient work on how to respond to the query. Some of the other aspects that need to be dealt with are tagging and prioritizing incoming emails, moving incoming emails to different locations, labeling emails with richer data as sender or receiver, Sending emails only at a particular time or context, Sending an email to only the right people.

II. LITERATURE SURVEY

This study provides a hypothesis that CBR methodology can be used to solve the email overload problem. Real-world email data sets have been analyzed to investigate the potential for applying CBR. The analysis shows that future questions could be mapped by utilizing similar previous questions and reusing the answers given to them[1].

This study sheds light on the opportunities for automating email processing. Email management consumes significant effort from senders and recipients. Some of this work might be automatable. A mixed-methods need-finding study done to learn: (i) what sort of automatic email handling users want, and (ii) what kinds of information and computation are needed to support that automation[2].

This study gives an insight into, a novel end-to-end system for automatically generating short email responses. The core of the system is state-of-the-art deep LSTM models that can predict full responses, given an incoming

email message. The responses are identified by a novel method for semantic clustering. A scalable system is developed by efficiently searching the target response space [3].

This study gives an insight into RPA. RPA is a completely software-based solution. The Process of mimicking human actions to perform a sequence of steps that lead to meaningful activity, without any human intervention is known as Robotic Process Automation. RPA refers to configuring a software robot which will do the job mostly which is repetitive [4].

The study provides key information on email classification. Documents' classification in general and emails' classification, in particular, utilize several natural language processing and data mining activities such as Text parsing, stemming, classification, clustering, etc for reasons such as Spam detection, subject, or folder classification, etc [11].

This study presents a novel email classification method for enterprises by analyzing the social behaviors of email users. A non-content based method that incorporates the privacy rights of employees without peeking into the details of email contents. In addition to protecting an enterprise from unnecessary disputes between employers and employees, the proposed method can reduce the efforts of processing needless messages of email managers [12].

III. SYSTEM DESCRIPTION

The proposed Email Client Automation is a robot that has a stand-alone capability for effective email communications and a strengthened functionality to intelligently categories the incoming email-based on the recommendation list. The user can write in his reply which can be sent to each of the clients (customer email ids) personally that fall under the same category as specified by the user. This process will still allow the user to have all control over the task. Further, this approach reduces the time and energy required to create a reply for an email from scratch and also saves the trouble to go through the whole inbox to find similar emails that would fall under the category as specified by the user.

The categorization scheme is user-defined and can have n number of categories. For example, for an online shopping application, the organization might receive an email regarding specification details of their product, or complaints, or from the shareholder, or their business relations. All these mails can be segregated into different folders such as customer complaints, business deals, inquiries, finance, etc.
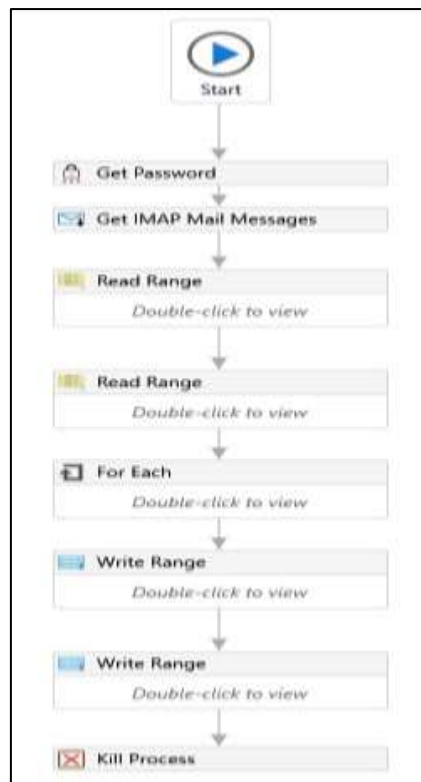
Fig. 1. Process flow diagram for Intelligent Email Segregation and Routing

A.   Email Segregation Framework

The framework is shown in Fig. 1 consists of three main sections Native email application server, Email processor engine, an analyzing module. Each of the three sections works in an integrated manner to effectively segregate the incoming email into specified categories and provides a fixed reply to each of the important emails.

I.   A Native Email Application server

It consists of a public-inbox where the inbound emails from outside the environment are received. This public-inbox receives emails that are meant to be sent to an Email Logger.

II.   Email Processor Engine

The email processor engine consists of Email Logger, Ingestion Layer, web scraping, Data Enrichment Module, RPA Algorithm, Segregated Mail Router each of these modules work in an integrated manner to achieve the intended task of email segregation.

- **Email logger:** It logs each mail along with its batch information (sender, receiver, timestamp) using the respective batch ID.
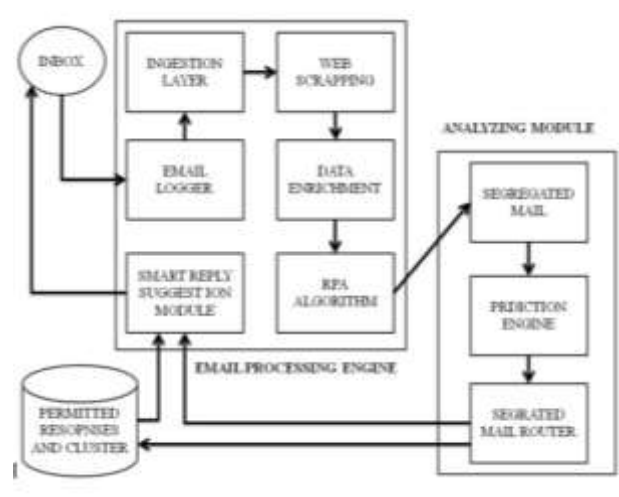
Fig.2. An Intelligent Email Segregation and Routing System Framework.

- **Ingestion layer:** It pulls the emails that are stored in the email logger. The pulled email comprises a header, subject, and body (content). The layer assigns a unique ID to each email it extracts and stores.
- **Web scrapping:** This function extracts the email body content. The extracted content, still preserving the unique ID, is appended to the other information contained in the email and stored in a data format that is made enabled for analysis in analyzing module.
- **Data enrichment module:** In this module, the irrelevant and noisy attributes are removed from the email data, to optimize the size of email data in the system.
- **RPA algorithm:** This module consists of an email services protocol that uses the email ID class label assigned for each mail to route them to respective
Subjects based on the user's recommendation list.
- **Segregated Mail Router:** This functionality directs the suggested reply back to the inbox where the reply is sent to the respective email ids.

III.   Analyzing Module

The system intends to segregate the inbound email onto a local folder based on its topic of discussion and then routes the segregated messages for decisive action. This module also includes a prediction module that performs feature selection and feature representation of the email content to predict the possible responses to an email under study using the permitted response and cluster.

B. Working

Fig. 1 shows the Process flow diagram for Intelligent Email Segregation and Routing System, the features are drawn out based on the software being used which is UiPath studio. UiPath software provides 4 activities which

are flow chart, sequence, state machine, and exception handling. So here the integration of all these activities is used to fulfill the requirements of the proposed system.

I.   Get IMAP Mail Messages

Email triggering is the foundation of any email automation process. For the same purpose, UiPath provides an activity window called as get IMAP mail messages as shown in Fig 3. By default, the workflow is configured to work along with the Gmail settings. When the workflow is running, it checks the Gmail for new messages. If no new email is found, then it will proceed to the next activity in the sequence. Once the email is read by the robot it will be marked as read. Here entire email is read as variable and this stored in the output variable that is named as mail collection.
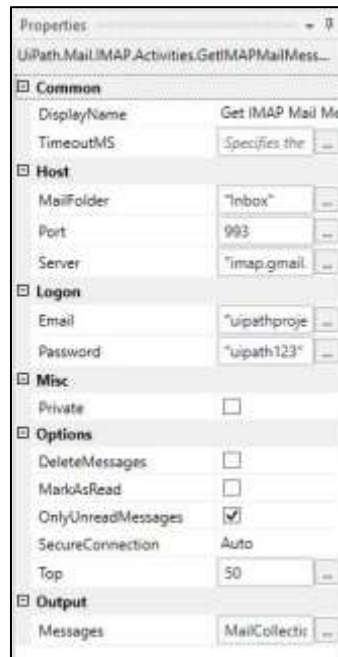

Fig. 3. Get IMAP Mail Messages Property Window

II.   Read Range

Read Range activity scans the data of an Excel range and caches it in a Data Table variable. The activity can only be used in the Excel Application Scope activity. The read range activity just extracts the data and stores it in a local data variable which is to be written into any document using the write range. Read Range property window is shown in Fig 4.

• **Range** - The range is used to set the range of elements to be read. If the range is not specified, the whole spreadsheet is read. If the range is specified as a cell, the whole spreadsheet starting from the specified cell is read. It supports only strings and string variables.

• **Sheet Name** - The SheetName is a parameter that provides the excel sheet name in which the read range is specified. Only Strings and string variables are supported.
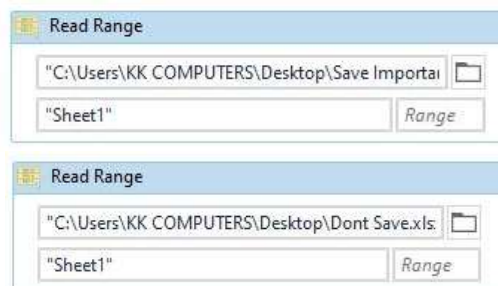

Fig. 4. Read Range Property Window

III.   For Each

A For Each activity lets the user iterate through data tables, lists, arrays, lists, or other types of collections. It processes each section of the data individually. For each property window is shown in Fig 5. The entire information of the inbox is stored in a single variable called as mail collection. All information in the mail collection will be transferred into the variable called item which can be used for further manipulation without harming the actual data.

Fig. 5. For Each Property Window

IV.   Write Range

The Write Range activity is used to write the data from a Data Table variable into a spreadsheet starting with the cell indicated in the Starting Cell field as shown in Fig 6. If the starting cell field is not specified then the data is written from the starting cell which is the A1 cell. If the excel sheet does not exist, a new one is created with the values specified in the Sheet Name property window. All cells within the specified range are overwritten. Changes are saved immediately.
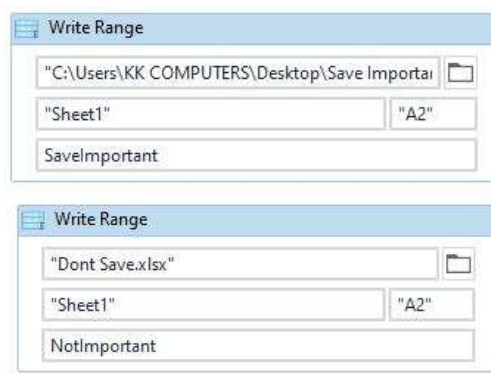

Fig. 6. Write Property Window

V.   Kill Process

The kill process kills Application Process. The Excel application is closed and unsaved changes from the Workbook will be lost. The next time when the document is opened, the Document Recovery option will be displayed on the left pane. Kill process activity window is shown in Fig 7.

Fig. 7. Kill Process Property Window

VI.   Close Application

The Close Window command is sent to the application window. The Microsoft Excel application will display a prompt asking for the unsaved changes to be saved. The activity will pause for 30 seconds (default timeout) before the application process closes gracefully. As soon as the timeout window is closed, the application process is killed (similar to the Kill Process activity). The Excel application is closed and the unsaved changes

are lost. The next time the document is opened, the Document Recovery options will be displayed on the left pane. The close application activity window is shown in Fig 8.
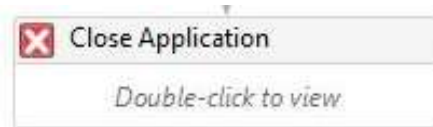

Fig. 8. Close Applications Property Window

IV. RESULTS AND DISCUSSIONS

The proposed system successfully segregates the incoming email into two categories namely important and non-important categories. The recommendation list for the same is filled in the main mail id excel sheet as shown in Fig. 9. These excel sheets contain the databases based on which the categorization takes place. The information extracted is copied and stored locally onto the excel sheet based on intended routing. Not only the sender and receiver data is extracted but the attachments of an incoming mail are also downloaded and parallelly the attachment names are updated on the excel sheets.

Information that is extracted based on the recordation is saved on to the save important excel sheet as shown in Fig. 10 and the information regarding the email ids that don't fall under the recommendation list are saved onto the don't save excel sheet as shown in Fig. 11.

The recommendation list can be changed and updated by the user at any time. For this study the incoming mails are classified into two, however, the code can be slightly modified to perform segregation for more than two categories depending on the user's needs.

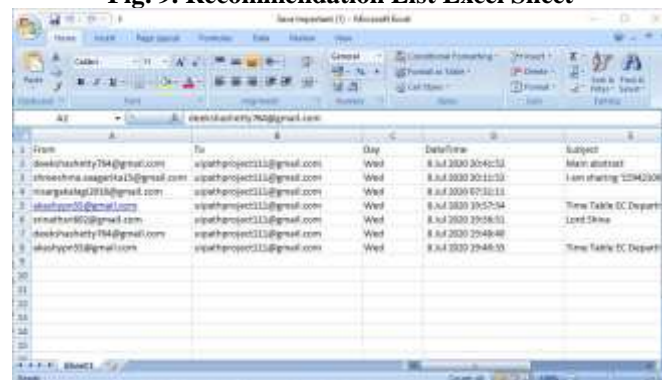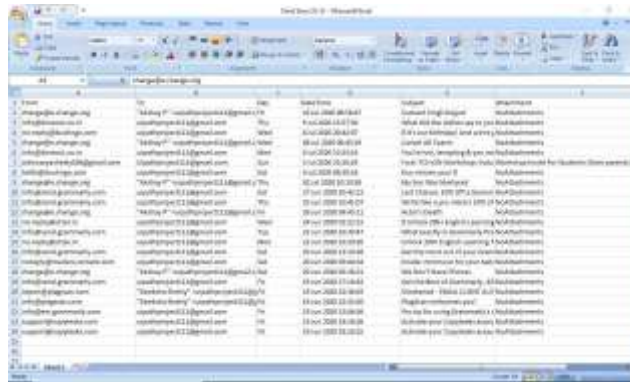**Fig. 9. Recommendation List Excel Sheet**

Fig. 10. Save Important Excel Sheet

Fig. 11. Don't Save Excel Sheet
Table 1. Email client automation with and without RPA

EMAIL AUTOMATION		
PARAMETERS	WITH RPA	WITHOUT RPA
EFFICIENCY	83.20%	78%
TIME CONSUMPTION	5%	23%
ACCURACY	87.20%	94.13%

V. EVALUATION MEASURES

Some of the parameters based on which the effectiveness of the Email Client Automation Robot is measured are
- Efficiency - It is the ratio of useful work performed by the system.
- Time consumption - It is the time taken by the system to process and throughput.
- Accuracy- It is a state of the system to be precise.

To study the effectiveness of the system a Gmail account was used wherein the incoming email was read and replied to using two cases

CASE1: E-mail processing without the implementation of the robot (i.e. manually).

CASE2: E-mail processing with the implementation of the robot (i.e. automation).

The activities such as opening the email using the user credentials, Checking the received mail, selecting important mails, reading the content of the mail, and replying to the mails individually are done manually in case 1 whereas in case 2 the robot does it for the user without any human intervention.

Efficiency

It was observed that it is difficult for a user to go through all the new mails entering his inbox, to segregate them, and to identify unwanted emails and reply to all important emails individually. But when it comes to the user who uses the e-mail automation, the e-mails which enter into inbox will get segregated based on the recommendation list into categories specified. This is far efficient compared to the first case 1 without automation and it requires the least user involvement for the work to be done manually.

Time consumption

With regards to case 1 it was observed that of the process without the automation, the user must do the work manually which requires much time, where the user needs to go through all the mails that have entered the inbox, read all the emails according to segregate based on importance and send a reply all based on the user's priority. In case 2 it was observed that of the process with automation since the task is automated it takes just a few minutes. Case 2 is comparatively less time consuming compared to the process in the absence of the RPA. The number of incoming mails, the amount of the contents in subject and body, and the size of attachments play a key role in determining time consumed to complete the task in both cases.

Accuracy

The level of accuracy is measured by how many replies are sent to the proper destination. In case 1 since the task is performed manually done the chances of sending a reply to the wrong destination id is less, the precise value is less and depends on user skills. But in the presence of the automation sending a reply to a wrong destination is the least.

Apart from these, the accuracy varies sometimes in both the conditions based on the number of the mails which bombard into inbox which fall under a different category. A precise reply is sent in case 1 which is rich in information case 2 lacks the richness in its information.
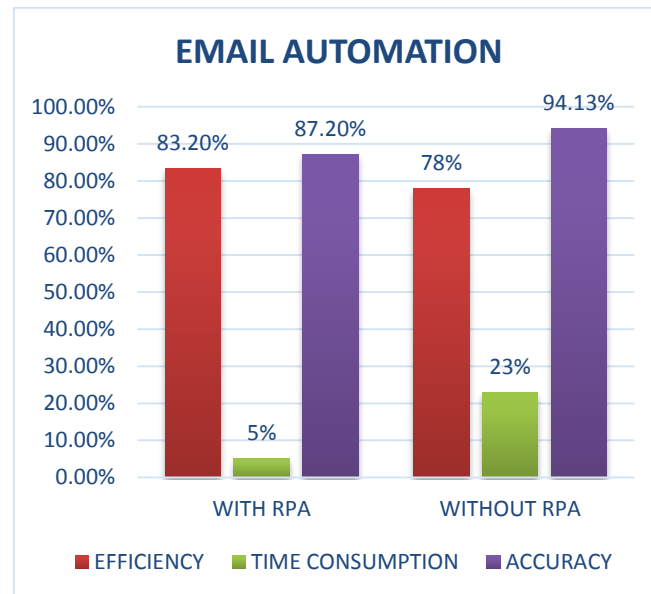
Fig. 12. A bar graph that depicts the performance metrics using efficiency, time consumption and accuracy

VI. CONCLUSIONS AND FUTURESCOPE

The proposed Email Client Automation is a robot that has a stand-alone capability for effective email communications and a strengthened functionality to intelligently categories the incoming email-based on the recommendation list. The user can write in his reply which can be sent to each of the clients (customer email ids) personally that fall under the same category. This approach is a background process that would still allow the user to have all control over the task. Further, it reduces the time and energy required to compose from the same task from scratch or scrape through the entire inbox to find similar queries.

Soon the following advancements can be made for an even more effective email client automation robot to ease the use email communication channel. The Intelligent Email Segregation and Routing System capabilities can be used to leverage email management functions like summarization, traffic management, business process discovery, etc. The system could also include the categorization of emails base on the subject matter of discussion to optimize the routing and allocation of emails. Artificial intelligence could be integrated to generate a reply based on the subject matter.

VIII. REFERENCES

[1] M. ASIF NAEEM, I. Wayan S. Linggawa1, Aftab A. Mughal2, Christof Lutteroth and Gerald Weber "*A Smart Email Client Prototype for Effective Reuse of Past Replies*" vol, 19, pp. 4-18. Feb. 2019.

[2] Soya Park, Amy X. Zhang, Luke S. Murray, David R. Karger, "*Opportunities for Automating Email Processing: A need – finding study*", MIT CSAIL, Glasgow, Scotland UK, 2019.

[3] Min Peng, Binlong Gao, Jiahuai Zhu, Jiajia Huang, Mengting Yuan, Fei Li "*High Quality Information extraction and query-oriented summarization for automatic Query-reply in social network*" Computer School of Wuhang University, no. 299 Bayi Road, Wuchang district, Wuhan, China, 2018.

[4] G. Ghosh, "*Automation with RPA (Robotic process automation)*" Application Development, Mphasis, Bangalore, India 2018.

[5] Anand V. Saurkar, Kedar G. Pathare, Shweta A. Gode, "*An overview on web scraping Techniques and tools*", International Journal on Future Revolution in Computer Science and Communication Engineering, ISSN: 2454-4248 Volume: 4 Issue: 4, 2018.

[6] Alexis Zamkow, "*Robotic Process Automation: Gearing up for greater integration*" Applied Innovation discover center Lead Toronto, Capgemini, 2016.

[7] Anjuli Kannan, Karol Kurach,Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado, Marina Ganea, "*Smart Reply: Automated Response Suggestion for E-mail*", San Francisco, CA, USA, 2016.

[8] Izzat Alsmadi, Ikdam Alhami "*Clustering and Classification of Email Contents*" Department of Computer Science, Boise State University, USA, vol, 119, pp. 19-32. Sep. 2015.

[9] Min-Feng Wangn, Meng-FengTsai, Sie-LongJheng, Cheng-HsienTang, "*Social Feature Based Enterprise Email Classification without Examining Email Contents*", vol, 16, pp, 21-28, Nov. 2015.

[10] Email automation mail trigger sample

[11] URL:https://99firms.com/blog/how-many-email-users-are-there/#gref

AUTHORS PROFILE

Mr. Akshay P N is a UG Scholar, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India.

Ms. Nisarga Kalagi is a UG Scholar, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India.

Deeksha Shetty is a UG Scholar, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India.

Mr. Ramalingam H M Senior Assistant Professor, Department of Electronics and Communication Engineering, Mangalore Institute of Technology and Engineering, Badaga Mijar, Moodabidre, Mangalore, Karnataka, India.