# CHAPTER 1

# INTRODUCTION

Automated invoice processing helps to save a serious amount of time and money, compared with manual paper invoice processing, creating efficiencies and increasing the accuracy of captured data. That's why Clear data is proud to offer this service to companies all over the World. With automation, the process was automatically triggered with the receipt of a vendor invoice. A UiPath Robot validated the eligibility of the claim against each line item of the invoice. Each line item of the invoice was audited against a predefined validation process. Exceptions were auto-routed and part-payments were released, and the vendor notified electronically. Automating manual process aggravation, will enable organization to streamline their accounts payable and expense management process with AI.

Technology has long enabled the automation of invoice processing from arrival to post. This means that at arrival of the invoice, the same accounts payable clerk will only need to scan the invoice into an automation software. The automation software then converts the invoice's scanned image into a text-researchable document. The different fields on an invoice can also be defined into the software so that it remembers which fields it should capture and register into the Excel Automation systems, for instance, the purchase amount, the quantity, the supplier name, the supplier code, and so on. The benefits of an automatic processing workflow may include reduced human error, on-demand reports, and data resilience.

In this project we are trying to perform few of the basic invoice processing operations via chat bots. Downloading the invoice attachments, Extracting the details and generating the bill with UiPath leads to bring the trust to the customer. The process usually begins when a supplier's invoice is received. Invoices can be sent via email, postal mail, fax, or EDI.

## 1.1 PROBLEM DEFINITION

The existing system provide only the data is extracted or captured from the invoice the data is sent into the system for automatic matching against the purchase order. This matching process can compare just the invoice data with that shown on the purchase order or be expanded to include a deeper level that looks at the receiving documents. Workflow steps can be configured such that the responsible person will then receive an email alert so that he or she can approve the invoice.

If there are other people involved in the approval workflow, email alerts to them will also be automatically generated. The typical workflow is a four-step process beginning with 1. Import of the images through scanning or email, 2. Identification of the vendor and business unit associated with the invoice, 3. Data extraction, and 4. Export of the extracted data and images. Without the AI it works for only a specific Templates. In the existing software it has data scrapping, screen scrapping and OCR for extracting the data.

## 1.2 OBJECTIVE OF THE PROBLEM

Our aim is to save the amount of time and money, creating efficiencies and increasing the accuracy of captured data. Validate the invoice data against data files, extract the data and generate the bill using ERP systems. Using AI Machine Learning package extract the data in any template, with the help of API key.

## 1.3 SIGNIFICANCE OF THE PROJECT

UiPath is a free, fully-featured and extensible version of our automation tool. This platform delivers the fastest and most reliable RPA that enhances business performance at unlimited scale. It is easy to use, highly responsive, and instantly scalable to allow you to build the process.

RPA tools are defined as non-integrated, stand-alone software for laptops and workstations. UiPath is recognized worldwide for product leadership and technical excellence. RPA is the largest in the industry, the most active in the automation field.

RPA represents a digital workforce that works side-by-side with employees to drive greater efficiency and eliminate almost any manual data-driven activity. Now information that was previously unattainable, unusable or time-consuming to collect and process is readily available to increase productivity, improve decision making and deliver a better customer experience.

**A UiPath feature has three main products:**

- **UiPath Studio-** It designs automation process using diagrams, which are visually appealing
- **UiPath Robot-** Executes the UiPath Studio Processes
- **UiPath Orchestrator –** It is an application that deploys, manages, and monitors processes/robots

UiPath Robots and Executors interact with huge amounts of data at the same time. We can run a process either on a single, or multiple robots or any specified number of robots and can group them in the environment.

UiPath Studio software automates back-office repetitive tasks. It is one of the tools used in automating business processes. It converts each task into UI automation, thereby making work easier and quicker. Multiple workflow designs are available in Studio. UiPath Studio comes with a debug component that easily locates problems within complex workflows. This easily verifies the execution activity and observes if there are any errors in the output.

## 1.4 OUTLINE OF THE PROJECT

Invoice processing in which generates the bill with the extracted data using Machine Learning Package. Login the Email and if it contains subject as invoice

then download that attachments and move to the defined folder. With the Machine Learning package extract the necessary details in any type of template using API key. Move the details to the Excel sheet and generate the output. After that process the report is transferred to the client.

# CHAPTER 2

# LITERATURE REVIEW

A literature review is a text of a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. Literature reviews are secondary sources, and do not report new or original experimental work.

In Robotic Process Automation (RPA) is the use of software with capabilities to handle high-volume, repeatable tasks that previously required a human to perform. It aims to use a computer to manipulate existing application software in the same way that a person works with those systems and the presentation layer to perform a specific task. This technology increases the productivity for human employees who no longer are tasked with boring work.

The Business Process: Vendor Payment Process - PO Invoices which was published on 2009 by Massachusetts Institute of Technology and that's about the Vendor invoices from purchase orders arrive at the mail desk. The Accounts Payable mail clerk sorts the invoices into categories. The processing and posting of these invoices, described below, depends on the category assigned to the invoice. The procedures are:

- Invoices arrive via US mail.
- Accounts Payable mail clerk sorts invoices into categories.
- If the invoice is for one of the following, go to Distribute Invoices:
- If the invoice is for one of the following, go to Enter an invoice:
- Invoices are distributed to Accounts Payable Operators.
- An Accounts Payable Operator determines the responsible person's name and MIT address using system task Display Mailing Address for Authorization of Incoming Invoices.

- The invoices are forwarded to the responsible individual in the department for approval.

The Invoice Processing Procedure (Accounts Payable) which was published on 2012 by University of Alberta and that’s about to manage the payment of external supplier invoices for materials, equipment and services purchased on behalf of the University of Alberta. The purpose of this procedure is to ensure that payments to external entities are accurate, appropriately supported and authorized, and recorded in the relevant fiscal year. The procedures are:

- original documents
- invoice types
- foreign currency invoices
- verification responsibility
- disputes
- corrections and adjustments
- yearend processing

The Fingate - iProcurement: How to Upload an Invoice was published on 2016 by Board of Trustees of The Leland Stanford Junior University. It is about Suppliers should mail invoices directly to Accounts Payable (suppliers should see Supplier Instructions for Submitting Invoices). The procedures are:

- Launch Oracle's Requisition and Purchase Order Query
- Find Purchase Order
- Check If Any Invoices Have Been Previously Applied or Paid
- Complete Upload an Invoice for Payment Processing Screen

The Invoice Imaging and Data Capture for Accounts Payable was published on 2008 by Pay Stream Advisors. It is about OCR and Data Capture are central components to a modern AP automation initiative.

**Advantages of the Automation process:**

- The Humans are under the era of committing the mathematical and parallax errors while in terms of the RPA the errors are very low
- The work can be done faster and efficient from man-hours and man-years to Minutes and Seconds

# CHAPTER 3

# SYSTEM ANALYSIS

System analysis is a Problem solving technique that decomposes a system into component pieces of purpose of studying how well those component parts work and interact to accomplish their purpose the following chapter provides the detail description of the existing system. It also provides an overview of the proposed system and feasibility of the Invoice processing Automation.

## 3.1 EXISTING SYSTEM

The existing system of Invoice processing comprises of functionality automating the work of retrieving emails, downloading attachments into a defined folder, and create bills in the accounting software. And they will be using the code to get the data.

The first challenge with automating invoices and receipts is that these documents often don’t come in a consistent, standard “template.” Every invoice could look different. Critical information such as the dollar amount or invoice number can be in a completely different place on different documents. In other words, OCR doesn’t know where to ‘look’ to find the necessary information. It has no longer necessary to create multiple layouts manually.

## 3.2 Drawbacks

### 3.2.1 Invoice Processing

- It only gets the data using the data scraping and screen scrapping.
- Time Consumption for data extraction.
- High operational costs per invoice on both sender and receiver side.

## 3.3 PROPOSED SYSTEM

The proposed system is invoice processing automation uses an artificial intelligent develop for getting the data in the invoice.

This new AI activity will enable UiPath Robots to read invoices and help automate customers invoice payable and expense compliance processes.

It is no longer necessary to create multiple layouts manually. Some companies have attempted to address this by creating a large number of static "templates" meant to address each invoice layout, one by one. These layouts help to point the OCR to the correct location on the page, so that it can find and extract the relevant information. While this method works when the number of layouts is small, it can quickly become unmanageable. It's simply not possible to build automation workflows for every possible invoice and receipt template your robot might encounter.

The documents your robot will encounter are often 'noisy' – meaning they have a lot of real-world complexity making them hard to read. An invoice may have been scanned on a low-quality office scanner. It may be 'skewed,' meaning it was scanned in at an angle. Sometimes it may have been scanned together with other non-invoice documents. All these factors can 'confuse' the robot and OCR, making it difficult to find the relevant information needed for processing into backend accounts payable systems.

## 3.4 FEASIBILITY STUDY

An analysis and evaluation of a proposed project to determine if it is technically feasible, is feasible within the estimated cost, and will be profitable. Feasibility studies are almost always conducted where large sums are at stake. A feasibility study aims to objectively and rationally uncover the strengths and weaknesses of an existing Invoive processing and threats present in the environment, the resources required to carry through, and ultimately the prospects for success in the invoice.

### 3.4.1 Tests of Feasibility

Feasibility study is conducted once the problem clearly understood. Feasibility study is necessary to determine that the proposed system in invoice is feasible by considering the technical, operational, and economical factors. By having a detailed feasibility study the management in the will have a clear-cut view of the proposed system of the invoice automation. Feasibility study encompasses the following things:

- Technical Feasibility
- Economical Feasibility
- Operational feasibility

#### 3.4.1.1 Technical Feasibility

A large part of determining resources has to do with assessing technical feasibility. It considers the technical requirements of the proposed project of invoice. The technical requirements are then compared to the technical capability of the invoice processing. The systems project is considered technically feasible if the internal technical capability is sufficient to support the invoice system's requirements. The analyst must find out whether current technical resources can be upgraded or added to in a manner that fulfils the request under consideration.

The essential questions that help in testing the operational feasibility of a system include the following:

- Is the project feasible within the limits of current technology?
- Does the technology exist at all?
- Is it available within given resource constraints?
- Is it a practical proposition?
- Manpower- programmers, testers & debuggers
- Software and hardware
- Are the current technical resources sufficient for the new system?

- Can they be upgraded to provide the level of technology necessary for the new system?

#### 3.4.1.2 Operational Feasibility

Operational feasibility is dependent on human resources available for the project and involves projecting whether the system will be used if it is developed and implemented. Operational feasibility is a measure of how well a proposed system in invoice processing automation solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis phase of bot development.

The essential questions that help in testing the operational feasibility of a system include the following:

- Does current mode of operation provide adequate throughput and response time?
- Does current mode provide end users and managers with timely, pertinent, accurate and useful formatted information?
- Does current mode of operation provide cost-effective information services to the business?
- Could there be a reduction in cost and or an increase in benefits?

#### 3.4.1.3 Economical Feasibility

Economic analysis could also be referred to as cost/benefit analysis. It is the most frequently used method for evaluating the effectiveness of a new system of the invoice bot. In economic analysis the procedure is to determine the benefits and savings that are expected from a candidate system and compare them with costs.

If benefits outweigh costs, then the decision is made to design and implement the invoice process. An entrepreneur must accurately weigh the cost versus benefits before taking an action.

Possible questions raised in economic analysis are:

- Is the system cost effective?
- Do benefits outweigh costs and system study?

# CHAPTER 4

# SYSTEM SPECIFICATION

## 4.1 FUNCTIONAL REQUIREMENTS:

### 4.1.1 E-Mail:

- The system should allow users to E-mail.
- The system shall inform the user if an answer is not available.
- The system shall inform the user about the unattached mail.
- The system shall notify the user about the bill generation with their requirements.

### 4.1.2 Searching:

- The system should search the user's invoice pdf in the E-mail.
- The system need to download the attachments when the subject is Invoice.

### 4.1.3 Feedback:

- The user should be able to leave feedback, which is comprised of a mail message to the system.

## 4.2 NON-FUNCTIONAL REQUIREMENTS

### 4.2.1 User Interface:

- The system shall maintain an easy to use interface across all functionality and for all users
- The client's user interface should be compatible with all commonly used browsers, such as Internet explorer, Firefox, Google chrome and Safari.

### 4.2.2 Scalability:

- The system shall be able to scale based on the number of users invoices using the system.

### 4.2.3 Security:

- The administrative system should be protected from unauthorized access.

- The database should be protected from attacks and unauthorized access.
- The interface should be protected from attacks.
- All passwords should be stored as a secure hash of the administrator password.

**4.2.4 Portability:**

- The system should run on a variety of systems that support the Robotic Process Automation.
- The system cannot process without internet access.

**4.2.5 Maintainability:**

- The system should be easy to maintain.
- There should be a clear separation between the interface and the business logic code.

**4.2.6 Exception handling:**

- Exceptions should be reported effectively to the user if they occur.

**4.2.8 Ethics:**

- The system shall not store or process any information about its users.
- Provide security to their invoices and process their bills.

## 4.3 SYSTEM CONFIGURATION

Memory : 8 GB RAM

Processor : AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx, 2.10 GHz

System type : 64- bit Operating System, x64-based pc

## 4.4 SOFTWARE REQUIRMENTS

Operating System : Windows 10

Tool : UI Path Community Edition

# CHAPTER 5

# SOFTWARE DESCRIPTION

A software requirements specification (SRS) is a description of a software system to be developed. It lays out functional and non-functional requirements, and may include a set of use cases that describe user interactions that the software must provide. Software requirements specification establishes the basis for an agreement between users and invoice processing bot on what the software product is to do as well as what it is not expected to do. Software requirements specification permits a rigorous assessment of requirements before design can begin and reduces later redesign. It should also provide a realistic basis for estimating product costs, risks, and schedules.

## 5.1 UIPATH

UiPath is a free, fully-featured and extensible version of our automation tool. This platform delivers the fastest and most reliable RPA that enhances business performance at unlimited scale. It is easy to use, highly responsive, and instantly scalable to allow you to build the process.

RPA tools are defined as non-integrated, stand-alone software for laptops and workstations. UiPath is recognized worldwide for product leadership and technical excellence. RPA is the largest in the industry, the most active in the automation field.

UiPath Robots and Executors interact with huge amounts of data at the same time. We can run a process either on a single, or multiple robots or any specified number of robots and can group them in the environment.

UiPath Studio software automates back-office repetitive tasks. It is one of the tools used in automating business processes. It converts each task into UI automation, thereby making work easier and quicker. Multiple workflow designs

are available in Studio. UiPath Studio comes with a debug component that easily locates problems within complex workflows. This easily verifies the execution activity and observes if there are any errors in the output.

### 5.2 FRONT END

The front end is designed using Ai package framework which includes a collaborative platform for bot creating and deploying. Machine Learning is the collaborative end to end bot platform use to get the details of the client. Here all the bot tools are integrated and it allows automatic detection of entities.

#### 5.2.1 Machine Learning package

Machine learning is a core sub-area of artificial intelligence; it enables computers to get into a mode of self-learning without being explicitly programmed. When exposed to new data, these computer programs are enabled to learn, grow, change, and develop by themselves.

##### 5.2.1.1 Features

- Machine learning can easily consume unlimited amounts of data with timely analysis and assessment.
- Machine learning algorithms tend to operate at expedited levels.
- Applying machine learning to practical applications and scenarios is simply vital.

##### 5.1.1.2 Advantages

- Sophisticated pattern recognition
- Intelligent decisions
- Self-modifying
- Multiple iterations
- Automation of tasks

## 5.3 BACK END

The back end is designed using Excel, whose primary function is to store data securely and retrieve it later, as requested by other software applications. With Excel automation accesses system and services under the control of a process automation server.

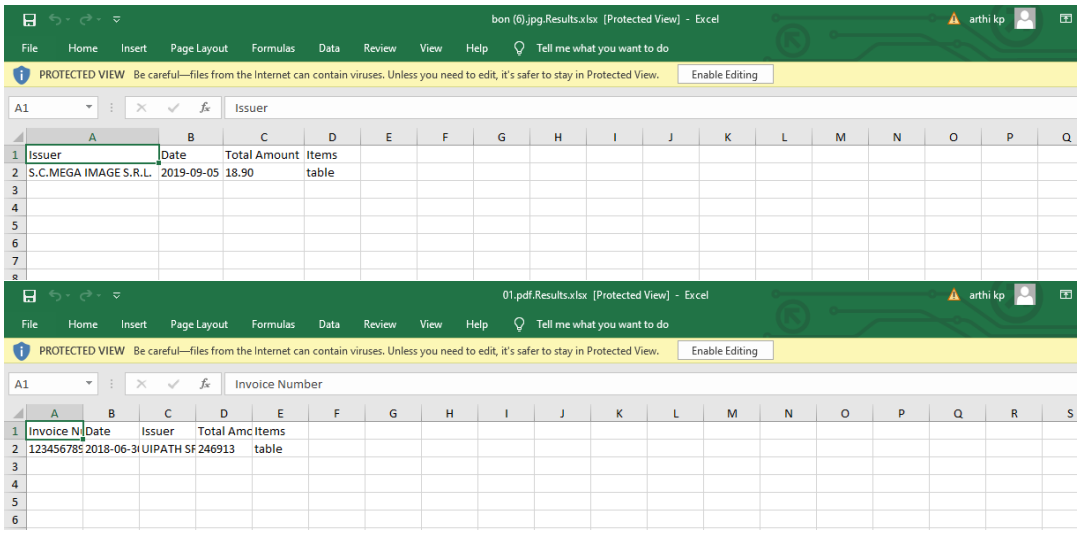



**Fig.5.3 Back End**

# CHAPTER 6

# PROJECT DESCRIPTION

Invoice processing automates the customer's invoice and give necessary details to the customer. This Invoice and Receipt AI is also trained to understand real world documents. It can automatically identify and extract a wide range of business values from your receipts and invoices, depending on the specific needs and requirements of your accounts payable and expense compliance processes. Even if your document includes 'noise,' the robot will still find the following relevant information, such as the vendor name or invoice number.

## 6.1 OVERVIEW OF THE PROJECT

This allows integrated human-in-the-loop validation of the extracted information. Most importantly, within these workflows, you can easily send your robots to input the extracted data into the correct accounts payable or expense management back-end systems to complete the end-to-end automation of these processes. These models are also self-learning and will improve over time as more users interact with them and use them in the workflows.

## 6.2 MODULE DESCRIPTION

### 6.2.1 Retrieving emails

The purchase order or without purchase order and other types of invoice were send to the bot and also that were downloaded and moved to a processed folder.

### 6.2.2 Reading invoices

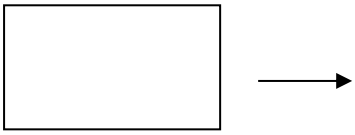
which can be found by downloading the UiPath.MachineLearningExtractor package in the beta feed of UiPath Studio, utilizes a 'template less' approach to processing invoices and receipts. This means that our AI can automatically determine the location of critical information that

needs to be extracted from the document, even if the format changes. It is no longer necessary to create multiple layouts manually.

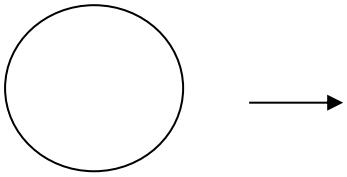
## 6.3 DATA FLOW DIAGRAM

Data flow diagram is used to describe how the information is processed and stored and identifies how the information flows through the processes. Data flow diagram illustrates how the data is processed by a system in terms of inputs and outputs. The data flow diagram also depicts the flow of the process and it has various levels. The initial level is context level which describes the entire system functionality and the next level describes each and every sub module in the main system as a separate process or describes all the process involved in the system separately.

Data flow diagram are made up of number of symbols,

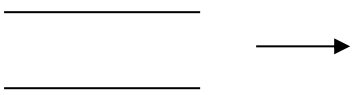


Square representing external entities, which are sources or destinations of data.



Circle representing processes, which take data as input, do something to it and output it.

Arrows representing the data flows, which can either, be electronic data or physical items.



Parallel lines representing data stores, including electronic stores such as databases or XML files and physical stores

### 6.3.1 DFD Level 0:

The users of the invoice are user and bot. The output of the system is reports.

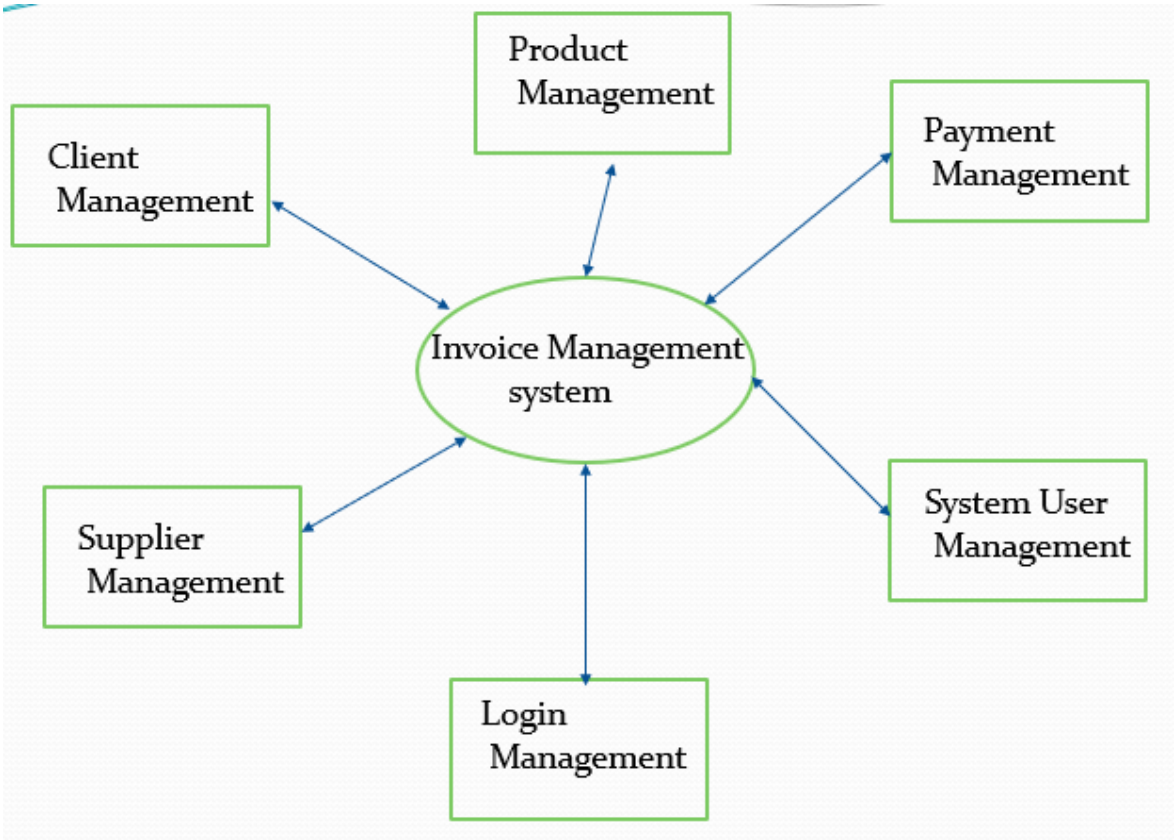



**Fig.6.1 DFD Level 0**

### 6.3.2. DFD Level 1:

In the DFD level 1, all the reports are been explained clearly. The user requests for an action, the different types of management like client, Product, Payment, Supplier, Login, System User can get the reports as same as the management.

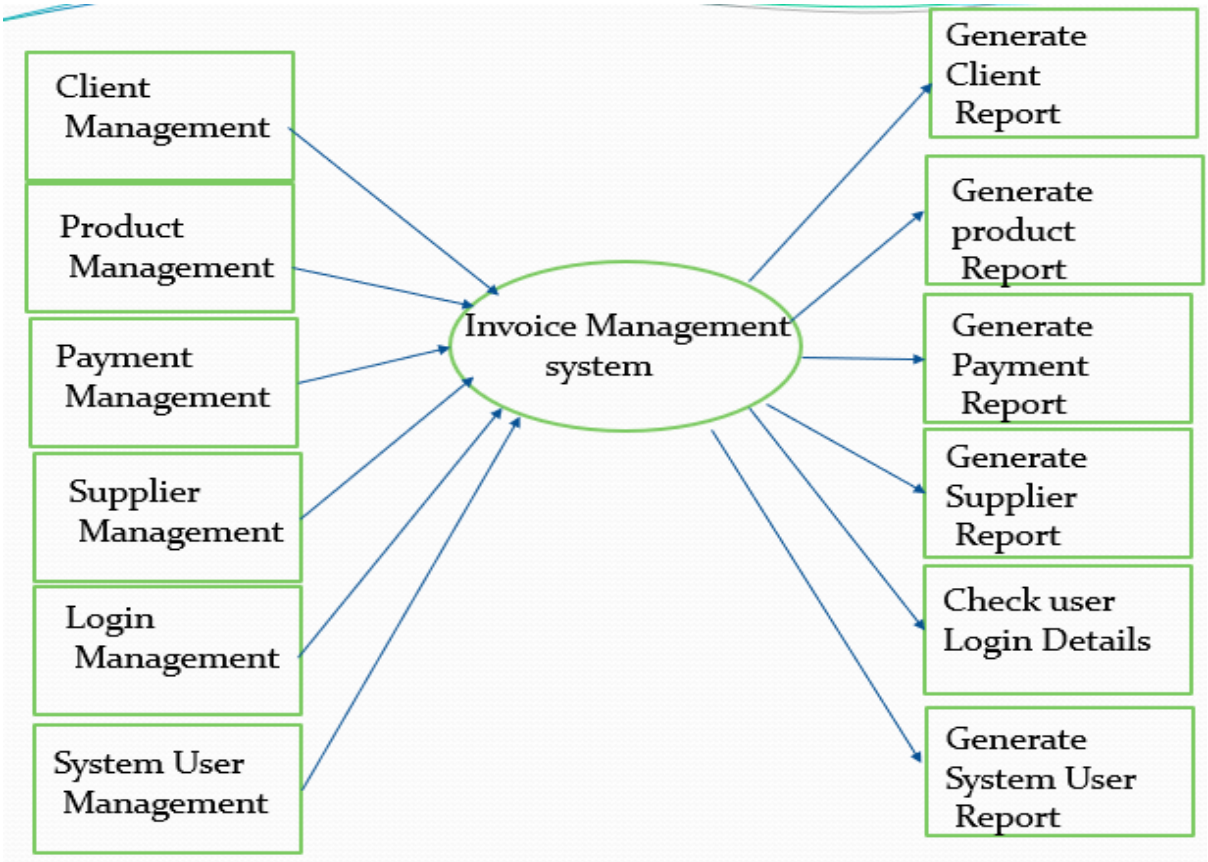



**Fig.6.2 DFD Level 1**

## 6.4 ER DIAGRAM

An entity–relationship model (ER model) describes inter-related things of interest in a specific domain of knowledge. An ER model is composed of entity types (which classify the things of interest) and specifies relationships that can exist between instances of those entity types. In software engineering an ER model is commonly formed to represent things that a business needs to remember in order to perform business processes. Consequently, the ER model becomes an abstract data model that defines a data or information structure that can be implemented in a database, typically a relational database.

Entity–relationship modeling was developed for database design by Peter Chen and published in a 1976 paper. Some ER modelers show super and subtype

entities connected by generalization-specialization relationships, and an ER model can be used also in the specification of domain-specific ontology.

### 6.4.1 Invoice Processing

The Invoice and Receipt AI can be accessed directly in UiPath Studio. This means we can drag-and-drop AI directly into your workflows and seamlessly integrate with other critical document processing capabilities such as Taxonomy Manager and Validation Station. This allows integrated human-in-the-loop validation of the extracted information. Most importantly, within these workflows, you can easily send your robots to input the extracted data into the correct accounts payable or expense management back-end systems to complete the end-to-end automation of these processes. These models are also self-learning and will improve over time as more users interact with them and use them in the workflows.

The ER Diagram for all the process flow is been given as follows.

The Diagram is been given for all the processes such as the process flow is been specified.

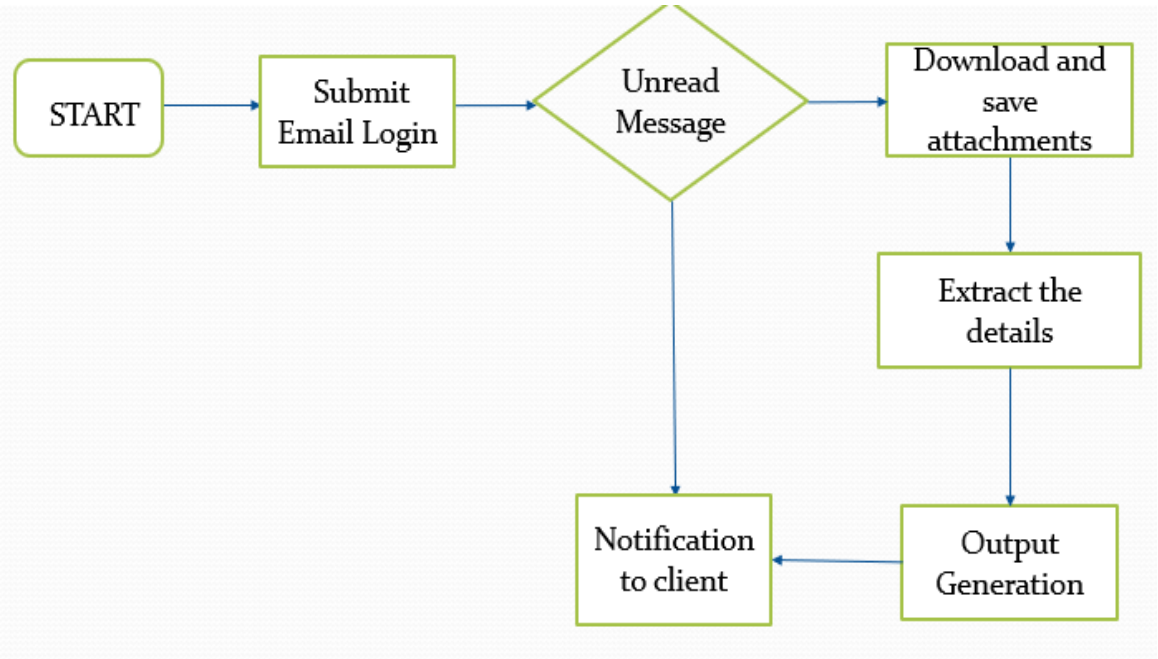



**Fig.6.3 Architectural Design**

## 6.6 INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

## 6.7 OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

# CHAPTER 7

# SYSTEM TESTING

System Testing is a level of testing that validates the complete and fully integrated software product. The purpose of a system test is to evaluate the end-to-end system specifications. Usually, the software is only one element of a larger computer-based system. Ultimately, the software is interfaced with other software/hardware systems. System Testing is actually a series of different tests whose sole purpose is to exercise the full computer-based system.

In system testing software test professional aims to detect defects or bugs within the interfaces and also within software as a whole. However, during integration testing of the application or software, the software tests professional aims to detect the bugs/defects between the individual units that are integrated together.

## 7.1 TESTING METHODS

Software Testing Type is a classification of different testing activities into categories, each having, a defined test objective, test strategy, and test deliverables. The goal of having a testing type is to validate the Application under Test for the defined Test Objective.

Software Testing Methodology is defined as strategies and testing types used to certify that the Application Under Test meets client expectations. Test Methodologies include functional and non-functional testing to validate the AUT.

- White box testing
- Basic path testing
- Stress testing
- Acceptance testing
- Black box testing

- Integration testing

**i)** Functional testing

**ii)** Non-functional testing

## 7.2 TYPES OF TESTING

### 7.2.1 Unit Testing

Unit testing is a level of software testing where individual units/ components of a software are tested. The purpose is to validate that each unit of the software performs as designed. A unit is the smallest testable part of any software. It usually has one or a few inputs and usually a single output. In procedural programming, a unit may be an individual program, function, procedure, etc. In object-oriented programming, the smallest unit is a method, which may belong to a base/ super class, abstract class or derived/ child class.

In this Invoice processing, every units of code is been tested and the correctness of every module is been ensured.

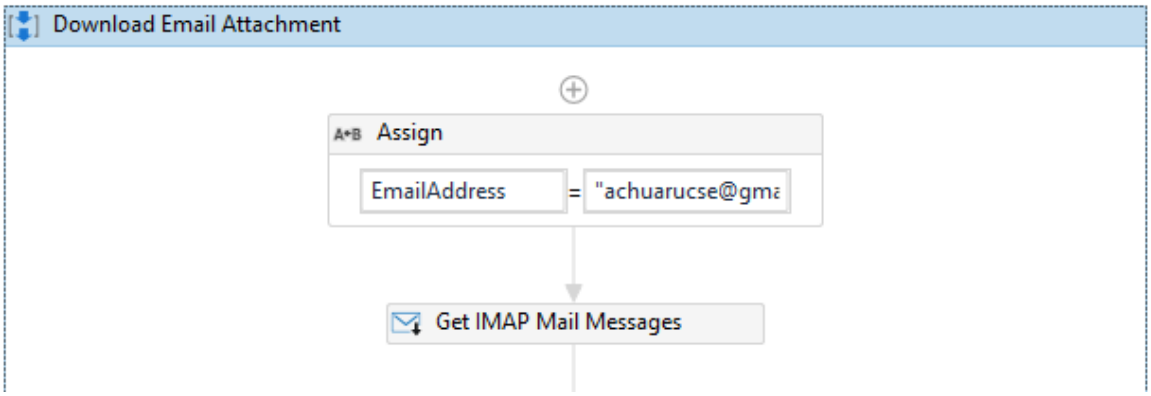



**Fig.7.2.1 Unit Testing**

### 7.2.2 Integration Testing

Integration testing (sometimes called integration and testing, abbreviated I&T) is the phase in software testing in which individual software modules are combined and tested as a group. It occurs after unit testing and before validation testing. Integration testing takes as its input modules that have been unit tested,

groups them in larger aggregates, applies tests defined in an integration test plan to those aggregates, and delivers as its output the integrated system ready for system testing. Integration testing is a level of software testing where individual units are combined and tested as a group. The purpose of this level of testing is to expose faults in the interaction between integrated units. Test drivers and **test** stubs are used to assist in Integration Testing.

In this bot, the units are been tested as a whole and the testing was successful.

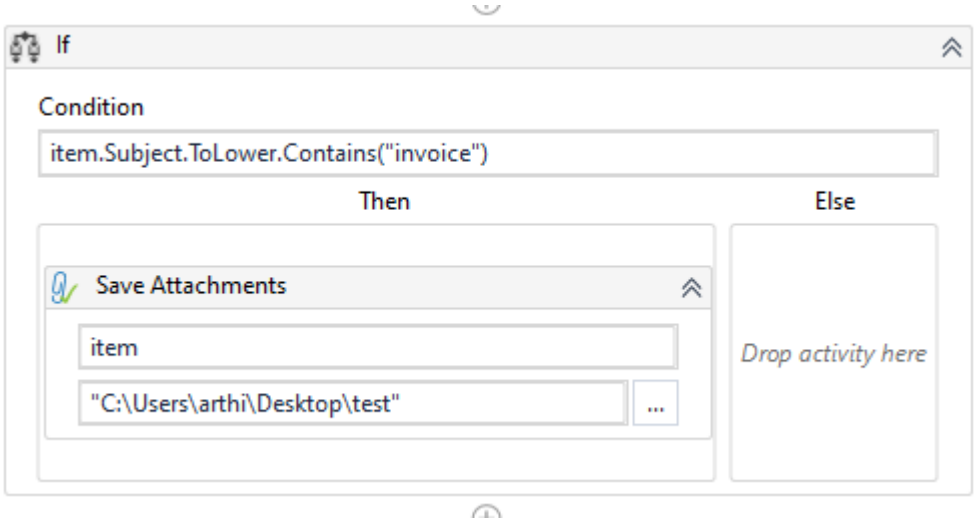



**Fig.7.2.2 Integration Testing**

### 7.2.3 Functional Testing

Functional testing is a quality assurance (QA) process and a type of black-box testing that bases its test cases on the specifications of the software component under test. Functions are tested by feeding them input and examining the output, and internal program structure is rarely considered (unlike white-box testing). Functional testing usually describes what the system does. Functional testing does not imply that you are testing a function (method) of your module or class. Functional testing is a type of software testing whereby the system is then tested against the functional requirements/specifications. Functions (or features) are tested by feeding them input and examining the output. Determine the output based on the function's specifications. Execute the test case.

Functional testing has many types:

- Smoke testing
- Sanity testing
- Regression testing
- Usability testing

### 7.2.4 Stress Testing

Stress testing a Non-Functional testing technique that is performed as part of performance testing. Stress testing is a type of Software Testing that verifies the stability & reliability of the system. This test mainly measures the system on its robustness and error handling capabilities under extremely heavy load conditions. Stress Testing is done to make sure that the system would not crash under crunch situations. During stress testing, the system is monitored after subjecting the system to overload to ensure that the system can sustain the stress. Reasons can include:

- to determine breaking points or safe usage limits
- to confirm mathematical model is accurate enough in predicting breaking points or safe usage limits
- to confirm intended specifications are being met
- to determine modes of failure (how exactly a system fails)
- to test stable operation of a part or system outside standard usage

The recovery of the system from such phase (after stress) is very critical as it is highly likely to happen in production environment.

In this Invoice Processing, whole of the modules are been tested and it has the safe usage measures.

### 7.2.5 Acceptance Testing

Acceptance Testing is a level of the software testing where a system is tested for acceptability. The purpose of this test is to evaluate the system's compliance with the business requirements and assess whether it is acceptable for delivery.

Formal testing with respect to user needs, requirements, and business processes conducted to determine whether or not a system satisfies the acceptance criteria and to enable the user, customers or other authorized entity to determine whether or not to accept the system.

In this process, the customer's acceptance is been monitored and it is been put into usage.

### 7.2.6 White Box Testing

White Box Testing is the testing of a software solution's internal coding and infrastructure. It focuses primarily on strengthening security, the flow of inputs and outputs through the application, and improving design and usability. White box testing is also known as Clear Box testing, Open Box testing, Structural testing, Transparent Box testing, Code-Based testing, and Glass Box are testing. It is one of two parts of the "box testing" approach of software testing. Its counter-part, black box testing, involves testing from an external or end-user type perspective. On the other hand, White box testing is based on the inner workings of an application and revolves around internal testing.

The term "white box" was used because of the see-through box concept. The clear box or white box name symbolizes the ability to see through the software's outer shell (or "box") into its inner workings. Likewise, the "black box"

in "black box testing" symbolizes not being able to see the inner workings of the software so that only the end-user experience can be tested.

In this automation, all the inner functionality is been tested and it is been correctly implemented.

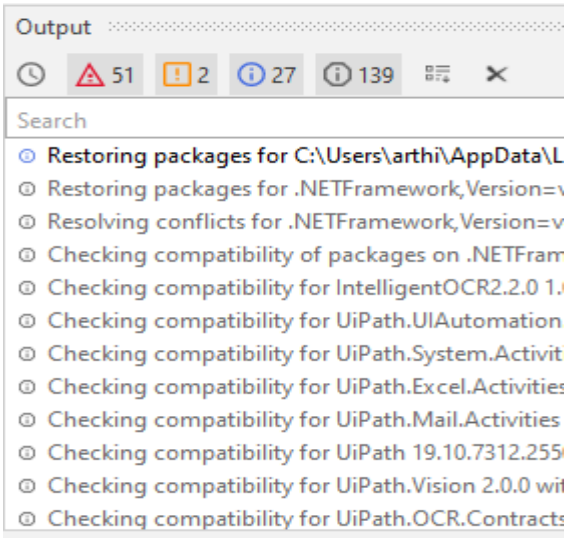



**Fig.7.2.6 White Box Testing**

### 7.2.7 Black Box Testing

Black box testing is a software testing techniques in which functionality of the software under test (SUT) is tested without looking at the internal code structure, implementation details and knowledge of internal paths of the software. This type of testing is based entirely on the software requirements and specifications.

In this process, the implementation part is been checked for its correctness.

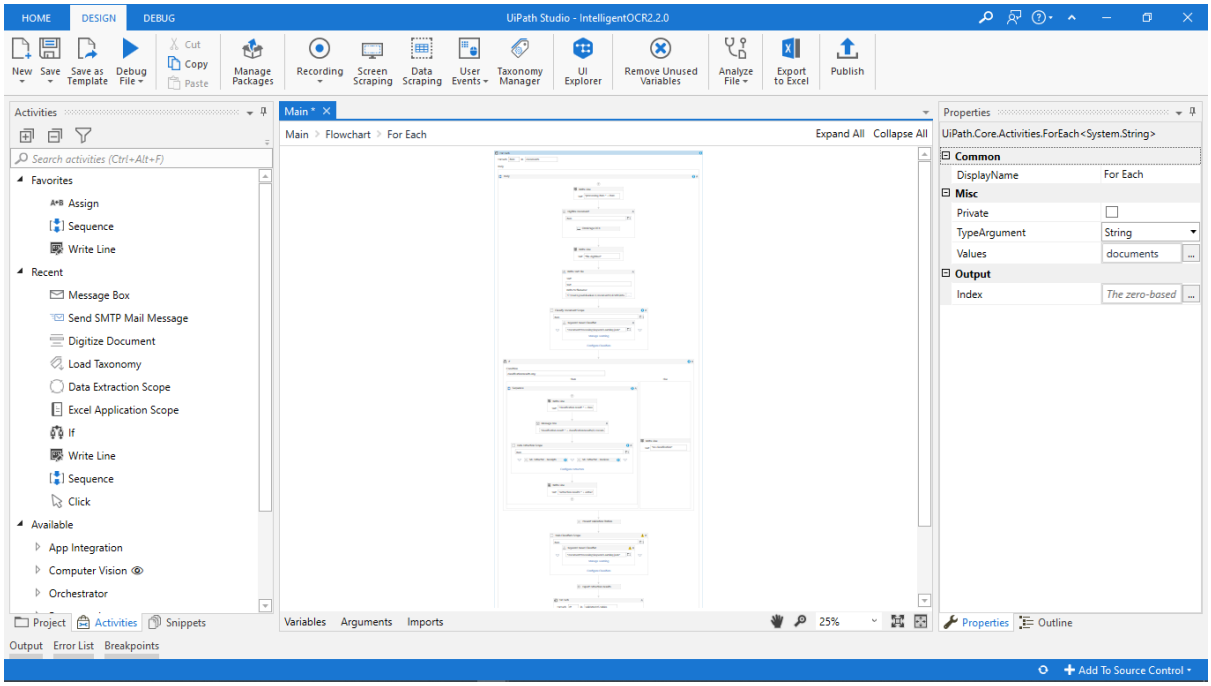



**Fig.7.2.7 Black Box Testing**

### 7.2.7.1 Methods of Black Box Testing

There are many types of Black Box Testing but following are the prominent ones -

- Functional testing - This black box testing type is related to functional requirements of a system; it is done by software testers.

- Non-functional testing - This type of black box testing is not related to testing of a specific functionality, but non-functional requirements such as performance, scalability, usability.

- Regression testing - Regression testing is done after code fixes, upgrades or any other system maintenance to check the new code has not affected the existing code.

## 7.3 TESTING STRATEGY

Test Strategy is also known as test approach defines how testing would be carried out. Test approach has two techniques:

- Proactive - An approach in which the test design process is initiated as early as possible in order to find and fix the defects before the build is created.

- Reactive - An approach in which the testing is not started until after design and coding are completed.

Test strategy calls for implementing two entirely different methodologies for testing this project. The chat bot includes a fair amount of manual UI-based testing.

In this banking bot, proactive approach is been used for testing. Since proactive approach is efficient it is been used in this bot.

# CHAPTER 8

# SYSTEM IMPLEMENTATION

## Robotic Process Automation (RPA) with UiPATH:

System implementation is the construction of the new system by considering the flow of activity and way to implementing it. Robotic Processing Automation known as RPA is an emerging technology that automates a process in computer using Software Robots.

Robot is a term that refers to software or applications that replicates the action of user and communicate with the system user interface. Process is said to be steps or sequence to do certain things. Automation is doing certain things automatically.

RPA is a process of creating software robots that could do certain process automatically, without human intervention.

**Example:**

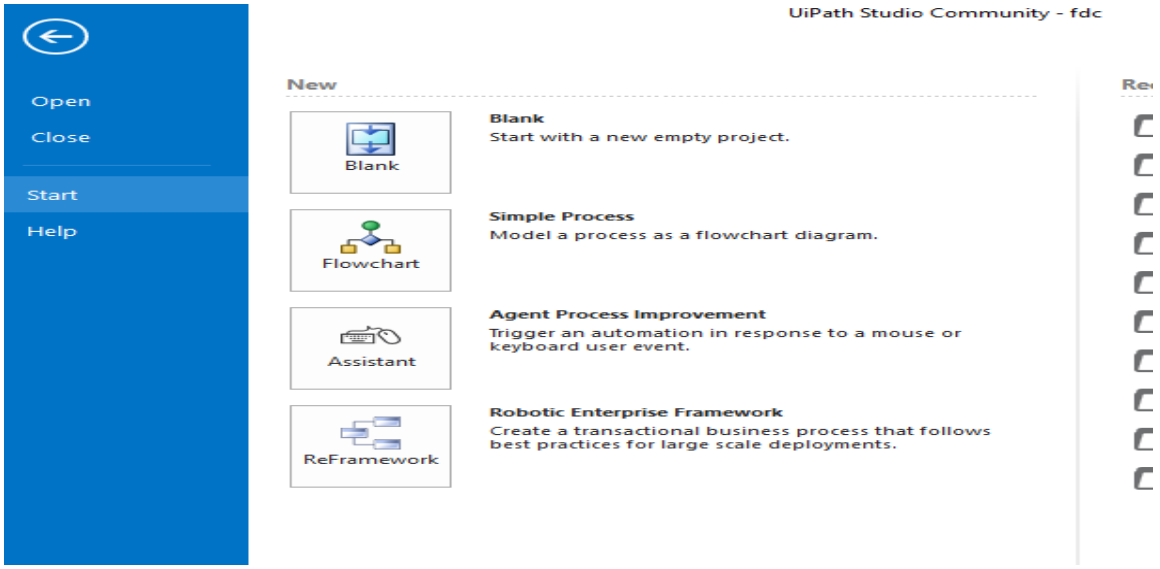
- Automatic Emails
- Data Scraping
- Automatic Document Creation
- Excel Automation

**Tools Available**

- UiPATH
- Blue Prism
- Automation Anywhere

**UiPATH:**

UiPath is the best RPA tool in terms of technology. The community edition of UiPath is free of cost and anyone can download and use it. UiPath is built using.NET framework. Hence, we can use functions of C# in it. UiPath studio is the IDE that is used to create Robotic Process Automation.




## 8.1 UIPATH STUDIO

This is how the start page of UiPath Studio looks like. You can select any of the options to create a process.

The process may be defined in two types:

- **Sequence**- Sequence is used when the tasks needs to be done in sequential manner.
- **Flow Charts** - Flow charts are used when there are multiple branches or conditions.

### 8.1.1 UiPath Activities:

UiPath has 300+ inbuilt activities. We can also import packages like Excel, Mail, and PDF in addition to the built-in packages. These activities are available in activities pane.

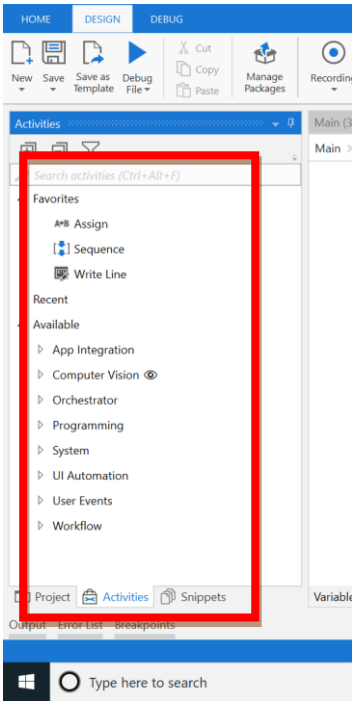



**Fig.8.1.1 Activities Pane**

### 8.1.2 Work Space:

UiPath work space is the place where we actually create sequences or flow charts. The tool allows the user to create variable and store values using the variables pane or using the properties pane. The sequences and flow charts are called as workflows in UiPath.

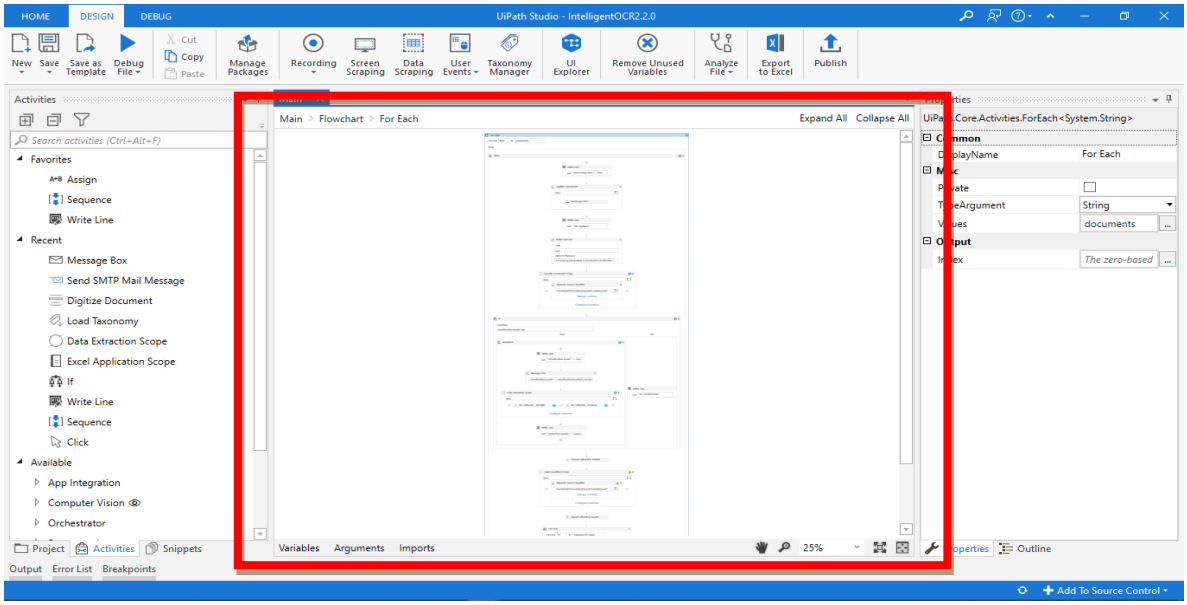



**Fig.8.1.2 Workspace**

### 8.1.3 Properties Pane:

In properties pane, we can give properties and values to components in work flows. Also we can create variables by hitting ctrl + K key in output textbox.

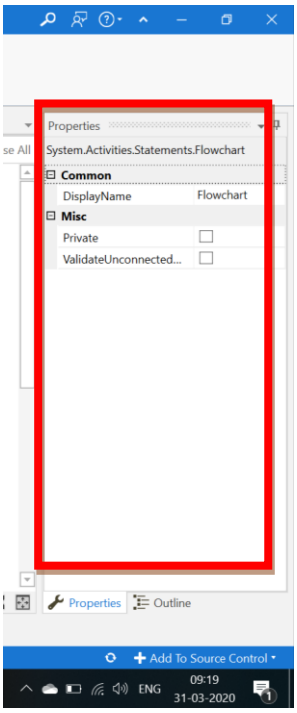



**Fig.8.1.3 Properties Pane**

These three are the important panes in UiPath. The other panes are Output Pane, where we could see the Outputs, Logs and Error Messages and Outline pane, which shows the outline of the project.

## 8.2 EMAIL-AUTOMATION:

In this automation it will login to the user's mail and download all the attachments when the subject has invoice. By using the Download Email Attachment activity, it will download with the specifications. And need to get the IMAP mail message to read the subject and process over the functions also set the server as "imap.gmail.com" and the host port as 993 and enable the options

like MarkAsRead and OnlyUnreadMessages to search for unread messages and download it and make as read. And give the Mail Folder as “Inbox”.

After that “ForEach” activity is used to scan one by one in the Email. And give the “ForEach” as “item” and “MailMessage”. Then in the body if the condition is based on the invoice it will proceed the further process. After finding the subject content as invoice it will download and save the file using “save Attachment” activity.

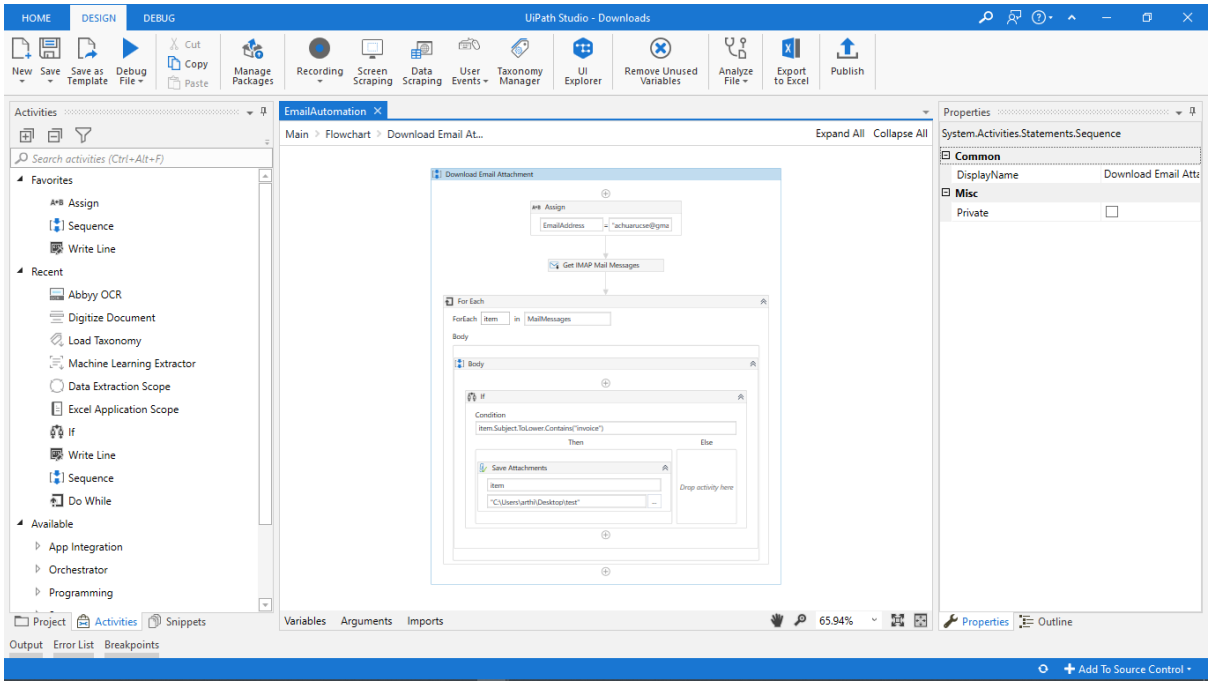



**Fig.8.2 Email-Automation**

## 8.3 SCRAPING DATA:

To get the details in the file go to manage packages and install the packages like UiPath.MachineLearningExtractor , UiPath.IntelligentOCR.Activities and required dependencies. If we haven’t yet defined a taxonomy for the documents we intend to process, it can be done using the Taxonomy Manager. After that drag the Load Taxonomy activity and store the taxonomy in a variable.

In the workflow drag in the Digitize Document activity, drag-and-drop an OCR Engine inside the Digitize Document activity. Define variables for both outputs: DocumentObjectModel and DocumentText.

Drag in the Data Extraction Scope activity to extract the data in the file. Populate input variables DocumentObjectModel, DocumentText, Taxonomy and Document Type Id. The Document Type Id is a string we can see in Taxonomy Manager if we click on the Document Type we need to extract. Define variable for output: Extraction Results.

Drag Machine Learning Extractor inside the workflow and drop it inside the Data Extraction Scope. Populate Endpoint input property with the URL of the endpoint you would like to use:

- https://invoices.uipath.com for extracting invoices
- https://receipts.uipath.com for extracting receipts
- a local custom endpoint in case you have one of these models hosted on premises

Click the "Configure Extractors" link within the Data Extraction Scope activity and Expand the Document Type we are interested in and populate, on the right side column, the names of the fields which correspond to the fields in your taxonomy.

Invoice available field names:

- "name"
- "vendor-addr"
- "billing-name"
- "billing-addr"
- "shipping-addr"
- "invoice-no"
- "payment-terms"
- "due-date"
- "po-no"
- "date"
- "net-amount"
- "tax"
- "total"
- "currency"

The item contains:

- "items"
- "line-no"
- "description"

- “item-po-no”
- “quantity”
- “unit-price”
- “line-amount”

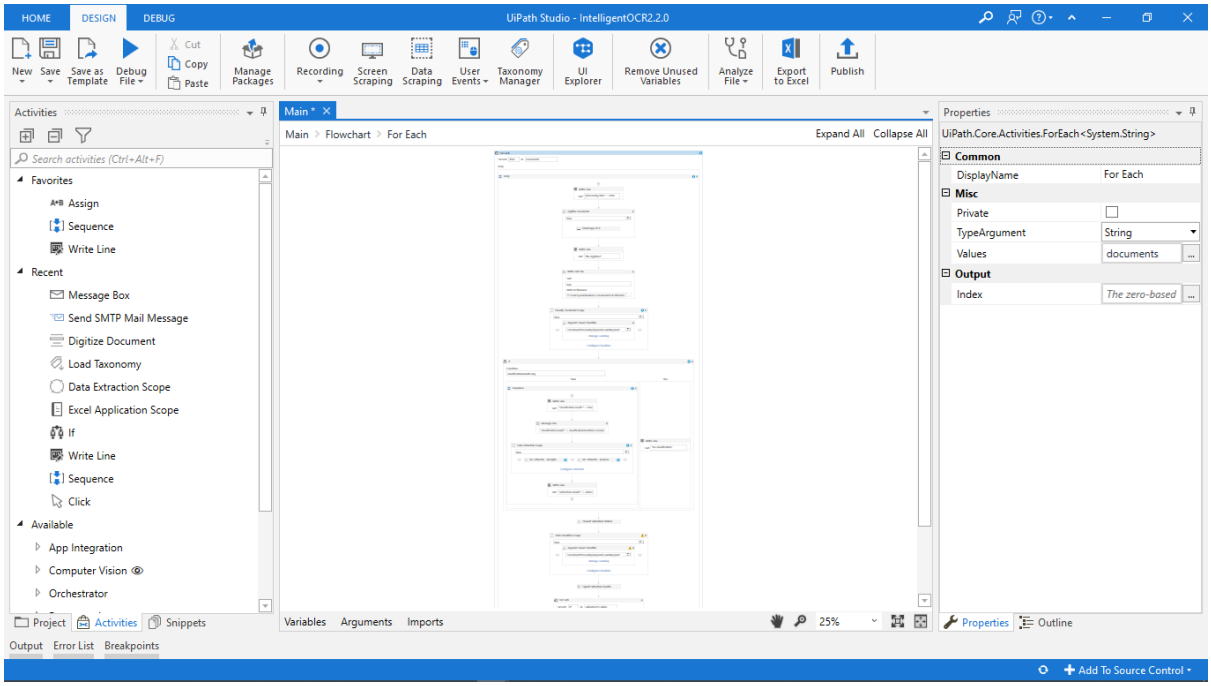



**Fig.8.3 Scraping Data**

## 8.4 NOTIFICATION :

After scrapping of data that details were stored in the Excel sheet. After that process using SMTP Mail Message activity it will notify to the user. The server will be “smtp.gmail.com” and the port number is 587. It will attach the processed file and notify to the user with current date and time. Assign time as a different variable and give the body and subject. Give the Logon details after that give the subject as ““report summary up to now” +time.ToString” , this can print current date and time with the subject. Then attach the required files in the Send SMTP Mail Message Activity, it can send the required data to the user.

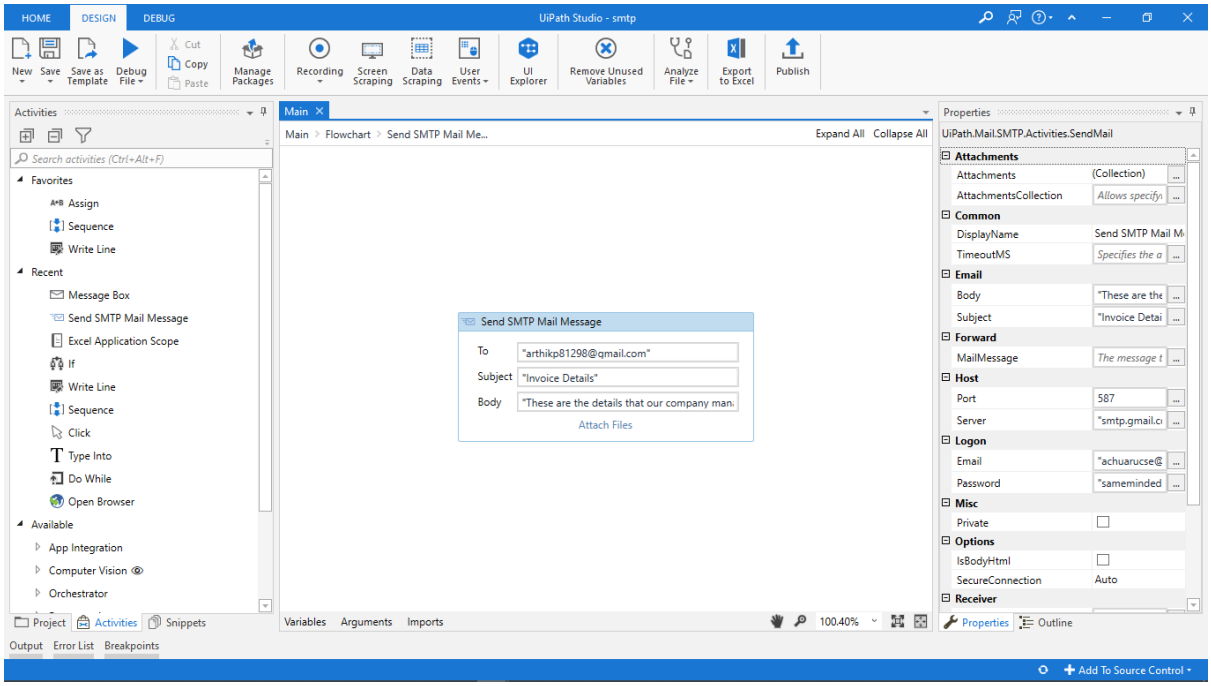



**Fig.8.3 Notification**

# CHAPTER 9

# CONCLUSION & FUTURE ENHANCEMENTS

## 9.1 CONCLUSION

Thus Invoice processing automation will be more efficient while it is been put into practice and it helps the company to easily perform the company's invoice. It offers attractive benefits for the financial services industry, particularly in automating back-office operation. The key is focusing on innovation within the company to cut unnecessary details and improve results. Our AI can automatically determine the location of critical information that needs to be extracted from the document, even if the format changes. It is no longer necessary to create multiple layouts manually. We are continuously improving our AI to handle the complexities found in real world documents that make them hard to read, such as noise or low resolution from printing and scanning. Drag-and-drop our new Machine Learning Extractor seamlessly into our existing document processing infrastructure provided by the IntelligentOCR activity package. Same infrastructure, new supercharged brain.

## 9.2 FUTURE ENHANCEMENT

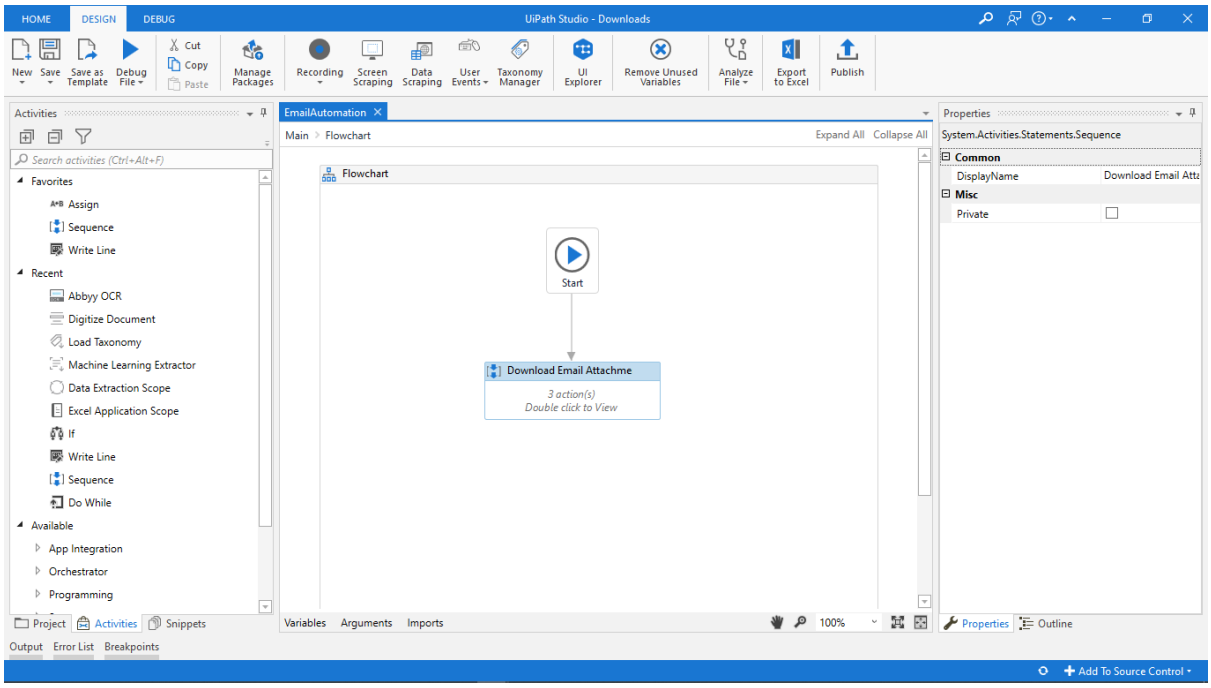
This project is focused on bot that performs the Invoice processing. Here messaging system is no applicable in uipath. The future enhancements can be done by using the Messaging mechanism and the bot can be deployed in various invoices and it can be made to implement in different languages. Also this bot can be hosted in website for each public to get the details. And can scrap the image invoice also.

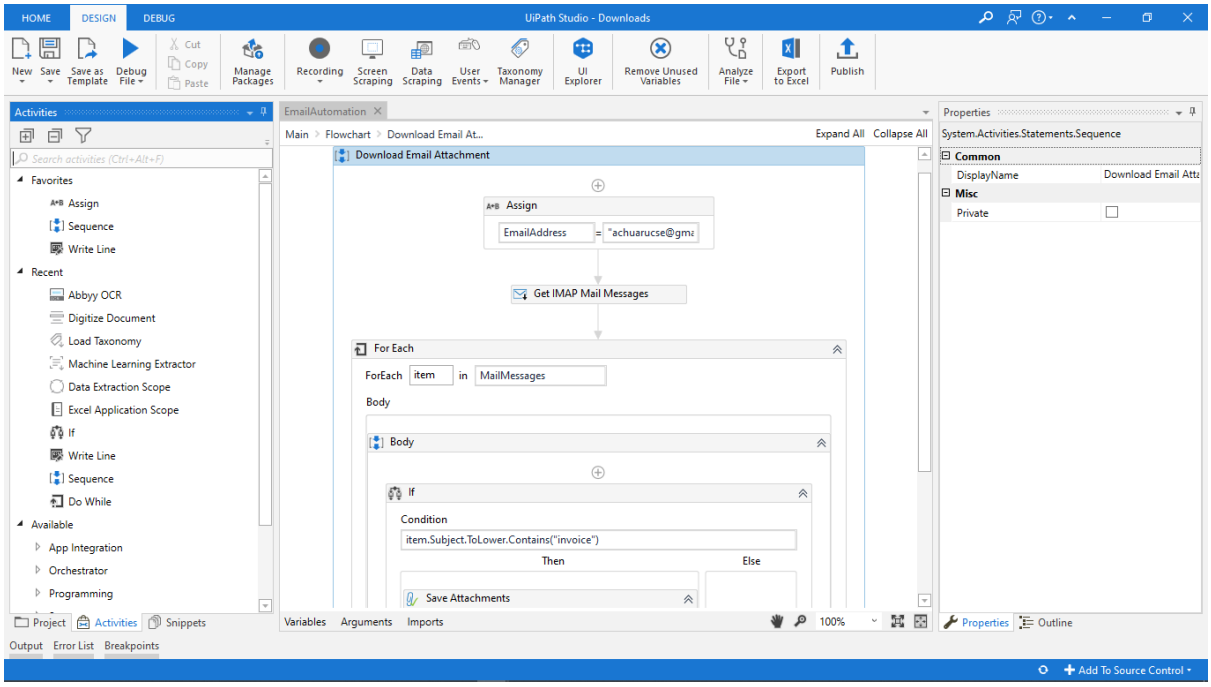
# CHAPTER 10

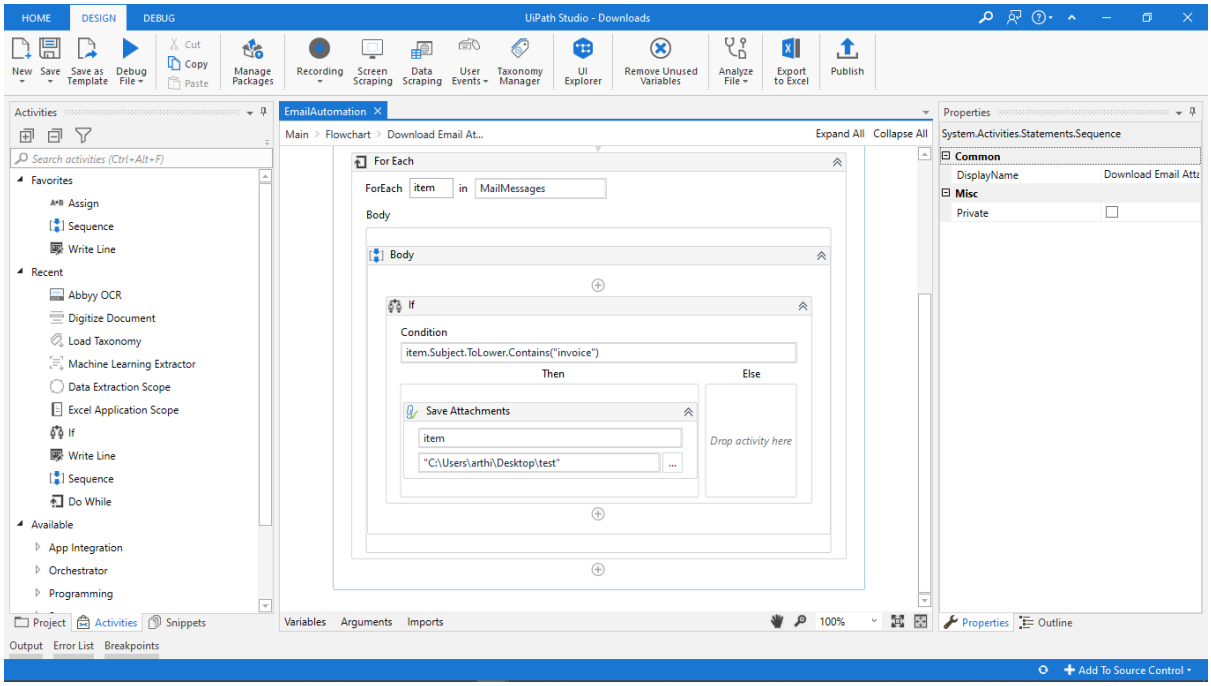
# APPENDIX

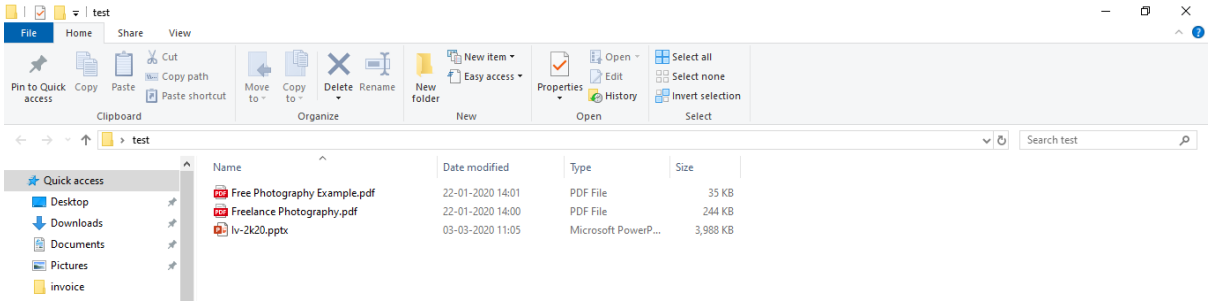
## 10.1 SCREEN SHOTS:

## E-mail automation:

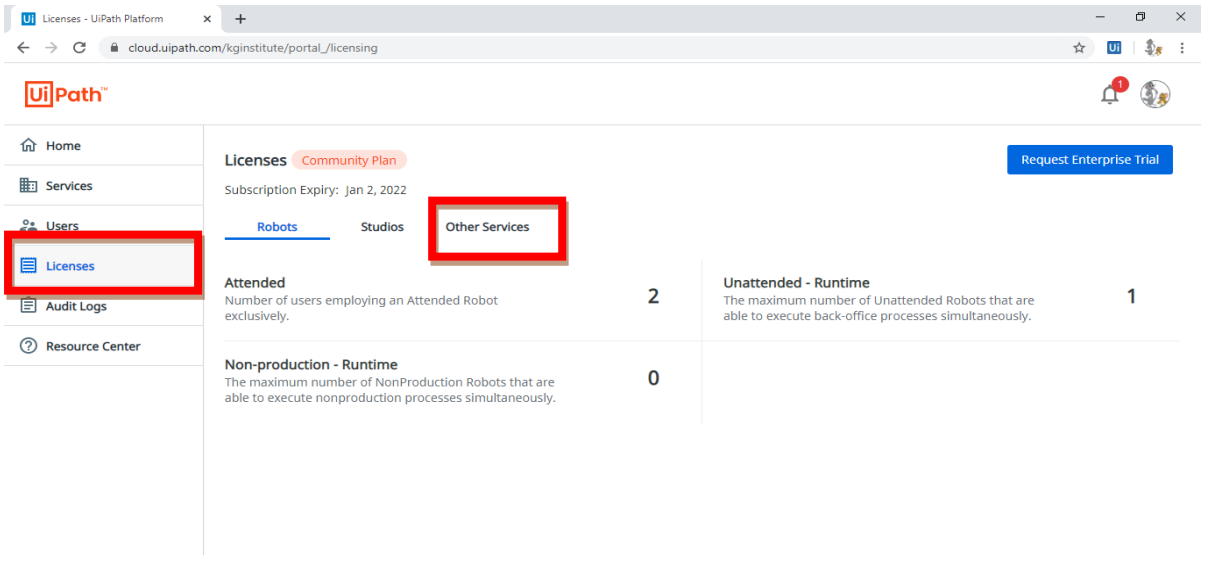


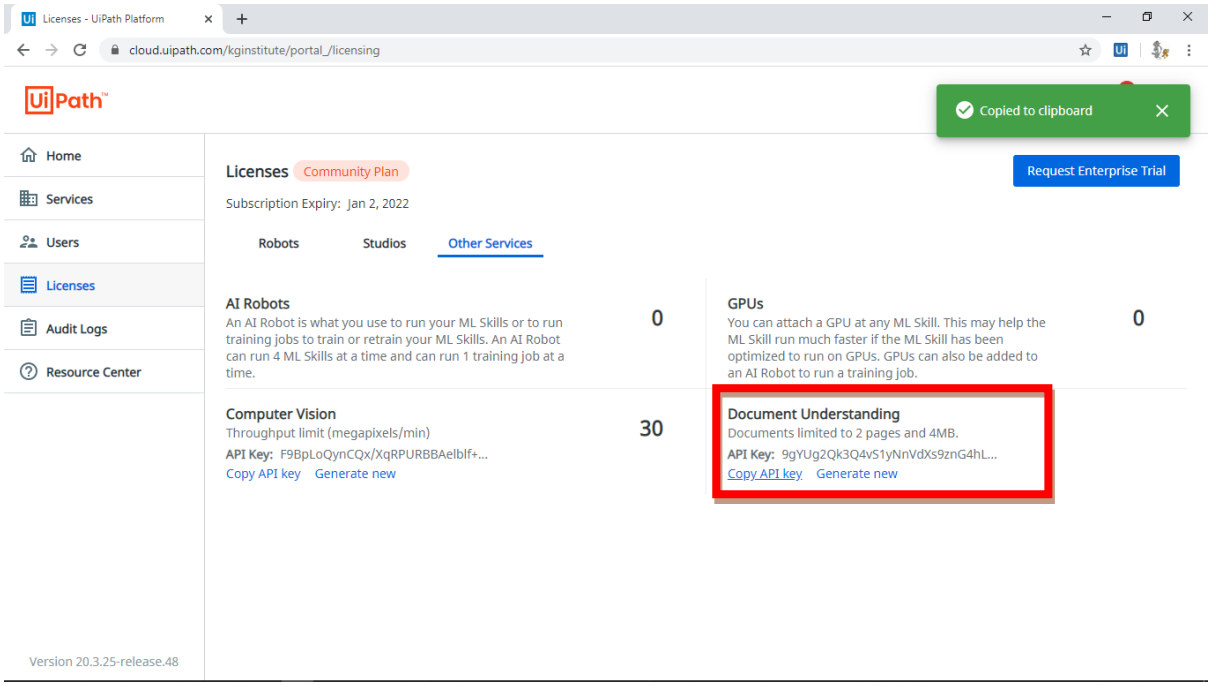





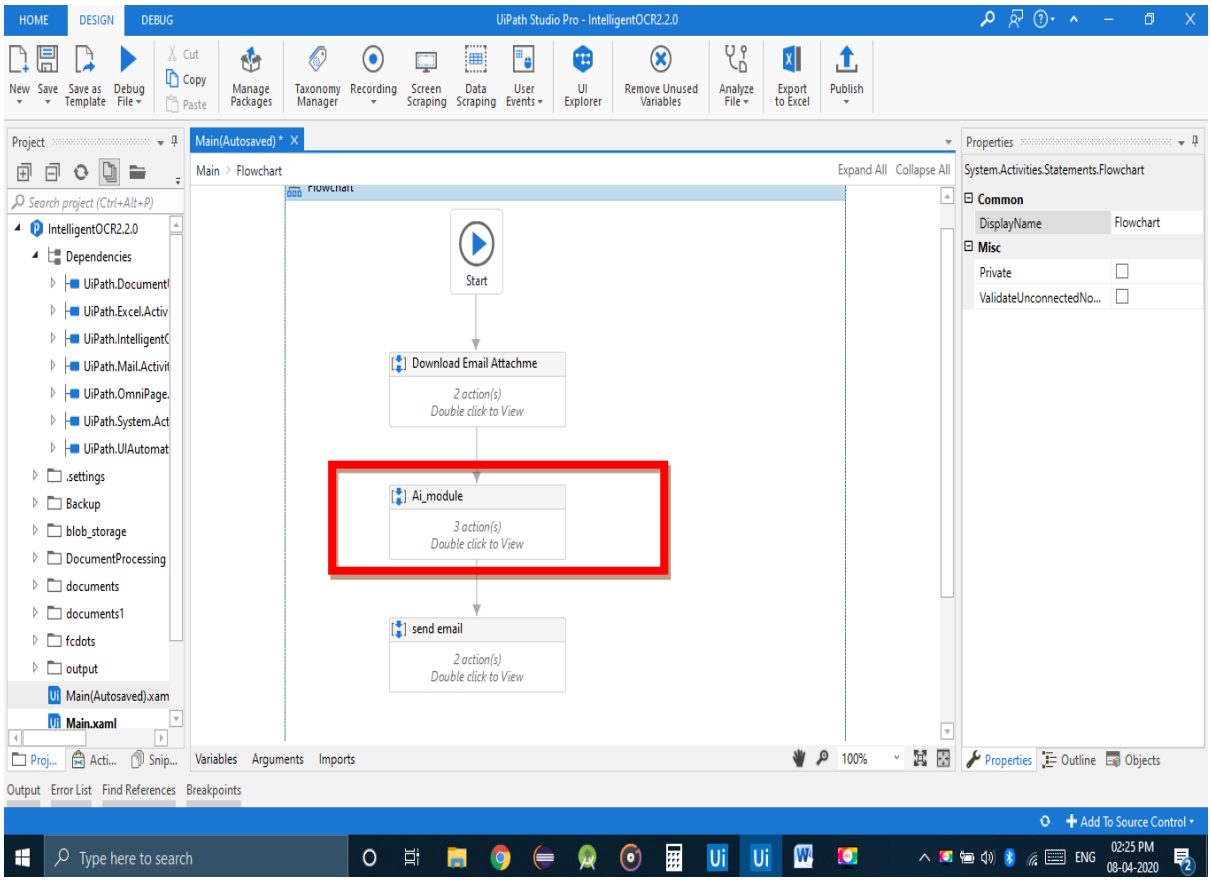



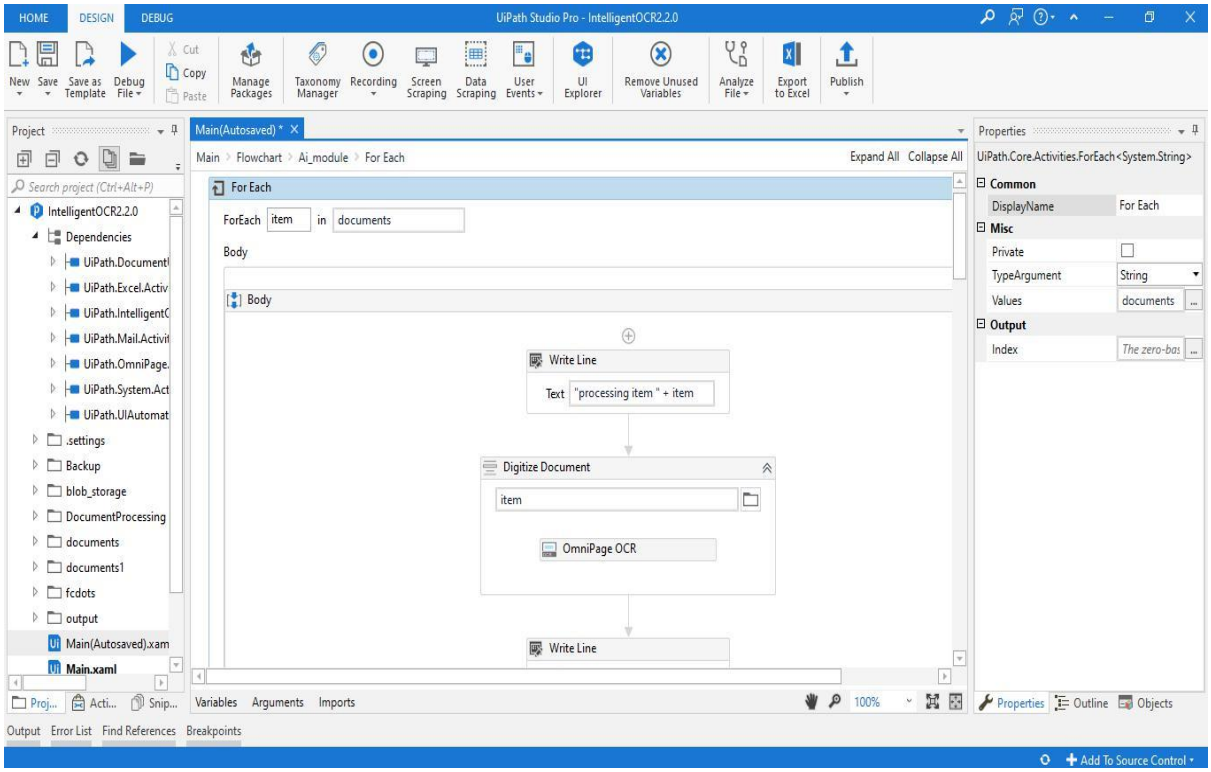

# API key Generation:

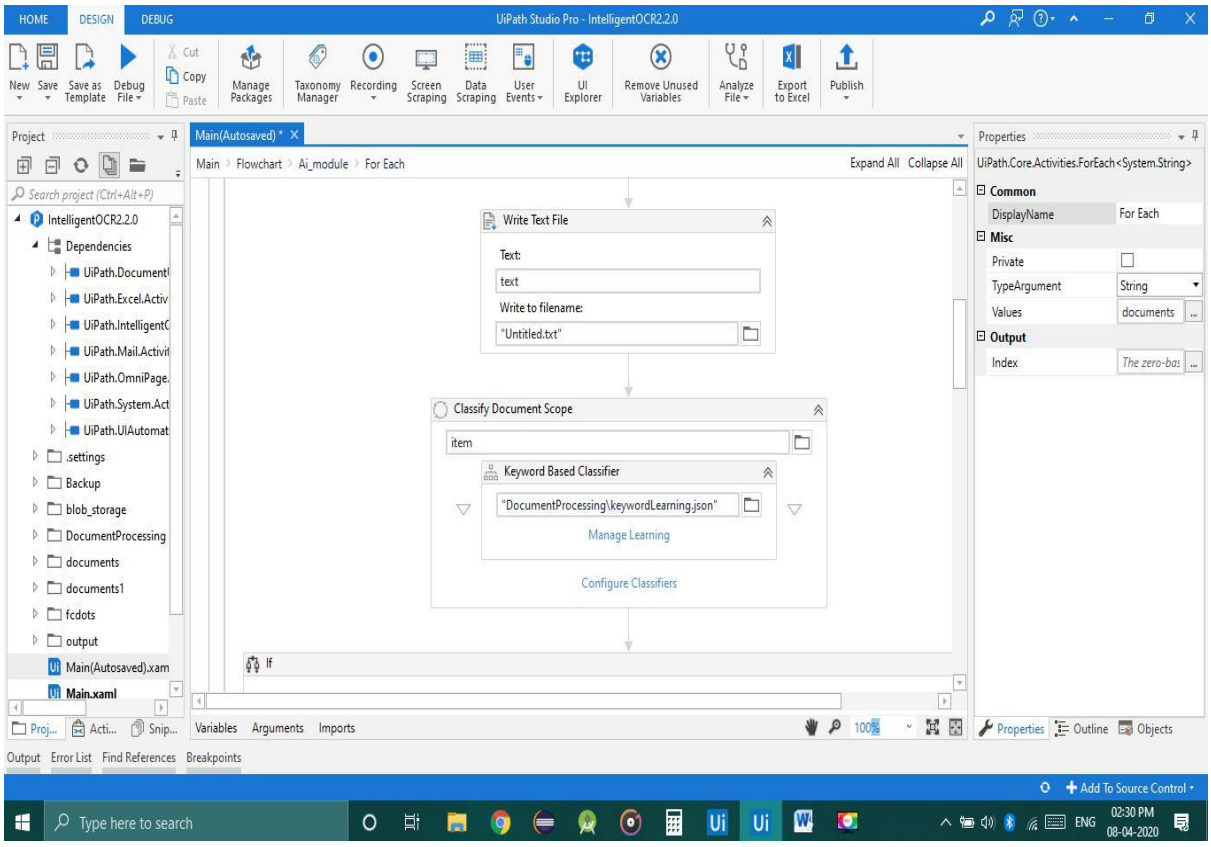


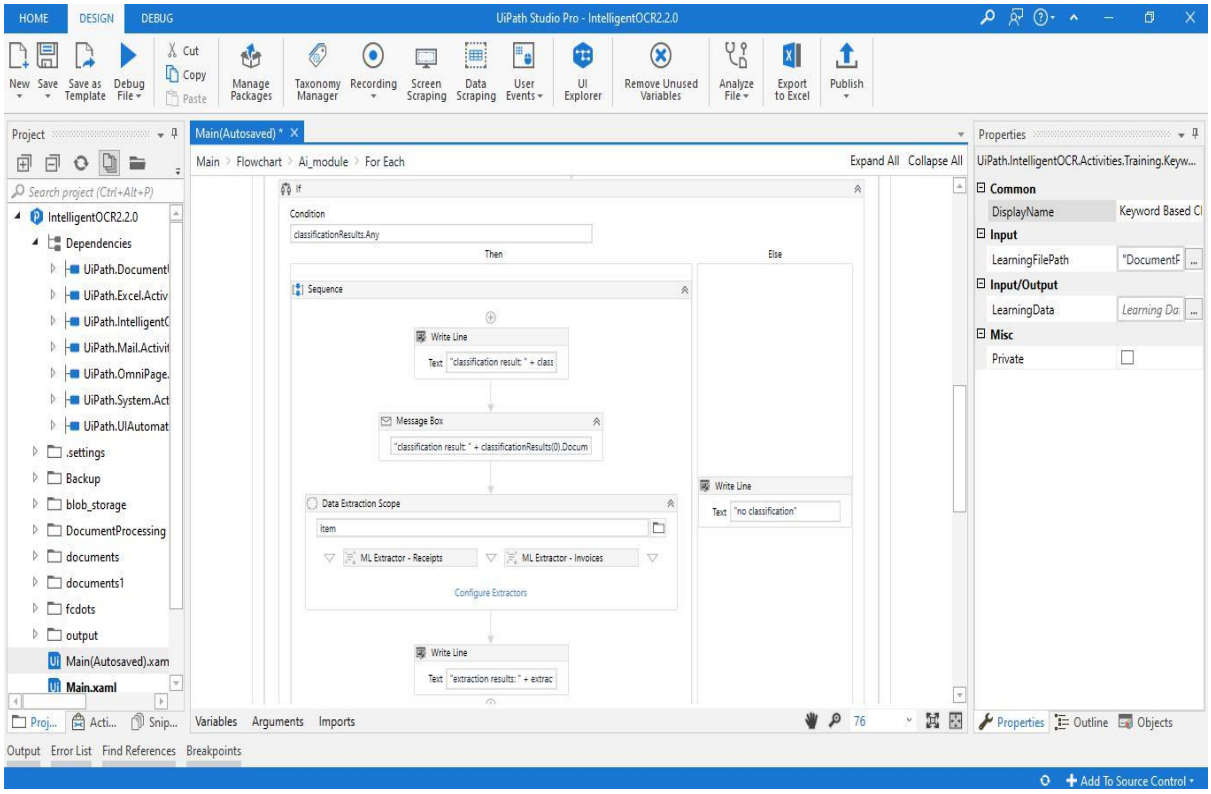


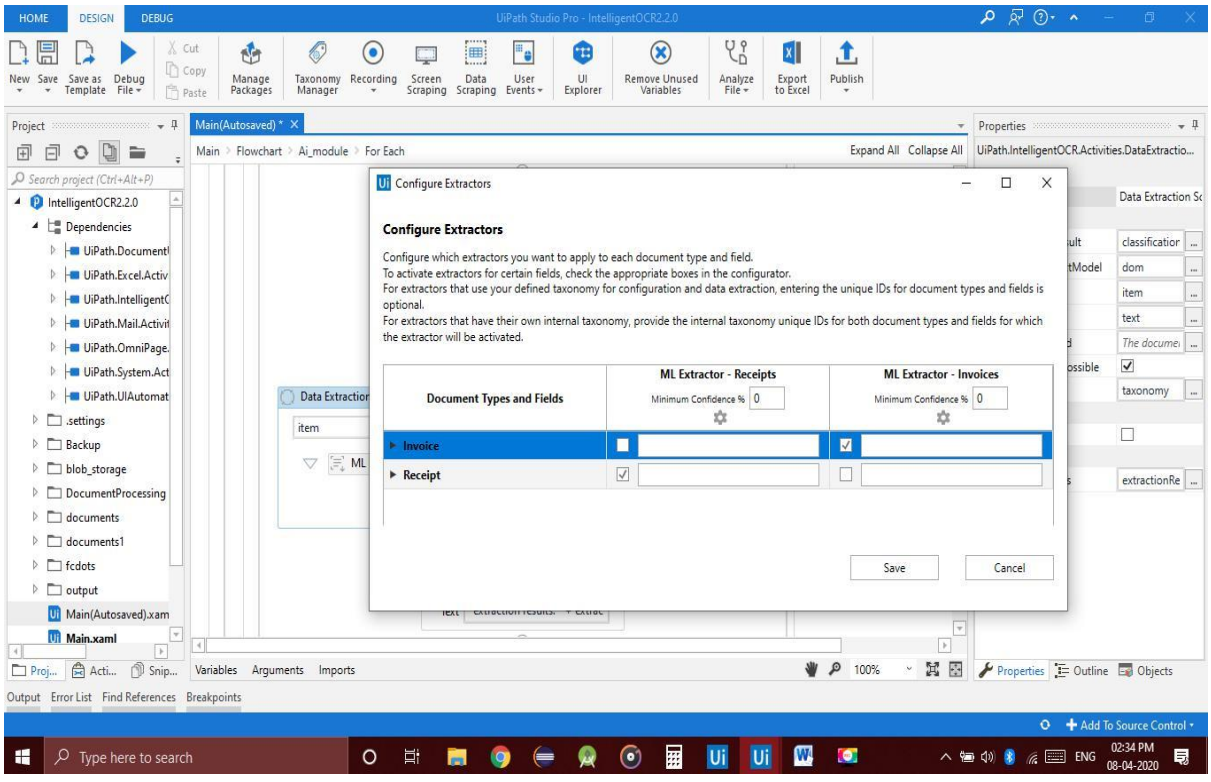

# Data Extraction using Machine Learning:

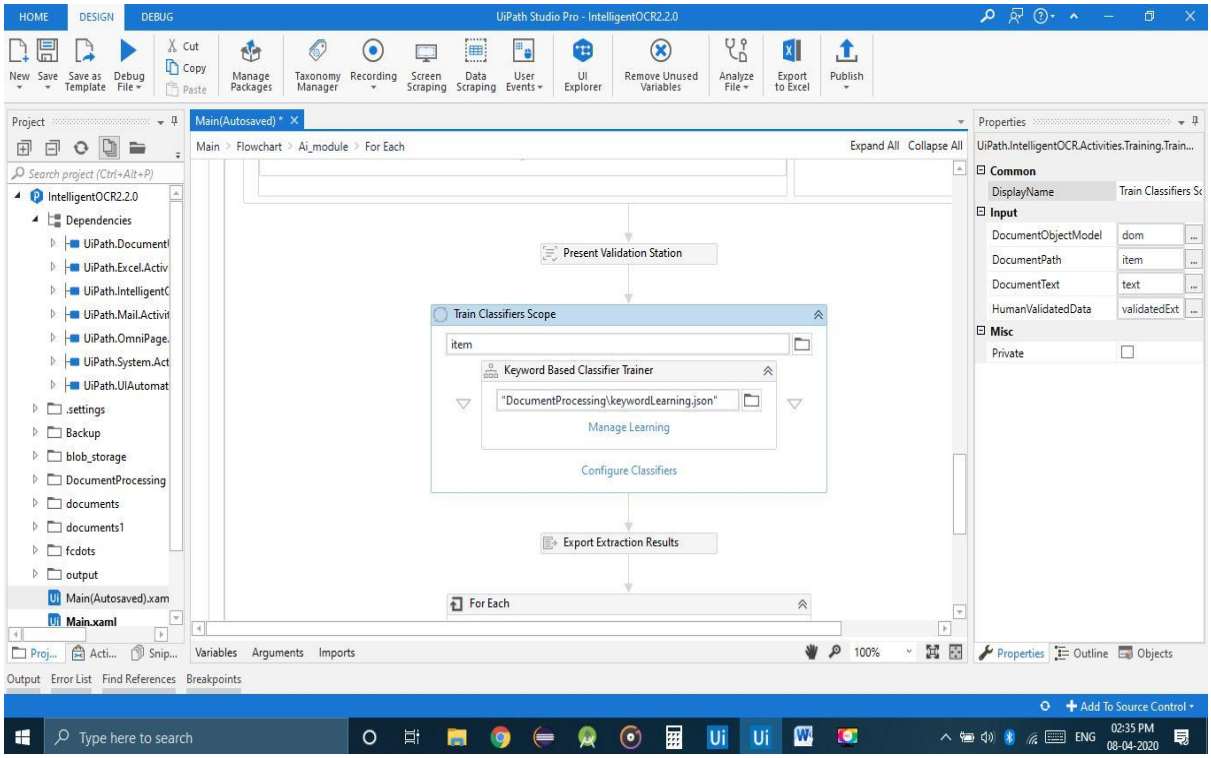


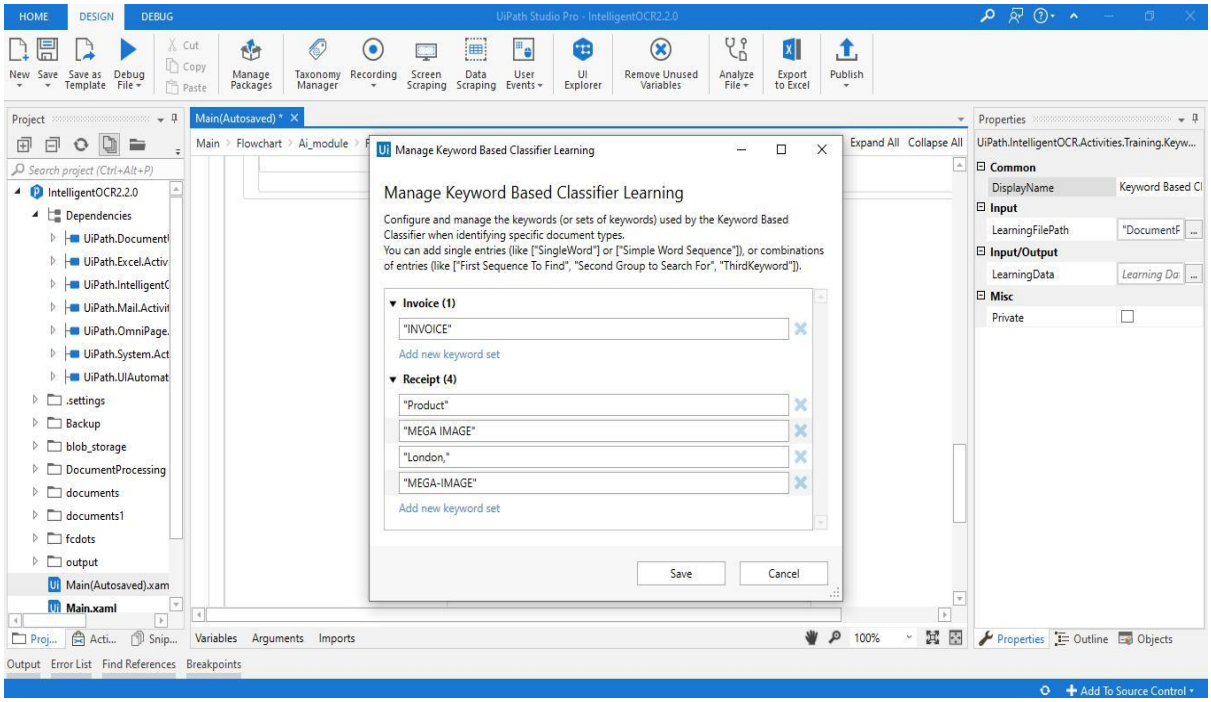

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Main(Autosaved) *
Main › Flowchart › Ai_module › For Each
Expand All
Collapse All
For Each
ForEach item in documents
Body
Write Line
Text "processing item " + item
Digitize Document
item
OmniPage OCR
Write Line
Properties
UiPath.Core.Activities.ForEach<System.String>
Common
DisplayName For Each
Misc
Private
TypeArgument String
Values documents
Output
Index The zero-bas
IntelligentOCR2.2.0
Dependencies
.settings
Backup
blob_storage
DocumentProcessing
documents
documents1
fcdots
output
Main(Autosaved).xam
Main.xaml
Variables
Arguments
Imports
100%
Properties
Outline
Objects
Output
Error List
Find References
Breakpoints
Add To Source Control

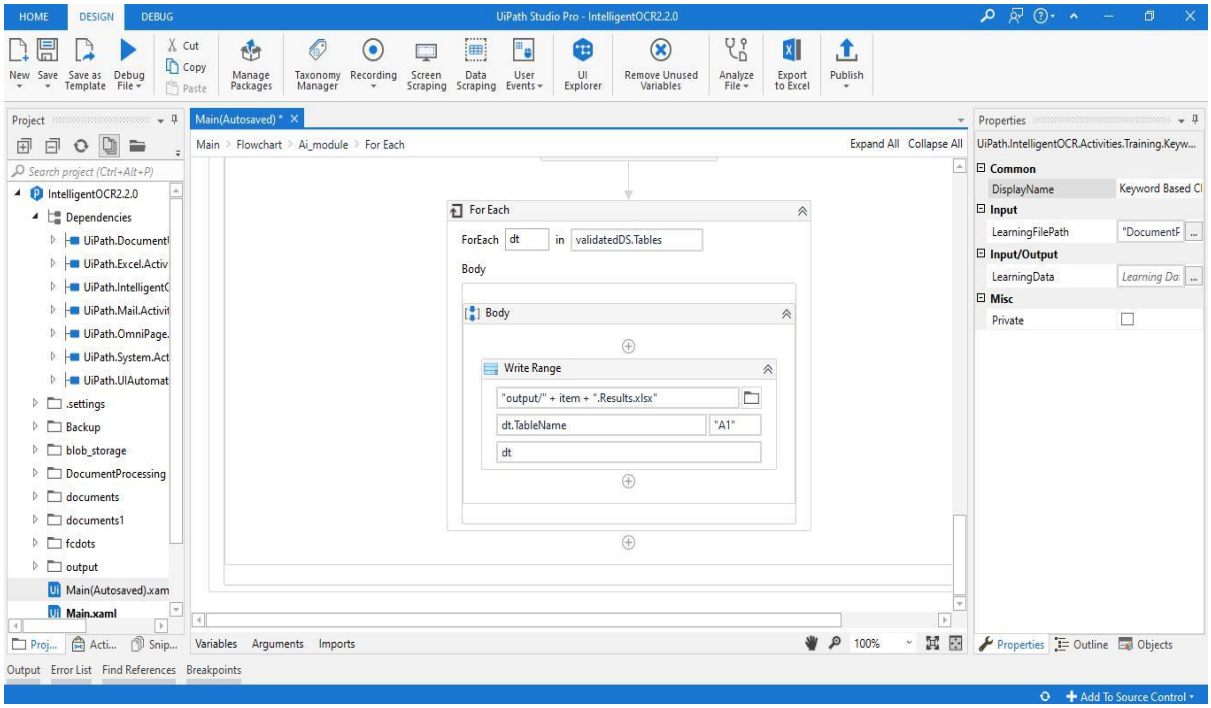

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Main(Autosaved) *
Main › Flowchart › Ai_module › For Each
Expand All
Collapse All
Write Text File
Text:
text
Write to filename:
"Untitled.txt"
Classify Document Scope
item
Keyword Based Classifier
"DocumentProcessing\keywordLearning.json"
Manage Learning
Configure Classifiers
If
Properties
UiPath.Core.Activities.ForEach<System.String>
Common
DisplayName For Each
Misc
Private
TypeArgument String
Values documents
Output
Index The zero-bas
IntelligentOCR2.2.0
Dependencies
.settings
Backup
blob_storage
DocumentProcessing
documents
documents1
fcdots
output
Main(Autosaved).xam
Main.xaml
Variables
Arguments
Imports
100%
Properties
Outline
Objects
Output
Error List
Find References
Breakpoints
Add To Source Control
Type here to search
ENG
02:30 PM
08-04-2020

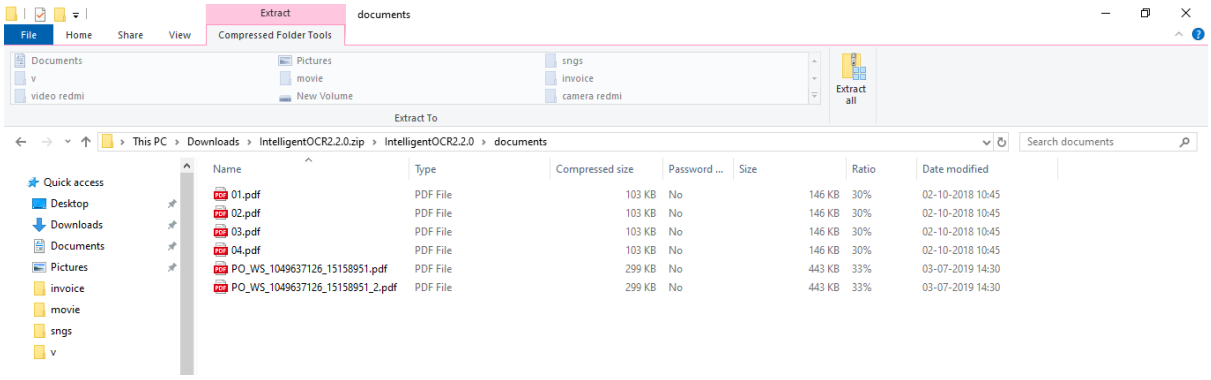

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Main(Autosaved) *
Main › Flowchart › Ai_module › For Each
Expand All
Collapse All
Condition
classificationResults.Any
Then
Else
Sequence
Write Line
Text "classification result: " + class
Message Box
"classification result: " + classificationResults(0).Docum
Data Extraction Scope
Item
ML Extractor - Receipts
ML Extractor - Invoices
Configure Extractors
Write Line
Text "extraction results: " + extrac
Write Line
Text "no classification"
Project
IntelligentOCR2.2.0
Dependencies
.settings
Backup
blob_storage
DocumentProcessing
documents
documents1
fcdots
output
Main(Autosaved).xam
Main.xaml
Properties
UiPath.IntelligentOCR.Activities.Training.Keyw...
Common
DisplayName
Keyword Based Cl
Input
LearningFilePath
"DocumentF
Input/Output
LearningData
Learning Da
Misc
Private
Variables
Arguments
Imports
Output
Error List
Find References
Breakpoints
Add To Source Control

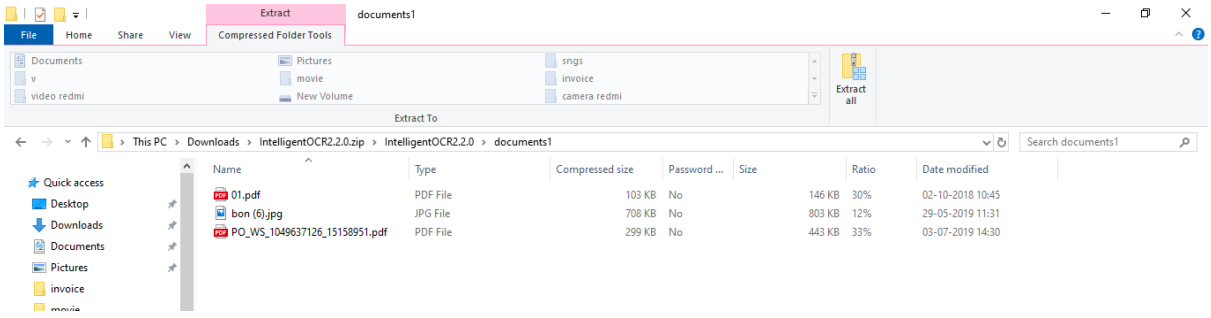

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Configure Extractors
Configure Extractors
Configure which extractors you want to apply to each document type and field.
To activate extractors for certain fields, check the appropriate boxes in the configurator.
For extractors that use your defined taxonomy for configuration and data extraction, entering the unique IDs for document types and fields is optional.
For extractors that have their own internal taxonomy, provide the internal taxonomy unique IDs for both document types and fields for which the extractor will be activated.
Document Types and Fields
ML Extractor - Receipts
Minimum Confidence % 0
ML Extractor - Invoices
Minimum Confidence % 0
Invoice
Receipt
Save
Cancel
Main(Autosaved) *
Main › Flowchart › Ai_module › For Each
Expand All
Collapse All
Project
IntelligentOCR2.2.0
Dependencies
Properties
UiPath.IntelligentOCR.Activities.DataExtractio...
Variables
Arguments
Imports
Output
Error List
Find References
Breakpoints
Add To Source Control
Type here to search
ENG
02:34 PM
08-04-2020

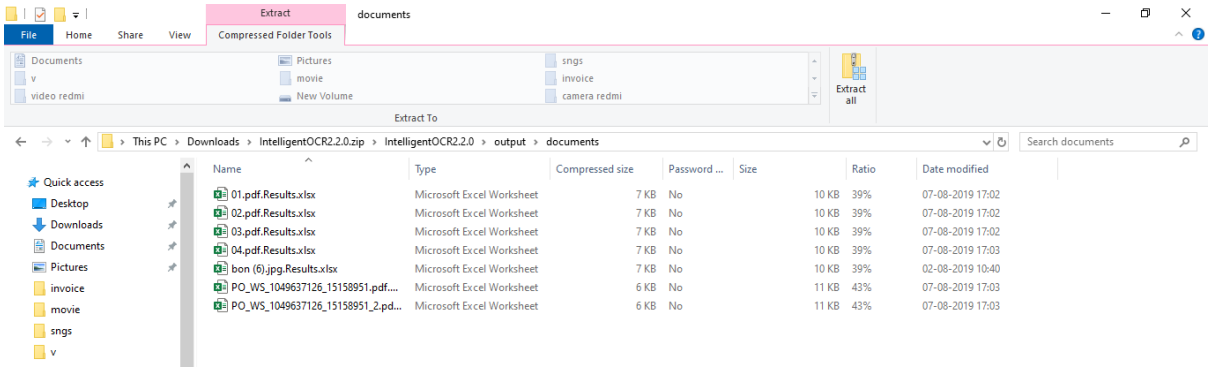

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Present Validation Station
Train Classifiers Scope
item
Keyword Based Classifier Trainer
"DocumentProcessing\keywordLearning.json"
Manage Learning
Configure Classifiers
Export Extraction Results
For Each

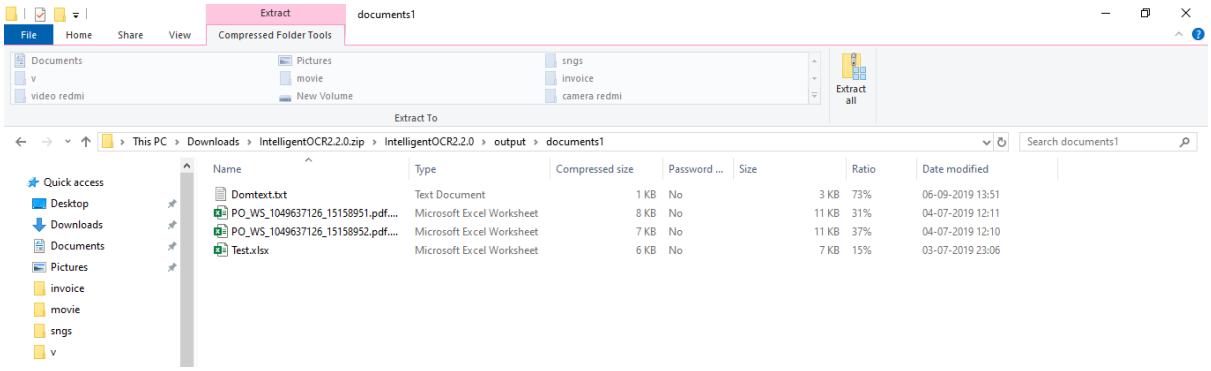

HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Manage Keyword Based Classifier Learning
Manage Keyword Based Classifier Learning
Configure and manage the keywords (or sets of keywords) used by the Keyword Based Classifier when identifying specific document types.
You can add single entries (like ["SingleWord"] or ["Simple Word Sequence"]), or combinations of entries (like ["First Sequence To Find", "Second Group to Search For", "ThirdKeyword"]).
Invoice (1)
"INVOICE"
Add new keyword set
Receipt (4)
"Product"
"MEGA IMAGE"
"London,"
"MEGA-IMAGE"
Add new keyword set
Save
Cancel


HOME
DESIGN
DEBUG
UiPath Studio Pro - IntelligentOCR2.2.0
Main(Autosaved) *
Main › Flowchart › Ai_module › For Each
Expand All
Collapse All
For Each
ForEach dt in validatedDS.Tables
Body
Write Range
"output/" + item + ".Results.xlsx"
dt.TableName
"A1"
dt
Properties
Common
DisplayName
Input
LearningFilePath
Input/Output
LearningData
Misc
Private
Variables
Arguments
Imports
Output
Error List
Find References
Breakpoints
Add To Source Control

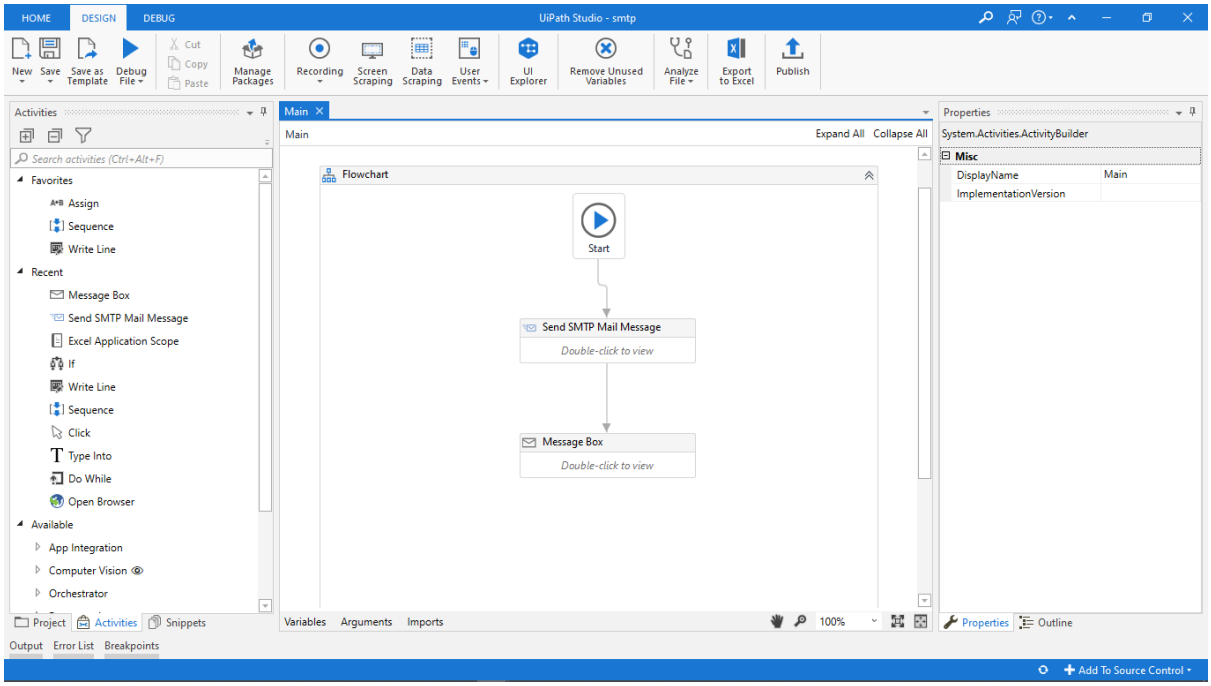

Extract
documents
File
Home
Share
View
Compressed Folder Tools
Extract all
Extract To
This PC › Downloads › IntelligentOCR2.2.0.zip › IntelligentOCR2.2.0 › documents
Search documents
Quick access
Desktop
Downloads
Documents
Pictures
invoice
movie
sngs
v
Name
Type
Compressed size
Password ...
Size
Ratio
Date modified
01.pdf
02.pdf
03.pdf
04.pdf
PO_WS_1049637126_15158951.pdf
PO_WS_1049637126_15158951_2.pdf

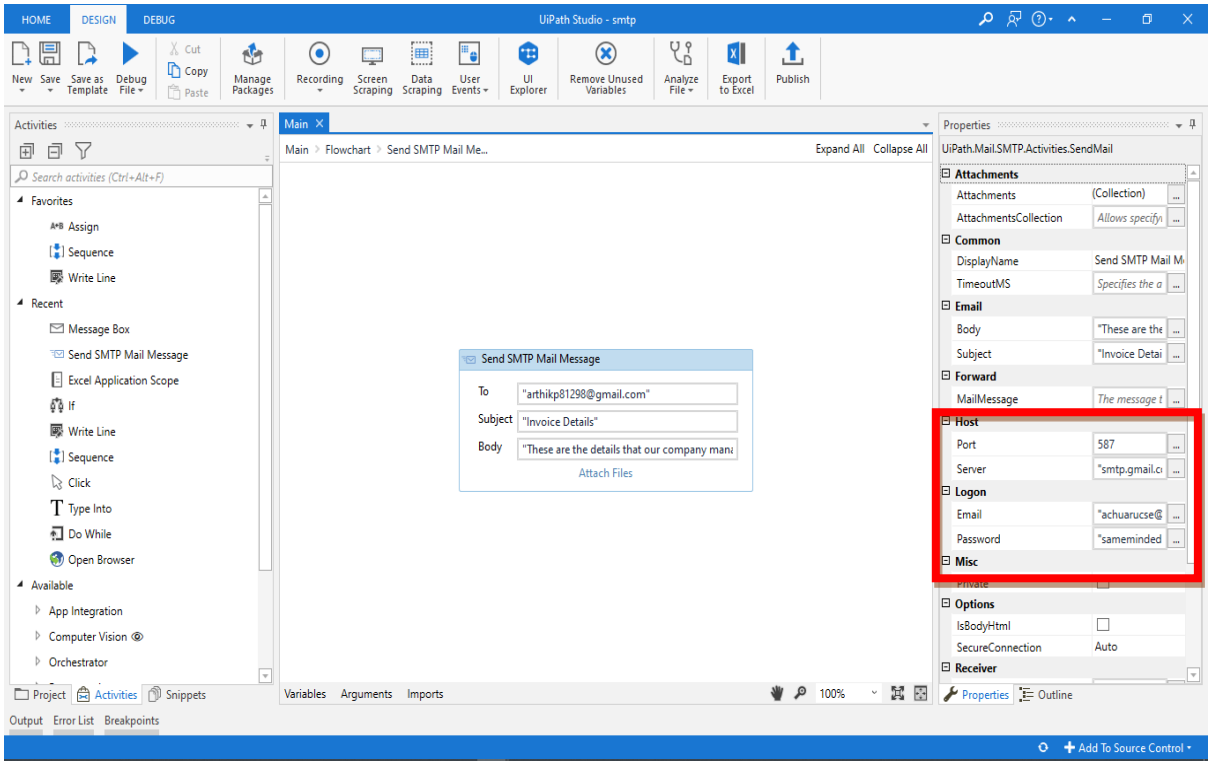

Extract
documents1
File
Home
Share
View
Compressed Folder Tools
Extract all
Extract To
This PC › Downloads › IntelligentOCR2.2.0.zip › IntelligentOCR2.2.0 › documents1
Search documents1
Quick access
Desktop
Downloads
Documents
Pictures
invoice
Name
Type
Compressed size
Password ...
Size
Ratio
Date modified
01.pdf
bon (6).jpg
PO_WS_1049637126_15158951.pdf

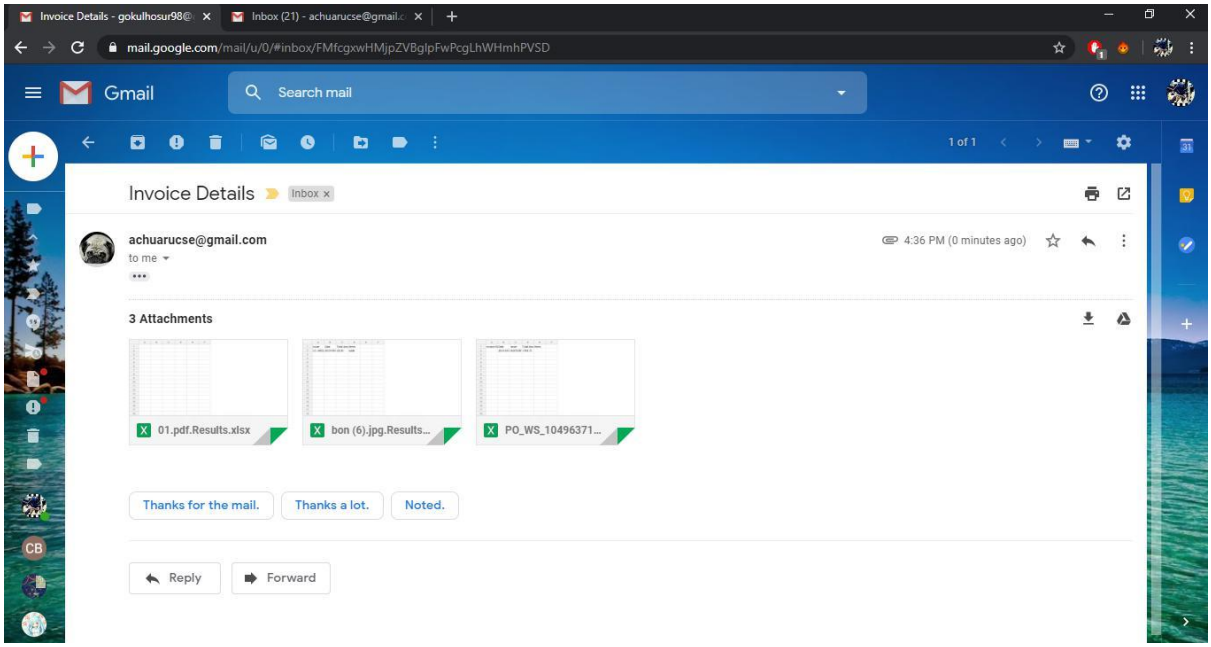



# Notification:




HOME
DESIGN
DEBUG
UiPath Studio - smtp
Activities
Main
Flowchart
Send SMTP Mail Me...
Expand All
Collapse All
Properties
UiPath.Mail.SMTP.Activities.SendMail
Send SMTP Mail Message
To
"arthikp81298@gmail.com"
Subject
"Invoice Details"
Body
"These are the details that our company mana
Attach Files
Host
Port
587
Server
"smtp.gmail.c
Logon
Email
Password
Misc
Options
IsBodyHtml
SecureConnection
Auto
Receiver
Variables
Arguments
Imports
Output
Error List
Breakpoints
Add To Source Control


Invoice Details - gokulhosur98@
Inbox (21) - achuarucse@gmail.c
mail.google.com/mail/u/0/#inbox/FMfcgxwHMjpZVBglpFwPcgLhWHmhPVSD
Gmail
Search mail
1 of 1
Invoice Details
Inbox
achuarucse@gmail.com
to me
4:36 PM (0 minutes ago)
3 Attachments
01.pdf.Results.xlsx
bon (6).jpg.Results...
PO_WS_10496371...
Thanks for the mail.
Thanks a lot.
Noted.
Reply
Forward

# CHAPTER 11

# REFERENCES


1. Google
2. Wikipedia- https://en.wikipedia.org/wiki/Invoice_processing
3. E-mail Creation- https://www.youtube.com/watch?v=VFGWBsw6 VlQ
4. Video tutorial https://www.youtube.com/watch?v=mvSGW4tTijU
5. Machine Learning Extractor - https://forum.uipath.com/t/receipt-and-invoice-ai-now-available-in-public-preview/125056
6. Uploading an invoice- https://web.stanford.edu/group/fms/fingate/finsystem/iprocure/howto/upload_invoice.html
7. Api key Generation - https://platform.uipath.com/kginstitute/por tal_/licensing
8. Tutorial - https://www.youtube.com/watch?time_continue=4&v=121WwvFGco4&feature=emb_logo
9. Vendor Payment Process - http://web.mit.edu/sapr3/windocs / bppcb01a.htm
10. Form - https://forum.uipath.com/