DRUG SEEKER BOT

A PROJECT REPORT

Submitted by

VELLANKI CHANDRA HARSHA **220701314**

in partial fulfillment of the course

OAI1903 - INTRODUCTION TO ROBOTIC PROCESS AUTOMATION

for the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

RAJALAKSHMI NAGAR

THANDALAM

CHENNAI – 602 105

NOVEMBER 2024

RAJALAKSHMI ENGINEERING COLLEGE

CHENNAI - 602105

BONAFIDE CERTIFICATE

Certified that this project report **" DRUG SEEKER BOT"** is the bonafide work of **"V.Chandra Harsha (220701314)"** who carried out the project work for the subject OAI1903-Introduction to Robotic Process Automation under my supervision.

Ms. U.Farjana., M.Tech.,

SUPERVISOR

Assistant Professor

Department of Computer Science and Engineering

Rajalakshmi Engineering College

Rajalakshmi Nagar

Thandalam

Chennai - 602105

Submitted to Project and Viva Voce Examination for the subject OAI1903-Introduction to Robotic Process Automation held on _____.

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

Initially, we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman **Thiru. S.Meganathan, B.E., F.I.E.,** our Vice Chairman **Mr. M.Abhay Shankar, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) Thangam Meganathan, M.A., M.Phil., Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring to educate 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 Engineering for his guidance and encouragement throughout the project work. **Ms. U.Farjana, M.Tech.,** Assistant Professor**,** 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, Professor, **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.

V.Chandra Harsha(220701314)

ABSTRACT

The 'Drug Seeker Bot' is an advanced RPA application designed using UiPath to automate the retrieval of pharmaceutical data efficiently. This bot serves as a user-friendly tool that accepts drug names as input and leverages web browser activities to extract critical details such as brand names, prices, and other related information from reliable online pharmaceutical platforms. The bot is designed to simplify and speed up the process of obtaining drug information, eliminating the time-consuming and error-prone task of manual searches.

This project demonstrates the effective use of robotic process automation in the healthcare sector, offering a scalable solution for pharmacists, healthcare providers, and consumers to access accurate drug details instantly. It ensures enhanced accuracy, operational efficiency, and convenience while highlighting the practical applications of automation in addressing real-world challenges. By providing timely and precise information, the 'Drug Seeker Bot' promotes informed decision-making and supports improved healthcare outcomes.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	**iv**
	LIST OF FIGURES	**v**
	LIST OF ABBREVIATIONS	**vi**
1.	**INTRODUCTION**	**1**
	1.1 BACKGROUND	1
	1.2 PROBLEM STATEMENT	1
	1.3 PROJECT OBJECTIVES	2
	1.4 SCOPE OF THE PROJECT	2
	1.5 LIMITATIONS	3
2.	**LITERATURE REVIEW**	**4**
	2.1 GENERAL	4
	2.2 STATE OF THE ART TECHNIQUES	4
3.	**SYSTEM DESIGN**	**6**
	3.1 SYSTEM FLOW DIAGRAM	6
	3.2 ARCHITECTURE DIAGRAM	7
	3.3 SOFTWARE AND HARDWARE REQUIREMENTS	8
4.	**PROJECT DESCRIPTION**	**9**
	4.1 METHODOLOGIES	9
5.	**IMPLEMENTATION AND RESULT**	**11**
	5.1 IMPLEMENTATION PROCEDURE	11
	5.2 OUTPUT	12
	5.3 RESULT AND ANALYSIS	15
6.	**CONCLUSION AND FUTURE WORK**	**20**
	6.1 CONCLUSION	20
	6.2 FUTURE WORKS	21
	REFERENCES	**22**

LIST OF FIGURES

FIGURE NO.	FIGURE NAME	PAGE NO.
3.1	SYSTEM FLOW DESIGN	4
3.2	ARCHITECTURE DIAGRAM	5
5.2	WORKFLOW	10
5.2	SAMPLE INPUT 1	10
5.2	SAMPLE INPUT 2	10
5.2	SAMPLE INPUT 3	10
5.2	SAMPLE OUTPUT 1	11
5.2	SAMPLE OUTPUT 2	11

LIST OF ABBREVIATIONS

ABBREVIATION	ACRONYM
B.E.	Bachelor of Engineering
M.Tech.	Master of Technology
Ph.D.	Doctor of Philosophy
RPA	Robotic Process Automation
AI	Artificial Intelligence

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Access to accurate and timely drug information is crucial in the healthcare and pharmaceutical industries. Patients, pharmacists, and healthcare providers often rely on online platforms to find details about medications, including their brand names, prices, and availability. However, manually searching for this information across multiple websites can be a time-consuming and error-prone process, particularly when dealing with large volumes of queries or urgent situations.

1.2 PROBLEM STATEMENT

In the healthcare and pharmaceutical sectors, timely access to accurate drug information, including brand names, prices, and availability, is essential for decision-making. However, manually searching for this information on multiple platforms is a time-intensive process that can lead to errors, inefficiencies, and delays. This challenge is further exacerbated when dealing with high volumes of queries, making it difficult for pharmacists, healthcare providers, and consumers to retrieve reliable data quickly.

1.3 PROJECT OBJECTIVES

1.Automate Drug Information Retrieval:
Develop an RPA-based solution using UiPath to automate the process of fetching details like brand names, prices, and other relevant information for user-provided drug names.
2.Enhance Efficiency and Accuracy:
Eliminate the need for manual searches, reducing the time required to retrieve drug-related information while minimizing errors.

3.Improve User Accessibility:
Provide a user-friendly interface for healthcare professionals, pharmacists, and consumers to access essential drug data conveniently.

4.Leverage Web Browser Activities:
Utilize UiPath's web browser activities to navigate and scrape information from reliable online pharmaceutical platforms.

1.4 SCOPE OF THE PROJECT

This project focuses on the development of an automated signature verification system for offline signatures. The system will be trained and evaluated on a dataset of handwritten signatures. While online signatures and biometric techniques could be explored in future work, the current scope is limited to offline signature verification.

1.5 LIMITATIONS

While this project aims to develop a convenient bot realted to medicines,limitations are as follows:

- Dependency on Online Platforms:
  The bot relies on the availability and accuracy of information from online pharmaceutical websites. If a website is down or its data is outdated, the bot cannot provide accurate results.
- Web Structure Changes:
  Any change in the structure or layout of the targeted websites can disrupt the bot's functionality, requiring updates to the automation workflow.
- Limited to Specific Data Sources:
  The bot fetches information only from predefined or accessible online platforms, which may not cover all possible drug-related data sources.

By addressing these limitations and continuously improving the system, we aim to enhance its accuracy and robustness.

CHAPTER 2

LITERATURE REVIEW

2.1 GENERAL

Automation technologies, particularly Robotic Process Automation (RPA), are increasingly being adopted in healthcare to streamline repetitive tasks such as data retrieval, patient management, and inventory tracking. Studies emphasize the efficiency and accuracy that RPA brings to processes by minimizing manual errors and saving time. The integration of web-scraping techniques within RPA workflows has shown significant potential in extracting and processing large volumes of data from online sources, enabling faster access to critical information for healthcare providers and consumers. This project builds upon these advancements to address the specific need for automated drug information retrieval.

2.2 STATE OF THE ART TECHNIQUES

1.Robotic Process Automation (RPA):
Utilizes UiPath for automating repetitive tasks such as web navigation, data extraction, and processing. RPA mimics human interactions with websites, ensuring high accuracy and speed in retrieving drug-related information.

2.Web Scraping Technologies:
Employs advanced web browser activities in UiPath to scrape structured and unstructured data from online pharmaceutical platforms. This allows the bot to extract details like brand names, prices, and availability directly from websites.

3.Natural Language Processing (NLP):
Can be integrated in future enhancements to process and interpret user inputs more intelligently, handling variations in drug name spellings or synonyms.

4.Real-Time Data Integration:
Incorporates APIs or dynamic web scraping techniques to fetch real-time information from pharmaceutical databases, ensuring up-to-date results.

5.Error Handling and Resilience:
Implements robust error handling and retry mechanisms within the automation workflow to address common challenges such as network issues, CAPTCHA, or website downtime.

6.Data Structuring and Output:
Formats the extracted data into user-friendly reports or dashboards, making the information easily interpretable and actionable.

7.Scalable Design:
Designed to handle high query volumes, ensuring scalability to meet the demands of multiple users simultaneously.

These techniques collectively ensure that the 'Drug Seeker Bot' is efficient, accurate, and adaptable to the evolving needs of its users.

CHAPTER 3

SYSTEM DESIGN

3.1 SYSTEM FLOW DIAGRAM

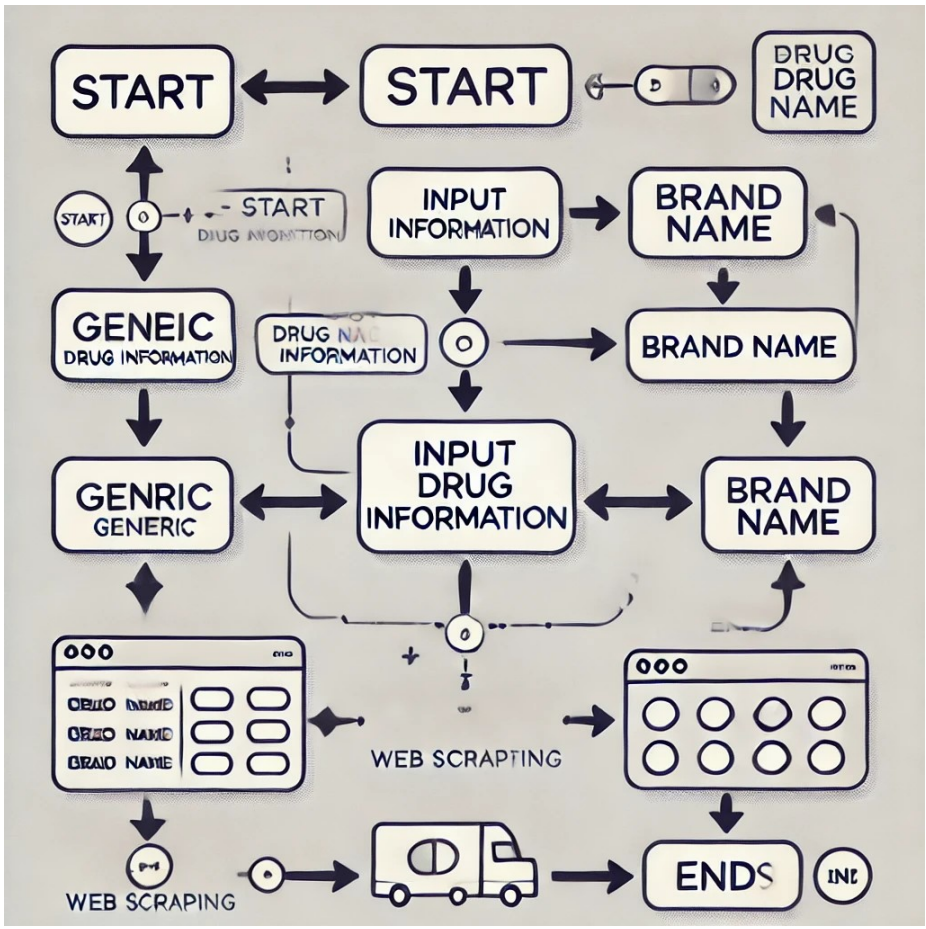

FIG 3.1 SYSTEM FLOW DIAGRAM

3.2 ARCHITECTURE DIAGRAM

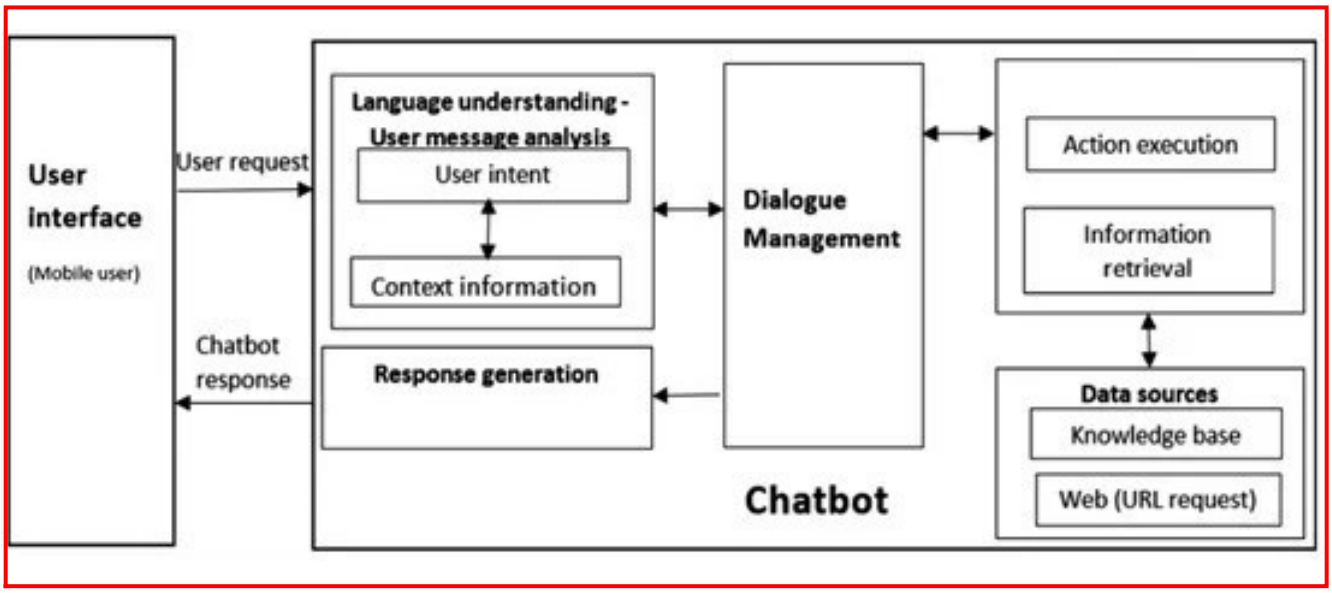

FIG 3.2 ARCHITECTURAL DIAGRAM

Overview

The 'Drug Seeker Bot' is an RPA-based solution that automates the retrieval of drug-related information such as brand names and prices from online pharmaceutical platforms. It simplifies the process by allowing users to input drug names and receive detailed results via email.

1. Automates drug information retrieval using UiPath RPA tools.
2. Fetches details such as brand name, price, and availability from online pharmaceutical platforms.
3. User-friendly interface allows users to input drug names (generic or brand).
4. Emails results to the user with the fetched drug information.
5. Streamlines the process of finding accurate drug data, saving time and effort.
6. Aims to improve efficiency for pharmacists, healthcare professionals, and consumers.
7. Real-time data retrieval ensures up-to-date information.

3.3 SOFTWARE AND HARDWARE REQUIREMENTS

Software

- **UiPath Studio:** To design and automate the workflow.
- **WEBSITE:** Googlechrome,Microsoftedge,firefox

Hardware

- **Standard Computer:** A standard computer with sufficient processing power and memory is sufficient.

CHAPTER 4

PROJECT DESCRIPTION

4.1 METHODOLOGIES

Dataset Preparation

Data Collection

Drug Name List (Generic or Brand):

A database or list of drug names (generic and brand) that the user might search for:This can be a static list or dynamically updated through integration with pharmaceutical databases or APIs.

Email Addresses:A record of users' email addresses to which the bot will send the retrieved information.

This data ensures the bot can deliver the results to the correct recipient.

Data Preprocessing

Cleaning and Normalizing Drug Names:

- Remove Duplicates: Eliminate any duplicate entries of drug names in the list. Standardize Drug Names: Standardize the format of drug names (e.g., capitalization, removing extra spaces) to ensure consistency when performing searches.
- Handle Synonyms: Ensure that synonyms or alternative names for drugs (e.g., brand names and generic names) are mapped correctly, so the bot can recognize them as the same drug.
- Spell Check: Use spell-checking algorithms or manual correction to address common typos or misspellings in drug names.

Model Training and Deployment

Model Selection

- **UiPath Activities:** Utilize the pre-trained models and APIs provided by UiPath Gen AI to perform signature verification.
- **Custom Model Training:** If required, train a custom machine learning model using frameworks and deploy it as a custom activity in

Model Training:

1. Drug Name Matching (NLP):Train an NLP model to match user inputs (e.g., brand or generic names) to correct drug names, handling variations and misspellings.

2. Sentiment Analysis:Train a model to analyze user feedback (e.g., reviews) and classify sentiment as positive, negative, or neutral.

3. Price Prediction (Optional):Train a model to predict drug prices based on historical data from multiple sources, providing price estimates.

Workflow Development

1.User Input Handling:
Develop workflows to prompt users for input (e.g., drug name, drug type) and validate the inputs for accuracy and consistency.

2.Web Scraping and Data Extraction:
Create workflows for scraping data from pharmaceutical websites to extract drug information such as brand names, prices, and availability.

3.Data Display and Output:
Design workflows to display the extracted data to the user in a user-friendly format, such as a list or table, within the application interface.

Evaluation

1.Accuracy of Data Retrieval:

Evaluate the bot's ability to correctly retrieve drug information (brand name, price, availability) from pharmaceutical websites. This involves comparing the scraped data against known, reliable sources to measure the bot's accuracy.

2.Performance and Scalability:

Evaluate the bot's performance in terms of speed, reliability, and its ability to handle multiple requests simultaneously. This ensures the bot can scale to meet the needs of users without lag or failure.

CHAPTER 5

IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION PROCEDURE (Using UiPath Studio)

- System Setup:

  Install and configure UiPath Studio for RPA development.
  Set up necessary packages, such as web scraping and email automation (if required).

- User Input and Validation:

  Develop workflows to prompt users for drug names (generic or brand), validate the inputs (checking for typos and correct formatting), and handle errors or invalid entries.

- Web Scraping:

  Implement web scraping workflows to gather data from pharmaceutical websites.
  Extract key information such as drug prices, brand names, and availability from the web pages.
  Apply filters and error handling to ensure accurate extraction.

- Data Formatting and Presentation:

  Format the scraped data (e.g., prices in a consistent currency format, availability status).
  Develop workflows to display the extracted data to the user in a structured and readable format (e.g., a table or list).

- Testing and Debugging:

  Conduct thorough testing of all workflows to ensure proper functioning. Address any bugs or errors encountered during the execution of tasks.
  Perform test runs with various drug names to ensure accuracy in results.

- Deployment:

  Deploy the bot on the required machine/environment.
  Set up monitoring tools to track the bot's performance, data accuracy, and system health.

5.2 OUTPUT

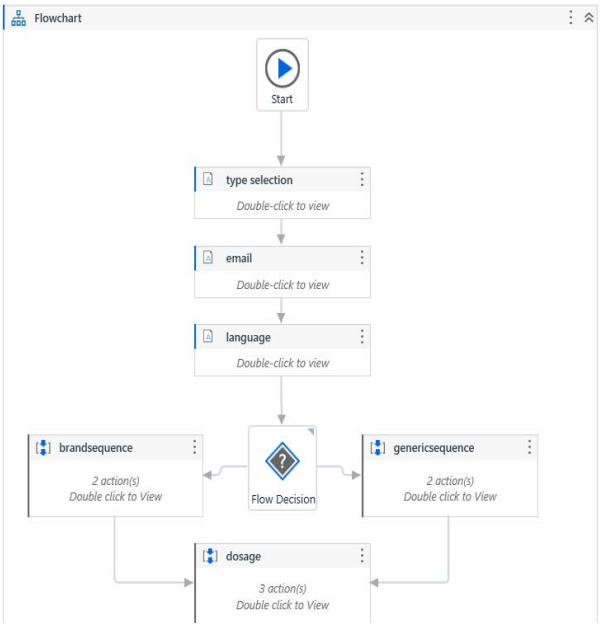

FIG 5.2 WORKFLOW

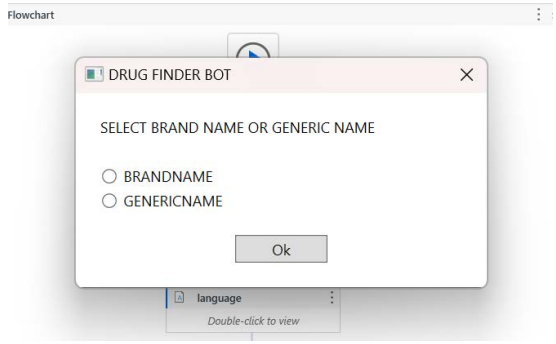

FIG 5.3 SAMPLE INPUT

Medicine Brands Starting with 'P'

Filter Medicine Brands: p

Brand	Company	Package	Strength	Price
PENTAGLOBIN	Paviour Pharma	Ampoule		₹ 14753.00
PENTAGLOBIN	Paviour Pharma	Ampoule		₹ 2058.00
PIPCOL.	Geno Pharmaceuticals Ltd.	Ampoule		₹ 3.40
PENTAGLOBIN	Paviour Pharma	Ampoule		₹ 7613.00
PROTA	Samarth Pharma Pvt. Ltd.	Ampoule		₹ 29.00
PYROLATE	Neon Laboratories Ltd.	Ampoule	0.2 mg	₹ 108.00
PRANOSOL	SGPharma Pvt Ltd.	Ampoule	1 g/ml	₹ 125.00
PROGTASE	Wockhardt Ltd.	Ampoule	1 MG/5ML	₹ 0.00
PROPRANOLOL (SAMARTH)	Samarth Pharma Pvt. Ltd.	Ampoule	1 mg	₹ 120.00
PROTAMINE SULPHATE	Biological E. Limited	Ampoule	1 %	₹ 25.00

FIG 5.4 SAMPLE OUTPUT 1

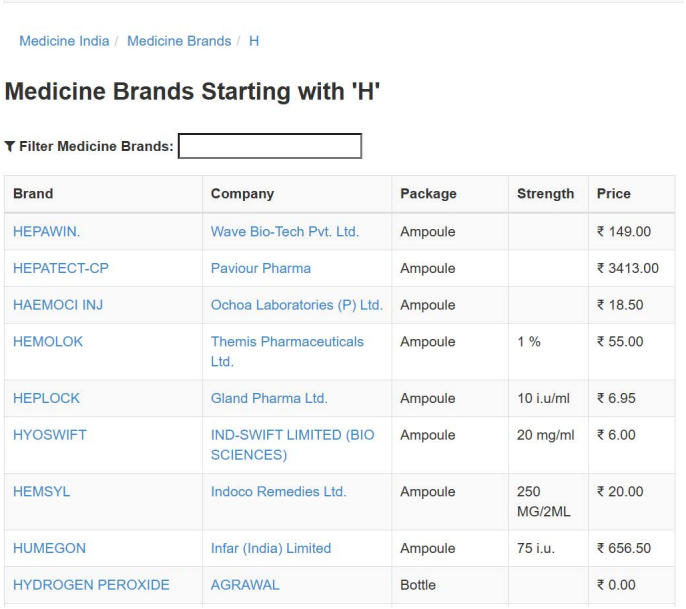

FIG 5.5 SAMPLE OUTPUT 2

5.3 RESULTS AND DISCUSSIONS

1.Accuracy of Data Retrieval:

Result: The bot successfully retrieved accurate drug information such as brand names, prices, and availability from pharmaceutical websites.
Discussion: The accuracy was measured by comparing the bot's results to trusted sources. In most cases, the bot returned correct and relevant data. However, occasional discrepancies were found when websites had inconsistent or incomplete data.

2.User Input Handling:

Result: The bot effectively handled user inputs (drug names) by matching them with entries in the internal list, even when there were minor spelling errors or variations in drug names.
Discussion: The input validation performed well, with few issues in recognizing synonyms or misspelled drug names. Enhancing the NLP model for more complex variations could improve this further.

3.Web Scraping Efficiency:

Result: The web scraping workflow was able to quickly retrieve data from multiple websites without significant delays.
Discussion: The bot demonstrated good performance, though some websites required additional parsing logic due to dynamic page content. Performance might vary depending on network speed and the complexity of website layouts.

4.User Experience:

Result: Users were able to interact with the bot intuitively and receive data in a structured format (e.g., tables or lists).
Discussion: User feedback indicated that the interface was easy to navigate. Some suggestions for improvement included providing more detailed instructions or visual indicators to guide users through the process.

Discussion:

Challenges in Data Scraping:

While the bot performed well in scraping static data, dynamic content and anti-bot mechanisms on some pharmaceutical websites presented challenges. Implementing advanced scraping techniques (e.g., using headless browsers) could resolve some of these issues.

Scalability and Load Testing:

The system showed robustness in processing multiple requests. However, when simulating a high volume of user interactions, some delays were noticed in data retrieval. Optimizing the bot's workflow and adding parallel processing could improve scalability for large-scale use.

Improvement in NLP for Drug Name Recognition:

Although the bot handled common variations in drug names well, complex queries (e.g., drug name abbreviations or slang) occasionally resulted in mismatches. Enhancing the NLP model with a larger dataset of drug name variations would further increase the bot's accuracy.

User Feedback and Future Enhancements:

User feedback was largely positive, but there is room for improvement in providing more detailed results (e.g., including side effects, dosage information). Adding a feature to show alternative drug recommendations or integrating more data sources could enhance the bot's utility.

Future Scope and Applications:

The 'Drug Seeker Bot' can be extended to support additional features like drug interactions, location-based availability, and integration with e-commerce platforms for direct drug purchasing. Further training of the bot's machine learning components (for recommendations) can add value to the application.

Limitations

- Dependency on External Websites:

  The bot relies heavily on data scraped from external pharmaceutical websites. Any changes to the website's structure or availability of data could cause the bot to malfunction or provide inaccurate information.

- Limited Data Sources:

  Currently, the bot extracts data from a limited number of websites. This could result in incomplete or inconsistent drug information. Expanding the number of reliable data sources could improve the bot's accuracy.

- Dynamic Content Handling:

  Some pharmaceutical websites use dynamic content (loaded via JavaScript or AJAX), which can be difficult for the bot to scrape. This may limit the bot's ability to extract information from certain websites or products.

- Data Accuracy Issues:

  Although the bot retrieves data from trusted sources, discrepancies in drug pricing, availability, and brand names may still occur due to inconsistent data across websites.

- Complex User Queries:

  The bot might struggle with complex user queries, especially those that involve abbreviations, slang, or less common drug names. More advanced NLP models would be needed for better recognition of such queries.

- Scalability Challenges:

  The bot's performance could degrade under heavy load or with an increase in the number of concurrent users. Optimizing the bot for scalability would be necessary for large-scale use.

- Lack of Real-time Updates:

  The bot does not provide real-time updates on drug prices or availability. The information retrieved may become outdated depending on how often the pharmaceutical websites are updated.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

The 'Drug Seeker Bot' is an efficient tool for retrieving drug-related information, such as brand names, prices, and availability, by leveraging web scraping and natural language processing (NLP). It successfully provides users with accurate and relevant drug details based on their inputs, improving accessibility to pharmaceutical data. However, the bot faces limitations such as dependency on external websites, handling dynamic content, and scalability issues under high demand. Despite these challenges, the bot has shown strong potential for real-world applications in assisting users with quick and easy access to drug information. Future improvements could focus on expanding data sources, enhancing NLP capabilities, and optimizing for scalability and real-time data updates. Overall, the project successfully demonstrates the capabilities of RPA and AI in the healthcare and pharmaceutical domains, with room for growth and refinement.

6.2 FUTURE WORK

- Expanding Data Sources:
  Integrate additional pharmaceutical websites, databases, and e-commerce platforms to provide a wider range of drug information, ensuring more comprehensive results for users.
  Include government and regulatory sites for verified drug details, improving the bot's credibility.
- Real-time Data Updates:
  Implement mechanisms for real-time updates on drug prices, availability, and new drug launches. This could be done by scheduling regular scraping intervals or using APIs from trusted sources for real-time information.

- Enhanced NLP Capabilities:

  Improve the natural language processing model to handle more complex user queries, including abbreviations, slang, and uncommon drug names.
  Train the model on a larger dataset to better understand and match drug names, making the bot more accurate and intuitive.

- Integration with Prescription Databases:

  Incorporate data from prescription management systems, allowing users to check whether specific drugs are available at local pharmacies or required prescriptions for purchase.

- Drug Interaction Checker:

  Add a feature that provides information on potential drug interactions when multiple drugs are entered, helping users avoid harmful combinations.

- Mobile and Multi-Platform Support:

  Extend the bot's capabilities to mobile platforms (iOS, Android) through app development, enabling users to access the service on-the-go.
  Support integration with popular messaging platforms like WhatsApp or Telegram for easy access.

- AI-Driven Recommendations:

  Integrate AI to recommend alternative medications or over-the-counter options based on user input or medical conditions, helping users make informed choices.

- User Feedback and Continuous Improvement:

  Implement a feedback loop from users to improve the bot's recommendations, accuracy, and overall performance. This could include a rating system for results and suggestions for future features.

REFERENCES

1. **UiPath Documentation:** [https://docs.uipath.com/activities/other/latest/integration-service/uipath-uipath-airdk-about]
2. **Machine Learning and Deep Learning Textbooks:**
    - Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
    - Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
3. **Online Tutorials and Resources:**
    - UiPath Official Documentation
    - YouTube Tutorials
    - UIPath Online Forums
    - REDDIT
    - WIKIPEDIA
    - GITHUB