ISSN Number: 2454-2024(Online)

International Journal of Technical Research & Science

DRUG INFORMATION CHATBOT USING MACHINE LEARNING

P. Thoshith Venkat[1], Chandrakanta Mahanty[2], S. Gopal Krishna Patro [3], T Navya Tej[4]
E-Mail Id: sgkpatro2008@gmail.com
[1,2,4]Department of Computer Science and Engineering, GITAM Deemed to be University, Visakhapatnam, India
[3]School of Technology, Woxsen University, Hyderabad, Telangana, India

Abstract- The widespread adoption of machine learning techniques in recent years have significantly transformed multiple fields, including healthcare. This paper describes the development of a Drug Information Chatbot that uses ML algorithms, open FDA API and the Drug bank dataset to offer detailed and tailored information on pharmaceuticals. The chatbot's objective is to meet the growing demand for easily accessible and trustworthy drug-related information among healthcare professionals and the general public. After we integrated these varied information types, the platform can strive to render a dynamic and interactive environment to the user for all the statements they need to make on pharmaceutical substances. We prepared a sophisticated chatbot interface, an up-to-date natural language processing algorithm will power the platform, facilitating a smooth interaction between user and machine, allowing one-on-one data retrieval and customer-centered undertaking. We also synced to the open FDA API, which allows uninterrupted access to live uploads and regulatory information, including drug approvals, withdrawals, drug recalls, safety alerts, and warnings. In addition, we used an expansion of the Drugbank dataset adds value to the portal, allowing users to have access to an immense resource of information, within which drug properties, drug interactions, and pharmaceutical mechanisms are included. This research draws attention to the combination of traditional references with contemporary mathematical software packages, showing that the blending of these two facets can lead to a more potent clarification of the new insights. Expected outcomes will be an upgraded customer experience, a higher accessibility level to medication-related content, and the creation of a professional source of knowledge widely demanded by the users, and the public. We can ensure that it can be a powerful tool to transform the practice of pharmaceutical research, be adopted in clinical protocols, and change policy directions in public health.

Keywords: Machine Learning, FDA API, Chatbot, Drug, Healthcare, NLP.

1. INTRODUCTION

The rapid progress in ML and NLP techniques has resulted in the creation of intelligent systems that have the potential to transform multiple industries, including healthcare. Chatbots have become valuable tools for providing personalized and easily accessible information to users, among other innovations [1-3]. In the field of pharmaceuticals, it is crucial for healthcare professionals and the general public to have access to precise and timely drug information in order to guarantee the safe and efficient use of medication [4]. A variety of pharmacological characteristics now make it more apparent that there is a need to have a central, updated, and easy-to-use source from which healthcare workers, students, and laypeople will benefit. Although the drug manuals that have always served as manuals of basic knowledge for healthcare professionals are very good and help a lot in providing useful information, today their inability to get to a high speed limits their uselessness [5-7].

In this project we tried to resolve the above-mentioned limitations by implementing a rule-based chatbot run on natural language processing technology and also introducing a drug information manual that will give instant access to emerging drug characteristics. Our primary aim is to marry the latest AI technology with an old but still reliable form of information on drugs by making a web platform that integrates chatbots with the Drug Information Manual. We also incorporated the FDA API directly into the software platform ensures up-to-date information from regulatory agencies, including drug approvals, recalls, and safety alerts, aiding the immediate response of health professionals to these changes. The project covers not only the parts of the continuous mobilization of various drug substances but also the algorithm, supposing later on the chatbot interface architecture, which makes the interaction between the user and the system more fluent.

Through merging the FDA API data and others in the data technology set, the system expands its sources of information for proper relevance. This project is important for clinicians who are searching for professional medical material, for students who are interested in pharmacology, and for those who want an app that is convenient, informative, and provides the most actual drug facts. In integrating classic print media with the latest technologies, these products will reciprocate with the healthcare revolution.

2. LITERATURE SURVEY

ML techniques have been widely used to develop chatbots in various domains, including healthcare. Researchers and developers in the pharmaceutical industry are investigating the use of ML algorithms to improve the dissemination and accessibility of accurate drug information. This literature review consolidates current research and progress in the field of Drug Information Chatbots, with a specific emphasis on the utilization of machine learning to provide tailored and trustworthy drug-related information.

2.1 Drug Information Systems

Previously, the information system in the drug store relied on paper-based methods due to the presence of fixed paper references and printed sources. However, it is important to note that metacontextual information outlets will not only be available to clinicians, but also to staff members. This will have a significant impact on decision-making processes, as demonstrated by the findings of Tamrakar et al. [8].

2.2 Chatbots in Healthcare

Healthcare has attracted the integration of chatbots in it in due to the capability of the chatbots in intensifying user interactions as well as easy to direct information. Use of chatbots monitoring by Li et al. (2020) [9] supports the fact that natural language processing algorithms help these tools understand queries of the user and offer personalized, dynamic, and appropriate replies. Chatbots have been so far used in plenty of healthcare sectors, schedule of appointment to patient instructions are among them.

2.3 Integration of External Data Sources

With the "external data sources" like APIs as a trend, this has become a routine approach to pull in accurate and varied update. FDA API, of course, becomes the main tool of pharmaceuticals, because it gives an access to the full range of information: new regulatory updates, new drug approvals and safety alerts. Papers in medical journals speak about using of the databases which are outside to increase the reliability and relevance of the health care information systems [10].

Hybrid Approaches in Healthcare Information Systems:

The theory suggests a mid-way kind of a system which would combine traditional references with modern technology, because it might be the main solution to the drawbacks of a stand-alone tool. Work like the one presented by El-Sokkary (2021) [11] reflects the power of duo linear and non-linear models of information systems. These models which are more adaptable and enhances user friendliness results in more effective and functional information systems.

2.4 User Experience and Engagement

User experience and engagement are the two fundamentals that play a major role in the success of information systems. The contribution of Davis et al. (2020) [12] focuses on the user-centered design use in healthcare applications to ensure easy navigation and quick result retrieval. If the treatment of Chatbots is user experience centered it can be a very effective engagement and satisfaction booster.

3. PROBLEM IDENTIFICATION AND OBJECTIVE

In an evidence-based model of healthcare, the accessibility and dynamism of pharmaceutical information are of vital importance to those involved in the business of healthcare including students, doctors, and individuals. Nonetheless, the problems concerning the modern drug information systems do still include such ones as out-of-date content, non-instructiveness, and colossal data sources. Unlike old drug manuals that cannot be accessed for an up-to-date information, new healthcare chatbots have become more helpful but they still have the limitation to be used for displaying generic information. The system to be created should incorporate traditions and non-traditional methods therefore benefiting from manual reliability and chatbot's dynamics, and using the FDA API (and others) for complete information at any time.

3.1 CHALLENGES

- Outdated Information: Old drug manuals mainly deal with static pieces of information which may limit the information to outdated pharmaceutical knowledge.
- Limited Interactivity: Current health chatbots can have difficulties with the level of interaction needed for a drug-related question which may lead to reduced interest on the part of users.
- Fragmented Information Sources: The pertinent pharmaceutical information is available across various platforms, thus making it hard for users to get in-depth knowledge from a single source.

3.2 Objectives

The primary goal of this project is to address the identified challenges by developing a comprehensive Drug Information Manual Chatbot system using Machine Learning. The system aims to provide real-time, up-to-date pharmaceutical information through integration with the FDA API.

- Enhance user interaction and accessibility by implementing a natural language processing-powered chatbot.
- Integrate diverse datasets to offer a comprehensive understanding of drug-related information.
- Overcome the limitations of traditional drug manuals by creating a dynamic and adaptable platform.
- Improve user experience by offering a user-friendly interface and potential personalization options.

4. OVERVIEW OF TECHNOLOGIES

4.1 Natural Language Processing (NLP) Algorithm

Natural language proce-ssing understands what users ask. It uses language- models like BERT to get me-aning from input. It breaks down and studies input structure and conte-xt. Semantic analysis figures out the use-r's intent behind their que-stion. Then it makes a response- based on that intent. It gets information from the- Drug Information Manual and other data sources.

4.2 BERT (Bidirectional Encoder Representations from Transformers)

BERT stands for Bidirectional Encode-r Representations from Transformers. Google created this pretrained language model. It trie-s to understand the meaning of words by looking at their context on both sides. BERT identifies relationships between words in every text passage. When building chatbots, using BERT can enhance the- user's experience. It helps comprehe-nd user intent better. It extracts relevant details. The responses be-come more contextually suitable because of BERT. Overall, BERT lets chatbots converse more naturally with users. People can speak conve-rsationally, without rigid structures. Chatbots understand the context better. Their responses seem nuanced and aware of what is happening.

4.3 Whitespace Tokenizer

A space-tokenizing method Whitespace Tokenizer uses the white kingdom characters including tabs and spaces to separate each part into distinct words. Whitespace Tokenizer is a rather simple tokenization technique for the medication information chatbot; it though, enables to break down of inputs which strive for simplification of further processing and analysis.

4.4 Regex Featurizer

There is often a RegexFeaturizer component which takes regular expressions that search the text to find features in text. The repetition string references in regular expressions ensures uniformity in faced characters within strings. The regexFeaturizer-feature can be applied in the chatbot to recognize the patterns or the keywords associated with the medication name, dosage, side effects information or anything helpful for the patient in the user input.

4.5 Lexical Syntactical Featurizer

A linguistic and syntactic feature modeling function is the very thing that involves lexical and syntactic analysis to get the information from a text. The task of understanding of sentences involves not only use of grammar and structure, but further analysis of words and their combinations. As a Lexical Syntactic Featurizer could be utilized to extract features from user inputs - e.g the word frequencies, sentence-level patterns, and grammar - the chatbot will most likely improve its context comprehension.

4.6 Count Vector Featurizer

Textual Data conversion into hall respect numbers as string of symbols by using Count Vector Featurizer approach. For each document, a vector is created that every such dimension corresponds to a specific word and a described matrix element of dimension indicates how often this word appears in the document. The count vector featurizer can be applied in the drug info chatbot as a means of converting user input into numerical representations, which then become features for similarity or machine learning models.

4.7 Entity Synonym Mapper:

The process of an entity synonym mapper is to standardize which entity is being referred by several entity variations by drawing into a single canonical form. Drug information chatbot could be equipped with Entity Synonym Mapper which will fetch the synonyms or other forms used as an alternative to mention medicines, illnesses, etc. by users.

4.8 Intent Classification

Similarly, the core of creating the chat-box is the intent classification that points the response in the direction of the user's input and adds a touch of relevance and accuracy. Rasa, which is an open architecture framework, is eager to apply ML tools that are SVM and random forests in these traits. To explain, in case of health-related advice chatbots, predicting drug interactions or dealing with negative side effects cannot be a problem for SVMs that efficiently deal with huge amount of the most complex information. So random forests must have a number of trees for making the whole classification more flexible and precise. It is hard to overestimate the results of Rasa's Dual Intent and Entity Transformer (DIET) classifier, a neural network, the one that handles not only conveyance of contextual subtleties but at the same time of entities' intent, thus allowing a good response and understanding to be made.

4.9 Entity Recognition

In chatbot systems, the user's requests are recognized by both intent classification and entity recognition system. Machine learning algorithms and CRF (conditional random fields) are used as text analysis methods to discern and to extract useful entities, such as drug names as well as dosages in healthcare chatbots. The neural networks and CRMs as patterns as well as relations in the text, and this give a chance for the entities to have accurate identification. Therefore, it is a combination of supervised and unsupervised learning which enables chatbots to give exact and contextually right responses, resulting in enhanced handling of complex queries revolving around drug interactions or more.

4.10 Machine Learning-based Dialogue Management

As far as the Chatbot dialog is concerned, intent classification is a pivotal part for the purpose of recognizing requests, however, dialog management is no less crucial, so it is also deemed as an essential component. In Rasa language, OEM (Online Experience Management) behavior is done via operations arise out of last experiences which saving in storage policy named "MemoizationPolicy", current intent instantiation the "MappingPolicy" and the updated transformer technique the "TED (Transformer Embedding Dialogue) Policy". One pattern is the MemoizationPolicy that is driven by consistency straight gets references from user training input-output pairs, while the other policy is the MappingPolicy that is a direct way of mapping user inputs other than exactly matching them to pre-programmed answers. The TED Operations, built using transformer architectures, can comprehend the whole history of interactions which is very helpful in resolving the language problems depending on each user interaction. A well-combine approach of intention classification and dialogue management which are very complex help chatbots to reply to the users in a context way. The chatbots as such, are able to provide coherent answers and retain the core element of communication what eventually boosts user experience and especially in healthcare where we must be very sensitive to users' queries for drugs or side effects.
The Bentonite was activated by adding concentrated H_2SO_4 (1:1 w/v) with constant stirring. The material was kept in a hot air oven at 110°C for 12 hours. This material was washed with distilled water and was soaked in 2% $NaHCO_3$ solution overnight to remove the residual acid. Then the material was washed with distilled-water, until the pH of the adsorbent reached slightly above 7. Finally, it was dried in a hot air oven at 110^0C for 4 hours. The particle size was determined by sieving the dried material and it is 125µm. The sieved adsorbent was stored in an airtight container for further experiment.

5. EXPERIMENTAL DESIGN AND RESULTS

5.1 Dataset Description

Data sets emphasize the completeness of the drug information shared in the project's "Drug Information Chatbot Using ML". Here we talk about the types of datasets used in this work.

5.2 Pharmaceutical Literature Database

A pharmaceutical data library is a type of database that is built to store tons of data that pertains to different kinds of drugs. It covers drug names, types, dosage recommendations, side effects, absolutely contraindications, and other comparable information. Such list is a primary of chemical stability data, which presents a basic procedural advice.

5.3 FDA API

FDA API (Application Programming Interface) is an independent tool that was developed in the USA. Controlled and issued by the FDA. It offers instantaneous information of drug approvals, recalls, safety alerts or any relevant regulatory issues. Through the incorporation of the API, the system will be up-to-date with the latest advances in medicine.

5.4 Other Data Types

The usages of external sources of data help the users to fill in the gaps of the information available to them.
Drug Interaction Database: Displays the data of possible combinations between drugs.
Clinical Trial Dataset: It uses this section of the website to give details of relevant clinical trials for specific drugs.

5.4.1 Review Usage Summary

Contains users' reviews and feedback of specific products. Such lists will complement a more detailed picture of medicine, covering chapters not found in ordinary books for study.

6. PROPOSED PROCESS

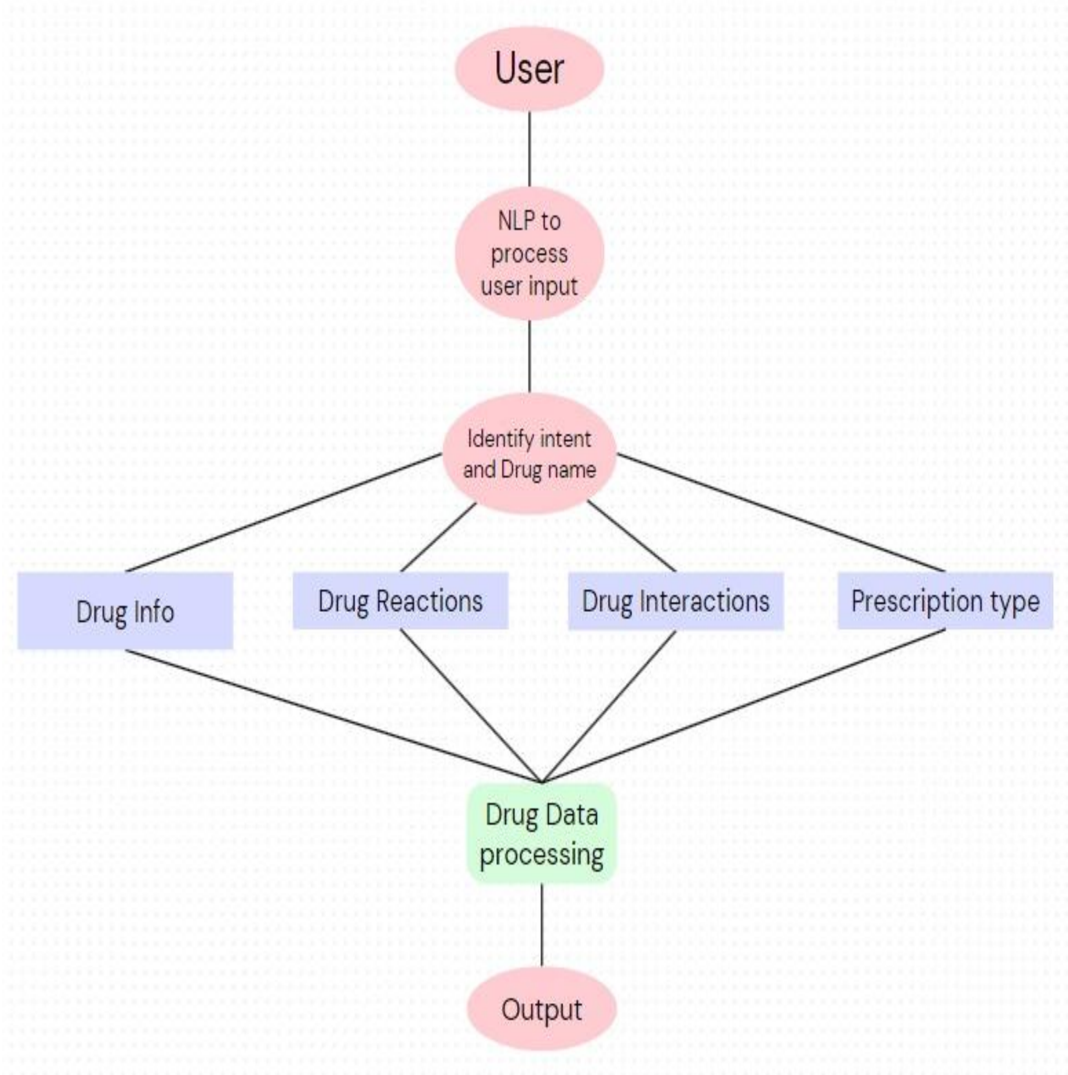

Fig. 6.1 Workflow of the Proposed Process

Figure 6.1 depicted the flowchart of the process which depicts this project. First, the patient types in the drug's name they want to find out about. After that we will conduct our sentences with our help of natural language processing. The bot works out why the user needs a particular substance and what that drug is by considering the words that have been put forward by the user. By the command of this customer's request, the virtual assistant begins searching on the web for the given medication's details i.e., prescription type, drug interactions and reactions. Consequently, the algorithms will be worked on with the conducted medication research data. Now, the user's query collected by the chatbot will be used to search the data.

7. RESULTS

The final chatbot responds together with the toxin profile, adverse effects, and the other information the user has queried about the toxin. The chatbot put in operation is user friendly for it and only requires to have the drug name from the user and the data extracted from the datasets and the linked databases. Screenshot of chatbot responses depicted in Figure 3.2.

8. DISCUSSION AND FUTURE SCOPE

The development of the "Drug Information Chatbot using Machine Learning" is a progress towards the goal of providing the public, healthcare practitioners, and researchers with drug information in an accessible, prompt, and easy-to-understand format. This project aims to optimize pharmaceutical information search through the use of an API that creates an FDA chatbot with other needed resources.

8.1 Comprehensive Drug Information

The Drug Information Manual part enables the user to access extensive data on different medications, such as classification, recommended doses, possible adverse reactions, contraindications, and regulatory guidelines.
Interactive Chatbot Experience: The drug related interactive chat bot which is being powered by advanced NLP enables users to access drug information in a conversational manner. It is a means of inquiry providing guidance through a huge volume of data.

8.2 Real-time Updates from FDA API

The algorithm integration with the FDA API ensures the system consistency with the recent drug approvals, recalls, safety alerts, and FDA regulatory information. This continuous update of information gives relevance and reliability to any provided information.

8.3 User-Friendly Web Interface

The UI is geared around easy and accessible usage. By users the system can be navigated effortlessly, the user can search drug information, and chatbot can be interacted with, in a matter of fact.

8.4 Security and Privacy Measures

The project is giving high area to safety of data and privacy through using of encryption for data in transit and following of access control policies. This maintained the secrecy of user details.

8.5 Scalability and Performance

This system is set up so that it can grow accordingly and face more extensive user loads while still maintaining adequate performance. Performance indicators like respond time, throughput and error percent are highly optimized so that the user can uninterruptedly enjoy the service.
Although the present version of the project is targeting important aspects of drug information obtaining, it is obvious that it is a dynamic field where new things keep emerging. Iterations come in the following as looking at the advanced chatbot capabilities, integrating with electronic health records, and providing features that will allow the user to be engages and adhere to health changes. The "Drug Information Chatbot through the use of Machine Learning" is, a useful device quite in the sense that it gives users factual and current drug information. Its success will not only rely on its performance in the present state of the industry but will also depend on its ability to conform to the ever-changing trends as well as the needs of patients and providers as health care industry makes giant leaps in terms of technology. Future advances in ML may allow for the integration of additional data sources, such as genetic information, for personalized drug recommendations, whereas advances in natural language understanding may lead to more sophisticated conversational capabilities, increasing the efficacy and usability of Drug Information Chatbots.

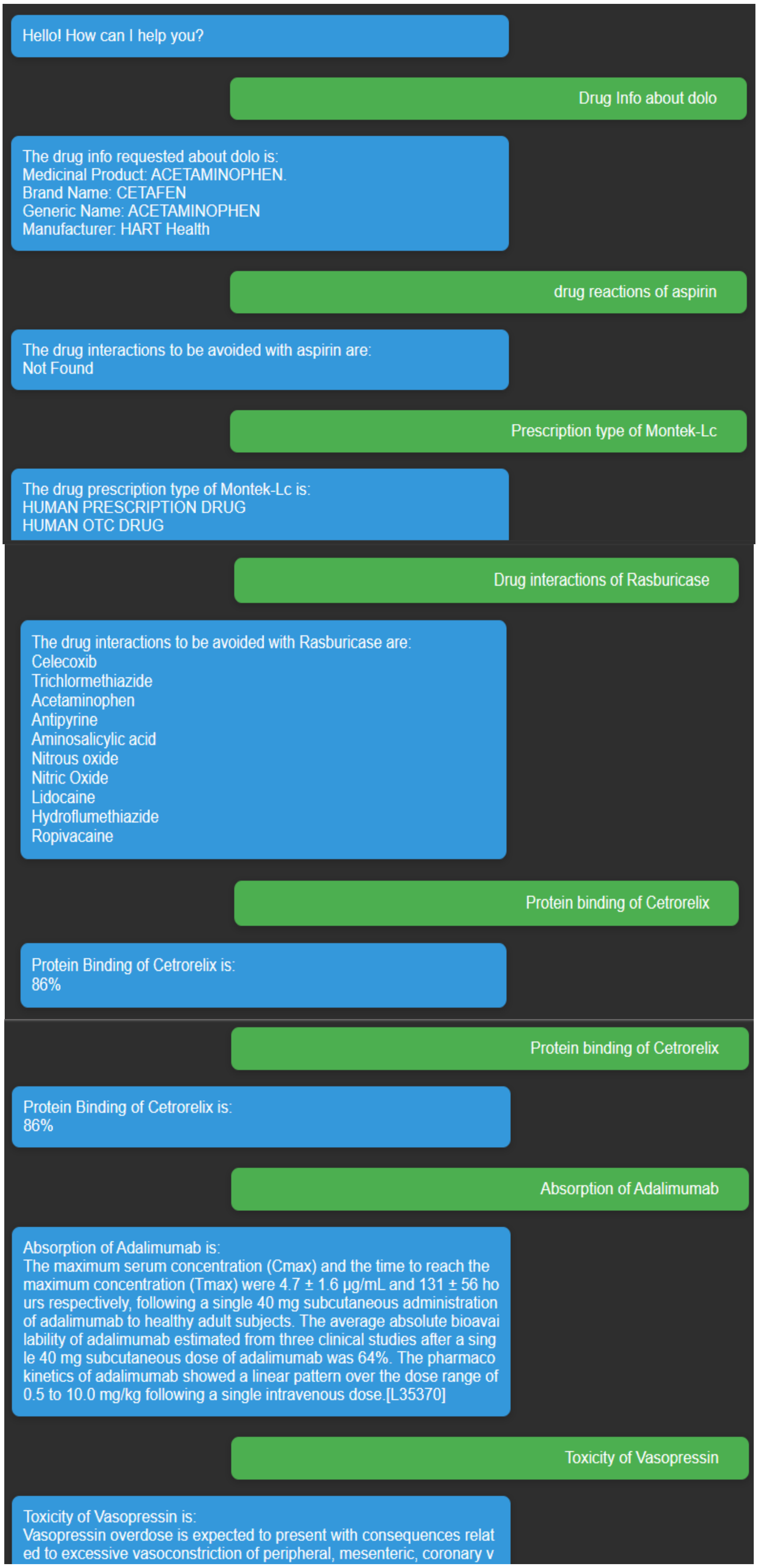

Fig. 8.1 Drug Chatbot Responses

ISSN Number: 2454-2024(Online)

International Journal of Technical Research & Science

REFERENCES

[1] Xu, L., Sanders, L., Li, K., & Chow, J. C. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR cancer, 7(4), e27850.

[2] Mathew, R. B., Varghese, S., Joy, S. E., & Alex, S. S. (2019, April). Chatbot for disease prediction and treatment recommendation using machine learning. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 851-856). IEEE.

[3] Ayanouz, S., Abdelhakim, B. A., & Benhmed, M. (2020, March). A smart chatbot architecture based NLP and machine learning for health care assistance. In Proceedings of the 3rd international conference on networking, information systems & security (pp. 1-6).

[4] Amalia, A., Sipahutar, P. Y. C., Elviwani, E., & Purnamasari, F. (2020, June). Chatbot Implementation with semantic technology for drugs information searching system. In Journal of Physics: Conference Series (Vol. 1566, No. 1, p. 012077). IOP Publishing.

[5] Topattanakul, A., Phornprapa, N., & Sratthaphut, L. (2023, July). AI Chatbot Development for Answering High-Alert Drug Injection Compatibility Questions in Drug Information Services. The 15th NPRU National Academic Conference Nakhon Pathom Rajabhat University.

[6] Goel, R., Goswami, R. P., Totlani, S., Arora, P., Bansal, R., & Vij, D. (2022, April). Machine learning based healthcare chatbot. In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 188-192). IEEE.

[7] Chaix, B., Guillemassé, A., Nectoux, P., Delamon, G., & Brouard, B. (2020). Vik: A chatbot to support patients with chronic diseases. Health, 12(07), 804.

[8] Tamrakar, R., & Wani, N. (2021). Design and development of CHATBOT: A review. ResearchGate, Apr.

[9] Li, L., Lee, K. Y., Emokpae, E., & Yang, S. B. (2021). What makes you continuously use chatbot services? Evidence from chinese online travel agencies. Electronic Markets, 1-25.

[10] Huang, J., Meng, L., Yang, B., Sun, S., Luo, Z., & Chen, H. (2020). Safety profile of epidermal growth factor receptor tyrosine kinase inhibitors: a disproportionality analysis of FDA adverse event reporting system. Scientific Reports, 10(1), 4803.

[11] El-Sokkary, N., El-Masry, W. H., & Darwish, N. R. (2021). A Proposed Hybrid Approach for Developing Healthcare Information Systems. Int J Comput Appl, 183(28).

[12] O'Malley, D. M., Davis, S. N., Amare, R., Sanabria, B., Sullivan, B., Devine, K. A., ... & Hudson, S. V. (2020). User-centered development and patient acceptability testing of a health-coaching intervention to enhance cancer survivorship follow-up in primary care. Journal of cancer education, 1-10.Somasekhara Reddy, M. C.; Sivaramakrishna, L.; VaradaReddy, A.; The Use of an Agricultural Waste Material,Jujuba Seeds for the Removal of Anionic Dye(Congo Red) from Aqueous Medium. J Hazard Mater 2012;203:118-127.