### A Project-I Report on

#### **UMIT Chatbot**

Submitted in partial fulfillment of Bachelor of Technology in

## Information Technology

by

Ms. Poorva Kalekar Ms. Komal Kamble Ms. Sakshi Khanvilkar

Under the guidance of **Prof.Prachi Dhannawat** 



Usha Mittal Institute of Technology S.N.D.T. Women's University Juhu Tara Road, Santacruz(West), Mumbai-400049 2022-2023

#### **CERTIFICATE**

This is to certify that **Poorva Kalekar**, **Komal Kamble**, **Sakshi Khanvilkar** has successfully completed phase-II of the project **UMIT Chatbot** in the partial fulfillment for the bachelor's degree in **Information Technology** during the year 2022-2023 as prescribed by SNDT Women's University.

**GUIDE** 

**HEAD OF DEPARTMENT** 

Prof.Prachi Dhannawat

Dr.Sanjay Shitole

**Principal**Dr.Shikha Nema

Examiner 1

Examiner 2

#### ACKNOWLEDGEMENT

We would like to express our profound gratitude to Prof. Prachi Dhannawat, and Dr. Shikha Nema (Principal) of SNDT University for their contributions to the completion of phase II of our project titled UMIT Chatbot. We would like to express our special thanks to our guide Prof. Prachi Dhannawat for the time and efforts she provided throughout the year. Your useful advice and suggestions were really helpful to us during the project's completion. In this aspect, we are eternally grateful to you. We would like to acknowledge that this project was completed entirely by us and not by someone else.

#### Submitted By:

Poorva Kalekar - 24 Komal Kamble - 26 Sakshi Khanvilkar - 30

# CONTENT

LIST OF FIGURES—	-5
ABSTRACT———————————————————————————————————	6
CHAPTER 1: INTRODUCTION——————	<del></del> 7
PROBLEM STATEMENT——————————————————————————————————	-10
SCOPE————————————————————————————————————	—11
CHAPTER 2: LITERATURE SURVEY————	—12
CHAPTER 3: EXISTING SYSTEM——————	-16
CHAPTER 4: PROPOSED SYSTEM—————	-17
CHAPTER 5: UML DIAGRAMS————————————————————————————————————	-19
CHAPTER 6: HARDWARE AND SOFTWARE RE-	
QUIREMENTS—	22
CHAPTER 7: IMPLEMENTATION—————	25
CHAPTER 8: APPLICATIONS—————	34
CHAPTER 9: FUTURE SCOPE———————	35
CHAPTER 10: CONCLUSION————————————————————————————————————	36
REFERENCES—	37

# List of Figures

Fig.1 Use-Case Diagram————	19
Fig.2 Class Diagram—————	20
Fig.3 Sequence Diagram————	21

#### **ABSTRACT**

Our main objective is to create a college website chabot that interacts with the users and answers their questions. Our user's requests via any other contact source like a phone call, email, etc are highly reduced. Most of the questions asked by the users are simple and repetitive hence addressing them by using the chatbot makes it simple for the user. There are a handful of questions that get asked a lot. These are standard FAQs related to admission, placements, etc. These can be answered in a few minutes by the chatbot. It will suggest a few subcategories to choose from and would recommend the college website if necessary.

#### **CHAPTER 1: Introduction**

A Chatbot (also known as a talbot, chatterbot, Bot, IM bot, interactive agent, or Artificial Conversational Entity) is a computer program or artificial intelligence that conducts a conversation via auditory or textual methods. Chatbots can be added to a buddy list or provide a single game player with an entity to interact with while awaiting other "live" players. If the bot is sophisticated enough to pass the Turing test, the person may not even know they are interacting with a computer. Chatbots are intelligent conversational computer programs that mimic human conversation in their natural form. A chatbot can process user input and produce an output. Usually, chatbots take natural language text as input, and the output should be the most relevant output to the user input sentence. Chatbots can also be defined as "online human-computer dialogue system(s) with natural language". Chatbots constitute therefore an automated dialogue system, that can attend to thousands of potential users at once. As consumers continue to move away from traditional forms of communication, chat-based communication methods are expected to rise. Chatbots are currently applied to a variety of different fields and applications, spanning from education to e-commerce, encompassing healthcare and entertainment. Therefore, chatbots can provide both support in different fields as well as entertainment to users this is the case for chatbots such as Mitsuku and Jessie Humani, "small talk" oriented

chatbots that could provide a sense of social connection. Chatbots appear, in fact, to be more engaging to the user than the static Frequently Asked Questions (FAQ) page of a website. At the same time, chatbots can simultaneously assist multiple users, thus resulting in more productive and less expensive compared to human customer supports services. In addition to support and assistance to customers, chatbots can be used for providing entertainment and companionship for the end user. Nonetheless, different levels of embodiment-the way chatbots are human-like and disclosure-how and when the nature of the chatbot is revealed to the user-seem impact users' engagement with and trust in chatbots. In recent years, with the commoditization and the increase of computational power, and the sharing of open-source technologies and frameworks, chatbot programs have become increas-Recent developments in Artificial Iningly common. telligence and Natural Language Processing techniques have made chatbots easier to implement, more flexible in terms of application and maintainability, and increasingly capable to mimic human conversation. However, human-chatbot interaction is not perfect; some areas for improvement are contextual and emotional understanding and gender biases. Chatbots are, in fact, less able to understand the conversational context and emotional linguistic cues compared to humans, which affects their ability to converse in a more entertaining and friendly manner. At the same time, chatbots tend to take on traditionally feminine roles which they perform with traditionally feminine features and often display stereotypical behavior, revealing a gender bias in chatbots' implementation and application. Since chatbots are so widespread and applied to many different fields, improvements in their implementations and evaluation constitute important research topics. The main contributions of this paper are: (i) extensive survey of the literature work on chatbots as well as the state of the art on chatbots' implementation methods, with a focus on Deep Learning algorithms, (ii) the identification of the challenges and limitations of chatbots implementation and application, and (iii) recommendation for future research on chatbots.

## PROBLEM STATEMENT

Chatbots utilize methods and algorithms from two Artificial Intelligence domains: Natural Language Processing and Machine Learning. However, there are many challenges and limitations in their application. In this survey, we review recent advances in chatbots, where Artificial Intelligence and Natural Language processing are used.

#### **SCOPE**

Our main objective is to create a college website chabot that interacts with the users and answers their questions. Our user's requests via any other contact source like a phone call, email, etc are highly reduced. Most of the questions asked by the users are simple and repetitive hence addressing them by using the chatbot makes it simple for the user. There are a handful of questions that get asked a lot. These are standard FAQs related to admission, placements, etc. These can be answered in a few minutes by the chatbot. It will suggest a few subcategories to choose from and would recommend the college website if necessary.

#### CHAPTER 2: LITERATURE SURVEY

Testing Chatbots are software agents used to interact between a computer and a human in natural language. Just as people use language for human communication, chatbots use natural language to communicate with human users. The main aim of their creation was to resemble a human being in the way they perform said interaction, trying to make user think that they are writing to a human. In this paper, we analyse some existing chatbots systems which are as follows:

- 1.Eliza is considered as the first Chatbot, which works on the pattern matching system. It is developed by Joseph Weizenbaum in 1964. ELIZA was one of the first chatterbots and one of the first programs capable of attempting the Turing test. ELIZA's creator, Weizenbaum, regarded the program as a method to show the superficiality of communication between man and machine, but was surprised by the number of individuals who attributed human-like feelings to the computer program, including Weizenbaum's secretary.[1]
- 2. A.L.I.C.E. (Artificial Linguistic Internet Computer Entity), also referred to as Alicebot, or simply Alice, is a natural language processing chatterbot —a program that engages in a conversation with a human by

applying some heuristical pattern matching rules to the human's input. It was inspired by Weizenbaum's classical ELIZA program. The program is unable to pass the Turing test, as even the casual user will often expose its mechanistic aspects in short conversations. ALICE was implemented by Richard Wallace in 1995.[5]

- 3. Cleverbot is a chatterbot web application that uses an artificial intelligence (AI) algorithm to have conversations with humans. It was created by British AI scientist Rollo Carpenter. Unlike some other chatterbot. Cleverbot's responses are not pre-programmed. Instead, it learns from human input: Human's type into the box below the Cleverbot logo and the system finds all keywords or an exact phrase matching the input. After searching through its saved conversations, it responds to the input by finding how a human responded to that input when it was asked, in part or in full, by Cleverbot.[2]
- 4. **Kuki**, formerly known as Mitsuku, is a chatbot created from Pandorabots AIML technology by Steve Worswick. It is a five-time winner of a Turing Test competition. Kuki claims to be an 18-year-old female chatbot from the Metaverse. It contains all of Alice AIML files, with many additions from user generated conversations, and is always a work in progress.[3]

5. **PARRY** was written in 1972 by psychiatrist Kenneth Colby, then at Stanford University. Parry is natural language program that simulates the thinking of a paranoid individual. This program was the first to pass the "Turing Test".

Sr No.	Paper Name	Authors	Purpose	Method/ Algorithm/Tools used	Result
1.	College Enquiry Chat-bot System	Prof. Ram Manoj Sharma [2]	The bot analyses users query and understands user messages.	Artificial Intelligence algorithms	Reduce the work load on the college's office staff and reduce the response time to a user's query.
2.	An Intelligent College Enquiry Bot using NLP and Deep Learning based techniques	P.Nikhila, G.Jyothi, K.Mounika, Mr. C Kishor Kumar Reddy & Dr. B V Ramana Murthy [3]	AIML is employed to make or customize alicebot that could be a chat-bot application supported ALICE free code.	AIML (Artificial Intelligence Mark-up Language)	The system will save time ,extra manpower of students.
3.	A Smart Chatbot	Harsh Pawar , Pranav Prabhu, Ajay Yadav, Vincent Mendonca , Joyce Lemos [6]	It uses SQL (Structured Query Language) for pattern matching which is been stored in program.	Using knowledge in database	It saves time and reduce the workload of students.
4.	College enquiry system	Nitesh Thakur, Akshay Hiwrale, Sourabh Selote, Abhijeet Shinde and Prof. Namrata Mahakalkar [10]	some emerging capabilities for evolving	NLP (Natural Language Processing)	It saves time and reduce the workload of students.
5.		Swati Khare	To minimize the time required to solve the queries of an user, reduce the work load	Using knowledgeable dataset and interpreter.	The system will save time ,extra manpower of students.

#### CHAPTER 3: EXISTING SYSTEM

Users need to personally visit the college and ask the college help desk if the student has any queries about the college. It takes a lot of time and money if the college is miles away from the student's native place.

#### Disadvantages:

- -The chatbot system is not known to people who do not have more knowledge.
- -Even if there exists a chatbot system, it is not accurate in providing the answers to our solution.
- -Also this process may lead to a communication gap between students and the college.

#### CHAPTER 3: PROPOSED SYSTEM

This application provides answers to the query of the user. Users have to query through the bot which is used for chatting. The system used is built in PyTorch, NLTK, and Neural network concepts. The answers are appropriate to what the user queries. The proposed system will have the following modules: A] Online Enquiry: Students can enquire about facilities and queries related to exams, academics, fee structure, etc. Students can also ask questions related to placement activities. B] Online Chatbot: The query will be answered on the basis of questions asked and the language model built and also the response media created. Given below is the system architecture of this chat-bot: The basic algorithm that will be implemented for working of this proposed system is as follows: Step 1: Start.

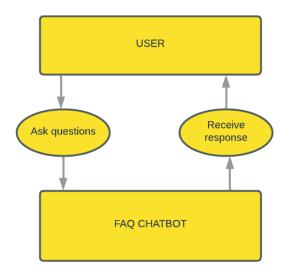
- Step 2: Get the input query from the user.
- Step 3: The query is pre-processed.
- Step 4: Fetch the remaining keywords from the query.
- Step 5: Match the fetched keywords with the keywords in the Knowledge base, and provide an appropriate response.
- Step 6: Further the Database module is used to call proper services using entity information to find proper data.
- Step 7: The keywords will be matched with the help of the keyword matching algorithm.
- Step 8: It returns the query response to the bot.

Step 9: Chatbot packages the data into a proper response for display by the client.

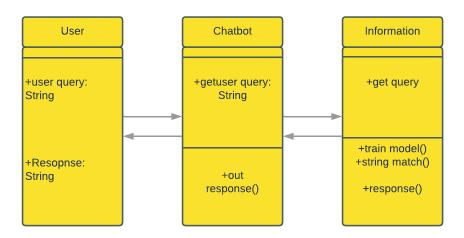
Step 10: Exit

# **CHAPTER 5: UML DIAGRAMS**

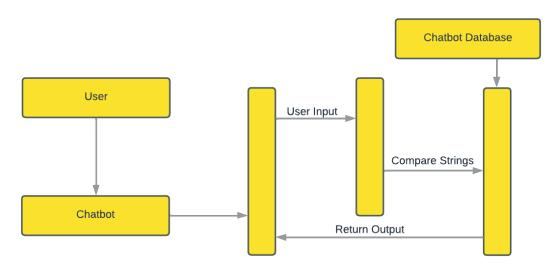
USE-CASE DIAGRAM



## CLASS DIAGRAM



# SEQUENCE DIAGRAM



# CHAPTER 6: HARDWARE AND SOFTWARE REQUIREMENTS

Hardware:

Processor: intel core5

Hard disk: 500GB

Monitor: 15"LED

RAM: 4GB

Input devices : Keyboard, Mouse

Software:

Operating system: Windows 10

Tool used: Visual Studio

Language used: Python: 3.8.5, HTML, CSS, Java

basic libraries:

NumPy: 1.19.5

NLTK: 3.2.5

PyTorch

Database: JSON

**Python:** Python frameworks are one of the most popular ways the language is used by web developers today. Several iterations of the language have come through, with one of the major transitions being the move from Python 2 to 3. Here we have used Python 3.8.5 version.

HTML: HTML is a markup language that defines the structure of your content. HTML consists of a series of elements, which you use to enclose, or wrap, different parts of the content to make it appear a certain way, or act a certain way. The enclosing tags can make a word or image hyperlink to somewhere else, can italicize words, can make the font bigger or smaller, and so on.

CSS: Cascading Style Sheets (CSS) is a stylesheet language used to describe the presentation of a document written in HTML or XML (including XML dialects such as SVG, MathML or XHTML). CSS describes how elements should be rendered on screen, on paper, in speech, or on other media. CSS is among the core languages of the open web and is standardized across Web browsers with W3C specifications.

Java: Java is a high-level, robust, object-oriented, and secure programming language. Java was developed by Sun Microsystems (which is now a subsidiary of Oracle) in the year 1995. James Gosling is known as the

father of Java. Before Java, its name was Oak. Since Oak was already a registered company, so James Gosling and his team changed the name from Oak to Java. Platform: Any hardware or software environment in which a program runs, is known as a platform. Since Java has a runtime environment (JRE) and API, it is called a platform.

**GUI:** A graphical user interface (GUI) is a digital interface in which a user interacts with graphical components such as icons, buttons, and menus. In a GUI, the visuals displayed in the user interface convey information relevant to the user, as well as actions that they can take. Today, it's hard to imagine computers without GUIs. But, there was a time when we did not even have a mouse cursor. Let's look at how GUIs came to be. In a GUI, visual elements represent actions users can take. objects that users can manipulate, and other information relevant to the user. Sometimes, GUI graphics are tied to real-world objects to represent their purpose and help users understand the functions available to them. For example, a directory (i.e., a group of files and/or directories) is represented by a folder icon, and the trash is represented by a trash can. The ultimate purpose of GUIs is to allow humans to interact with a device's underlying code by separating us from the technical details and presenting a simplified interface to t- the user. This process is called abstraction and is a key concept when designing and building interfaces. GUI Examples. You don't have to look far for examples of graphical user interfaces you're using one right now. Most websites, web apps, software programs (e.g., Microsoft Word and Apple Music), web browsers (e.g., Chrome and Internet Explorer), and operating systems (e.g., Windows, macOS, iOS, and Android) have GUIs.

NumPy 1.19.5: NumPy 1.19.5 is a short bugfix release. Apart from fixing several bugs, the main improvement is the update to OpenBLAS 0.3.13 which works around the Windows 2004 bug while not breaking execution on other platforms. This release supports Python 3.6-3.9 and is planned to be the last release in 1.19.x cycle.

NLTK: 3.2.5: NLTK 3.2.5 release: September 2017, Arabic stemmers (ARLSTem, Snowball), NIST MT evaluation metric and added NIST international tokenize, Moses tokenizer, Document Russian tagger, Fix to Stanford segmenter, Improve treebank detokenizer, VerbNet, Vader, Misc code and documentation cleanups, Implement fixes suggested by LGTM.

PyTorch: PyTorch is an optimized tensor library for deep learning using GPUs and CPUs. Features are classified by release status: Stable: These features will be maintained long-term and there should generally be no major performance limitations or gaps in documentation. We also expect to maintain backward compatibility (although breaking changes can happen and notice will be given one release ahead of time). Beta: These features are tagged as Beta because the API may change based on user feedback, because the performance needs to improve, or because coverage across operators is not yet complete. For Beta features, we are committing to seeing the feature through to the Stable classification. We

are not, however, committing to backwards compatibility. Prototype: These features are typically not available as part of binary distributions like PyPI or Conda, except sometimes behind run-time flags, and are at an early stage for feedback and testing.

JSON: JSON stands for JavaScript Object Notation. JSON is a lightweight format for storing and transporting data. JSON is often used when data is sent from a server to a web page. JSON is "self-describing" and easy to understand. The JSON format is syntactically identical to the code for creating JavaScript objects. Because of this similarity, a JavaScript program can easily convert JSON data into native JavaScript objects. The JSON syntax is derived from JavaScript object notation syntax, but the JSON format is text only. Code for reading and generating JSON data can be written in any programming language. JSON file ends with a .json extension but not compulsory to store the JSON data in a file. You can define a JSON object or an array in JavaScript or HTML files.

#### CHAPTER 7: IMPLEMENTATION

Each conversational intent contains:

- -a tag (a unique name)
- -patterns (sentence patterns for our neural network text classifier)
- -responses (one will be used as a response)

And later on we'll add some basic contextual elements. NLP Concepts used are-

Tokenization- Splitting a string into meaningful units (e.g.: words, punctuation characters, numbers etc) Stemming-Generate the root from the words. the crude heuristic that chops off the ends of words. NLTK is a toolkit built for working with NLP in Python. It provides us with various text-processing libraries with a lot of test datasets. A variety of tasks can be performed using NLTK such as tokenizing, parse tree visualization, etc... Use the pip install method to install NLTK in your system: pip install nltk Once the intents.json file is loaded we start creating the training data. We tokenize our pattern We only want unique words so with the help of lower and stem we will ignore any repetition and punctuation marks. Next, we will import the bag of words with the tokenized sentence and all the words.

We create three linear layers according to the diagram. Our loss decreases with every epoch which means our examples are not very complex and our neural network is good for the purpose.

A Feed Forward Neural Network is an artificial neu-

ral network in which the connections between nodes do not form a cycle. The feed-forward model is the simplest form of a neural network as information is only processed in one direction. We create three linear layers using pytorch. Our loss decreases with every epoch which means our examples are not very complex and our neural network is good for the purpose.

For frontend we have used GUI i.e. graphical user interface. In a GUI, visual elements represent actions users can take, objects that users can manipulate, and other information relevant to the user. Some common visual elements are:

- the mouse cursor, which moves and manipulates other elements
- buttons, which users may click or tap to initiate an action
- toolbar and ribbons, essentially groups of buttons
- icons, small images that represent information or interactive components
- menus, lists of clickable items
- scrollbars, for scrolling down the page

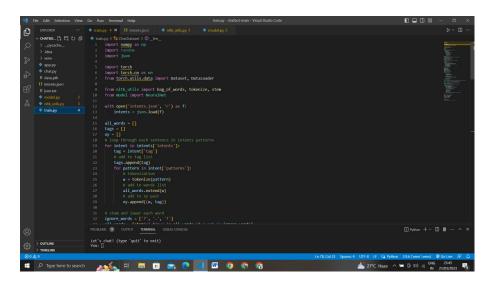
The languages that we have used include HTML, CSS and Java. In modern web development HTML and CSS really cannot live without each other. Both are necessary to achieve a modern experience. It is technically possible to make an entire website without using a CSS style sheet but it would be very impractical. More so, when it comes time to update or re-brand something it will be a massive headache.

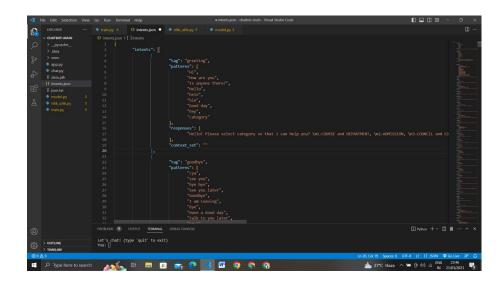
CSS works on a system of "selectors" like the name of the

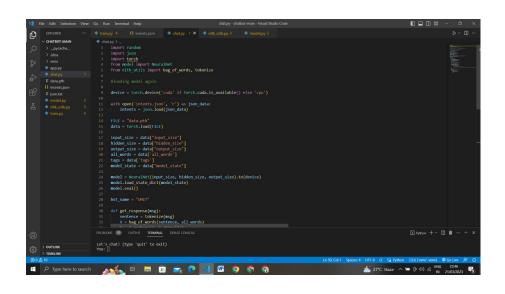
HTML element or an attribute like a class or ID. Using CSS selectors allows us to apply the same style attribute to all members of a class or instance of an ID without needing to copy the style stuff to each element one by one. The term cascading is what we are most interested in here, basically, there is a hierarchy and inheritance to CSS that allows for maximum code reuse and efficiency. This is really where the magic of CSS comes in and is critical to understanding how to develop your brand's style.

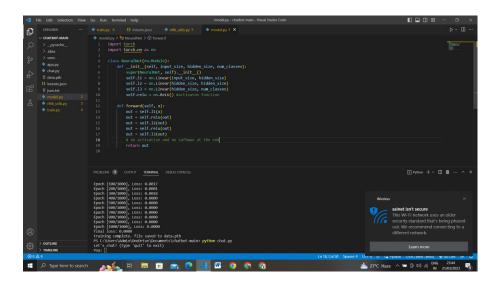
When it comes to web design and web development you will need a strong understanding of all three of these technologies to efficiently conceive, translate, and encode ideas to the web. The big takeaway should be to use HTML with its element tags for structure and content. Use CSS selectors on elements and like attributes like class, id, and in-line style to change how things look efficiently. And finally, any interaction with the server or browser or processing of logic will need to be done in Java.

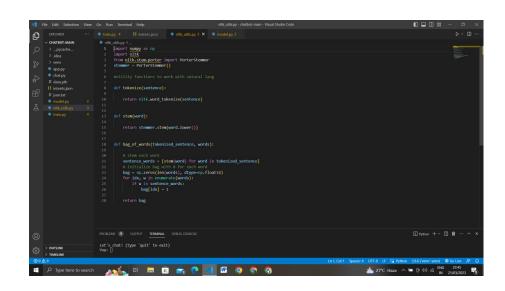
# **CODING**



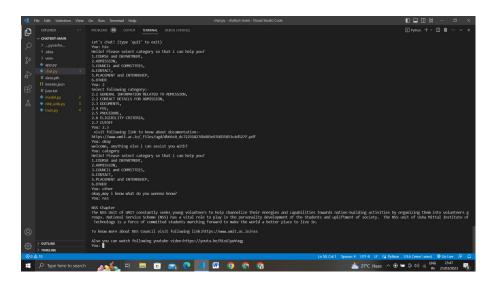


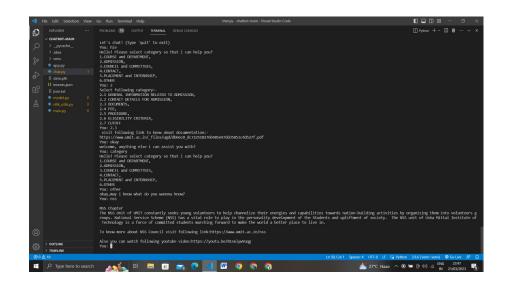


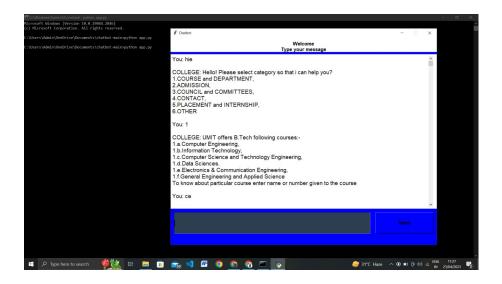


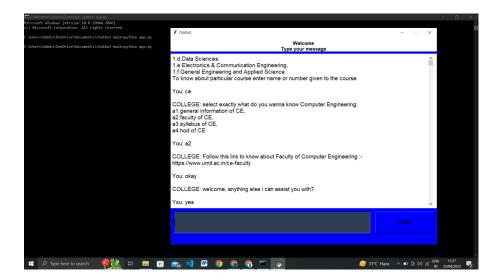


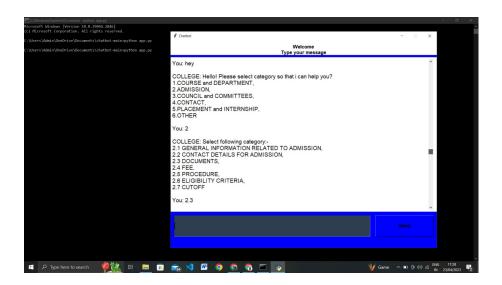
## **OUTPUT**

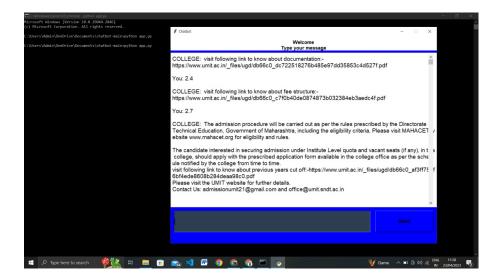












## **CHAPTER 8: APPLICATIONS**

- College inquiry chatbots help students to the right sources of information.
- Not only college inquiry chatbots any chatbot will provide them an instant as well as accurate response.
- Enhance Artificial Intelligence Based Chat Bot -The system will be used in most of the colleges around the country and it can be used in various firms and business-related industries.

#### CHAPTER 9: FUTURE SCOPE

The presented project can also be published on the google assistant platform as it is powered by Google. The google assistant production for chatbots has certain terms and conditions. That is where the limitations of our project come into the picture. The database of our proposed project is limited to our college and just some common information about our campus it still does not have all the thorough information about the whole university, so that is something we need to work on. And another thing is we lack a live website for our project currently, that is something we really need to work on. The moment we get a live website for our project it will be ready to be published on google Assistant and that is going to be a very big success for us and the project.

#### **CHAPTER 10: CONCLUSION**

This work presents the use of several latest trend technologies such as Artificial Intelligence, Machine Learning, and Python for developing a fully functional chatbot. Our main aim was to create a bot for our university which will be able to solve all the campus-related queries and that is what we did successfully. Moreover, this chatbot architecture is the same as the basic architecture of an interactive agent in general. The machine learning layer is the first layer where the input is received and understood by the machine. The inside core layer after the ML layer in the AI layer will fetch the desired response for the request query. The main contribution of this work is the development of the generic architecture for the creation of personality models for the chatbot and the application.

#### REFRENCES

- Bayan Abu Shawar and Eric Atwell, 2007 "Chatbots: Are they Really Useful?": LDV Forum GLDV Journal for Computational Linguistics and Language Technology. http://www.ldvforum.org/2007 $_Heft1/Bayan_Abu-Shawar_and_Eric_Atwell.pdf$
- Kerly, A., Hall, P., and Bull, S. 2007. Bringing chatbots into education: Towards natural language negotiation of open learner models. Know.-Based Syst. 20, 2 (Mar. 2007), 177-185.
- Lane, H.C. (2006). Intelligent Tutoring Systems: Prospects for Guided Practice and Efficient Learning. Whitepaper for the Army's Science of Learning Workshop, Hampton, VA. Aug 1-3, 2006.
  - http://en.wikipedia.org/wiki/Chatterbot
- ALICE. 2002. A.L.I.C.E AI Foundation, http://www.alicebot.org/
- P.Nikhila, G.Jyothi, K.Mounika, Mr. C Kishor Kumar Reddy and Dr. B V Ramana Murthy on , "Chatbots Using Artificial Intelligence", International Journal of Research and Development, Volume VIII, Issue I, January/2019, ISSN NO:2236-6124, pp 1-12.

-Sagar Pawar, Omkar Rane, Ojas Wankhade, Pradnya Mehta, Web Based College Enquiry Chatbot with Results", International Journal of Innovative Research in Science, Engineering and Technology, ISSN(Online): 2319-8753, ISSN (Print): 2347-6710, Vol. 7, Issue 4, April 2018, pp 3874-3880.

-Harsh Pawar, Pranav Prabhu, Ajay Yadav, Vincent Mendonca, Joyce Lemos, "College Enquiry Chatbot Using Knowledge in Database", International Journal for Research in Applied Science Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98, SJ Impact Factor: 6.887, Volume 6, Issue IV, April 2018, pp 2494-2496.