A Project Report on

AI And Web-Based Interactive College Enquiry Chatbot

Submitted in partial fulfillment of the requirements

in

Computer Engineering

by

Rohan Parkar (17102022)
Yash Payare (17102064)
Keyur Mithari (17102014)
Jitesh Nambiar(17102017)

Under the Guidance of
Prof. Jaya D. Gupta

Department of Computer Engineering
A.P. Shah Institute of Technology G. B. Road, Kasarvadavli, Thane(W),
Mumbai-400615
UNIVERSITY OF MUMBAI
Academic Year 2020-202

Approval Sheet

This Project Report entitled **"AI and Web-Based Interactive College Enquiry Chatbot"** Submitted by **"Rohan Parkar"(17102022), "Yash Payare"(17102064), "Keyur Mithari"(17102014), "Jitesh Nambiar"(17102017)** is approved for the partial fulfillment of the requirement in **Computer Engineering** from University of Mumbai.

_____ _____

(Prof. Jaya D. Gupta) (Prof. S.H.Malave)

Guide Head of Computer Engineering Department

Place : A.P.Shah Institute of Technology, Thane

Date :

CERTIFICATE

This is to certify that the project entitled "**AI and Web-Based Interactive College Enquiry Chatbot**" submitted by "**Rohan Parkar"(17102022),** "**Yash Payare"(17102064),** "**Keyur Mithari"(17102014),** "**Jitesh Nambiar"(17102017)** for the partial fulfillment of the requirement for award of a degree Bachelor of Engineering in **Computer Engineering**, to the University of Mumbai, is a bona fide work carried out during the academic year 2020-2021.

_____

(Prof. Jaya D. Gupta)

Guide

_____ _____

(Prof. S.H.Malave) (Dr.Uttan D.Kolekar)

Head of Computer Engineering Department Principal

_____

External Examiner

Place : A.P.Shah Institute of Technology, Thane

Date :

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

Rohan Parkar (17102022)

(Signature)

Yash Payare (17102064)

(Signature)

Keyur Mithari (17102014)

(Signature)

Jitesh Nambiar (17102017)

Date:

Project Conception and Initiation

1.1 Abstract

Students have to visit colleges to collect various information like College fees, Term Schedule, College Activities etc. during their admission process or as per their daily needs. This process is very tedious and time consuming, also it requires manpower in providing required information to visitors. A chatbot is the best tool which provides quick way to interact with the users. It is very helpful to the users as it allows to enter questions in natural language and desired information is obtained easily to the user. The project deals with user's request in form of question-based message and processes it to deliver a desired response in form of message. It avoids the process of visiting colleges and gathering related information as per the needs. The model is trained using concept of Natural Language processing (NLP) in Artificial Intelligence (AI). Python Language is utilized for the development of Chatbot. User can ask College-related questions, then the query is applied as an input to algorithm, which processes the message and displays the corresponding response to the user.

Humans react to others based on their mood and emotions. Whereas chatbots are bound by some rules, resulting them to treat a customer in the most polite and perfect way. Students can ask questions to chatbot at any time of the day and get reply very quickly. Chatbots can simultaneously have conversations with thousands of people at any time of the day. A chatbot can work 24x7 without getting tired. It is subjected to minimal errors thus increasing the productivity.

1.2 Objectives

1. To create an interactive Chatbot using Natural Language Processing (NLP) in
   Artificial Intelligence (AI).

2. To assist the students regarding college academic calendar, college fees, college timetable, etc.

3. To give students information about ongoing events and activities in college.

4. To provide 24/7 assistance for students queries.

5. The system will reply using an effective GUI which implies that as if a real person is talking to the user.

1.3 Literature review

AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT) paper proposed by Neelkumar P. Patel, Devangi R. Parikh, Prof. Darshan A. Patel, Prof. Ronak R. Patel. The project deals with user's request in form of question-based message and processes it to deliver a desired response in form of message. It solves the process of visiting colleges and gathering related information as per the needs, as it is time consuming. Also, the user can communicate to admin office with telephone number provided but doesn't receive a positive feedback. The project is a web-based chatbot. Graphical User Interface (GUI) is much similar to messaging application, which provides a friendly environment to the user as they are much aware of operating messaging applications. The user types a question and on performing submit, the message is preprocessed and the most relevant information from the database is provided as a response in similar way of messaging. Developing a chatbot solves the problems that can arouse in gathering required information. It can be accessed from anywhere at any time. In various websites, users are not able to find the required information on website which in turn end up closing the websites, which can be fulfilled by using chatbot. Presently, there are various chatbots available like ALICE bot which uses AIML (Artificial Intelligence Mark-up Language) and program Eliza. Such chatbot performs pattern matching which requires particular patterns to be matched. Hence, a chatbot named "UNIBOT" is developed. It delivers efficient and relevant response to the user corresponding to their entered message. The interface is effectively interactive. The time of response is minimal. It requires less memory and database hits are very less. The front-end is developed using HTML, CSS and jQuery. Ajax is used to call and get response from PHP file, whereas, jQuery is used to display the messages to the user.

Implementation of a Chat Bot System using AI and NLP proposed by Tarun Lalwani, Shashank Bhalotia, Ashish Pal, Shreya Bisen, Vasundhara Rathod designed Chat bots that are the source of answers to the users questions in any particular domain where it is operating. This is a project on chat bot for college inquiry system using AI and NLP. The bot is capable of answering admission related queries, viewing user profiles and retrieves attendance and grade/ pointers, getting information about examinations to be held and fetching particulars about placement activities. It aims to analyze user's queries and answer them, to save the time of the user, update about the college activities, etc. The bot first pre-processes the input to standardize it. The bot also validates the user with the help of user-id and password. There are two types of users, admin

and the user (students, teachers, etc.). If the user is trying to make a normal conversation with the bot, the input is mapped to an appropriate pattern in Artificial Intelligence Modeling Language (AIML) files and If the pattern is not available in AIML files, a random response is sent suggesting “Invalid Input”. Also, the question need not to be exactly same as in the database. The input is matched to find a similar pattern and if found, the answer corresponding to that query is posted. To create knowledge base AIML files are created. When user converses with our chat bot, the input is matched to patterns listed in AIML files and corresponding answer is returned as response. Information extraction from the input text is done by extracting keywords. Appropriate Lemmas of the keywords were found using Lemmatization and POS tagging, to group together the different inflected form of the words. Also, a log file is maintained, containing inputs which the chat bot was not able to answer. It can be accessed by the admin to check where the bot fails and use it as a feedback.

Review of AI Chatbot proposed by Prashanth S, Rakshith Gowda N, Sourabh Kakade, Gouramma, Vivek Sharma designed a dialog-exchanging system generating a meaningful and empathetic conversation between human and a computer is called a chatbot which processes the natural language input that can be either a speech or a textual form and gives the response in the same language and expression as of the human. The research is based on the survey of various chatbots and it is shown how the various chatbots differ from each other’s. To make chatbot the various technology has been used to make. A chatbot can be considered as a question-answer system where experts provide knowledge for solicitation of users. A chatbot is a software designed to simulate a conversation with a human partner. The college inquiry chat-bots will be built using artificial algorithms that analyze user's queries and understand the user's message. The User can ask the question any college-related activities through the chat-bot without physically available to the college for inquiry. The System analyzes the question and then answers it to the user. The user can ask any question related to college he asks about annual day, college fees, faculty details, sports day, etc. It helps the student to be updated on what is happening in the college. The system replies using an effective Graphical User Interface as if a real person is talking to the user. Natural language processing technologies are used for parsing, tokenizing, stemming and filtering the content of the complaint. This chatbot can be used by any user like students, faculties, parents to fetch the information from the institution website. AIML and NLP are used for creating chatbots. AIML is Artificial Intelligence Mark-up Language (AIML) by using NLP to answer user questions. Predefined knowledge base helps develop a response to the query. Chatbot for the educational sector, where users (a student or parents) can ask a query regarding college admission, about college information and other things related to academics. Chatbot is implemented to meet the academic needs of the

visitors. Sometimes the user will ask the query not in a proper way so that the A.I chat bot will not find the proper answer for that, it will give an answer based on the keyword what the user had given or it will not answer the question. It will test the queries to gather the information that is required to answer the question. It will predict the missing data from the query so that the Chatbot should understand and reply it properly. First it will identify the missing data then it will do a querying same to provide an accurate response.

Online Chatting System for College Enquiry using Knowledgeable Database proposed by Bathe Pooja Prashant, Malusare Sonali Anil, Kolpe Monika Dilip designed by a chatterbot or Chatbot aims to make a conversation between both human and machine. The machine has been embedded knowledge to identify the sentences and making a decision itself as response to answer a question. This project is mainly targeted at colleges and the synchronization of all the sparse and diverse information regarding regular college schedule. Generally, students face problems in getting correct notifications at the correct time, sometimes important notices such as campus interview, training and placement events, holidays and special announcements. Smart Campus tries to bridge this gap between students, teachers and college administrators. Therefore, in the real-world scenario, such as college campus, the information in the form of notices, oral communication, can be directly communicated through the android devices and can be made available for the students, teachers directly for their android devices and the maintenance of application will be easier in later future because of the use of architectural MVC which separates the major works in the development of an application such as data management, mobile user interface display and web service which will be the controller to make sure for fast and efficient maintenance of application.

A Web Based College Enquiry Chatbot with Results proposed by Sagar Pawar, Omkar Rane, Ojas Wankhade, Pradnya Mehta designed a chatbot (also known as a talkbot, chatterbot, Bot, IM bot, interactive agent, or Artificial Conversational Entity) is a computer program which conducts a conversation via auditory or textual methods. The chatbot has information stored in its database to identify the sentences and making a decision itself as response to answer a given question. In this paper bigram is used for calculating the sentence similarity. The machine has been embedded with the knowledge to identify the sentence and making a decision itself as response to answer a question . In this study of paper, the statistics of pattern matching on the text data and statistics of compressed pattern matching on compressed form of same text data are compared . Here the information repository is in the form of a connected graph where the nodes contain information and links interrelates the information nodes. The design

semantics includes AIML (Artificial Intelligence Markup Language) specification language for authoring the information repository such that chat bot design separates the Information repository from natural language interface component . The paper describes a novel method is proposed where bigram is applied to quantify the text and improved information gain algorithm are used to create appropriate feature during text categorization. The paper illustrates the implementation and semantic enhancement of domain-oriented question answer system based on pattern matching chatbot technology developed within industrial project FRASI . In this paper a modular knowledge base is equipped with the conventional architecture. It helps in building a specific module that deals with a particular feature of the conversation. This enhances agent's interaction capabilities . In this paper conversation of agent based on modular knowledge representation is proposed. It has a dynamic and flexible behavior. Because of the modularity of the architecture, it allows a concurrent and synergic use of different techniques making it possible to adapt to the specific characteristics of the domain. It has set of modules which is automatically triggered through a component . OntBot uses appropriate mapping techniques to transform ontologies and knowledge into relational database and then use that knowledge to drive its chat. Uses rule matching to match the sentence . Here the growing problem of malicious chatbots are taken into consideration and provide a supporting evidence to distinguish between human and chatbots . Traditional chatbot knowledge base is hard constructed and time consuming. Automatic chatbot knowledge acquisition from online is used here. It uses rough set and ensemble learning for that purpose .

1.4 Problem Definition

The need for college enquiry system arises due to various reasons which include: the slow nature of college website, an outsider would not know where to search for a particular piece of information, difficult for the person outside college's domain to extract information.

The smart solution for all the drawbacks lends to the need of the system. The college enquiry Chatbot which will provide the response by summarizing the query and then output answers, it also provides selective information what the user wants. A college Chatbot will dispense all answers relating to domains such as admission, examination cell, notice board, attendance, placement cell and other miscellaneous domains.

The major features of the chat bot are:

1. College admission related queries could be answered through it.
2. Viewing user profiles and retrieves attendance and grade/ pointers.
3. College students can get information about examinations to be held.
4. College students can fetch particulars about placement activities.

College enquiry Chatbot will act as a fast, standard and informative widget to enhance college website's user experience and bestow users with righteous information. The bot will analyze user's queries and understand users' message and then reply accordingly. It uses AI & NLP. This way users' time and efforts will be saved and s/he will be equipped with effective answers.

1.5 Scope

- In future we can include both text and voice-based queries. The users will have to give voice input and the system will give the text output for better conversation systems.

- It is often impossible to get all the data on a single interface without the complications of going through multiple forms and windows. This problem can be solved by chat bot by providing a common and user-friendly interface to solve queries of college students and faculties and parents.

- Accuracy can be increased by providing more data to the database and training the neural model.

- Also, after successful execution of chat bot in college domain, we can implement it in other domains like medical, forensic, sports, etc. It will be beneficial in all the fields as without spending much time, we are accessing the relevant information and that too without any sorting.

1.6 Technology stack

Software's Used:

1. **Anaconda**: Anaconda is an open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment.

2. **Pytorch**: PyTorch is an open-source machine learning library based on the Torch library, used for applications such as computer vision and natural language processing, primarily developed by Facebook's AI Research lab.

3. **Visual Studio Code**: Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

4. **Flask**: Flask is a micro web framework written in Python. It is classified as a micro framework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.

Algorithms:

1. **Natural Language Processing (NLP)**: Natural language processing is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.

1.7 Benefits for environment and society

- Chatbots will continue to operate every day throughout the year without requiring to take a break thus requiring less manpower.

- Chatbots on the other hand can simultaneously have conversations with thousands of people at any time of the day.

- Humans react to others based on their mood and emotions. Whereas chatbots are bound by some rules, resulting them to treat a customer in the most polite and perfect way.

- Chatbots can do repetitive tasks without errors unlike humans. This helps people save time and increase productivity.

Project Design

2.1 Proposed System

The proposed methodology makes use of both qualitative and quantitative perspectives, and includes a broad array of approaches such as literature reviews, expert opinions, focus groups, and content validation.
Online Enquiry:
Students can enquire about facilities and query related to exams, academics, fee structure, etc. Students can also ask questions related to placement activities.
Online Chatbot:
The result can be showed in the form of images and card format or in text format. The query will be answered on the basis of questions asked and the language model built and also the response media created. Users that want to enquire about the college at the time of admission or any competition held in the college can query to the chat-bot. The user can query about the college related activities through online with the help of this web application. The chatbot provide 24/7 assistance for students queries. The chatbot helps the student to be updated about the college activities. The college enquiry chatbot is built using Machine Learning and Artificial Intelligence that analyses user's queries and understands them. This chatbot is a web application which provides answer to the query of the students. The user can chat using any format there is no specific format the user has to follow. The user needs to type the query in the message box on the chat window of the Chatbot GUI. The query is pre-processed. E.g., suppose there is this query what are the project domains for CSE fourth year major projects. So, we are going to remove these stop words like are, the using pre-processing technique. The chatbot uses Convolution Neural Network (CNN) machine learning algorithm to learn the database (.json file). The chatbot finds for the most matched query from the database and the response is searched. The chatbot provides response to the user as a message on the chat window.

2.2 Design(flow of modules)

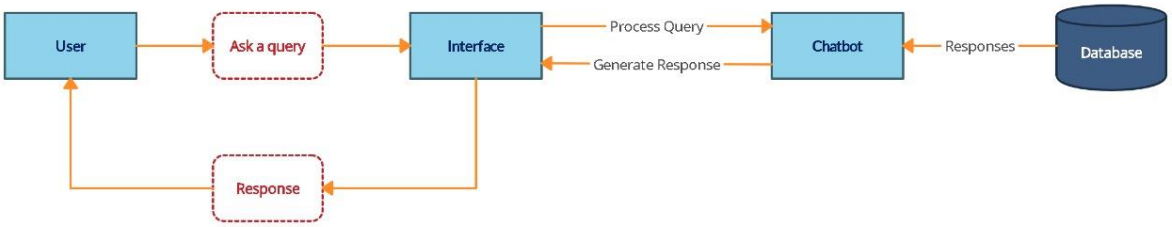

Figure 2.1: Data-flow Diagram

2.3 Class Diagram

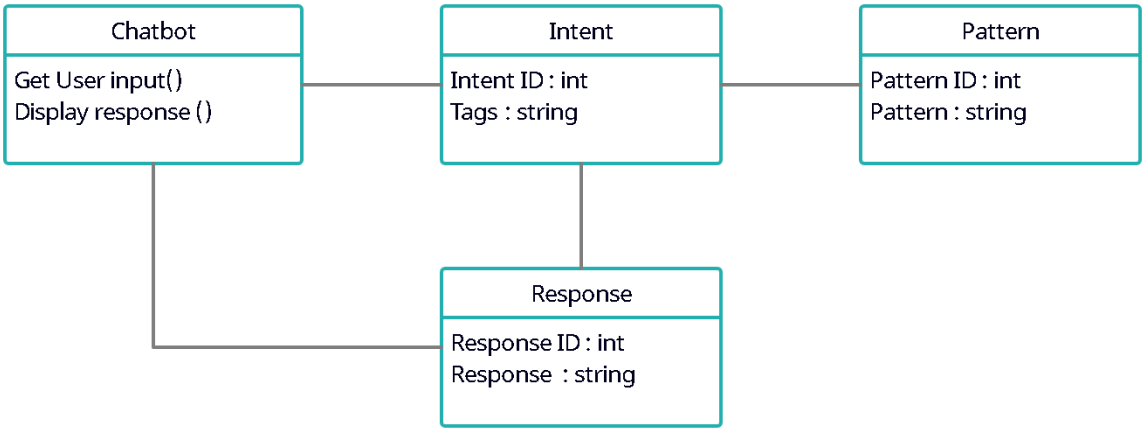

Figure 2.2: Class Diagram

2.4 Modules:

AI Chatbot

A chatbot is a computer program that simulates human conversation through voice commands or text chats or both. Chatbot, short for chatterbot, is an Artificial Intelligence (AI) feature that can be embedded and used through any major messaging applications. The chatbot functions through machine learning using an artificial neural network. Chatbots typically take the form of chat client, leveraging natural language processing to conduct a conversation with the user. A machine learning model is trained using datasets. The user's message will be forwarded by API to trained model where trained model will process on the requested message and pick the appropriate response from the database and the response is forwarded to the user by API.

Natural language processing (NLP)

Natural language processing (NLP) facilitates human-to-machine communication without humans needing to "speak" Java or any other programming language as it allows machines to obtain and process information from written or verbal user inputs. In NLP there are various steps such as:

- **Tokenization**
  Tokenization is a way of separating a piece of text into smaller units called tokens. Here, tokens can be either words, characters, or sub words. The formatting is based on spaces.

- **Stemming**
  Stemming is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers. The input to the stemmer is tokenized words.

- **Bag of Words**
  Bag of Words model is used to pre-process the text by converting it into a bag of words, which keeps a count of the total occurrences of most frequently used words. This model can be visualized using a table, which contains the count of words corresponding to the word itself.

Database

The database of the chatbot is in .json format. JavaScript Object Notation (JSON) is a standard text-based format for representing structured data based on JavaScript object syntax. It uses human-readable text to store and transmit data objects consisting of attribute–value pairs and array data types. The .json file contains the intents i.e., the goal the user has in mind when typing in a question or query to the chatbot application. Every intent has a tag name saved in a variable named ‘tag’ (for example: greeting, committees, location, contact, etc.). Also, with respect to every tag there are number of queries a user can ask under the heading ‘patterns. There are different patterns in which a user can ask a single query. All these patterns are included in this ‘patterns’' variable. The responses to these queries are stored in the ‘response’ variable. The chatbot returns this response to the user as a message if the query asked is the most matched with the pattern from the patterns variable. The database is created studying the college website and analyzing what queries a user can ask and what responses he or she will expect. Also, google forms have been used to collect data for the database.

Planning for next semester

Planning :

- **Creating Dataset**
  Creating a custom database with college information. Gathering various questions related college using google forms.

- **Creating AI chatbot**
  Training a Neural Network model using machine learning on the Dataset created.

- **Connecting Chatbot with frontend**
  Using Flask to connect the connect the model with web application.

- **Testing Accuracy**
  Increasing accuracy by adjusting the learning rate and adjusting other parameters of the model.

- **Result and Analysis**
  Reviewing the final results of the product and analyzing it.