

Multimedia Chatbot using Classification

Priya Sharma¹, Shivani Sabnis², Yash Mane³, Abhinay Chauhan⁴

^{1,2,3,4}B.E. Student, Computer Science, Ramrao Adik Institute of Technology, Maharashtra, India

Abstract - With the ever-proliferating need for information, the office staff often find themselves burdened with additional responsibilities of repeatedly answering the common queries. Chat bots in the digital world have made every organization curious to use them as a major tool to interact with their customers. The system, Multimedia chatbot, for answering various queries of different committees in the college has been developed by classifying the questions into various committee category questions. The proposed system aims at alleviating the setbacks faced by the students by developing a chatbot system as an application that aids students in getting updated with college activities. The chatbot is built using voice recognition techniques and Part-Of-Speech (POS) tagging which is a natural language processing technique that analyze user's queries and understand user's input. The chatbot includes voice assistance, which is accomplished using Google Speech API; and responds with the appropriate available format like audio, images, text. The chatbot has been made by using AIML (Artificial Intelligence Markup Language), JAVA and NLP (Natural Language Processing). Our chatbot effectively answers college related queries so the students get the desired information quickly without having to go through the unnecessary hassle. Students can provide input through voice or by text and can use this application for enquiring anytime in the college premises.

Key Words: POS tagging, NLP, AIML, JAVA, Voice assistance.

1. INTRODUCTION

With burgeoning domains of cutting-edge technology and information, prospects are that chatbots can bridge the gap between the reservoir of data and the latest generation programs which perform human conversation intelligently. The Multimedia chatbot is a system designed for students to fire queries pertaining to academic information as well as general information about the college. In behest of the user's query, the system interprets the question and delivers the result. It is not required for the users to follow some sequential format for firing the queries. NLP (Natural Language Processing) is used that primarily programs the computer to surveys and process large natural language data. The knowledge base contains question and answers. For every question a keyword or series of keywords are present which acts like classification words for that particular question. As the data-mining and machine-learning techniques have improved, chat bots have begun to be more practical in daily life applications. One of the examples being automatic telephone answering systems. This project, titled 'Multimedia Chatbot' streamlines a

pathway towards creating a chatbot application for our college which can accept user queries as text input as well as audio input, cognize and fathom the context and connotation of the queries, look up the pertinent data in the knowledge base and administer a relevant answer as output-all through an interactive and easy to understand chat interface. The chatbot encompasses an admin panel which has complete control over the system, enabling them to manage the knowledge base, keep the records of chat logs etc. Chatbot can pull data database and present it in the most appropriate format. The project overcomes the disadvantage of getting all the data on a single page without any complications of visiting multiple pages by providing a user-friendly interface that uses natural language processing to interact and solve queries of college students. The system analyses user's queries which are answered using AIML and NLP. Users can directly enter the chat screen and can ask any college related queries and in response get the answer from the chatbot with an attractive graphical user interface whereas the admin will have its user id and password for security purpose to manage database and chatlogs. Admin can know about the questions that were unanswered by the chatbot on viewing the chatlogs and can get the knowledge base updated. The system helps the users to be updated about college activities. The system analyses student's queries by using NLP's Grammatical Tagging algorithm.

2. LITERATURE SURVEY

COLLEGE ENQUIRY SYSTEM [1]: This System is a web application which provides answer to the query of the student very effectively. Students can input their query in the bot to receive the answers. If the answer is found invalid, then system declares the answer as invalid. These invalid answers are modified or else removed from the database by the admin of the system. Chatbot helps the students to get the answers to their queries. This keeps the students notified with the college activities. The system has an online notice board. On this notice board, any Text notices or PDF documents can be displayed. This helps the user to be updated with the important notices. Not much time will be wasted by the user to search for the important notices. The answer to the query is answered on the basis of the users queries and the knowledge base. The important keywords are fetched from the keywords and the answer to those keywords are searched in the knowledge base. If the match is found, the relevant answer is provided to the user or the default message is shown to the user that Answer to this query is not available at the moment, please revert back after some time. The Keyword Matching algorithm will be used to match the keywords from the knowledge base [1].

CHATBOT FOR COLLEGE MANAGEMENT SYSTEM USING AI [2]: The User can ask any college-related queries to the chatbot with the advantage of not personally going to the college for enquiry. The System examines the question and then provides the answer to the user. With the help of AI, system can provide answers to any queries asked by the students. This is done using an effective GUI as if a real person is talking to the user to use the system, user first requires to register and then login to it. This system uses two algorithms. First, Porter Stemmer Algorithm, which is a process for removing suffixes from words in English Second, Word Order Similarity between Sentences. Sentences having exactly the same words but differing in ordering may mean different. It is simple for humans to process word order information. Due to this, many preexisting methods do not handle this type of information. So, this system established a method which stores the word order information into account when calculating sentence similarity [2].

ENQUIRY CHATBOT USING KNOWLEDGE IN DATABASE [3] – This system answers to student questions that is related to college. In this system, students are first expected to select the category regarding which they want to question the bot. Once, the category is selected, student inputs the query which is then checked in the database for its presence. If the query is matched in database, the corresponding answer is forwarded to the student as an output. These queries can be related to admission, faculty details, etc. Students won't have to personally go to the college for enquiring. The chatbot's development is done using Microsoft bot framework, which is using Microsoft cognitive service i.e. LUIS.ai for training language model which is used to identify intent of the user and fetch an appropriate response. The System makes the use of built in Artificial Intelligence to provide the answer go the query provided by LUIS.ai. The response from the system is given in an effective GUI manner such that it appears as if a real person is communicating. First, bot will transfer the query to the LUIS.ai. which further forwards its reply to the system after which it fetches for the value from the knowledge base supplying the answer to student question [3]?

Summary of related work:

Table -1:

System	Year	Advantages	Disadvantages
College Enquiry Chatbot	March 2107	1. Open Source 2. Effectively Categorize Products	1. Text-Only 2. No Self-Learning
Chatbot For College Management System	2017	1. Open Source 2. Use of Worldnet Dictionary 3. Automatically Prioritize	1. Text-Only 2. Time Efficiency

		Query1	
College Enquiry ChatBot Using Knowledge In Database	April 2018	1. Open Source 2. Uses LUIS.ai for training language model	No Self Learning New Language Delay

The systems studied here serves a stepping stone in developing our own comprehensive chatbot solution. From the case studies referred, a complete study of the existing chatbot system is undertaken. Here, the problem with 'college enquiry chatbot' is that if the query is not found in the database, the query is declared invalid. In our system, we will try to overcome it-whenver user submits a complaint, the query is detected. Then it is checked if the query resides in the database. If answer found in database, it is displayed on user's screen. If a particular query is not found in the database, a pop up asking for the user's email id comes. Such situation is handled by admin. Once he answered the question the answer is provided to that user. And that question -answer pair is stored in database so that whenever that question will be asked, the corresponding answer will be found in database and answer will be given as an output. Therefore, answers need not be entered manually by the admin anymore.

In case of 'chatbot for college management system', it is not integrated with text-to-speech and speech-to-text software's. Our system is a speech-recognized chatbot. Voice recognition will be accomplished using a library called Google Speech. Speech recognition can be broken into two groups with five subsets in total; speakers and speech styles. Speakers make up single speaking and speaking independently where speech styles include isolated word recognition, connected word recognition and continuous word recognition.

In addition to all this, our system will also comprise of chatlog section where the admin can see all the questions asked by the user. The questions unanswered can be viewed from this chatlog and the database is updated accordingly. The literature survey gave a brief overview of the systems already set in place. Comparisons between these systems gave a clarity and better understanding of our project. Advancement in technology has brought in many such systems which are already in place today, as briefed in literature survey. But all of these systems have their own set of limitations. Our system multimedia chatbot aims to provide a comprehensive solution to these limitations by yielding better accuracy in terms of performance and computational time compared to the preexisting systems and developing a scalable, robust and user-friendly web application.

3. DESIGN

The design of the system is represented in the pictorial format by the figure 3.1.

The Flowchart shows how the user will interact with this desktop application by entering their query in the text box provided on the front end of the system. Once they submit this query, the query will be checked in the knowledge base if the exactly same question is present. If present, the answer will be displayed on the screen. Otherwise, the query will be sent for pre-processing using NLP' POS-tagger technique. This pre-processed query is again checked in the database to provide the answer to the user. If the answer is not found, the query will be checked in the AIML.

There are situations where chatbot don't have a response. For such circumstances, a log file is created where all the queries asked by the user are added including the unanswered queries. Admin handles such situation. Admin will check for these unanswered queries and update the knowledge base with the answers to these queries so that whenever user enters the same query next time, he will get the answer. Also, when the answer to any query is unavailable, a window pop-up asking for the user's email id and phone number. When the admin updates the knowledge base, the response is sent to the user who asked the certain question to their entered email id. In this way, the chat bot accuracy is improved.

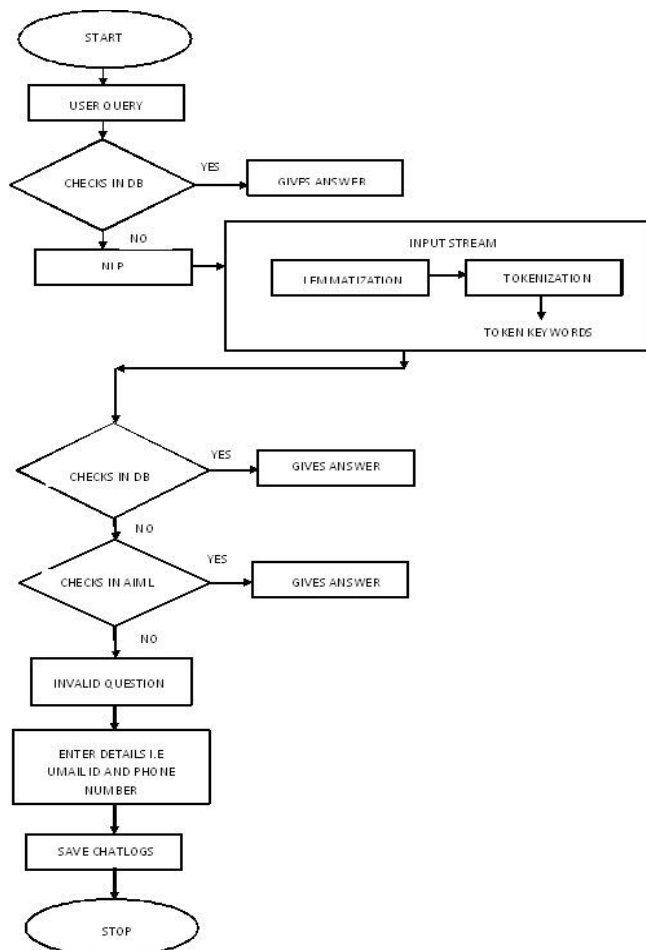


Fig -1: Flow chart depicting Navigation of the working of chatbot

4. IMPLEMENTATION

Multimedia chatbot converses with students. The query fired by the student is processed using NLP. The grammatical tagging algorithm is implemented or working of this proposed system is as follows:

Grammatical Tagging Algorithm:

INPUT: Sentence String (str)

OUTPUT: Response String (res)

STEP 1: Take query as input from user.

STEP 2: Check for response using rule-based approach.

STEP 3: If response string found.

STEP 3.1: Return response string.

STEP 3.2: STOP

STEP 4: Else

STEP 4.1: str.trim(); //removes extra spaces.

STEP 4.2: String g1[]=str.split(); //tokenizes the query

STEP 4.3: Repeat the following steps till g1.length();

STEP 4.3.1: String res=MaxenTagger.tagString(g1[]); //labels each token as nouns, verbs, adverbs, adjectives, etc.

STEP 4.4: If res.endsWith(NN,NNS,NNP) or If res.endsWith(VBG,VBD,VBN). // tags for noun and verb.

STEP 4.4.1: Match res with keywords in database.

STEP 5: If match found, return res string as an output to the user.

STEP 6: STOP

The front end for the chatbot is developed using jQuery. Java Swing is used for building the UI components. The knowledge base for the chatbot is developed using MySQL. The responses for the normal greetings are generated using the custom AIML files. The POS tagger is used for processing the queries. The appropriate answers are provided to the user in the form of text or voice.

Natural language processing (NLP) is a subfield of artificial intelligence, computer science which is concerned with the interaction between computer and human (natural languages).

In our system, as user input will be in English, to let machine understand this language we used Natural Language Processing.

The technique used in NLP is pos tagger. A Part-Of-Speech Tagger (POS Tagger) reads the text and allots parts of speech to every word, i.e., noun, verb, adjective, etc. Pos includes 36 tags. A part-of-speech tagger, or POS-tagger operates a series of words, and attaches a part of speech tag to each word. The given text is divided into tokens. Example: text = word tokenize ("When is Udaan conducted?") postag(text) [(('When', 'WRB'), ('is', 'VB')) (('Udaan', 'NN'), ('conducted', 'JJ'))] Once this is done, the words categorized are checked in database to match with the keywords. If the match is found, the corresponding answer is displayed. [4]

Artificial intelligence Markup Language is one of the first tools used to design ALICE. An XML based markup language, AIML is designed to create artificial intelligent application like chatbot. AIML has a data store which is used for answering questions which are not related to college. For example, "What is your name?". Basically, AIML includes responses to normal greetings. Using AIML we will create our custom AIML file. This is done by invoking the bot by using the writeAIMLFiles () function. A java program called AddAIML.java is created for this purpose. The program execution adds the custom entries to the bot's brain.

In java we will be using jQuery for creating user interface. Java Swing is used to design window-based applications. It is entirely written in java. Java Swing supplies platform-independent and lightweight components. The javax.swing package is used for providing classes for java swing API such as JButton, JTextField, JTextArea, JRadioButton, JCheckBox, and JMenu.

5. RESULT

The system has setup a chat-based application which would serve the users by answering any queries they might have regarding various daily college activities. Be it a simple question such as "Where is CSRoom" or a slightly complex one such as "Steps for issuing bonafide certificate".

The administrator panel is controlled by the admin who has total control of the system. Admin is able to update the knowledge with the latest details from time to time.

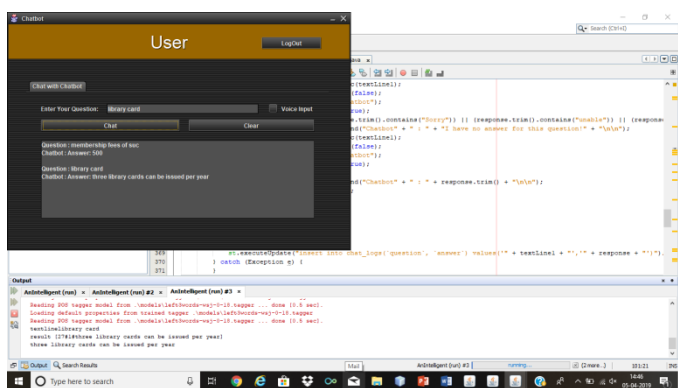


Fig 5.1 CHAT SCREEN

Fig 5.1 shows the chat screen which is used by the users to interact with the chatbot and ask any college related queries.

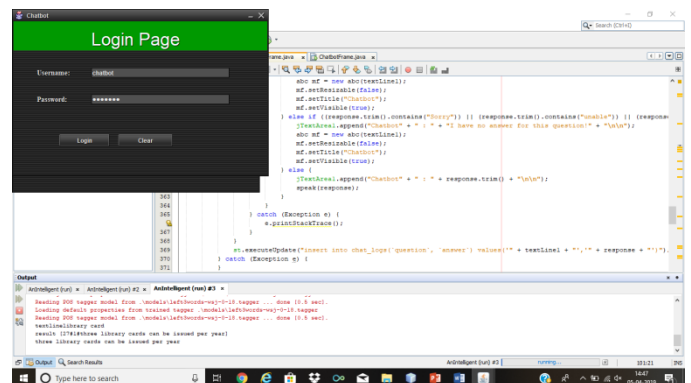


Fig 5.2 ADMIN LOGIN PAGE

Figure 5.2 shows the LoginPage where the admin logs in using credentials. By logging in, the admin can view the chatlogs, update the database, etc.

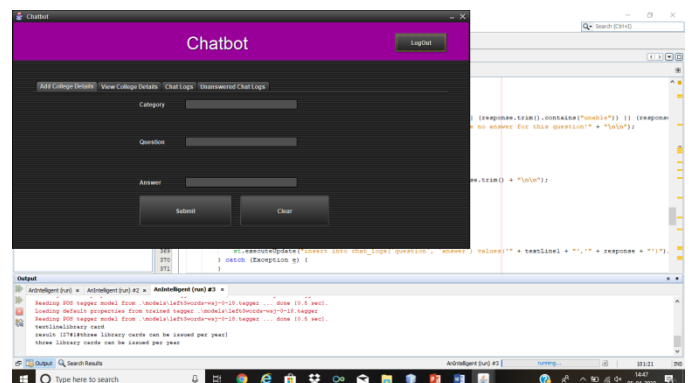


Fig 5.3. ADMIN PANEL FOR UPDATING QnA

Figure 5.3 shows the admin panel. once the admin is logged in, this panel is used to add, update, view the database and answered/unanswered chatlogs.

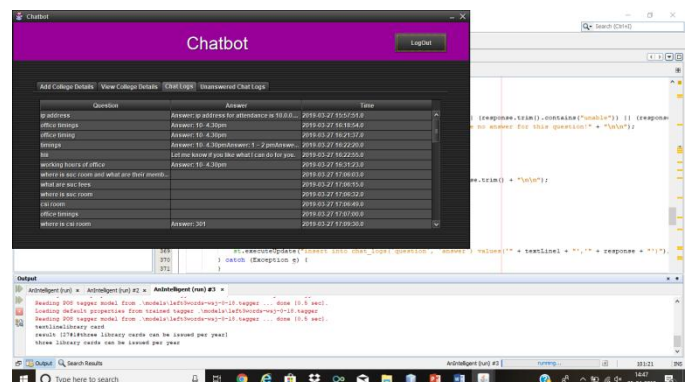


Fig 5.4 QnA

Fig 5.4 shows the Q n A section which shows all the questions and corresponding responses stored in database.

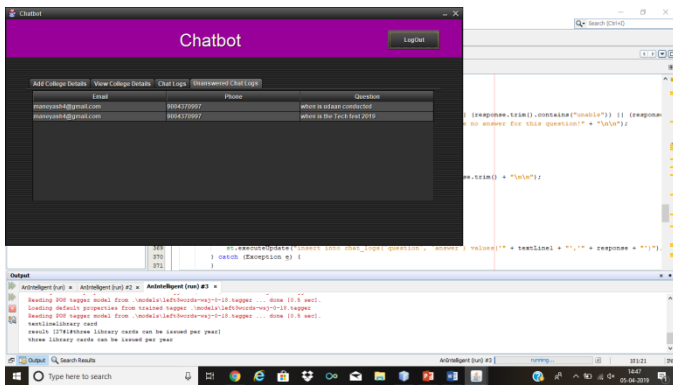


Fig 5.5 UNANSWERED CHAT LOGS

Figure 5.5 shows Unanswered chatlogs. This is the tab in admin panel where the admin can view the unanswered chatlogs.

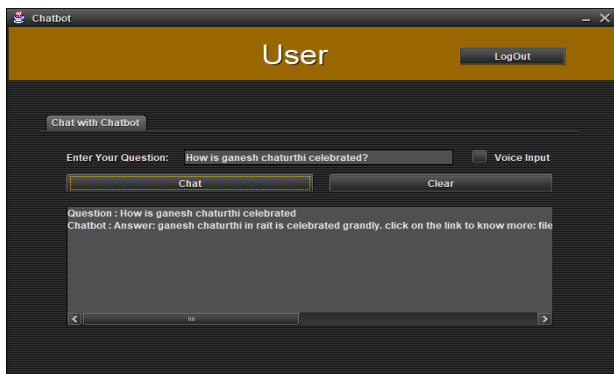


Fig 5.6(a). CHAT INTERFACE



Fig 5.6(b). ANSWER IN VIDEO FORMAT

Figure 5.6(a) shows the query asked by the student and the answer displayed in the form of link which when clicked shows the video for answering the query in 5.6(b). Also, it is the same for image answer.

6. FUTURE SCOPE

The system involves answering queries related to college. It holds a combination of decision tree and NLP approaches to better understand the user's intent to offer smoother

interactions. A future approach for this chatbot could be including recognition of language and also, it's translation facilities. Further, the system could also be designed for text analytics. Additionally, Content moderation could also be added so as to ensure that the system does not get misused.

7. CONCLUSION

Multimedia Chatbot is designed as an ultimate virtual assistant, helping to execute all kinds of interactions that are currently being performed by the college staff, circumventing all the unnecessary hassles faced by the students. Pos tagging is used to classify and solve the problem. A dynamic database is created using SQL which is capable of updating the knowledge base when the answer is not found. A more user accessible chat system using Artificial Intelligence is created by the introduction of a simpler input method using not only text but also voice, creating and catering for a more personal and convenient experience. The system's dialogue interface is being constantly perfected, so as to correctly interpret user's natural language and also for bettering their answering capabilities.

REFERENCES

- [1] College Enquiry Chatbot http://ijritcc.org/download/browse/Volume_5_Issues/March_17_Volume_5_Issue_3/1492062497_13-04-2017.pdf
- [2] Chatbot for College Management System https://www.researchgate.net/publication/327667155_Chat-Bot_For_College_Management_System_Using_AI
- [3] College Enquiry Chatbot Using Knowledge in Database <https://www.ijraset.com/files/serve.php?FID=15814>
- [4] POS tagging https://www.ijarcsse.com/docs/papers/Volume_7/4_April2017/V714-0_140.pdf