Chatbot: A User Service for College

Jayshree Ghorpade-Aher[1], Simran Kukreja[2], Tejal Karpe[3], Saloni Kakkad[4], Riddhi Kontamwar[5]

[1, 2, 3, 4, 5]*Dept. of Computer Engineering, MITCOE, Pune, India*

Abstract: A chatbot is a software that is used to interact between a computer and a human in natural language. The chatbot application provided in this paper helps the users to access the college related information along with the functionality of college map from any location with internet connection. This system reduces the work of college administration and the staff in case of answering all the basic queries of the students. The proposed system makes use of AIML along with NLP in order to provide responses to the user. It has made the use of client-server communication for providing faster responses to the users.
Keywords: College Chatbot, Natural Language processing (NLP), Artificial Intelligence Markup Language (AIML).

I. INTRODUCTION

A chatbot is an artificial intelligence (AI) software that can simulate a conversation with a user in natural language through messaging applications, websites, mobile apps, etc. It is often described as one of the most advanced and promising expressions of interaction between humans and machines. However, from a technological point of view, a chatbot only represents the natural evolution of a Question Answering system leveraging Natural Language Processing (NLP). Several applications of chatbot uses Artificial Intelligence Markup Language to chat with user. One of the primary goal of chatbot is to resemble an intelligent human and make it difficult for the receiver of the conversation to understand the real working of the internal application. The following facts are important during designing a Chatbot and are needed to be considered.

1) *Selection of Operating System and Platform:* Windows can be used. XAMPP and Tomcat is used as a server. Application is built in Eclipse.
2) *Creating Chatbot:* For creating chatbot, a program is written in Java Programming Language and Python is used for implementing NLP.
3) *Pattern Matching:* In computer science, pattern matching is the act of checking a given sequence of tokens for the presence of the constituents of some pattern. AIML (Artificial Intelligence Markup Language) is used for pattern matching.
4) *Conversation:* The conversation follows a basic English language and interacts in an easy to read manner.

In this system, the user can ask the questions related to any college-related activities through the chatbot without physically being available in the college for inquiry. The system analyses the question and then it responds to the user based on the match of the query. With the help of AIML, the system answers the query asked by the students. Hence, this chatbot helps the organizations to ensure quality service provision and customer satisfaction with less human efforts.

In this paper, working of chatbot is described thoroughly. The section II gives the Literature Review and section III explains the Proposed System of Chatbot. The section IV explains various algorithms used in the creation of chatbot. Lastly, the conclusion is given in section V.

II. LITERATURE REVIEW

"K.Bala" et al Question Answering system selects the most appropriate answers by using linguistic features available in natural language techniques. The proposed approach simplifies the chat-bots realization which uses two solutions. First one is the ontology, which is exploited in a twofold manner: to construct answers very actively as a result of an deduction process about the domain, and to automatically populate, off-line, the chat-bots KB with sentences that can be derived from the ontology, describing properties and relations between concepts involved in the dialogue. Second is to preprocess of sentences given by the user so that it can be reduced to a simpler structure that can be directed to existing queries of the chat-bots.

"Bayu Setiaji", et al. proposed the evaluation of sentence equivalence is completed with diagram that splits the input sentence in to two parts. The data of chatbot are deposited in the database. The database is appointed as information storage and predictor is used for storing the function and perform pattern matching. This application can be developed by using programming language of Pascal and Java. "Godson Michael Dsilva" et al. proposed the practical smart chatbot for customer care by using Software as a Service by analyzing the messages of each application server, users to check if it is vigorous or not. If its actionable then an automated Chatbot will initiates conversation with that user and help the user to resolve the issue by providing a human way interactions using LUIS and cognitive services. To provide a highly robust, scalable and extensible architecture, this system is implemented on AWS public

cloud. "Emanuela Haller" et al. gives method of recognizing the reality in texts and giving the past content for developing a conversation which is used in middle-school CSCL scenarios. The system take a plain text as input and output is a qualified user that is capable of answering all type of questions. The purpose is to provide a generic solution to this problem.

"Agnese Augello" et al. proposed Question Answering (QA) systems can be recognised as information accessing systems which try to answer to natural language queries by providing answers instead of providing the easy listing of le links. This system mainly chooses suitable answers by using attribute available in natural language techniques. The aim of this paper is to give helpful information regarding products which is useful for consumers to obtain what they want exactly.

"Ercan Canhasi" et al. Propose gives chatterbot which is based on AIML (Artificial Intelligent Markup Language) structure for training the model and uses Microsoft voice synthesizer for recognition of the word spoken by the user. Natural language processing used for understanding and Microsoft speech recognition is used in speech recognition and speech synthesis for speech to text and text to speech so people get along with it easily.

III. PROPOSED SYSTEM

1) *Train Common Questionnaires Dataset:* Initially the system can load the common questionnaires dataset and train the bot accordingly. The questions are stored in database under different tag categories like hostels, mess, groceries, etc.

a) *Tag wise Question and Response:* Whenever the user uses the system, firstly they need to mention their query. From the query the system can extract the tag word and then according to the question tag system gives the response to the user.

For example:

{"intents": [
{"tag": "hostel",
"patterns": ["Where is the hostel", "Are there any hostels nearby", "Hostel near me"],
"responses": ["Hostel is off campus as well as in the campus", "Yes, following is the list of hostels", "List of hostels"],
"context_set": ""
}]}

2) *Input Questions:* The user can interact with the system anytime and from anywhere. Through this module user can ask their queries to the system. The system performs pre-processing on the input query and then gives the response. The user can ask their college related common queries.

3) *Preprocessing:* The system can understand the users query by using Natural Language Processing (NLP). It can apply different techniques for understanding the sentence and word. They can analyze the query using two different ways like,

a) *Sentence Tokenization:* In this techniques system divides the sentence into several tokens. It splits the large raw text into several sentence to extract more meaningful information.

b) *POS Tagging:* This algorithm is used for detecting the parts of speech. It detects whether the word token is noun, verb or adjective.

c) *Word Tokenization:* In this technique the sentence or data is split into several words.

For eg.

i) "Where is atm near me".

ii) This sentence is split into words as follows,

iii) ['Where', 'is', 'atm', 'near', 'me'].

d) *Word Lemmatization:* Lemmatization is a more methodical way of converting all the grammatical forms of the root of the word. Lemmatization uses context and part of speech to determine the form of the word and applies different normalization rules for each part of speech to get the root word (lemma).

e) *Word Similarity:* By using this technique the system can find the similar words. WordNet dictionary is used for finding the synonyms.

WordNet Dictionary- WordNet is a combination of dictionary and thesaurus. It groups English words into sets of synonyms called 'synsets'. It provides short definitions and usage examples, and records a number of relations among these synonym sets or their members.

4) *Show Nearby Hostels:* User can search nearby hostels using this application. Application will track your current position and display nearby hostels based on current location. This makes use of the GPS system which is integrated in the bot itself.

5) **Show college map:** Users can view/search all nearby buildings in the campus. GPS system provides options to navigate the user through the campus with ease. For example, if user wants to search the library then he can do so by using this application.

IV. ALGORITHMS

A. *Natural Language Processing:*

We want a computer to communicate with users in natural language. we would not force users to learn a new language. This is particularly important for casual users and those users, such as managers and children, who have neither the time nor the inclination to learn new interaction skills. Many of the problems of AI arise in a very clear and explicit form in natural language processing and, thus, it is a good domain in which to experiment with general theories. So we are using NLP in our project.

Porter Stemmer Algorithm: It is a process for removing the common morphological and flexional endings from words in English. Following are the steps of this algorithm:

1) Gets rid of plurals and -ed or -ing suffixes
2) Turns terminal y to i when there is another vowel in the stem
3) Maps double suffixes to single ones: -ization, -ational, etc.
4) Deals with suffixes, -full, -ness etc.
5) Takes off -ant, -ence, etc.
6) Removes a final –e

B. *Artificial Intelligence Markup Language:*

AIML is an XML based markup language meant to create artificial intelligent applications. AIML makes it possible to create human interfaces while keeping the implementation simple to program, easy to understand and highly maintainable. AIML is a form of XML that defines rules for matching patterns and determining responses. It has class of data object called an AIML object which describes the behavior of computer programs. AIML consists of words and wildcard symbol *. It contains of units called topics and categories. There are 3 types of categories:

1) *Atomic category:* Consists of patterns which has no wildcard symbols
2) *Default category:* Consists of Patterns who have some wildcards. These match any input.
3) *Recursive category:* Contains <srai> tag. <srai> is a multipurpose tag, used to call/match the other categories.

a) *Basic Form of AIML*

<category>

<pattern> What are the facilities provided by the college</pattern>

<template> The college provides gym, indoor and outdoor games</template>

</category>

b) *Tags in AIML*

<aiml> tag: Each AIML file begins with this tag and is closed by </aiml> tag

<category> tag: The basic units of an AIML dialog are called categories. Each category is a fundamental unit of knowledge. A category consists of (i)an user input, (ii) a response to user input, presented by the chatbot, and (iii) an optional context.

<pattern> tag: This tag contains a possible user input. There is just a single <pattern> in each <category> tag, and it must be the first element to be set. Also, the words are separated by single spaces, and wildcards can replace parts of a sentence.

<template> tag: This tag contains possible chatbot answers to the user. It must be within the scope of a <category> tag, and be placed after the <pattern> tag. Most of the chatbot information is bounded by this element

This work chooses the chatbots based on AIML. The motivation is the fact that the usage of Pattern Recognition, in conjunction with AIML, presents features such as:

i) Easy of implementation, since AIML is a XML (eXtensible Markup Language) based markup language, the tags make the implementation of dialogues easier.

ii) There are computational systems that help developers in the chatbot codes creation and Web deployment for users access.

iii) There is a high level of reuse, since a significant amount of chatbot projects is developed under a free/open software license.

Steps in building an AIML rule-based chatbot

1. Install AIML modules.
2. Create a standard startup file
3. Creating AIML Files
4. Including random responses in AIML files.
5. Write a python program with brain module for faster response.

V. WORKFLOW

The below diagram shows the processing of input query and how that query is identified according to categories by the use of AIML. It finds for a match of the query in the template that is the database and then retrieves the response which is then sent as an answer to the user. If there is no answer for the input query, then it sends the default response to the user.
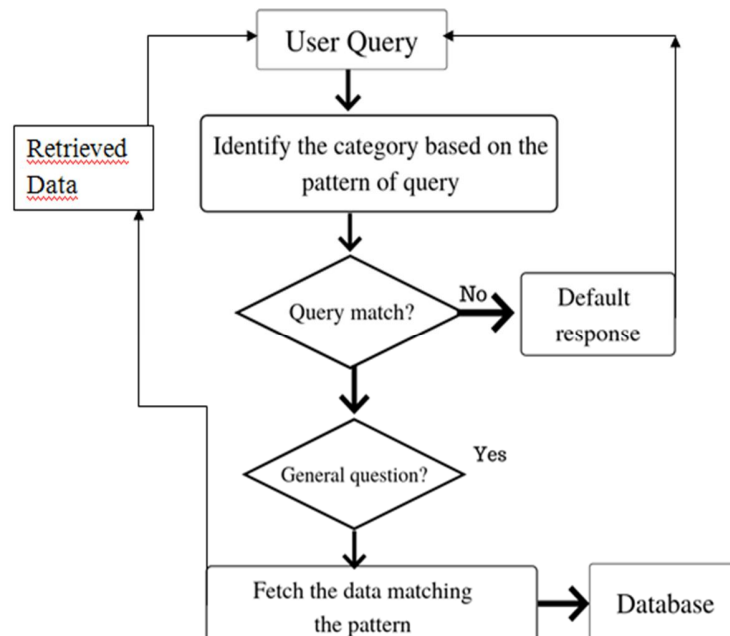

Figure 1: Query processing workflow

VI. RESULTS
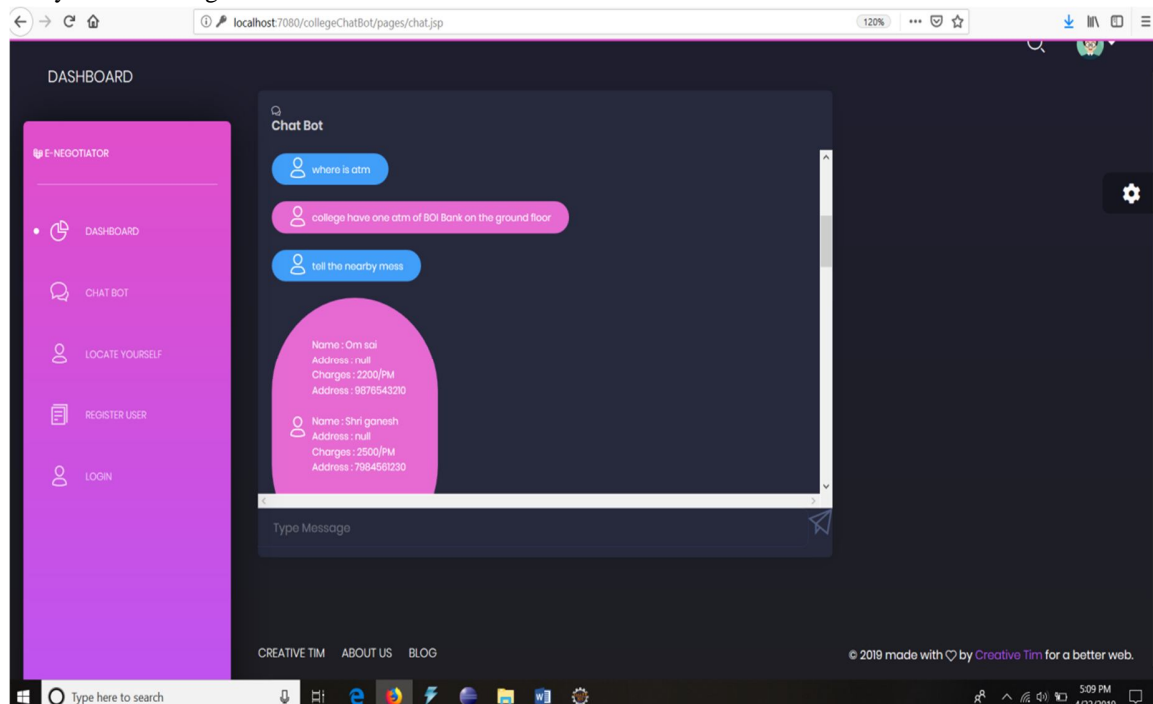
The proposed systems working is as shown below:
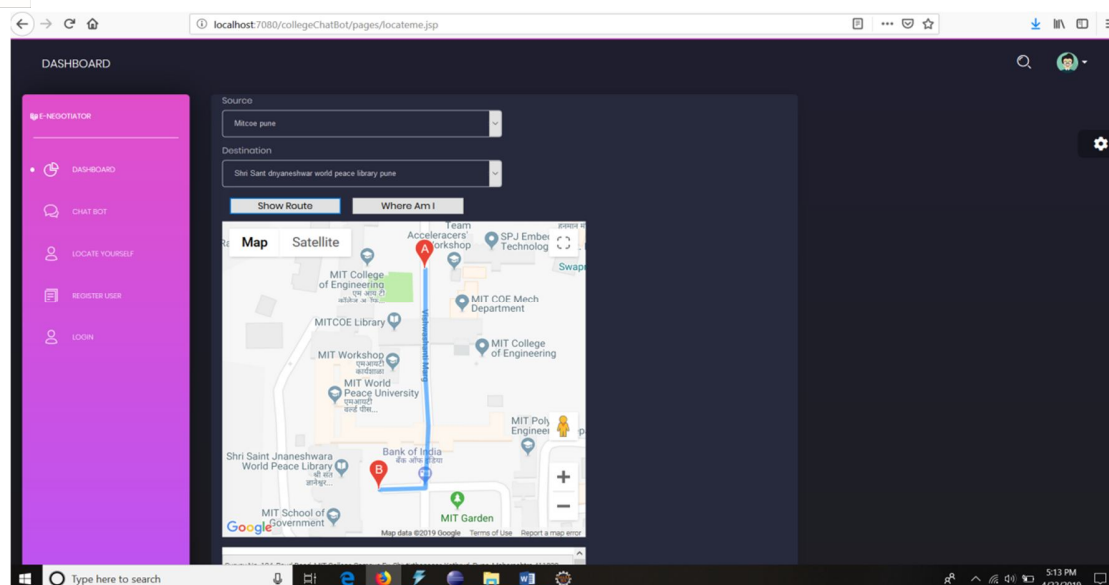

Figure 2: Chatbot UI

Figure 3: Map functionality

VII. CONCLUSION

Hence, it can be analyzed that the development and improvement of chatbot design grow at an unpredictable rate due to different methods and approaches used to design a chatbot. They help us by saving time to answer the questions of an user efficiently. In the proposed system the client-server communication enhances the speed of chat between user and the bot. Since there are many different designs and approaches for creating a chatbot, researchers need to interact and must agree on a common approach for designing a Chatbot and utilizing it in various fields.

REFERENCES

[1] M. Rahman, A. A. Mamun and A. Islam, "Programming challenges of chatbot: Current and future prospective," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017,pp.75-78. doi: 10.1109/R10-HTC.2017.8288910

[2] B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling," 2016, 7th International Conference on Intelligent Systems, Modeling and Simulation (ISMS), Bangkok, pp. 72-77.

[3] Quote, Riza Berkan founder of exClone Inc., https://emerj.com/ai-podcast-interviews/when-will-chatbots-reach-human-level-sophistication/

[4] B. R. Ranoliya, N. Raghuwanshi and S. Singh, "Chatbot for university related FAQs," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi,2017,pp.1525-1530.doi: 10.1109/ICACCI. 2017.8126057

[5] E. Haller and T. Rebedea, "Designing a Chatbot that simulates a Historical Figure," 2013 19th International Conference on Control Systems and Computer Science, Bucharest, 2013, pp. 582-58

[6] G. M. D'silva, S. Thakare, S. More and J. Kuriakose, "Real world smart chatbot for customer care using a software as a service (SaaS) architecture," 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), Palladam, 2017, pp. 658-664

[7] A. Augello, G. Pilato, A. Machi and S. Gaglio, "An Approach to Enhance Chatbot Semantic Power and Maintainability: Experiences within the FRASI Project," 2012 IEEE Sixth International Conference on Semantic Computing, Palermo, pp. 186-193

[8] M. C. Surabhi, "Natural language processing future," 2013 International Conference on Optical Imaging Sensor and Security (ICOSS), Coimbatore, 2013, pp. 1-3.

[9] H. N. Io and C. B. Lee, "Chatbots and conversational agents: A bibliometric analysis," 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore,2017,pp.215-219. doi: 10.1109/IEEM.2017.8289883

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)