INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Impact Factor: 7.542

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Admission Enquiry Chatbot for College Website

Rutuja Labba, Rudvi Manjrekar, Shweta Mishra, Arathi B

UG Student, Dept. of C.E., SIES Graduate School of Technology, Navi Mumbai, India

Assistant Professor, Dept. of C.E., SIES Graduate School of Technology, Navi Mumbai, India

ABSTRACT: Students have a lot of questions regarding the admission process and answering these manually can be a strenuous task for the admission cell. Employing chatbots in college's website supersedes the manual toil. College Enquiry Chat bot is for providing answers to students' enquiries during the admission process. It will be a hybrid model incorporating two approaches sequentially, first by using the AIML language and then by using the ML algorithm. Data retrieved from the college authorities about the queries which majorly contain the FAQs from past years are used for pattern matching in AIML as well as it will be stored in a database. Most probably the user's query will be satisfied by the prompts given by our system but in case where the user feels he/she is not satisfied with the prompts, then they can give their query as input and the system will fetch the answer for the query using ML algorithm. The query goes through a data cleaning process using NLP. The system provides appropriate answers to the queries by predicting the intent of the query as recognised from the trained data stored in the system and printing the answer retrieved for the corresponding intent.

KEYWORDS: Chatbot; AIML; Machine Learning; NLP; Logistic Regression; Multinomial Naive Bayes

I. INTRODUCTION

Chatbots are nowadays so common that almost every other website has a chatbot. There are types in it which are Rule based, Retrieval based or Generative based chatbots. These can be developed by AIML, NLP Algorithms, Deep Learning techniques or by Machine learning algorithms. Dialog Flow, Pandorabots, Rasa stack and Chatterbot are a few frameworks for developing chatbots. Since we are developing a chatbot for the college needed for admission enquiry which is need specific domain, we will be starting of with AIML which provides query prompts and incorporating ML algorithm in the latter part for when the user is not satisfied with the prompts, where they have to type the query and it returns a reply based on the intent predicted by the algorithm.

The motivation was that during the admission process, students have a lot of doubts related to the college and facilities provided by it. These queries are answered by the admission cell or office. The people think it is better to visit the college for enquiry rather than searching for the information on the website. The ratio of students to the faculties is large and hence it becomes arduous to revert to each one in appropriate time. It might happen that some questions are lost or remain unanswered in the sea of queries or there might be a delay in replying. Designing a chatbot would decrease the load on faculties significantly and students would get their queries solved without any latency..

The scope of the chatbot that we are designing should be efficient enough to fulfill its purpose of admission enquiry. The system should satiate the queries asked by users with proper responses, and if the response does not exist for a particular query, then it should be notified to the administrator so that the data is up-to-date. Once it's completely developed, it should be deployed on our college website for potential students of our college to use and clarify their doubts.

II. RELATED WORK

In [1] a chatbot developed with the NLP approach. They have taken user query as an input in a single sentence form, tokenized the sentence and used cosine similarity or bi-gram algorithm for finding similarity score of sentences. This chatbot will only take as input or accept questions asked without a break and answer accordingly. Multiline questions with newline or tab characters in between lead to misleading answers. So, whatever the user's query is, it will consider only the portion before the tab or newline for generating the answer. In [2] a chatbot developed by using AIML as well as NLP which follows both rule and retrieval-based approach. In [3] a chatbot developed using the knowledge base of ALICE, AIML or plain pattern matching using DFS algorithm. In [4], the implementation of a chatbot for college enquiry system which is developed using the knowledge base of ALICE bot is studied. This chatbot needs queries to be specified in a particular pattern. Input in any other way of sentence formation than the one which is specified in the

database gives a default error reply. In [5] a chatbot developed using the AI algorithm for pattern matching and Java Applets for the Graphical interface. In this type of chatbots, since it considers each sentence as a new query it does not save context. So every query which the user asks should contain the context in it and also should have all the necessary keywords for generating the proper output. In [6] a chatbot developed by using AIML as well as LSA where template based and greetings are answered by AIML and service based used LSA to satisfy the user's query. In [7] a chatbot developed by using the existing bot framework like Azure bot. This chatbot includes an additional feature of sentiment analysis along with active learning. It has lack of flexibility while using a Chatbot framework. In [8] a chatbot developed using plain pattern matching or keyword matching.The chatbot uses codeigniter which is a PHP framework. The queries, keywords and weight related to them, and their answers are stored in the database. It has lack of flexibility while using query prompts. In [9] a chatbot developed with NLP approach using NLTK library, TF-IDF for vectorization and cosine similarity for finding similarity score of sentences.

III. PROPOSED ALGORITHM

A large portion of the chatbots keep either a rule-based approach or retrieval-based approach whereas this chatbot is utilizing both which is giving greater adaptability to users making it easy to work with and user friendly. Increasing the odds for the client to get their query answered.

In the retrieval-based part of the chatbot we are utilizing two calculations Multinomial Naive Bayes and Logistic Regression rather than a solitary calculation expanding the efficiency of the chatbot and keeping away from bogus outcomes to be returned.

The queries which were unanswered due to insufficient data are stored and sent through email to the administrator so that it can be updated in the database for future use.

We have used AIML for matching the pattern of queries. If this method is not sufficient, we perform intent classification by Multinomial Naive Bayes and Logistic Regression.

The input text is cleaned using regular expression, stopwords are removed using NLTK English Stopwords Corpus and lemmatization is done with the help of WordNetLemmatizer. Count Vectorizer is used to tokenize the sentences in the dataset and build a vocabulary of keywords out of it. TF-IDF is the term frequency into inverse document frequency which is used to assign a numerical value to the keywords in vocabulary. Pipeline is used to chain together the Count Vectorizer, TF-IDF transformer and ML classifier to automate the process.

We have used the MultinomialNB classifier and Logistic Regression classifier from python scikit-learn library. The fit and predict functions are used to train a dataset and predict the intent based on trained data. The function predict_proba is used to check whether it returns a desirable probability.
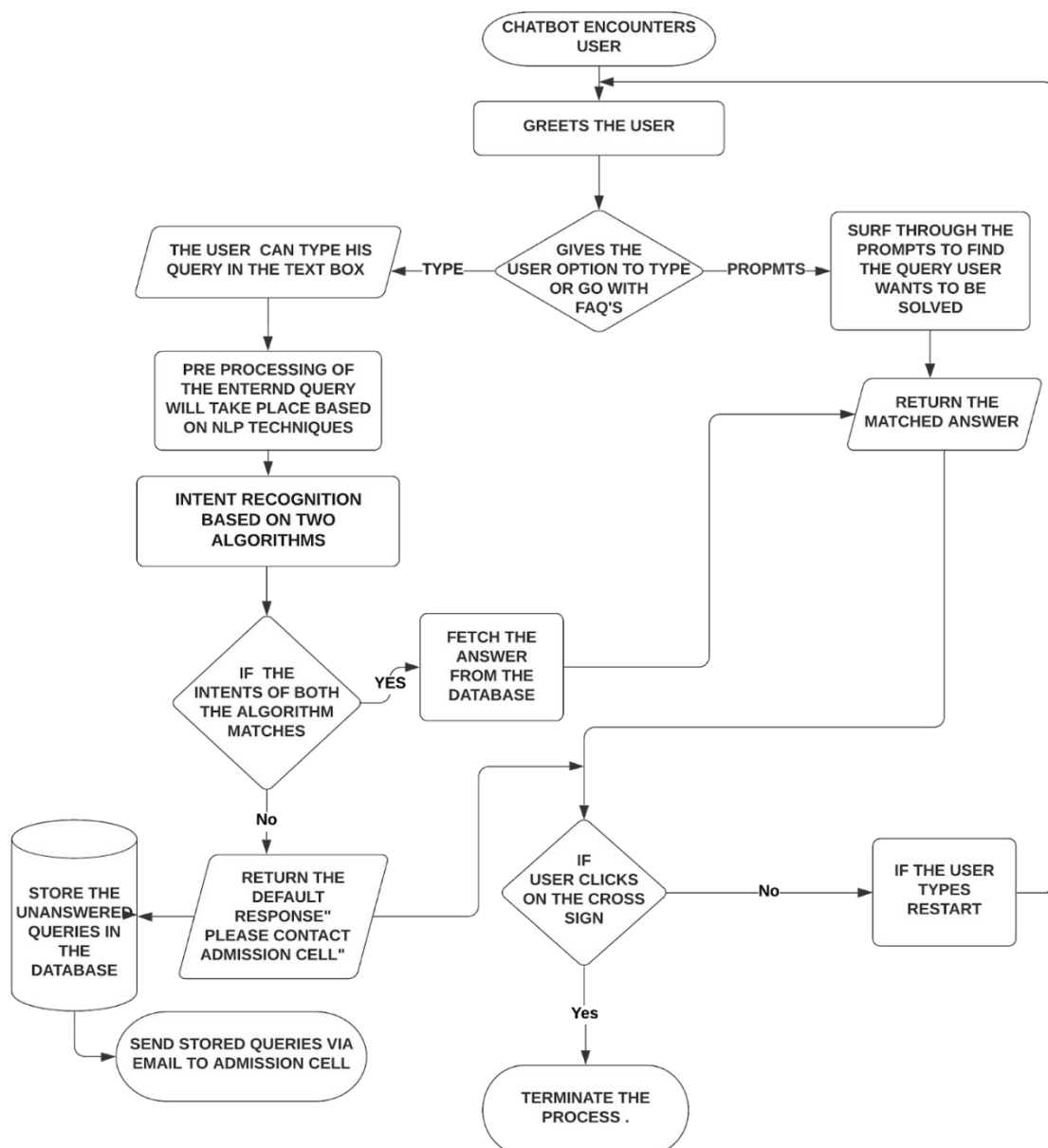
The smtplib and email libraries are used to send the email to the admission cell with the content containing the unanswered questions of the users.

IV. METHODOLOGY

The Methodology is shown using a flowchart above, it states that as soon as the chatbot encounters the user it performs the following steps.

- Greets the user and then give option based on the FAQ'S
- The user will then select from the options i.e., the user will select either to type his query or to surf through the prompts.
- If the user selects to surf through the prompts and chooses a query from the options provided the chatbot will give a response based on the AIML (Artificial Intelligence Markup Language) code written for the rule-based part of the chatbot which will lead to the output.
- If the user selects to type his query as the input text is given
  - The entered text input will be processed as:
    - Special characters and digits are removed from the query.
    - Stopwords are removed from the query.
    - Tokenization is the term used to describe the process of converting the normal text strings into a list of tokens.
    - Lemmatization is the process in which grouping of different inflected forms of a word is done. It also roots several words into one common root but the output is a proper word.
  - The cleaned text will then be used and intent recognition will take place based on two algorithms
    - Multinomial Naive Bayes
    - Logistic regression

o The process checks that the number of keywords should be greater than one. The intent predicted from these two algorithms are checked if equal or not. It also checks whether there are sufficient unanswered questions in the database to be sent through email to the administrator.

o If intent recognised by both the algorithm matches output will be given as the answer declared against that intent in the database.

o If the intent recognised by both the algorithm does not match the question will be stored in the database for improvement and will return as "please provide more information" or "Try contacting the admission cell for your doubt to be solved.
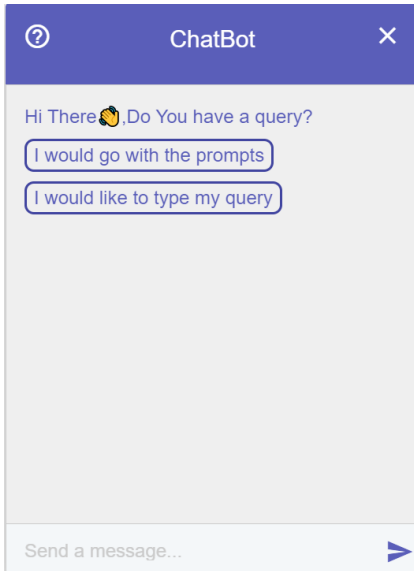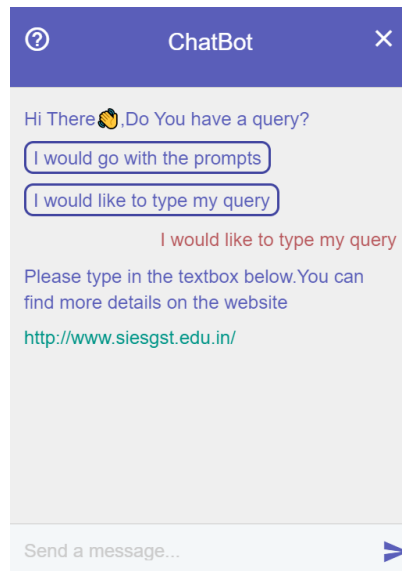
V. RESULTS

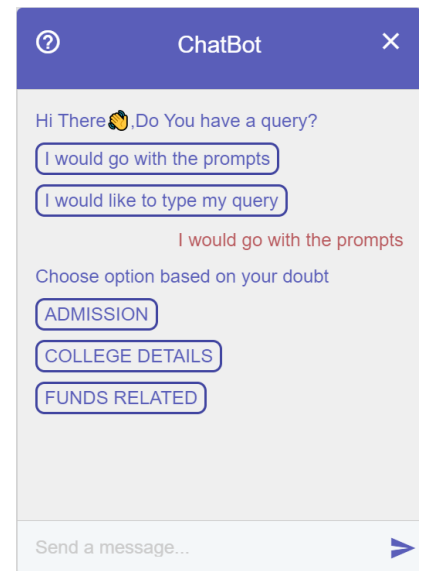Fig.1. Chat Window 1

Fig.2. Chat Window 2

Fig.3. Chat Window 3

The above image shows the chat window present during interaction. The user on opening the screen receives a greeting with two prompts for further processing. The further directions are shown when the prompt is clicked. The question mark icon on hover gives help to user on the working. The cross-sign icon is for exiting the chat window.

	Precision	Recall	f1-score	support
accuracy	-	-	0.78	80
macro-average	0.53	0.52	0.51	80
weighted average	0.86	0.78	0.79	80

Table 1. Classification Report - Multinomial Naive Bayes

	Precision	Recall	f1-score	support
accuracy	-	-	0.89	80
macro-average	0.81	0.77	0.76	80
weighted average	0.94	0.89	0.89	80

Table 2. Classification Report - Logistic Regression

VI. CONCLUSION AND FUTURE WORK

The chatbot on greeting, will provide query prompts to the user which most probably will answer the users doubt. In cases where the user is not satisfied with the prompts, they can send the query as a typed text input to the chatbot. A database is maintained that stores the unanswered queries for the admission cell to update the same. For future scope, the chatbot can be integrated with voice features so that users can interact more easily with the chatbot. The chatbot has a domain specific topic that is to provide solutions to only admission related queries. The scope can be increased such that the students that are already admitted to the college can also use it for daily updates.

REFERENCES

1. Sagar Pawar, Omkar Rane, Ojas Wankhede, Pradnya Mehta, "A Web Based College Enquiry Chatbot with Results", International Journal of Innovative Research in Science, April 2018
2. Karanvir Singh Pathania, "College Enquiry Chat Bot", May 2019
3. Drishti Malik Vibhor Sharma, Monika Goyal, "An intelligent behaviour shown by chatbot system", International Journal of New Technology and Research, 2017.
4. Balbir Singh Bani, Ajay Pratap Singh, "College Enquiry Chatbot Using A.L.I.C.E", International Journal of New Technology and Research 2017.
5. Ms Menal Dahiya, "A Tool of Conversation: Chatbot", International Journal of Computer Sciences and Engineering 2017.
6. B. R. Ranoliya, N. Raghuwanshi and S. Singh, "Chatbot for university related FAQs", 2017 International Conference on Advances in Computing, Communications and Informatics.
7. Ujaliben Kalpesh Bavishi, "Implementing a College Enquiry Chatbot", Spring 2019
8. Ms.Ch.Lavanya Susanna, R.Pratyusha, P.Swathi, P.Rishi Krishna, V.Sai Pradeep, "College Enquiry Chatbot" , International Research Journal of Engineering and Technology, 2020
9. TarunLalwani, Shashank Bhalotia, Ashish Pal, Shreya Bisen, Vasundhara Rathod, "Implementation of a Chatbot System using AI and NLP", International Journal of Innovative Research in Computer Science & Technology, May 2018
10. ParulPandey, "Building a Simple Chatbot from Scratch in Python",Analytics Vidhya, September 17,2018. [Online].
11. N. P. Patel, D. R. Parikh, D. A. Patel and R. R. Patel, "AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)
12. S. Das and E. Kumar, "Determining Accuracy of Chatbot by applying Algorithm Design and Defined process," 2018 4th International Conference on Computing Communication and Automation (ICCCA)

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING