Design a Chatbot for Chennai Corporation Using Logistic Regression Algorithm

Jeberson Retna Raj, Salman, Senduru Srinivasulu

*Abstract: Chatbots are the famous nowadays in business because of its service offered to the community at large. They provide support of 24*7 for business in terms of customercare, helpline, planning, analyzing and decision making. In this paper, a chatbot for Chennai corporation is proposed. This chatbot helps the citizens in providing the responses for their queries related to civic problems. There is no such system is available to handle the public grievances automatically. This system handles the public query and the relevant suggestion and responses will be given promptly. The chatbot receives the text or voice input and processed. The voice recognition module used to recognize the voice query and the voice to text convertor used to convert the voice data into text format. The matchmaking process used to match the input query with the available data set and the relevant responses is generated. If no match is for the query, the matchmaker will find the relevant response from online sources. The output channel equipped with the text to voice converter which converts the text data into voice and it will be delivered to the end user. The naïve bayers and logistic regression algorithm is implemented for classifying the query and the performance is compared. The result shows that the logistic regression algorithm outperform well with the precision and recall values.*

Keywords :Chatbot, NLP, Logistic Regression.

I. INTRODUCTION

Intelligent systems are very famous nowadays which helps the human community for reducing the manual work. Chatbot is familiar in business organization for customer support, planning, analyzing and decision support. Chatbot for text and voice are extensively used in 24*7 customer support for business houses. These intelligent systems are combined with Artificial Intelligence and Machine learning techniques. Cognitive computing is familiar in businesses for training the computer systems the same way human think, understand and respond. Human Computer Interaction focuses the design process of the how the human and machine interaction can be done with ease of use. Companies implemented their own Chatbot applications to meet their business needs. For example, the Chatbot applications of Apple's siri, IBM Watson, Google go, and the humanoid Sofioa. All these systems are working based on the cutting edge technology which includes AI and ML. Chennai is one of the modern metropolitan city in India for its culture and

tradition. The city municipal is divided into more than 200 zones and it is well managed by the city administration. The municipal system works based on the day to day needs of the city and the people to extend the services. The civic problems like drainage, street light, roads, health, drinking water, sanitization, garbage etc., are received manually and the respective department responded with the service. This is a time consuming process which may delay the service to the general public. There is no such automated grievance redressal system currently available in Chennai municipal to assist the public complains. To alleviate these difficulties, an automated Chatbot is introduced to assist the city residents. The system is equipped with a well-designed interface for posting the complaints and through phone one can able to get proper assistance and guidance to their grievances. The public can register their queries through text as well as voice to the system. The system receives the input of text or voice data and it will be processed. If it is the text input, the system has sent it to the matchmaking process and the relevant response is given as an output. If the input is voice data, the voice recognition system process the input and the voice-to-text converter component will convert the voice data into text. Furthermore, the converted text input will be forwarded to the matchmaking module for finding the appropriate responses. The naïve bayes classification algorithm is used to classify the input query and the appropriate responses will be given. The response will be again provided the reverse process of input, the text-to-voice component used to transfer the voice data into text and finally the response will be delivered to the user. Initially, the system is trained with the input query of English and later it can be extended to the Tamil language. The no matching query will be treated as a new query and it will be stored into the database separately for training the appropriate responses later.

A. Problem statement

The chatbot system is not known to people who are not more into the technology. Even if there exist a chatbot system, it is not much accurate in proving the answer or solutions. Users need to manually visit to the organization to get their queries answered by the institution help desk. This process consumes lot of time as well as money as the customer needed to visit college if its miles away from home. Also, this process may lead to communication gap between student and college.

Revised Version Manuscript Received on 10 September, 2019.
  JebersonRetna Raj*, Department of Information Technology, School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India.(Email: jebersonretnarajr@gmail.com)
  Salman, Department of Information Technology, School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India.
  SenduruSrinivasulu, Department of Information Technology, School of Computing, Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India

B. Objectives

The Chatbot application can be designed using the following objectives:

a) To design a Chatbot application for Chennai corporation for handling both text and voice queries

b) To evolve a matchmaking framework equipped with artificial intelligence and machine learning to relate the input request and its appropriate responses.

c) To introduce the Natural Language Processing technique to process the input request and generate appropriate responses.

C. Literature survey

Numerous system has been presented in literature regarding chatbot application. Weizenbaum J et al presented a chatbot ELIZA and the mental issues faced with the system is described [1]. Thomas et al presented a route direction bot for showing way in a single layout [2]. Rashmi et al implemented a inquisitive chatbot for providing addresses for a complex queries. The missing data is predicted for improving the accuracy[3]. Siddharth Gupta presented a collaborator based chatbot for promoting product in a e-commerce based website. The chatbot recommend the relevant items to the customer[4]. The other chatbot applications like Co-Sense, SearchTogether, Coagmento, CollabSearch are used to identify the potential customer segment and product trends in the market. They help the community users for reviews, critics and social support [5]. Efron et al designed a bot for classifying the queries of scientific categories in Twitter and naturally responded with the available data. SearchBuddies intended for creating indexed list for facebookqueries[6].

II. METHODOLOGY

The Chatbot for Chennai municipal grievance redressal system can be implemented with the help of Artificial Intelligence and Machine Learning techniques. The system receives input text / voice data as input and processes it. The speech recognition module employed with the voice-to-text converter will convert the voice input into text data and forwarded to matchmaking component. In matchmaking process, the input text query is mapped with the trained responses using keyword based searching and matching. The machine learning algorithms used to test the input data with the trained responses and identified the suitable response for an input query. The matchmaking manager retrieved the matched responses and provided as a response statement to the output channel. The output component again converts the text output into the voice command with the help of text-to-voice response. The detailed architecture of the proposed system is shown in figure 1.
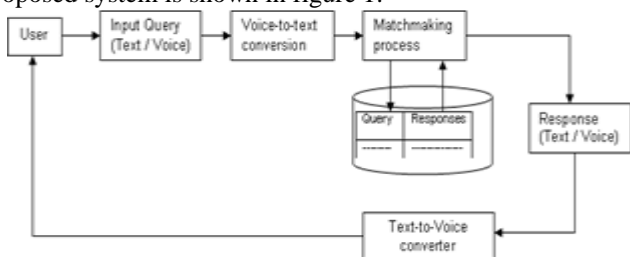

Figure 1: Proposed architecture of Chatbot

The proposed Chatbot has the following modules:

Step-1: Train the dataset
Step-2: Reading the input data (text/voice)
Step-3: Speech recognition process
Step-4:Conversion of voice-to-text using Natural Language Processing
Step-5: Matchmaking of input query with the search the relevant response
Step-6: Conversion of text-to-voice data based on the nature of the query
Step-7: Deliver the responses
Step-2 to step-7 repeatedly continues until no more queries

A. Train the dataset

Chatbot is the technology where they need to fetch the results from the databases as they have trained. For fast and the great accuracy response we need to create the dataset according to the problem domain. These datasets perform major role to generate the response for the user. Accuracy will be increased if training of dataset is more and more. There is a recommendation that creating a dataset for chatbot, always create YML file for training dataset. The civic grievances related responses which includes help line numbers, point of contact, remedies etc.

After creating the datasets, chatbot need to be train then only they can generates the response for the user's query.

YML file which consists of lot of conversations according to the problem which the chatbot need to be train to produce the response. Now, again Accuracy will be increased if training of dataset is more and more.

B. Reading the input query

This is the main stage where the chatbot is ready to answer for the problem domain as it has trained. User will generate the query in the form of text or speech etc. This step takes an input from source like console, API, voice, etc. It will get input a query from the source and store in the variable "query" and preprocessing will be done to convert the voice data into text data.

C. Speech recognition process

If the input is voice data, the speech recognition component synthesis the voice and the noise will be remove by the suitable preprocessing algorithm.

D. Conversion of voice into text using Natural Language Processing

The Natural Language Processor (NLP) which receives the input voice data and appropriate text data will be generated. The nltk tool kit of python is used to process the input data. The detailed flow diagram is shown in figure 2.
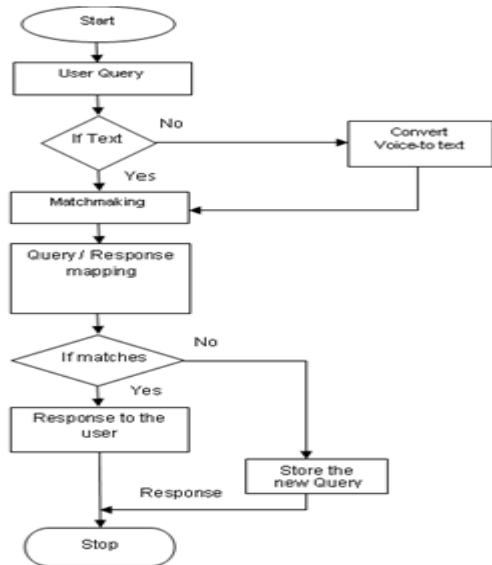
Figure 2 : Flow diagram of Chatbot

E. Matchmaking of input query with the relevant response

The matchmaking process employed with the help of machine learning algorithms. The query and relevant responses are trained with the help of listener library in python. The matchmaking process has dived into searching and matching process. The searching process takes the user query as input and it performed keyword based matching. The matching process used to identify the relevant matching responses in the database. After converting the query into text now the Bot is ready to answer it. Searching is the most important form of data structures. Search is a crucial part of how a chatbot quickly fetches the desired response. There are some examples that helps the chatbot to select a response include the following generalities:

a) The Equivalence of an input query to known query
b) Frequency in which the same known reactions occur
c) The probability of an input statement to fit a range that is a part of a known statement

The logic adapter in ChatterBot library use naïve Bayesian classification algorithm to determine if an input query meets a particular criteria or not.

F. Conversion of text-to-voice

The text-to-voice response converter will convert the responses into voice and it will be delivered to the target user. If the query has no matching with the responses then the query will be stored in a separate database for train the future responses for the new query.

G. Deliver the response

Final task of chatbot is to show the response to user with more accuracy and the same format that user sent the query. Before giving the response to the user there are some need to be changes according to the nature of query, if query is voice that the given response which is in the form of text need to be converted into voice as well in this phase.
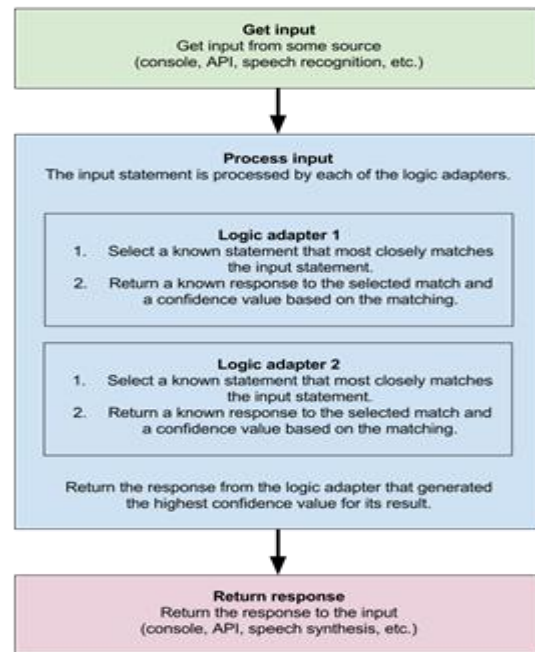
Process interaction of Chatbot

H. Train the Chatbot

Generally, Training is the most import part of chatbot without the training of data chatbot will not able to generate response. A given list of conversations files will be used to train the chatbot. The file which are going to train will be consists of conversation against the human problems and according to the problem domain.

For example of training dataset:
Class: greeting
"hello"
"how are you "
"good morning"
"good evening"
For some input sentence classification:
Input: "hi good morning"
"hi" (no matches)
"good" (class: greeting)
"morning" (class: greeting)

I. CHATTERBOT

Chatterbot is third party library which help to implements the Chabot easily. Searching is the most important form of data structures. Search is a crucial part of how a chatbot quickly fetches the desired response.

```
pseudocodechatbot(query,db.sqlite)
{
        qr="t"
        if (type(query)!=str)
        {
        query=recogniser(query,language="en-
in")
                qr="s"
        }
        while (true)
        {
                if (query!="bye")
                {
        result=get_response(db.sqlite)
                }
                else
                {
                        print("bye")
                        break
                }
        }
        if (qr=="s")
        {
        speak(result)
        }
        else
        {
        print(result)
        }
}
```

III.EXPERIMENTAL RESULTS

The prototype of Chabot for Chennai Corporation is implemented using python language with the supporting libraries. The system receives the input query through text and voice. The input query is preprocessed and the stemming algorithms removed stop words and unnecessary white spaces in the query. The voice query is processed by the voice recognition system and it used to identify the human voice and captures the voice data. If the input query is voice, the voice to text conversion process used to convert the voice data into text. Thereafter, the preprocessing algorithm applied with the text data. The naïve bayers classification algorithm used to classify the query and the matchmaking process helps to identify the responses from the database. If the response is not available in the database, the matchmaking engine searches the responses in Wikipedia sources for the right response. If the query matches with the trained responses in the database, it will be forwarded to the output channel for delivering to the end user. The output channel contains the converter which converts the input text into voice and delivered to the user. If the input query is text, the matchmaker identified the relevant response and directly it will be delivered to end users. The sample query interface and the response is shown in figure.
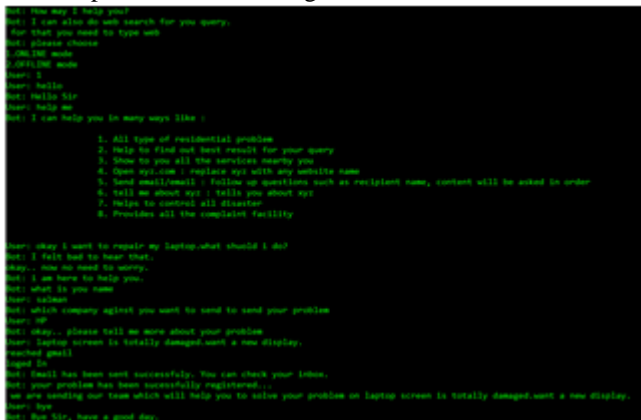

Figure: Chabot query interface and its responses

The proposed system compared with other algorithm with logistic regression and the performance is shown in table 1

Algorithm	Precision (%)	Recall (%)
Naïve bayes	80	75
Logistic regression	90	87

In this application, the logistic regression outperforms the naïve bayers algorithms. The sharp difference with the logistic regression is due to the missing data prediction. This additional feature further increases the accuracy of the system handling the query and its responses.

IV.CONCLUSION

The Chatbot application is successfully implemented for Chennai Corporation to handle the public query. It can work in 24*7 without any manual intervention. The general public query is processed and the relevant responses are generated to the immediate need of the user. The Chabot provides guidance to the user based on the trained responses. If no match of the query it will be stored in the database for training the system in future. Both text and voice query has been successfully handled the relevant responses are delivered to the user. The system further extended to support the registration of public grievances and complains in their locality. The system classifies the complaints and sends to the respective departments automatically. This will reduce the gap of time consumption for providing better service to the public.

REFERENCES

1. J. Weizenbaum, "Elizaa computer program for the study of natural language communication between man and machine," Communications of the ACM, vol. 9, no. 1, pp. 36-45, 1966.
2. N. Thomas, "An e-business chatbot using aiml and lsa," in Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on. IEEE, 2016, pp. 2740-2742.
3. S. Reshmi and K. Balakrishnan, "Implementation of an inquisitive chatbot for database supported knowledge bases," S¯adhan¯a, vol. 41, no. 10, pp. 1173-1178, 2016.
4. Siddharth Gupta, Deep Borkar, Chevelyn De Mello, SaurabhPatil, "An E-Commerce Website based Chatbot", Siddharth Gupta et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2) , 2015, 1483-1485
5. Shamsi T. Iqbal and Brian P. Bailey. 2008. Effects of Intelligent Notification Management on Users and Their Tasks. In CHI. ACM, 93-102.
6. Brynn M. Evans and Ed H. Chi. 2008. Towards a Model of Understanding Social Search. In CSCW. ACM, 485-494.