# Voice bot: Customized Voice-Chat bot using Natural Language Processing

**R Sharanya, Rithin S Reddy, S Asha and Naveen Chandra Gowda**

School of CSE, REVA University, Bangalore, India

**Abstract**.

The Chatbots have a demanding requirement in today's automated world in fields like Marketing, Supporting Systems, Education, Healthcare, Entertainment, etc. Though chat bots are still in their relative infancy, technologically they have existed for decades. The bots provide ease of communication and good customer service. There are many chatbots present in the institutional environment but it still lacks in terms of excessive manual power expanded regardless of its automation due to limiting database (lack of content in the database) or unsatisfactory responses and also being inconsiderate towards disabilities. The Voice Bot tackles these situations by providing an immense database with multiple responses and giving cacographers and disabled users an interactive experience through speech recognition.

**Keywords**. Speech recognition, institutional environment, disabled users, cacographer, Voice Bot

## 1. INTRODUCTION

Chatbot is considered to be an instant messaging platform and an effective auto-response system for a two-way dialogue with human users. The bots are existing and are growing with the increase in demand. They are used across industries like healthcare, travel industry, banking, human resource, e-commerce, and personal services to provide information in conversation mode.

The first bot to be made was ELIZA and was developed by Joseph Weizenbaum on the basis of a keyword matching technique [1]. The bots have adopted new technologies over the upcoming years like the usage of Automatic Speech Recognition model which is dependent upon the voice of the speaker [2]. The general purpose chatbots are difficult to design as they require a more complex knowledge base to meet the needs of different domains, hence developing it in a particular realm is preferred [6]. One of the biggest domains where they are utilized is the scholastic environment, for example, the institutional websites or course sites where the interaction is provided to the students or users to solve queries or report progress.

Chatbots, though being a new way of interaction in the academic front like online platforms and devices still lack in areas concerning the customer satisfaction which is achieved through a good database of responses, the observed responses of the existing

chat bots are not the user desired outputs and in some cases very confusing to understand. At the end of the day, manual power needs to be expanded to persuade and resolve the issues on the user end. It falls behind in terms of automation and manpower, archive content, unsatisfactory responses, and consideration shown towards disabilities.

The usage of chatbot can happen through text or voice inputs, depending upon the context it can be deemed less or more valuable. Instead of understanding and going with the chatbot creator's view, it can apply an user focused approach on how the chatbots are received and the way people experience it in their daily lives.

The problem in the context is the common scenario seen in academic environments where students do cumbersome and taxing manual tasks that require human involvement in most of the cases, majority of the current existing chat bots provide only the textual interaction and a limited database for the query response and do not take disabilities into account, hence it's confined to solving a small set of user queries. The convenience could be improved if the system is not only text-based but also voice-based and trained along with a customizable database. This is the problem addressed by the paper.

The Voice Bot eradicates most of the human intervention by helping users with tasks or queries with respect to the scholastic domain, it provides information to users when they require it and gives the option of querying and chatting in both voice and text, making it flexible and user-friendly while providing an interactive, immersive experience in an automated environment. It also improves productivity by reducing the crowd at the help desk and human efforts by giving quick responses and being easily accessible. With the enabling of voice-based interaction, it serves the purpose of being useful to disabled and cacographic users.

The paper is sectioned into the following distributions as follows: The I part is about introduction, II section explains the Literature survey with the related works, III section provides information on the proposed work through the methodology, algorithm, and system design, the IV section describes expected results and experimental analysis, the V section talks about limitations, VI is the conclusion and the future work section and VII is the reference used for the paper.

## 2. LITERATURE SURVEY

### 2.1 RELATED WORK

Over the past decade, chat bots have garnered a lot of attention in most of the online domains, e.g. E-commerce platforms, travel sites, online ticket booking sites, government sites, College websites, company websites, etc. A number of selected research papers between the years 1966 and 2021 are examined and described below.

The introduction of the chatbot happened through ELIZA, the first bot made on the pattern matching domain. ELIZA was created by Joseph Weizenbaum by utilizing the keyword matching methodology. The process involved the reading of the provided user input and searching for specific keywords based on which the answer was retrieved. In case the required word is not present then the conversation between the user and bot continued. The input queries are analyzed based on the rule of decomposition which utilizes the keywords and the generation of response are made by reconstruction rules in accordance with selected decay law. The basic issues with the author's bot are the response generation in keywords absence, script editing provisions, identification of keywords, etc.

The [4, Fig. 1] shows an example of a sample queries asked by the user.

```
ELIZA: What other reasons might there be?
YOU: Well, I don't know.
ELIZA: I'd like to know more about your opin-
ion.
YOU: I don't like my family so much.
ELIZA: That is quite interesting.
YOU: Yeah.
ELIZA: I'd like to know more about your opin-
ion.
YOU: My family members aren't friendly, so I
can't talk with them.
ELIZA: I see.
YOU: Do you have family?
ELIZA: Why do you think that?
YOU: I don't know who you are.
ELIZA: Why do you ask?
YOU: I'm sorry.
ELIZA: That's okay. Please continue.
```

*Fig. 1: A Sample Conversation with ELIZA from Munira Ansari, Saalim Shaikh, Mohammed SaadParbulkar, Mumbai, India, "Intelligent Chatbot", IJERT, NREST – 2021 Conference Proceedings.*

In Paper [2] the purpose of this paper was to introduce the Automatic Speech Recognition model, which is a stand-alone speaker, word-based and low-key vocabulary by recognizing phonemes, phoneme collections and words. It has developed an app that serves as a Basic English language course. The proposed paper is for the benefit of the disabled. The proposed work is 78 percent accurate according to the studies made.

In Paper [3] the author aims to help reduce organizational dependency on people and reduce the need for othersystems for various processes. The proposed chatbot identifies the user context that triggers a specific response objective. As it responds to a flexible response the response you want will be created for the user. The proposed system has used ML(Machine Learning) algorithms to learn Chatbot by meeting the responses and requests of various users.

In Paper [4] the chatbot queries used to resolve users' queries, provide information to users as they need it, improve service time quality and make customers happy by providing intelligent solutions. Students can connect to the chatbot on the web via a laptop or smart phone. Students can ask different questions about the details of admission in the native language and both can answer their questions with appropriate answers. The proposed app was easily accessible to customers and provided feedback to users anywhere at any time.

The author of the paper [5] explains about the types of chatbots used and the examples and specific enhancements that can be made. Many institutions are attracted to technology to improve their customer support when they encounter technologies like chat bots that enter into as flawless service assistants in such situations. One of the possible solutions discussed on paper is the emergence of a voice-based chat bot that

allows for two conversations between a client and a computer using the human voice as a means of communication. This voice-based bot can help become a valuable asset for the client.

In this paper [6], the reviews include a selected count of papers focused on Chatbot design strategies over the past decade.

The contribution of each study is analyzed and identified. This paper presents a survey on the designed techniques used in the chat bot on the domain of speech conversion between the two entities human and computer. The author concludes by remarking that chat bots are deprived of a common solution, researchers need to document the improvements done on the bot in order to improve the speech interaction between the entities.

The end objective of the project is to demonstrate chatbot using the deep learning models. Recurrent Neural Network and other sequential learning models are not able to achieve this task. Also the Natural Language Processing model undertakes a lot of data which can be considered difficult to gather when it comes to small scale implementations. Hence the paper implements SLM (sequence learning model) along with LSTM (Long short-Term Memory) [7].

The paper[8] presents various constraints presented in chat bot like the problems of the consumer seen in general usage, for example the lack of customer service, limited data storage and the bad optimization of the results. The author compares the two models which are used for chat bot on various factors like sentence responses provided, accuracy measurement, the LSTM and BiLSTM (Bidirectional Long Short-Term Memory). The major focus area of this paper is on finding the best suited and most effective model for the chat bot to work efficiently.

## 2.2  GAPS

➢      ELIZA [1] has technical problems such as keyword identification, limited content, and the provision of editing of its own documents.

➢      In paper [2] the proposed system was developed to have external sound defects and a little vocabulary in the system.
➢      In paper [3] the system provides two types of inputs but is lagging behind in providing two types of outputs, providing textual responses.
➢      Paper [4] provides a web-based chatbot with machine learning that is used to generate responses and text output.
➢      Paper [5] explains about the future enhancement as the voice based chat bot which can prove to be a beneficial client communication asset.
➢      Paper [6] reviews a selected number of chatbot design strategies.

The proposed system works on the provision of keyword retrieval keywords on the website, providing comprehensive content in the archive, and the provision of document editing and site customization according to user needs. The system uses audio filtering with a power threshold in the speech recognition python. Provides dual input and output type for speech-to-text conversion and text-to-speech conversion.

### 3. PROPOSED WORK

The system proposed is based on Artificial Intelligence, the Voice Bot utilizes the python programming language to provide an interface for the user to interact with as the GUI (Graphical user interface). The bot uses multiple API's (Application Programming Interface) for the speech recognition and text to voice conversion like the Speech Recognition and REST (Representational State Transfer), Django etc. For getting responses from the back end it utilizes SQL (Structured Query Language).

### 3.1 SYSTEM MODEL
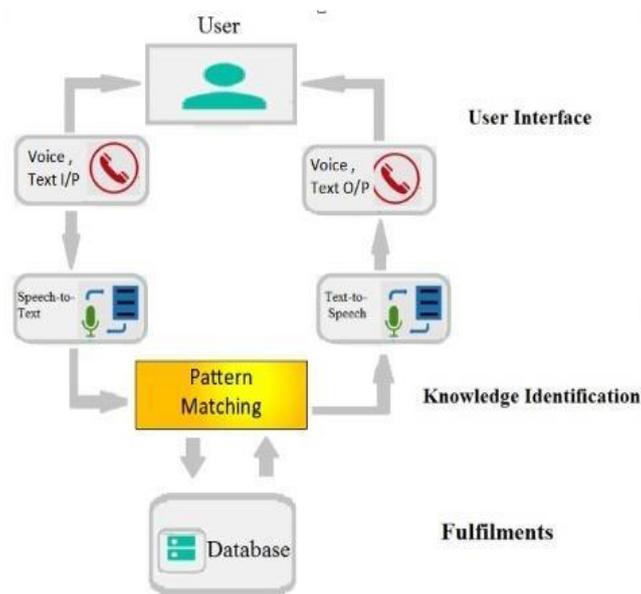


Fig 2. System Architecture of Voice bot

The Fig. 2 explains how the voice chatbot handles the queries from the user to give responses through speech and text with the help of text-to-speech and speech-to-text conversion and fetching responses from the database after pattern matching.

### 3.2. MODULES

### 3.2.1 VOICE TO TEXT DIAGRAM

The text to speech conversion system utilizes the speech recognition package where feature extraction takes place and the acoustic and language modules are applied to give out the textual output in the notepad. The user's input in the speech or voice format is converted to text and printed in the notepad as shown in Fig. 3.
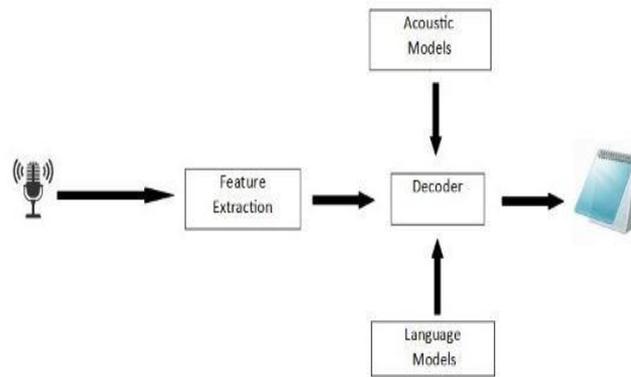
6



Fig. 3: Speech to text converter process in notepad

3.2.2. VOICE CHATBOT

The Voice chatbot makes use of the speech to text converter and also implements text to speech to provide the response. It uses Natural language processing with speech recognition and pyttsx3 (python text to speech).The Fig. 4 displays the working of the Voice chatbot.
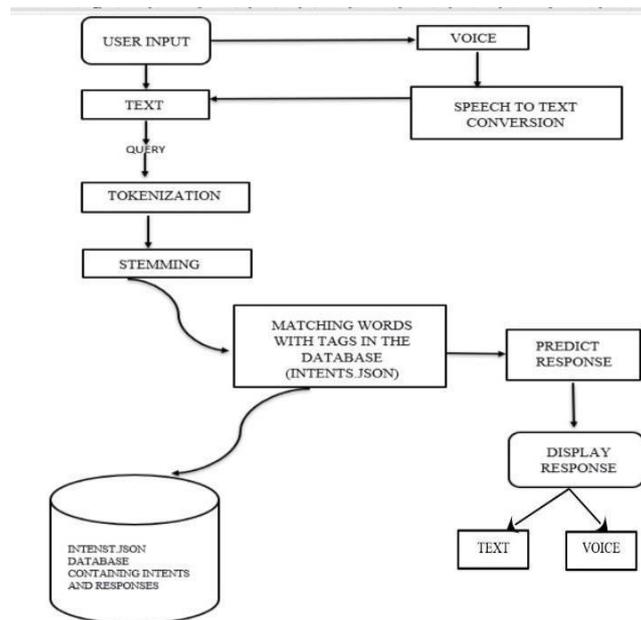


Fig. 4: The Working model of completed Voice bot

## 4. IMPLEMENTATION

### 4.1 ALGORITHMS

Lemmatization and POS (Part of Speech) tagging, using WordNet: Deleting of the data in the input text is done by removing keywords. The appropriate Lemmas keywords were obtained using the Lemmatization and POS tags, combined with a different type of translated word. WordNet from the Python package "NLTK (Natural Language Processing)" was used for this purpose.

4.2 NATURAL LANGUAGE PROCESSING

NLP (Natural Language Processing) is used as one of the most important concepts in bot design. Machines use sophisticated algorithms to separate any text content in order to extract meaningful information from it. The data collected is then used to further teach the machines the natural language skills.
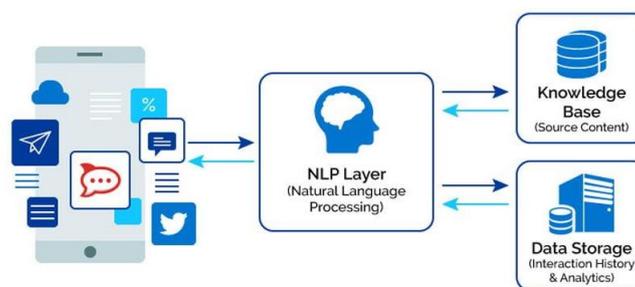


Fig5: NLP Working from KollaBhanu Prakash, "Chatterbot implementation using Transfer Learning and LSTM Encoder-Decoder Architecture", Volume 8.No. 5, IJETER, May 2020.

The Natural Language Processing layer utilizes both the data storage and the knowledge base in order to retrieve appropriate and accurate responses to the given query submitted by the user as shown in [7, Fig.5].

Incorporating NLP into chat bots gives the meaning of having higher human presence. When a chat bot is developed and used, this is a common use, hence there are questions asked regarding it. It seems so intertwined with human behavior that customers can try to trick and discard the chat bot. It can be solved by adding automatic answers, but that is often a failure as it is almost impossible to wait for what queries to be answered and how they will be solved. The goal of Natural Language Processing is to build a type of system that can understand and give meaning to a text and also perform tasks automatically like classification, translation, etc.

4.2.1 STEPS OF NLP ALGORITHM

1.     Lexical Analysis – It is the first step in the Natural Language Processing responsible for the source code scanning in order to convert a string of characters into meaningful expressions, this phase divides the input text into sections, words and sentences. Tokenization, Lemmatization and Stemming techniques come under the Lexical Analysis.

2.      Syntactic Analysis – This phase of NLP is mainly used for parsing and providing relationships between the different words and also to check their arrangement and grammar.
3.      Semantic Analysis- This analysis is based on the meaningful representation of the words and it focusses on the literal meaning present behind them.
4.      Discourse Integration – Integration of speech depends and is implemented based on the meaning of the sentences that come before it.
5.      Pragmatic Analysis – This final phase focusses on information extraction from the provided text, it plays by a set of cooperative dialogue rules.

4.3 LSTM (Long Short Term Memory) ALGORITHM

Long Short Term Memory (LSTM) networks are a type of recurrent neural network required in complex problem domains like machine translation, speech recognition, and more .LSTMs are a complex area of deep learning.

The LSTM model network is called as the model that has influenced the past and is known to display the capability to learn from sequential data. LSTM consists of three gates-input, forget, output gates [7].

$f_{m+1} = (^{(h \to f)}h_m + \theta^{(x \to f)}x_{m+1} + b_f)$ forget gate (1)

$i_{m+1} = (^{(h \to i)}h_m + \theta^{(x \to i)}x_{m+1} + b_i)$ Input gate (2)

$c_{\tilde{m}+1} = \tanh(\theta^{(h \to c)}h_m + \theta^{(w \to c)}x_{m+1})$ update candidate (3)

$c_{m+1} = f_{m+1} \odot c_m + i_{m+1} \odot c_{m+1}$ memory cell update (4)

$o_{m+1} = (^{(h \to o)}h_m + \theta^{(x \to o)}x_{m+1} + b_o)$ Output gate (5)

$h_{m+1} = o_{m+1} \odot \tanh(c_{m+1})$ Output (6)

The above equations represent the LSTM model [8]. The input gate checks if there is a need to allow the fresh data or not, forget gate deals with the unnecessary details. The proposed framework approach consists of various steps, such as raw data selection, pre-data processing, feature extraction and NN preparation.

## 5.      RESULTS AND DISCUSSION

Speech recognition systems are tested using two factors: Accuracy and speed. Platforms used for voice chatbot functionality Windows 10, Python 3.8 IDLE, and PyCharm 2021.3. The test requires the following specification: 64-bit operating system with Intel core Processor 5.1.60GHz, OS version - Windows 10 One Home Language, Ram - 8.00 GB.

The simulation of a voice robot is performed in two parts with respect to the accuracy of word recognition by pronunciation in English. Voice recognition module is said to boast about 75% - 85% accuracy under noise and greater than 85% in good conditions. The Table. 1 and Fig. 6 pie chart show test results of the conducted simulations.

| Simulation (Eng.-in) | Total words | Detected Words | Undetected Words | Accuracy |
|---|---|---|---|---|
| 1 | 80 | 75 | 5 | 93.75 % |
| | 90 | 84 | 6 | 93.33 % |
| | 100 | 92 | 8 | 92.00 % |
| | 110 | 97 | 13 | 88.18 % |
| | 120 | 105 | 15 | 87.50 % |
| Total | 500 | 453 | 47 | 90.6% |

*Table. 1: Manual Analysis Simulation 1*

Manual Word Analysis Simulation 01
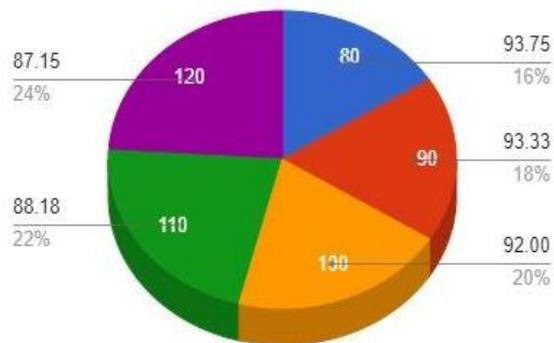


*Fig. 6: Pie Char 1 for manual analysis 1*

Second simulation was performed with a bit lower net speed and some introduction of background noise. The Table. 2 and the Fig. 7 pie chart below show the results.

The percentage is a bit less compared to previous simulation due to the presence of noise affecting the accuracy.

| Simulation (Eng.-in) | Total words | Detected Words | Undetected Words | Accuracy |
|---|---|---|---|---|
| 2 | 80 | 72 | 8 | 90.00% |
| | 90 | 82 | 8 | 91.11% |
| | 100 | 93 | 7 | 93.00% |
| | 110 | 95 | 15 | 86.36% |
| | 120 | 101 | 19 | 84.36% |
| Total | 500 | 443 | 57 | 88.6% |

*Table. 2: Manual Analysis Simulation 2*
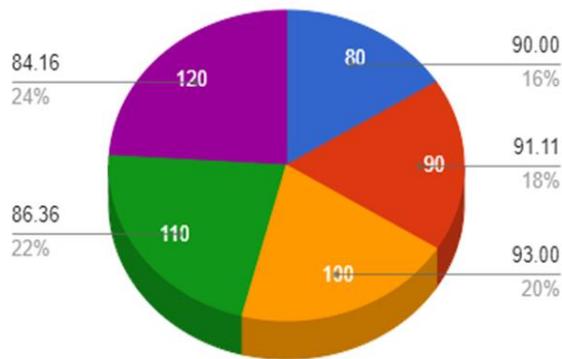
**Manual Word Analysis Simulation 02**



*Fig. 7: Pie Chart 2 for manual analysis 2*

## 6.    OUTPUT RESULT

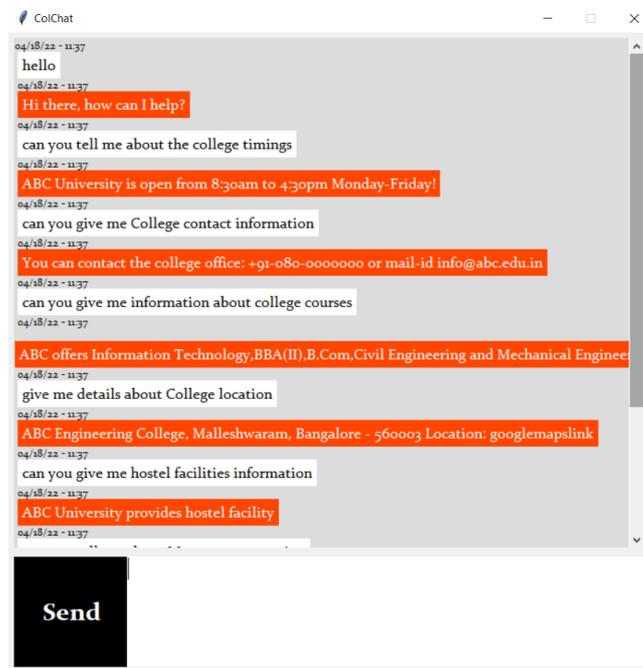The outputs considering three scenarios are listed below:

Fig. 8: Accurate Responses of the Voice bot

In Fig. 8 for any given input by the user or the customer as a query, the bot processes it and searches in the database, accurate responses are obtained in both textual and voice format to the user in the GUI (Graphical User Interface) and the speakers.
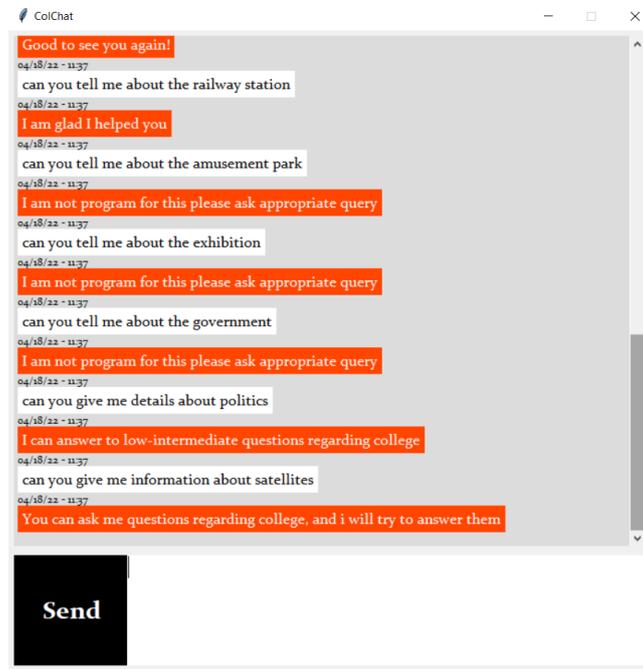
Fig. 9: Incorrect Response of the Voice bot

In Fig. 9 when the user asks queries not related to the domain the bot's response is as shown in the figure. Since the keyword is not found in the database the bot responds saying 'I am not programmed for this please ask appropriate query".
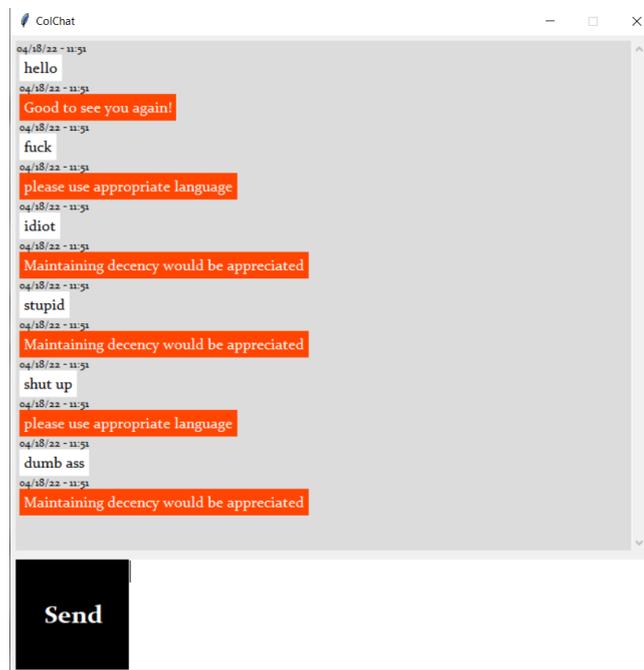
Fig. 10: Indecent Queries of the Voice bot

In Fig. 10 when the bot is misused by the user by cursing, it responds with either "Please use appropriate language or "maintaining decency would be appreciated", which re the input given in the database for the usage of inappropriate queries. The bot can handle unethical behavior of users while taking the input from them. The database is equipped to handle such situations and if required can be improved with more content.

## 7. CONCLUSION

In today's world of automation the voice bot is well suited to its voice and text resources. The paper uses Voice Chatbots and Machine learning concepts, as well as NLP for input and feedback. The main goal is to reduce human interaction, increase automation, provide audio filters and provide in-depth information to disabled users and photographers by input and output in two ways. Objectives were introduced and applied to the bot.The bot can be made accessible to users on many web platforms on institutional websites, the accuracy of the bot can be increased through audio filtering functions. By combining a lot of pronunciation it can be used by different people. It can be upgraded to an app for frequent and easy access. It will be useful in all areas as without wasting much time, it gets to the right information and that too without filtering. It can be developed to receive queries in two languages and provide respond in the same input language as the user asks. It can be improved in the future by solving obstacles and taking it to the top in a web-based environment.

14

## 8.    REFERENCES

[1]    Joseph Weizenbaum, Massachusetts Institute of Technology, "ELIZA – A Computer Program For the Study of Natural Language Communication Between Man And Machine", Volume 9 Number 1, 1996. A.G. OETTINGER, Editor.

[2]    Vishal Bhargava, Nikhil Maheshwari, Department of Information Technology, Delhi Technological University (Formerly DCE), Delhi. "An Intelligent Speech Recognition System for Education  System", 2008.

[3]    SangeetaKumari, Zaid Naikwadi, AkshayAkole, PurushottamDarshankar, Vishwakarma Institute of Technology, Pune. "Enhancing College Chatbot Assistant with the Help of Richer Human-Computer Interaction and Speech Recognition" Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020)

[4]    Munira Ansari, Saalim Shaikh, Mohammed SaadParbulkar, Talha Khan, Anupam Singh, M.H. SabooSiddik Polytechnic, Mumbai, India, "Intelligent Chatbot", IJERT, NREST – 2021 Conference Proceedings.

[5]    Bhakti karkera1, Flavia Gonsalves,  "Chatbots in the world of artificial intelligence", IRJET, Volume 5, June 2018.

[6]    Sameera A and John Woods, "Survey on Chatbot Design Techniques in Speech Conversation Systems", International Journal Of Advanced Computer Science and Applications, vol. 6, No. 7, 2015.

[7]    KollaBhanu Prakash, YVR Nagapawan , N Lakshmi Kalyani , V Pradeep Kumar, "Chatterbot implementation using Transfer Learning and LSTM Encoder-Decoder Architecture", Volume 8.No. 5, IJETER, May 2020.

[8]    PrasnurzakiAnki, AlhadiBustamam, "Measuring the accuracy of LSTM and BiLSTM models in the application of artificial intelligence by applying chatbotprogrammes", Volume 23, No. 1, July 2021.

[9]    Neelkumar P. Patel, Devangi R. Parikh, "AI and Web-Based Human-Like Interactive University Chat bot (UNIBOT)", IEEE,2019.

[10]    Urmil Bharti, Deepali Bajaj, HunarBatra, Shreya Lalit, Shweta Lalit, AayushiGangwan, "Med bot: Conversational Artificial Intelligence Powered Chat bot for Delivering Tele-Health after COVID-19",IEEE ,2020.

[11] Prof.K.Bala,MukeshKumar,SayaliHulawale, SahilPandita, "Chat-Bot For College Management System Using A.I", International Research Journal of Engineering and Technology(IRJET)

[12] Hiremath et. al., "Chatbot for education system", International Journal of Advance Research, Ideas And Innovations In Technology, vol 4, Issue 3.

[13] Voice-based Chatbots – a Revolution in Customer Relations, Author: Thomas Saint Hilaire.

[14] Levy D., 2009. DO-MUCH-MORE Chatbot Wins 2009 Loebner Prize for Computer Conversation, News from the Benelux Association for Artificial Intelligence, Vol. 26, No. 4, p. 67-78, available at http://www.unimaas.nl/bnvki/archive/2009/26.4. pdf.

[15] Maughan R., 2002. Conversational Agents, School of Computer Science and Information Technology, University of Nottingham Pirner J., 2003. About Jabberwock, available at http://www.abenteuermedien.de/jabberwock

[16] Roberts F., Gülsdorff B., 2007. Techniques of Dialogue Simulation, Lecture Notes in Computer Science p. 420- 421.

[17] Shieber S. M., 2006, Does the Turing Test Demonstrate Intelligence or Not?, AAAI-06, Boston, Ma.

**Biographies**

Photograph
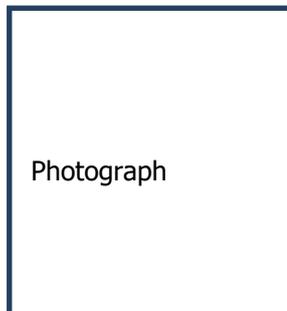
**Author Name** received the bachelor's degree in computer engineering from Cankaya University in 2010, the master's degree in computer engineering from Gazi University in 2014, and the philosophy of doctorate degree in Electrical-Electronics & Computer Engineering from Duzce University in 2017, respectively. He is currently working as an Assistant Professor at the Department of Computer Engineering, Faculty of Engineering, Duzce University. His research areas include mobile security, deep learning, and social network analysis. He has been serving as a reviewer for many highly-respected journals.