

Cognitive Intelligence System for Institutional Management Using Speech Recognition

Antony Vigil¹, Barathwaj C², Saurav Shekhar³, Keerthivasan M⁴, Indreish Yuvaraj⁵

¹ Assistant Professor (Sr.G), Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, India.

^{2, 3, 4, 5} Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, India.

Abstract – We create a database related to a college institution like staff details, college events dates, functions etc. For this, first we have to create an android application for using speech recognition and a database that contains all the college details. This paper presents the design and development of an intelligent voice recognition chat bot that not only responds to voice but also undergoes an OTP (one time password) and an email verification process. So using the chat bot android application, we can give a speech input asking for college details, and as soon as all the security levels are passed (OTP and Email verification) it gives the desirable college details in text format. By introducing an artificial brain, the web-based bot generates customized user responses, aligned to the desired character. Questions asked to the bot, which is not understood is further processed using a third-party expert system (an online intelligent research assistant), and the response is archived, improving the artificial brain capabilities for future generation of responses. The application is developed with an view of acting as an online friend for the user, so that it would be comfortable for the application user to interact and use the application. Thus, this becomes an advantage for the application user. This chat bot is developed with features that overcome the existing disadvantages of the chat-bots in the market. The drawbacks of those chat-bot's are changed as an advantage for this chat-bot.

Index Terms – AI, Chat bot, OTP, ML.

1. INTRODUCTION

At present the application of intelligent chat robots is spreading like wild fire, besides the basic function of intelligent reply, it does have three layers of security including email verification, OTP (One Time Password) lastly but most importantly the speech recognition. Many bot framework provided simple and easy to use for developing the personalized intelligent chat bot. Here we are introducing new features to the framework. By utilizing such platform and combining the voice recognition and artificial intelligence we can develop a new platform which will be very different from the rest and much enhanced. Our chat bot plays a major role in college management system, where it follows a tree structure for searching the details from the database. The cloud is very secure in our chat bot and modifying the details in database undergoes two levels of security such as OTP (one

time password) which is been sent both to the modifier and the details of the person which is been modified. And then another level of security in the mail, where after modifying the modified person details, the message is been sent through E-Mail to the modified person. Compared to the other chat bots which are available in the market which uses different types of algorithm for fetching the details from the database, our chat bot uses the most powerful algorithm which plays a major role in fetching operation which is faster and secure.

2. BACKGROUND

The advent of chat bots has created a new dimension to AI research. Chat bots are intelligent interfaces that can make conversation in a coherent manner. They were made to replace human beings as chat agents. The conventional chat bots use AIML, a pattern matching XML parser, as a response system. AI community has tried to construct chat bots like ALICE and ELIZA. These are generic chat bots. There are others like chat bots for cultural heritage (Pilato et al., 2005), security training (Kowalski et al., 2013), blind high school students (Bigham et al., 2008). There has been a discussion on the utility of chat bots (Shawar and Atwell, 2007). All these chat-bots function similarly. If we consider Alice, it basically generates an AIML file from a dialogue corpus with patterns and templates (Shawar and Atwell, 2003).

The limitation begins with the presence of a corpus which assumes all knowledge comes from previous dialogue done by human agents. Secondly generating AIML from a corpus cannot guarantee a coherent chat because there is a fear of getting repetitive statements, which will worsen the user chat experience. There are two defects in existing chat bots - lack of intelligence and similar responses for repetitive statements due to the pattern matching nature of AIML. There are works which try to understand the semantics (Augello et al., 2009). But the responses are generated using statistics. In this case, the responses are memorized. Also, people have tried to capture semantics through different ways of knowledge representation (Pilato et al., 2012). This seems an over approach because humans have a generalist approach towards

understanding concepts. Different knowledge representation has an inherent undecided ability problem of application in which context.

3. INTELLIGENT CHAT BOT

A chat bot like ALICE lacks AI in the true sense. Firstly, if a bot wants to mimic a human, it should not memorize every conversation in database. An intelligent chat bot must be powered by AI/NLP to reply coherent messages at least from the business point of view. ALICE has no way to understand the intent of the user. That is why these existing chat bots are not good enough for replacing human beings. They are over reliant on AIML. There is a demand for an interface between AIML and the chat bot such that it understands chats, filters entities and then generates text in AIML readable format in a smart way. Thus, it can convert its knowledge to a response custom made for AIML. So AIML will become a layer that is at the lowest level of the chat bot brain and is used just an interface to reply in a natural language. That's how we have designed o which smartly extracts all sorts of information such as name, intent, mail, city, etc and generates a coherent response to user.

Secondly, an intelligent bot will never respond in the same way if user is sending same text. A random response as customary in AIML is not the right way to deal with repetitions. The bot should have a counting mechanism to know that there is repetition in incoming messages and if this behavior continues further chats should be routed to a human agent in a polite manner or the chat terminated in case of completely automated bots. The ability to count can be achieved only if the bot has a system to understand the intent of the client and if it maintains a state diagram or a milestone status with a counter for each state. In this manner, if it encounters the same state on two or more consecutive occasions, it can safely route the user to a human agent or terminate. Thirdly, response cannot be generated based on probability. That will cause havoc because sentence generation should be accurate and AIML, that is self curate and not generated by pattern extraction algorithms, is the proper way to respond. System cannot entirely depend on AIML, but it cannot get rid of AIML. AIML is a very powerful device when it comes to generating responses in natural language. The three important criterion of an intelligent chat bot are:-

- Understanding rather than memorization
- Ability to handle repetitive queries
- AIML based response mechanism

4. PROPOSED SYSTEM WORKING

In this section, we would give a brief elaboration about the working of our system. The chat bot we developed is for helping the college's to manage the essential events and other important aspects with ease. The process of our chat bot is based upon the method of "talk reply and talk reply". The

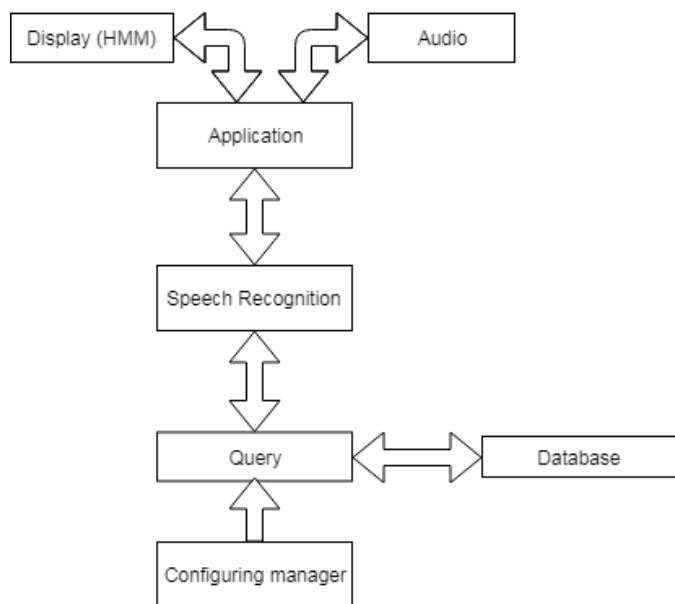
algorithm we have used for developing the chat bot makes it more users interactive. The chat bot provides help to the user 24*7 with assistance. With the help of speech recognition system it is capable of recording the voices of the user and storing them in the back end for future purposes. This mechanism makes the user to interact more with the chat bot and therefore making the user feel like having a conversation with an online friend. It is been developed as an Android application and can be installed in any phone. The chat bot is developed as a supportive hand to the user and the databases of the application are stored in the cloud. It is secured and very reliable to use and store data to the cloud. The databases are stored in the cloud in the hierarchical view manner. Since, we developed a chat bot for college purpose it makes the administrator of the institution to control all the events of the college.

The queries that are searched in the application would be stored in the form of a tree structure and thus making the processing of the query to be quick and time saving. The application is made to store the details and data about the user in the template manner and is easy to proceed when a query is asked related to the user. The user of this application can be any student and also the staffs of the college. The students and the staffs or whosoever uses the application would get a unique user id to log in and to utilize the application. The student user can use and access only his application and cannot view the other is content.

This is to enhance to the security of the databases that are entered in the application. On the other hand, if the application is being installed and used by staff then he or she will be able to view and edit any information regarding any student. In the same way they can't view or edit other staff's data. The administrator i.e., the highest authority of the institution would likely be the administrator of the institution and who can view and edit the detail about any student or staff. Hierarchical view ensures data security. In order to make the database and the application more secure the data are protected via generating One Time Password (OTP) and an email while the data are edited.

5. COMPONENTS METHODOLOGY

1. Start the application
2. Get the input(audio format)
3. Voice recognized through Gaussian Mixture Model.
4. User asks the query needed.
5. Availability of the query will be searched in the databases.
6. If available, the output will be displayed on the screen of the application (voice converted to text by HMM algorithm).



1. System architecture

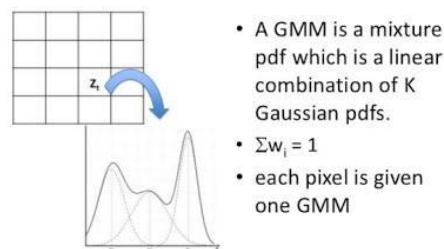
6. PROPOSED ALGORITHM

Natural Language Processing (NLP) is used to develop the ideal chat bot, that can have a conversation that is as natural as possible and that it is indistinguishable from a normal one between humans. The simpler older chat bots, are the chat bots that employ heuristics with pattern recognition, rule based expression matching or very simple machine learning. The important aspect is that these systems are good at comparing a fixed set of rules. If you look at the simpler chat bots, any response (provided it was correct grammar beforehand) is void of any grammatical error. This is of course due to the pre-written sentences in the repository. It might however be unable to handle any input it does not recognize because of human grammatical errors or not matching sentences. The newer smarter chat bots are the exact opposite, if they are well “trained” they can recognize the human natural language and can react accordingly to any situation. However, the big disadvantages is that these natural responses require a great amount of learning time and data to be able to learn the vast amount of possible inputs. The training will prove if the bots are able to handle the more challenging issues that are normally obstacles for simpler chat bots. Depending on the question, these can be long or short conversations. Longer conversations tend to have deeper meanings and multiple questions that the chat bot would have to consider in its extrapolation of the total picture. Ultimately the tasks that NLP should be able to handle are in the following summary. The following task can be text, speech and even image related.

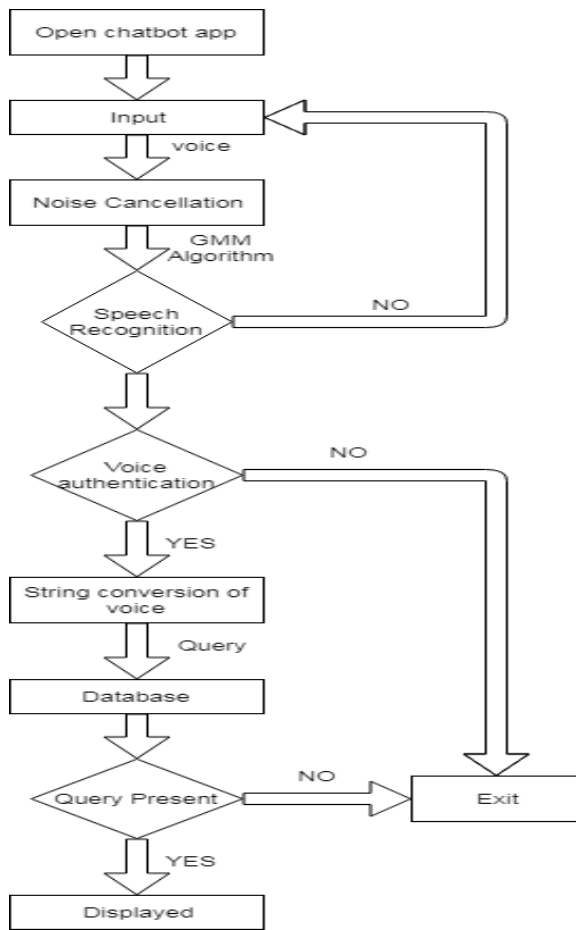
GMM Algorithm (Gaussian Mixture Models): A Gaussian mixture model is a probabilistic model that assumes all the

data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. One can think of mixture models as generalizing k-means clustering to incorporate information about the co variance structure of the data as well as the centers of the latent Gaussians.

Gaussian Mixture Model (GMM)



HMM Algorithm (Hidden Markov Models): Introduction Hidden Markov models are widely used in science, engineering and many other areas (speech recognition, optical character recognition, machine translation, bio informatics, computer vision, finance and economics, and in social science). Definition: The Hidden Markov Model (HMM) is a variant of a finite state machine having a set of hidden states, Q, an output alphabet (observations), O, transition probabilities, a, output (emission) probabilities, B, and initial state probabilities, Π . The current state is not observable. Instead, each state produces an output with a certain probability (B). Usually the states, Q, and outputs, O, are understood, so an HMM is said to be a triple, (A, B, Π). Formal Definition: Hidden states $Q = q_i, i = 1, \dots, N$. Transition probabilities $A = a_{ij} = P(q_j \text{ at } t+1 | q_i \text{ at } t)$, where $P(a | b)$ is the conditional probability of a given b, $t = 1, \dots, T$ is time, and q_i in Q. Informally, A is the probability that the next state is q_j given that the current state is q_i . Left-right models with a handful of states are used to describe diaphones or triphones. States have hundreds to thousands mixture components (or nowadays we eschew GMMs and use other emission models). This is already a ridiculously large model space, so there are tied-states - that is to say some states in the model share the same emission probabilities with other states in the model. Anything that pools together the parameters from several parts of the model ensures that more training data will be used.. The lexicon maps a word into a sequence of phones. While there are approaches to automatically learn it, this data primarily comes from linguists. Then word models are concatenated together to build a language model. The result is an extremely big FST, which is not even fully composed in memory from its component - but only traversed on the fly during recognition.



2. Flowchart

7. WORKING ILLUSTRATION

A small sample illustration about the chat bot is likely all the details about the institution and the members of it will be pre loaded in to the chat bot system. Before editing the information at first a One Time Password (OTP) will be

generated to the staff whose details are to be edited and in the similar manner an OTP would be sent to the Administrator too. If the OTP is not sent successfully then email verification would be sent to their respective mail ID's. By this the data can be secured and the chat bot would save and display all the actions of the user that he/she searched about, would be showed to the administrator in the activity log of the user.

8. CONCLUSION

Through voice-based chat bot, we are trying to redefine the voice-based chat bot in a well-defined manner. The novelty lies in a manner such that it is not just a response generator but an artificial intelligence which finds the best solution for the queries. The process is done in a hierarchical manner which makes the chat bot very effective and an extra method of security is added, which makes the chat bot ahead compared to another chat bot. In the future, we will retain the milestone format but we need to scale Chat bot to handle the entire college management and user.

REFERENCES

- [1] 2016 7th International Conference on Intelligent Systems, Modeling and Simulation., Chat bot Using Knowledge in Database Human-to-Machine Conversation Modeling.
- [2] Y. Bin, P. Cunlin, and L. Dan, "Chinese Text Feature Extraction Method Based on Bigram," Proc. IEEE of 2013 International Communications, Circuits and Systems (ICCCAS), 2013, pp. 342-346, doi: 10.1109/ICCCAS.2013.6765352
- [3] S. Ghose and J. J. Barua, "Toward The Implementation of A Topic Specific Dialogue Based Natural Language Chat bot As An Undergraduate Advisor," Proc. IEEE of 2013 International Conference on Informatics, Electronics Vision (ICIEV), 2013, pp. 1-5, doi:10.1109/ICIEV.2013.6572650.
- [4] C. Erdogan, H. Nusret Bulus, and B. Diri, "Analyzing The Performance Differences Between Pattern Matching and Compressed Pattern Matching on Texts," Proc. IEEE of 2013 International Conference on Electronics, Computer and Computation (ICECCO), 2013, pp. 135-138, doi:10.1109/ICECCO.2013.6718247.