International Journal of Research in Engineering, Science and Management
Volume-2, Issue-3, March-2019
www.ijresm.com | ISSN (Online): 2581-5792

AI based Chatbot for Education Management

S. Brindha[1], K. R. Dharani Dharan[2], S. J. Julian Samraj[3], L. Maria Nirmal[4]
[1]*Professor & HoD, Department of Computer Networking, PSG Polytechnic College, Coimbatore, India*
[2,3,4]*Student, Department of Computer Networking, PSG Polytechnic College, Coimbatore, India*

***Abstract*: AI based digital assistants like Google Assistant, Alexa are frequent in current days. Education is essential for human civilization. With the increase in the number of students in each institution faster methods must be found for efficient use of data produces in educational institutions. Statistics and analytics of educational database include student activity monitoring, result analysis, student and faculty details, academic calendar events, etc. It would be much simple and efficient when the digital assistant is introduced. This can be achieved by using a voice-controlled digital assistant that can provide results to complex analytical speech queries. This digital assistant uses speech to text, text to speech conversion algorithms. Which uses a well-trained keyword filtering algorithm. There will be a secured database where the student academic data will be stored.**

***Keywords*: AI, speech recognition, NLTK, model database, chatbot, speech engine**

1. Introduction

AI-based digital assistants like Google Assistant, Alexa are being increasingly used in current days. In education, the need for digital assistants is essential to managing statistics and analytics of educational database, which includes student activity monitoring, result analysis, student and faculty details, academic calendar events, etc. It would be much simple and efficient when the digital assistant is introduced. This can be achieved by using a voice-controlled digital assistant that can provide results to complex analytical speech queries. Hardware includes a combination of microphone and speaker connected to the raspberry pi. As well as monitor for displaying text-based results.

2. System architecture

Far-field voice recognition setup, using 2X microphone combination. The audio will be first preprocessed to remove any noise and crosstalk. The audio is tracked for a pre-specified keyword. If the keyword is present, the audio is then converted to text by a well-trained sequence to sequence-based RNN. The algorithm we use is Baidu's end-to-end system. The Natural Language Toolkit is commonly known as NLTK, it is a suite of libraries and programs for symbolic and statistical natural language processing (NLP) for English written in the Python programming language. Example: "What is today's attendance of DCN?" for this sentence, the program will classify the words by tagging them and to keep them in the order it tokenizes the words. NLTK is one of the leading platforms to build Python programs to work with human language data. It is simple, easy-to-use and interfaces with over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum. NLTK is platform independent and is available for Windows, Mac OS X, and Linux. In all, NLTK is a free, open source, community-driven project. The reply techniques will include a text-based as well as a speech reply. This is done using an RNN where it will take multiple micro audio tracks, stitch them together. The resulting audio will resemble human-like tones and modulation.

The database is created using Mongo DB. MongoDB is used to store data in a flexible manner, JSON-like documents, meaning fields can vary from document to document and data structure can be changed over time. It is a cross-platform document-oriented database program. It can model maps to the objects in our application code, making data easy to work with. Peer-peer queries, indexing, and real-time aggregation provide powerful ways to access and analyze the data. MongoDB is used because it works on a distributed database, so high data availability, horizontal scaling, and geographic distribution are built in and easy to use. MongoDB is free and open-source. It is the powerful suite of services that allow teams to safely expose the data from the frontend; build backend logic, third-party service integrations, or APIs; and run code in response to data changes all without servers. Botpress is used to develop Chabot. It consists of the necessary messaging channels, the necessary backend integrations for the bot and can be interacted with and creating the main dialog flows. Botpress entails analyzing the conversations, refining the conversational experience by building new flows and to make it smarter by continuously improving the NLU.

A. System requirements

The project consists of a remote server and a client which works on a TCP/IP network. The server which is a high-end GPU powered machine will run the module. The client can be anything from smart speakers to web apps or even a personal computer. The devices used will fetch responses from the backend server.

3. Voice recognition

Voice recognition software on computers requires that analog audio is converted into digital signals, known as analog-to-digital conversion. For a computer to decipher a signal, it must have a digital database, or vocabulary, of words or syllables, as well as a speedy means for comparing this data to signals. The speech patterns are kept on the hard drive and loaded into memory once the program is run. A comparator checks these data patterns against the output of the A/D converter -- the action is known as pattern recognition.

Practically, the scale of a voice recognition program's effective vocabulary is directly associated with the random-access memory capability of the pc within which it is installed. A voice recognition program runs repeatedly, quicker if the whole vocabulary is loaded into RAM, as compared with looking the disc drive for a few of the matches. Processing speed is crucial, as well, as a result of it affects how fast the pc will search the RAM for matches.

Fig 1. Voice recognition process

- *Voice Input:* With the help of microphone audio is input to the system, the pc sound card produces the equivalent digital representation of received audio.
- *Digitization:* The process of converting the analog signal into a digital form is known as digitization, it involves both sampling and quantization processes. Sampling is converting a continuous signal into a discrete signal, while the process of approximating a continuous range of values is known as quantization.
- *Acoustic Model:* An acoustic model is created by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word. It is utilized by a speech recognition engine to acknowledge speech. The software package acoustic model breaks the words into the phonemes.
- *Language Model:* Language modeling is used in many natural language processing applications such as speech recognition tries to capture the properties of a language and to predict the next word in the speech sequence. The software package language model compares the phonemes to words in its inbuilt dictionary
- *Speech engine:* The job of speech recognition engine is to convert the input audio into text; to accomplish this it uses all sorts of data, software algorithms, and statistics. Its initial operation is digitization as mentioned earlier, that is to convert it into an acceptable format for additional processing. Once the audio signal is in proper format it then searches the best match.

4. Text analysis

Text Analysis begins with a normalization process in which text is broken up into its component parts; chunks of text are broken into sentences, sentences into words and punctuations, and words into their respective phonemes. Sentence tokenization involves sorting out typical sentence break punctuation, like the punctuation mark, question mark, colon, semi-colon, or period. The challenge is that these punctuations do not always mark the end of a sentence; to solve this problem, machine learning classifiers are trained on text to learn the difference between sentence-ending and non-sentence-ending punctuations. Normalization also includes the conversion of non-standard words into the natural language.

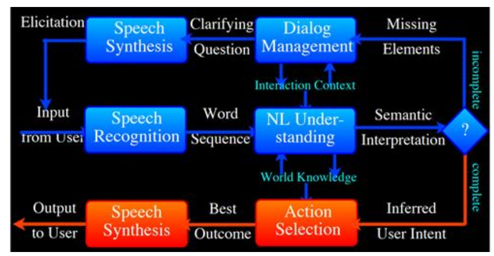

Fig. 2. Text analysis

Examples include numbers, abbreviations, and acronyms which can be pronounced in several different ways. Finally, normalization involves the disambiguation of homographs, or words that are pronounced differently and contain different meaning based on context. For these words (e.g. “I used by the paper” vs. “I put my paper to use”), part-of-speech (POS) analysis can be run to tag the part of the speech of the homograph, and in doing so, disambiguate its pronunciation. The second step in text analysis is pronunciation. This involves access to a pronunciation lexicon which includes mappings between words/phonemes and their pronunciations, as well as name lexicons, which map names to their pronunciations. The character-to-phoneme mapping process was initially rule-based, but has since progressed and now uses advanced algorithms to generate the most accurate mappings based on context.

5. Response generation

Response generation is arguably the most central component of Chabot architecture. As input, the Response Generator (RG) receives a structured illustration of the spoken text. This conveys data concerning who is speaking, the dialogue history, and the context. As output, the RG generates a response to deliver to the user, that it will deliver to the Dialogue Manager (DM). The response selector has access to a few key elements it will use to form its decision about what to say:

- An information database / data corpus, which can dissent in content based on implementation.
- A dialogue history corpus, which will only exist in more complex models.

International Journal of Research in Engineering, Science and Management
Volume-2, Issue-3, March-2019
www.ijresm.com | ISSN (Online): 2581-5792

- An external data source, which provides the bot with intelligence (e.g. A dog is an animal). This latter “common sense intelligence” is usually achieved by permitting bots to access and retrieve documents from search engines.

6. Project result

The result for the speech recognition using Baidu’s deep speech algorithm was not successful as the accuracy for Indian accent was greater than lesser than 20% as in Fig. 1. So, the usage of CMU Sphinx is done which had a greater accuracy of 60%. Online Speech recognition API like Google’s STT and IBM Watson services have accuracy more than 95%. So, Google’s free Speech Recognizer has been implemented for the project.

Fig. 3. Speech recognition output

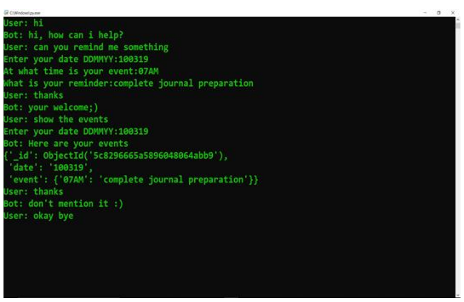

Fig. 4. Final bot result

For the Speech synthesis sapi5 has been implemented so the local resources of the working system can be used. The final output can say store different event details and can provide them on demand which is demonstrated in Fig. 4.

7. Conclusion

This paper presented the implementation of AI based Chabot for education management.

References

[1] Bird, Steven; Klein, Ewan; Loper, Edward; Baldridge, Jason (2008). "Multidisciplinary instruction with the Natural Language Toolkit" (PDF). Proceedings of the Third Workshop on Issues in Teaching Computational Linguistics, ACL.
[2] Kaur, R., & Sharma, D. (2016). An Improved System for Converting Text into Speech for Punjabi Language using eSpeak. International Research Journal of Engineering and Technology, 3(4), 500-504.
[3] https://botpress.io/docs/build/overview/
[4] Jin, Jiong, et al. “An information framework for creating a smart city through internet of things.” Internet of Things Journal, IEEE 1.2 (2014): 112-121.
[5] https://github.com/ServiceStack/ServiceStack/wiki/Postman
[6] Schmidhuber, Jürgen (2015). "Deep learning in neural networks: An overview". Neural Networks.
[7] MongoDB. "MongoDB Stitch JavaScript SDK". GitHub.
[8] CHATBOT: Architecture, Design, & Development by Jack Cahn
[9] J. Jia, “The Study of the Application of a Keywords-based Chatbot System on the Teaching of Foreign Languages”, Report of University of Augsburg, Augsburg,
[10] Dmitry Ilvovsky, “Chatbot with a Discourse Structure-Driven Dialogue Management,” National Research University Higher School of Economics Moscow, Russia.