AN AI ENABLED TOY FOR INTERACTIVE COMMUNICATION WITH KIDS

[1]Dr. T. S. Mastan Rao, [2]Gourishetty Sai Sampreethi, [3]Merugumalla Rachana Shree, [4]Poola Sukumar Babu

[1]Professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India.

[2]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India.

[3]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India.

[4]JNTUH, Computer Science and Engineering, CMR Technical Campus, Medchal, Telangana, India.

Abstract

Now a days through the advancements in the Natural language processing technology there are so many voice assistants developed to speak and interact with the people like Google Assistant, Alexa, Cortana, Siri, etc. This voice assistant central unit is its knowledge base, this helps the voice assistant to give an intelligent answer/ output to the user. Many parents due to work from home or with some other reasons unable to spend time with their children so they are sending them to play school or they are kept and they play alone, because of this there growth and mental health may be effected negatively. So with the help of knowledge base and speech recognition, pyttsx3, nltk.corpus, nltk.tokenize, we can develop a voice assistant which will be user friendly for kids and interact efficiently with kids. Then it helps to grow in a positive way so then the society will grow automatically the nation and world grows in a healthy way.

Keywords: knowledge base, speech recognition, tokenize, stop words, intents.

I.INTRODUCTION

The project title "An AI Enabled toy for interactive communication with kids" an artificial model with an artificial knowledge which will interact with children to help them in various works for example telling rhymes, alphabets, tables, etc. These voice assistants helps the working parents, baby care centre and play school to engage their children. It can take care of kids better than the care centre. The child will grow with a healthy mind. So, automatically the society will be healthy and nation will be developed. Keeps children happy through interaction. It is user friendly. The disadvantages of existing systems are they are costly. These are not explicitly designed to children. The digital voice assistant feature such as turning on/off the lights, fans, etc can may make kids lazy. Kids speak differently from adults and voice assistants does not always understand them. Children will be happy when they have a friendly interaction. And when we speak with kids it should be in soft and smooth tone and vocabulary should be simple and explanation should be simple and understanding to them. What the content the voice assistant is delivering to the children should make them feel and understand some morals, motivation, a correct point of view, unbiased, attractive, simple, etc it should keep these all things in a sentence formation for the output. Currently there are so many helpful algorithms, solutions and technologies to implement the voice assistant by following the above constraints. We have some libraries in Natural Language Processing like nltk, speech recognition, pyaudio, pysxtt3, etc. we humans don't have a constant mindset and constant mood we will be affected by our surroundings that's why the adults may be sometimes cannot communicate properly but machines won't and this will be a good thing to interact with kids. Every child's childhood is so crucial period at this time they learn what is good and bad, they will observe every interaction so a good way of communication is necessary to make the children grow in a good path and mindset so these voice assistant helps them. Every interaction should convey a useful content and should be meaningful.

II.LITERATURE SURVEY

Many people implemented voice assistants in different ways and to support and help in different scenarios for the different algorithms and approaches they have used some are as follows: [4] to support blind people and help them in various works developed by Abdolrahmani, A., Kuber, R., Branham, S. M. (2018). Siri Talks at You: An Empirical Investigation of Voice Activated Personal Assistant (VAPA), [5] developed a chatbot to help school kids to learn and resolve their doubts in various subjects simply and understandably, [6] from interaction and trust perspective evaluation based on information quality, system quality, trust, intention, personal innovativeness. [7] using Google Dialogflow and machine learning CNN algorithm, they have developed a voice assistant where firstly, they have trained the model with the intents and entities in the google Dialogflow console where there is an already in-built NLP in this console. And this model is also helpful in object detection, using the CNN algorithm and speech recognition. [8] using the Recurrent Neural Network algorithm for processing the audio, which is converted to the spectrogram, this spectrogram is going to convolute, and RNN will observe the dependencies of the audio from past and present like this it learn what part of the audio is necessary and then it formats the required output.[9]A wav2letter++ is the fastest opensource system that is 2 times faster than other optimized frameworks for neural networks speech recognition and is entirely written in C++. This framework gives a highly efficient performance. [10] keyword spotting is a major component of human technology interfaces. Maximizing the detection accuracy at a low false alarm (FA) rate, while minimizing the foot print size, latency, complexity are the goals keyword spotting towards achieving them we study Convolution Recurrent Neural Networks CRNN. Inspired by large scale state of the art speech recognition system, we combine the strength of Convolutional layers and recurrent layers.[11] Keyword spotting (KWS) is essential for enabling speech-based user interactions on smart devices. For a good user experience, KWS requires a real-time response and high accuracy. Recently, neural networks have become an attractive choice for KWS architecture because of their superior accuracy compared to traditional speech processing algorithms. Due to its always-on nature, the KWS application has a highly restrained power budget and typically runs on tiny microcontrollers with fixed memory and compute capability. The design of neural network architecture for KWS must consider these constraints. In this work, we perform neural network architecture evaluation and exploration for running KWS on resource-constrained microcontrollers. We train various neural network architectures for keyword spotting published in the literature to compare their accuracy and memory/compute requirements. We show that it is possible to optimize these neural network architectures to fit within the memory and compute constraints of microcontrollers without sacrificing accuracy. We further explore the depthwise separable convolutional neural network (DS-CNN) and compare it against other neural network architectures. [12] Describes an audio dataset of spoken words designed to help train and evaluate keyword spotting systems. Discusses why this task is an interesting challenge, and why it requires a specialized dataset that is different from conventional datasets used for automatic speech recognition of full sentences. Suggests a methodology for reproducible and comparable accuracy metrics for this task. Describes how the data was collected and verified, what it contains, previous versions and properties. Concludes by reporting baseline results of models trained on this dataset. [13] Neural machine translation is a newly proposed approach to machine translation. Unlike the traditional statistical machine translation, the aim of neural machine translation is to build a single neural network that can be jointly adjusted to maximize translation performance. The recently proposed models for neural machine translation often belong to the encoder-decoders family and consist of an encoder that encodes a given sentence into a fixed-length vector from which a decoder in turn generates a translation. In this paper, we believe that the use of a fixed-length vector is a bottleneck in improving the basic encoder-decoder architecture performance efficiently and propose to extend this by using a model which automatically (soft-) search for parts of a given sentence that are relevant to predicting a required word, without having to form these parts like a hard segment explicitly. With the help of this new approach, we achieve a translation performance comparable to the existing state-of-the-art phrase-based system for the task of English-to-French translation. Furthermore, qualitative analysis reveals that the (soft-)alignments found by the model agree well with our intuition. [14] Speech recognition is a sequence prediction problem involving sequential data. Besides employing different deep learning approaches for frame-level classification, sequence-level discriminatory training has been proved crucial to achieving state-of-the-art performance in large vocabulary continuous speech recognition (LVCSR). However, keyword spotting (KWS), one of the most common speech recognition tasks, almost only benefits from frame-level deep learning due to the hardship of getting competing sequence hypotheses. The few studies on sequence discriminative training for KWS are restricted to fixed vocabulary or LVCSR-based methods and have not compared to the state-of-the-art deep learning-based KWS techniques. This paper proposes a sequence discriminative training framework for fixed vocabulary and unrestricted acoustical KWS. Sequence discriminative training for both sequence-level generative and discriminative models was examined systematically. By introducing word-independent phone lattices or non-keyword blank symbols to create competing hypotheses, feasible and efficient sequence, discriminative training techniques are proposed for acoustic KWS. Experiments showed that the proposed approaches obtained consistent and considerable improvement in fixed vocabulary and unrestricted KWS tasks compared to previous frame-level deep learning-based acoustic KWS techniques. [15] Deep Neural Network–Hidden Markov Model (DNN-HMM) based methods have been used successfully for many always-on keyword spotting algorithms that detect a wake word to trigger a device. The DNN predicts the state probabilities of a given speech frame, while HMM decoder combines the predictions of DNN's multiple speech frames to compute the keyword detection score. The DNN, in prior methods, is trained independently of the HMM parameters inorder to minimize the cross-entropy loss between the predicted probabilities and the ground-truth state probabilities. The mismatch between the end metric (detection score) and the DNN training loss (cross-entropy) is the main source of sub-optimal performance for the keyword spotting task. We resolve this loss-metric mismatch with a novel end-to-end training strategy that learns the DNN parameters by optimizing the detection score. At last, we make the HMM decoder (dynamic programming) differentiable and back-propagate through it to maximize the score for the keyword and minimize the scores for non-keyword speech segments. Our method does not require any change in the model architecture or the inference framework; therefore, there is no overhead in run-time memory or compute requirements. [16] We describe the architecture of an always-on keyword spotting (KWS) system for battery-powered mobile devices used to initiate an interaction with the device. An always-available voice assistant needs a carefully designed voice keyword detector to satisfy battery-powered devices' power and computational constraints. We employ a multi-stage system that uses a low-power primary stage to decide when to run a more accurate (but more power-hungry) secondary detector. [17] Learning to store information over extended time intervals by recurrent backpropagation takes a very long time, mostly because of insufficient, decaying error backflow. We briefly review Hochreiter's (1991) analysis of this problem, then address it by introducing a novel, efficient, gradient based method called long short-term memory (LSTM). Truncating the gradient where this does not do harm, LSTM can learn to bridge minimal time lags in excess of 1000 discrete-time steps by enforcing constant error flow through constant error carousels within special units. Multiplicative gate units learn to open and close access to the constant error flow. LSTM is local in space and time; its computational complexity per time step and weight is O. 1. Our experiments with artificial data involve local, distributed, real-valued, and noisy pattern representations. In comparisons with real-time recurrent learning, back propagation through time, recurrent cascade correlation, Elman nets, and neural sequence chunking, LSTM leads to many more successful runs, and learns much faster. LSTM also solves complex, artificial long-time-lag tasks that have never been solved by previous recurrent network algorithms.[18] Keyword spotting refers to detection of all occurrences of any given keyword in input speech utterances. In this paper, we define a keyword spotter as a binary classifier that separates a class of sentences containing a target keyword from a class of sentences which do not include the target keyword. In order to discriminate the mentioned classes, an efficient classification method and a suitable feature set are to be studied. For the classification method, we propose an evolutionary algorithm to train the separating hyper-plane between the two classes. As our discriminative feature set, we propose two confidence measure functions. The first confidence measure function computes the possibility of phonemes presence in the speech frames, and the second one determines the duration of each phoneme. We define these functions based on the acoustic, spectral and statistical features of speech. The results on TIMIT indicate that the proposed evolutionary-based discriminative keyword spotter has lower computational complexity and higher speed in both test and train phases, in comparison to the SVM-based discriminative keyword spotter. Additionally, the proposed system is robust in noisy conditions.[19] In recent years, Artificial Intelligence (AI) has shown significant progress and its potential is growing. An application area of AI is Natural Language Processing (NLP). Voice assistants incorporate AI using cloud computing and can communicate with the users in natural language. Voice assistants are easy to use and thus there are millions of devices that incorporates them in households nowadays. Most common devices with voice assistants are smart speakers and they have just started to be used in schools and universities. The main motive of this paper is to study the capabilities of voice assistants in the classroom and to present findings from previous studies.[20] We evaluated the effects of a voice output communication aid (VOCA) and naturalistic teaching procedures on the communicative interactions of young children with autism. A teacher and three assistants were taught to use naturalistic teaching strategies to provide opportunities for VOCA use in the context of regularly occurring classroom routines. Naturalistic teaching procedures and VOCA use were introduced in multiple probe fashion across 4 children and two classroom routines (snack and play). As the procedures were implemented, all children showed increases in communicative interactions using VOCAs. Also, there was no apparent reductive effect of VOCA use within the naturalistic teaching paradigm on other communicative behaviours. Teachers' ratings of children's VOCA communication, as well as ratings of a person unfamiliar with the children, supported the contextual appropriateness of the VOCA.[21] Researchers taught a 5-year-old boy suffering from autistic disorder and severe language delay to use requests for assistance conditionally. A within-participant multiple-probe design across three functional tasks was implemented in order to evaluate the child's acquisition and use requests for assistance conditionally during intervention with each task. Results indicated an initial acquisition of requests for assistance followed by a brief period of over-generalization. As completing a task independently increased, requests for assistance decreased correspondingly. The participant's use of requests for assistance conditionally and independent task completion was sustained across time.[22] The article conveys applications of voice recognition technology based on artificial intelligence in the educational methodology. The author presents a comparative analysis of living examples of artificial intelligence in the educational process. Artificial intelligence used in specialized software makes the educational process more convenient for students and teachers. A description of an application, "Academic phrase bank," was developed by the author. The application consists of two specializing actions for Google assistant. The application allows for increasing academic vocabulary, train of creating grammatically correct academic expressions, and memorizing templates of academic phrases. In active mode, this application helps create correct phrases of academic English and improve the ability to understand English speech.[23] An educational platform employed with AI can be adjusted to meet the students need and consider their personal features (for example initial level of knowledge, speed of learning, current interests etc) to identify the most effective way of studying. Interaction of the students with a smart platform reveals their weaknesses and strengths and motivates them to additiomaking itg, so it makes possible to create a system of adaptive learning.[24] There are some popular AI-based technologies in the educational process: adaptive knowledge assessment, spaced repetition, adaptive feedback, virtual assistants, formal verification of creative works, etc. Let us consider some of these technologies' features and the mechanisms of their implementation with AI.[25] Computer-aided assessment using AI [4] implies using specialized applications that allows oneself to test and check various types of creative work automatically, such as essays, etc. Computer-aided systems analyze student answers, provide feedback, finding factual or semantic errors. The data gathered from students' answers are using for prepare an individual learning path adapted to the current knowledge of students.[26] Natural User Interface (NUI) describes interaction between a person and a computer with intuitive actions that do not require special training of users. The purpose of using a user-friendly interface is to hide the complexity of a system as much as possible. Even if the user is inexperienced or requires a rather complicated interaction, it will be able to interact with the system by NUI. Examples of interfaces with NUI include interfaces based on touch, gestures, body movements (Kinect), voice, etc. Neurocomputer interfaces are expected to be created soon.[27] VUI is becoming widely used as a communication for many various intelligent devices. For example, all modern smart speakers do not have a graphical display and keyboard, so you can communicate with them only by VUI. A virtual chatbot is a program that imitates the real interaction with a person. Interfaces of modern chatbots based on VUI. These types of interfaces are implemented like text chat in Facebook Messenger or Skype Instant Messaging for dialog with an intelligent device. There are standard development tools for creation VUI. For example, a script describing the interaction between users and a script of interaction between a user and a computer have are markably similar structure. The only exception is that the chatbot interface can display images and hyperlinks, while the VUI can only play sounds.[28] The growing prevalence of artificial intelligence and digital media in children's lives allows them to interact with novel non-human agents such as robots and voice assistants. Previous studies show that children eagerly adopt and interact with these technologies, but we have only limited evidence of children's distinction between artificial intelligence and humans. This study investigated the communication patterns and prosocial outcomes of interactions with voice assistants.[29] Voice control is a major growing feature that change the way people can live. The voice assistant is commonly being used in smartphones and laptops. AI-based Voice assistants are the operating systems that recognize human voices and respond via integrated voices. This voice assistant will gather the audio from the microphone and then convert that into text, later it is sent through GTTS (Google text to speech). GTTS engine will convert text into audio file in English, then that audio is played using the play sound package of python programming Language.[30] Intelligent Personal Assistants are poised to become the primary and most significant Human-Computer Interface in the near future. This is attributed to advancements in Artificial Intelligence, Machine Learning, the Internet of Things, Natural Language Processing, and Data Sciences. A range of generic as well as specialized IPAs is being researched and developed by the industry. Despite security and privacy issues, the adoption and utility of these agents are increasing. This paper proposes a singular adaptive multi-role IPA (SAM-IPA) that goes beyond scheduling and search facilities to handle multi-dimensional IoT as well as application data. SAM-IPA will not only act as a singular HCI responsible for delegating multifarious tasks on behalf of the user but will also increase awareness via IoT and adapt based on ML over BigData. The proposed SAM-IPA will leverage the application interface communication mechanisms and technology potentials in the foreseeable horizon, drawing concrete findings while identifying research areas needing deliberation.

III.PROPOSED SOLUTION

So many people have developed different kinds of voice assistants for interactive communication. But this project aims to develop an efficient model to understand and interact with children. For this we developed the model specially thinking about the children ideology and in simple way they can understand. Firstly, we have built a knowledge base where it has the whole information it needed and the experience it needed. To make an artificial model intelligent we input knowledge base through which our voice assistant not only becomes intelligent but also it can give an intelligent answer to the user query. Voice assistant firstly takes a voice input from the user with the help of a microphone next through the user voice frequency some electronic signals are going to create by this it helps to convert this voice into text. Firstly, it breaks the audio into small units then it will search the appropriate word to each audio like this it is going to transcribe the audio into text. Then from this text we need to extract the important words called intents. So there will be so many unimportant words like is, that, this, to, etc. called stop words for this we used a library nltk.corpus from where we import stopwords, before this we need to separate each word in the query means forming a list then we can easily remove stopwords. Then we perform stemming operation on each important intent stemming means reducing the words into its root form. Then next process is lemmatization which means every word has so many synonyms and vocabulary so that from this to reach a base word we perform lemmatization and the word is called lemma. Next step is to find the required information from the knowledge base to give useful output. Knowledge Base is like the human brain where the total information is stored and to give an intelligent response the knowledge base is the main reason and through this only the model is going to act intelligently. We search Knowledge Base with the intents like where are the intents are present in the base and gather the information then this gathered information which is in text format is going to convert into voice with the help of speech recognition module implementation this process is going to done both text to voice and voice to text.

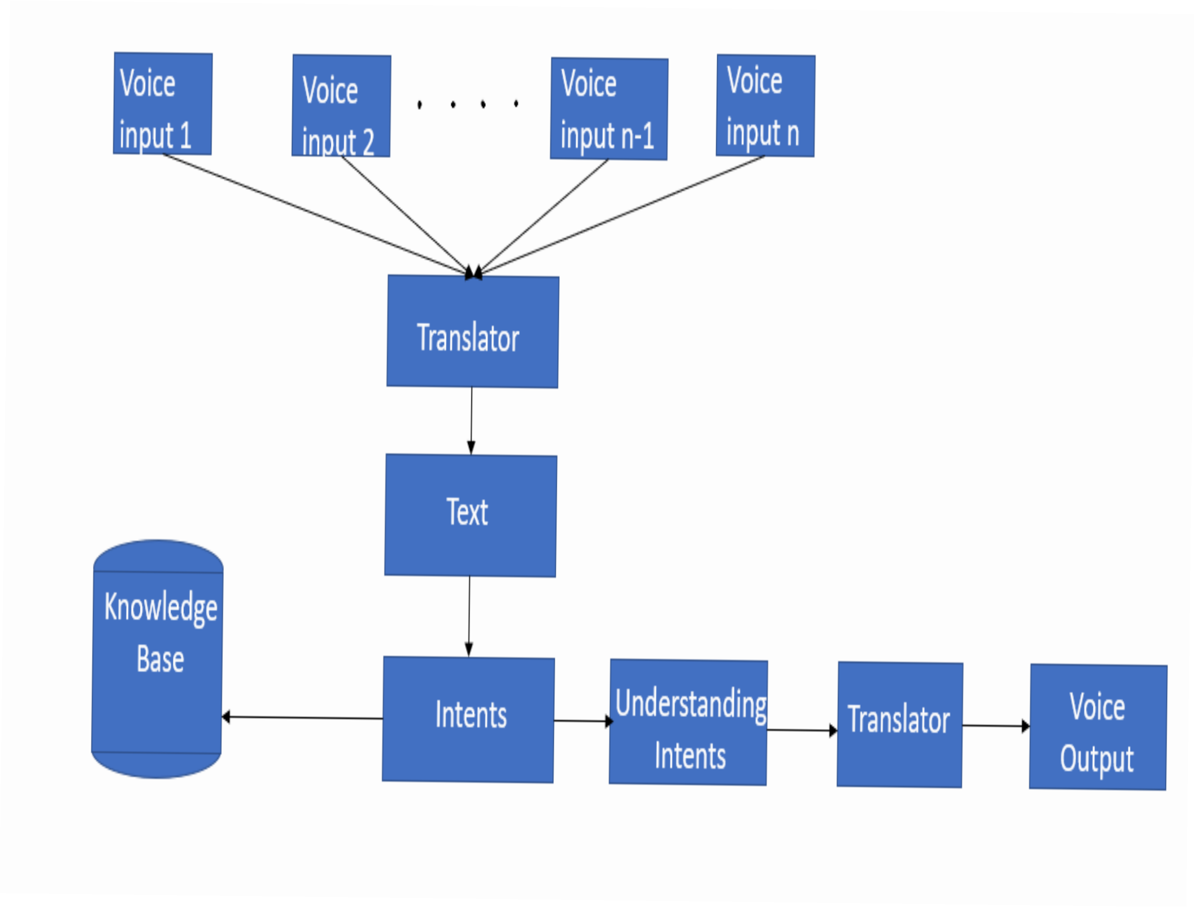
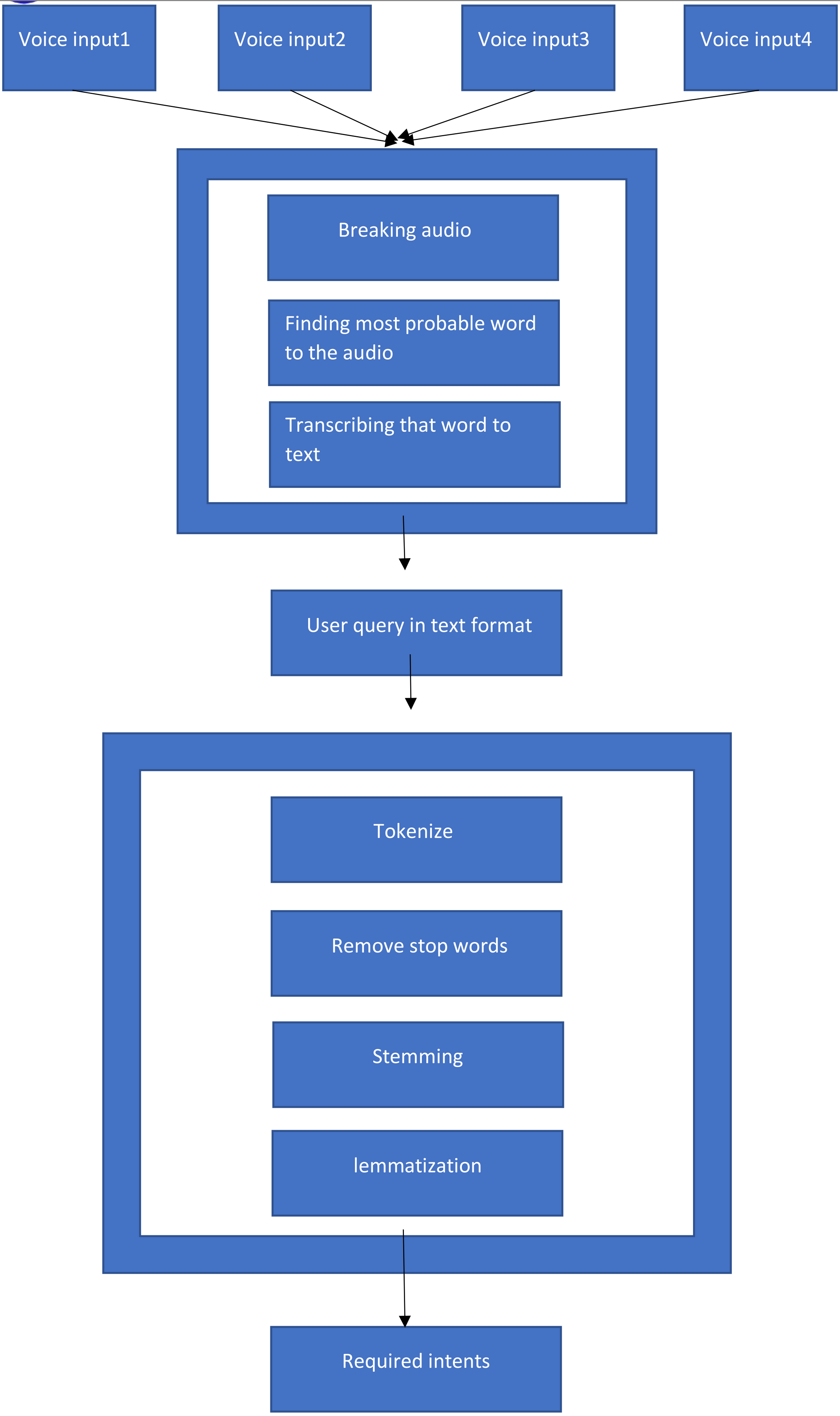


Figure 1: Architecture of voice assistant


Voice input1
Voice input2
Voice input3
Voice input4
Breaking audio
Finding most probable word to the audio
Transcribing that word to text
User query in text format
Tokenize
Remove stop words
Stemming
lemmatization
Required intents

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

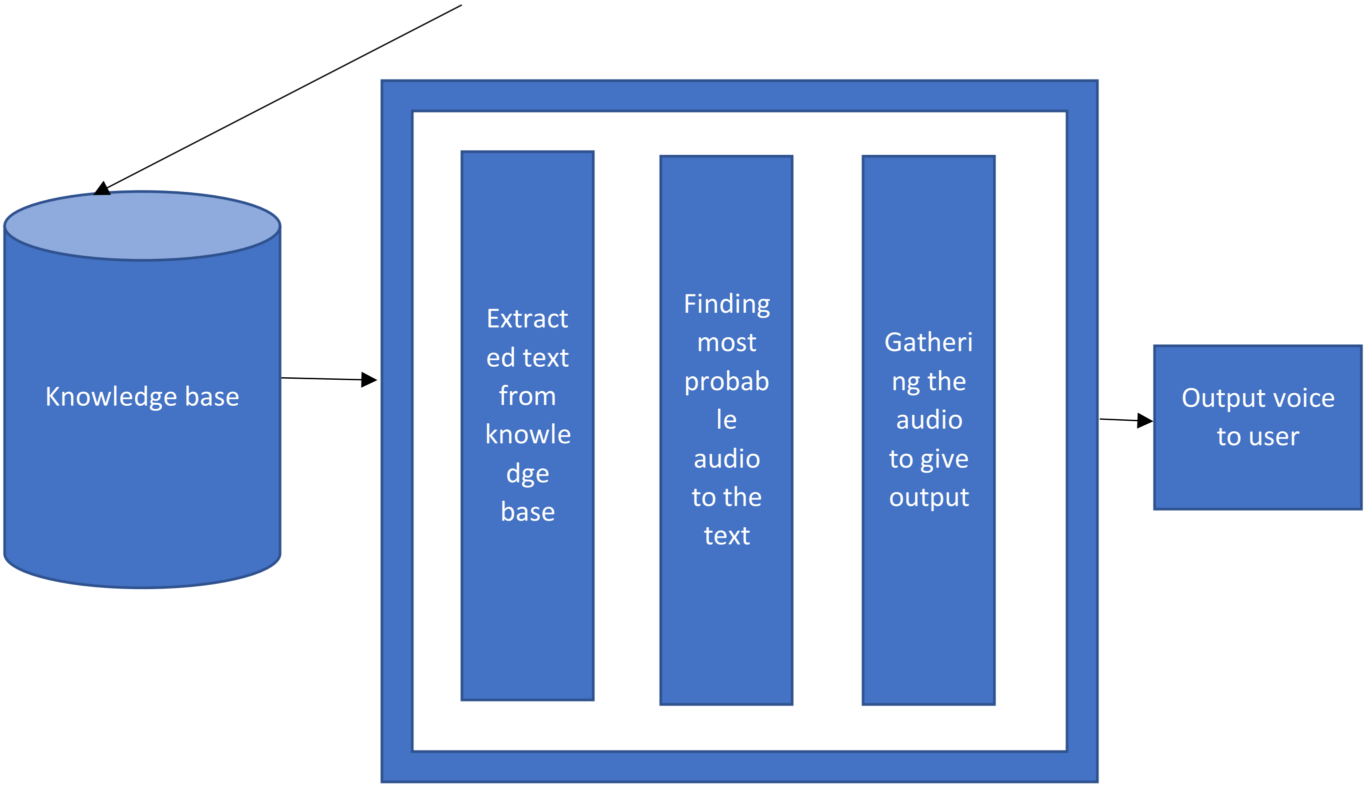


Figure 2: Detailed Architecture of voice assistant

Execution steps are as follows

Step 1: Open the web interface and click the button to give voice input.

Step 2: Give the query input as your voice. The microphone takes the input.

Step 3: Now the model converts the voice to text and tokenizes the query and removes the stop words and extracts intents and with these it will search the knowledge base and gathers the information.

Step 4: Converts this gathered information from text to voice and outputs the information to user.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

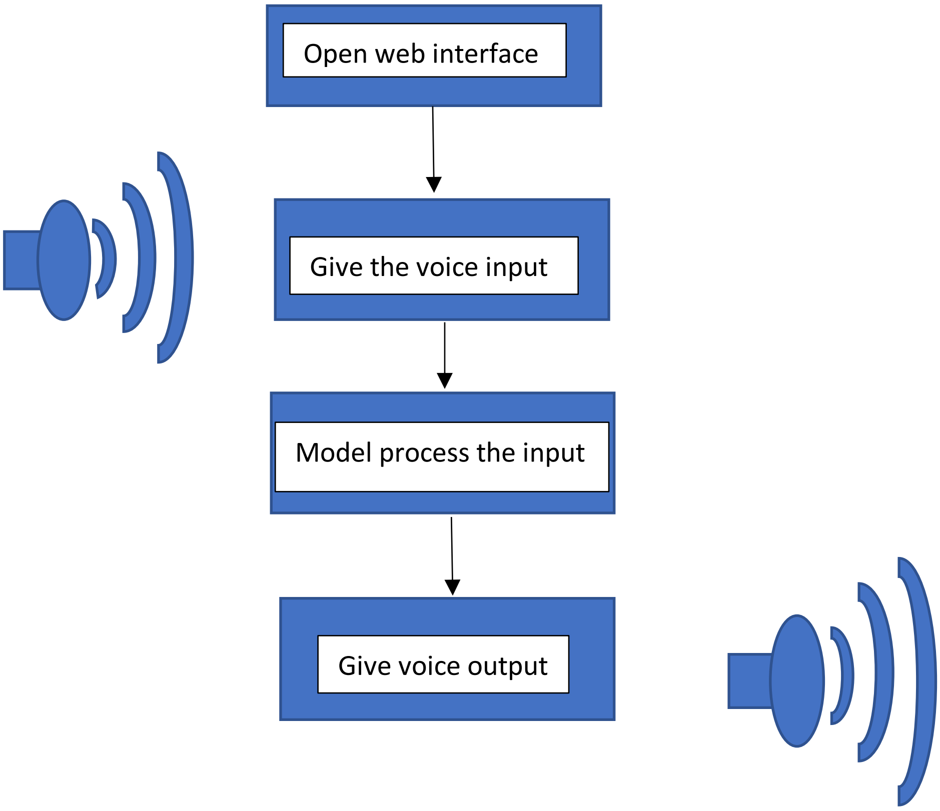


Figure 3: Execution flow of the voice assistant

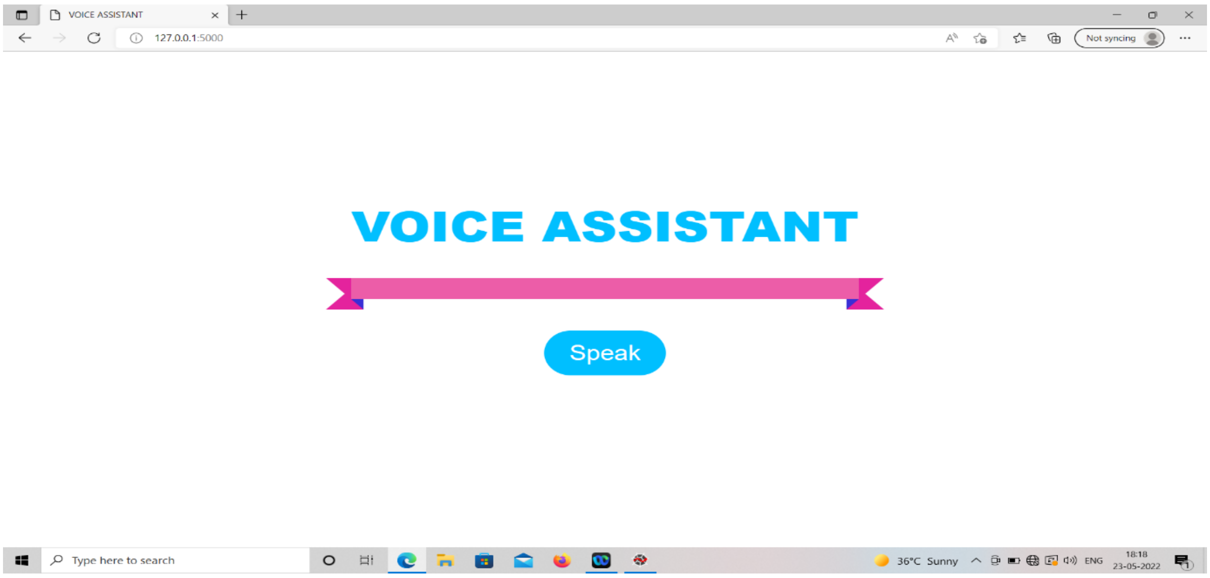


Figure 4: Web Interface

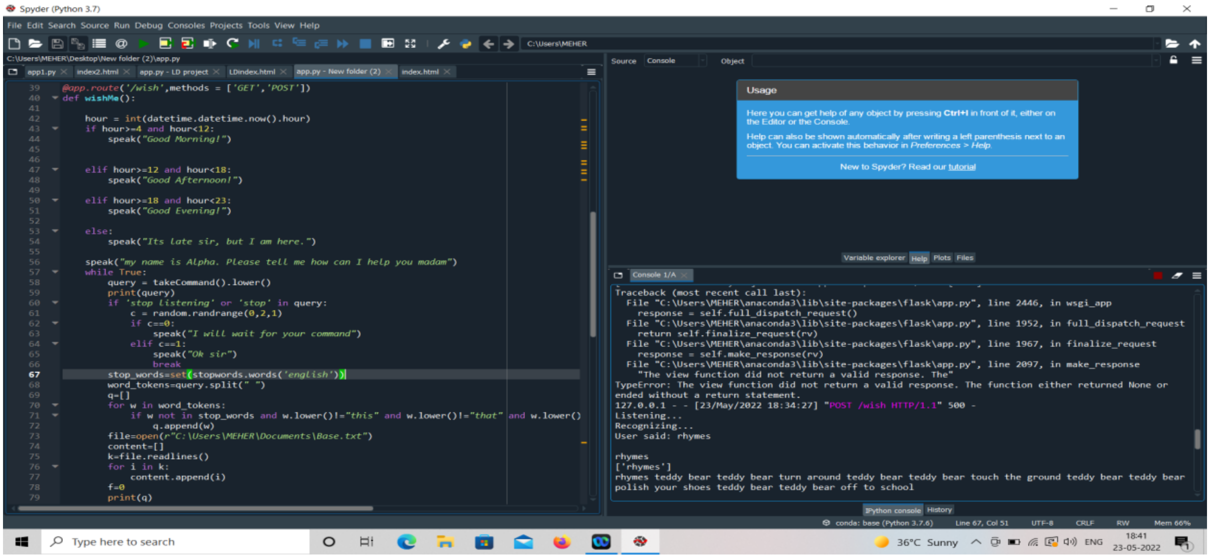


Figure 5: output for rhymes

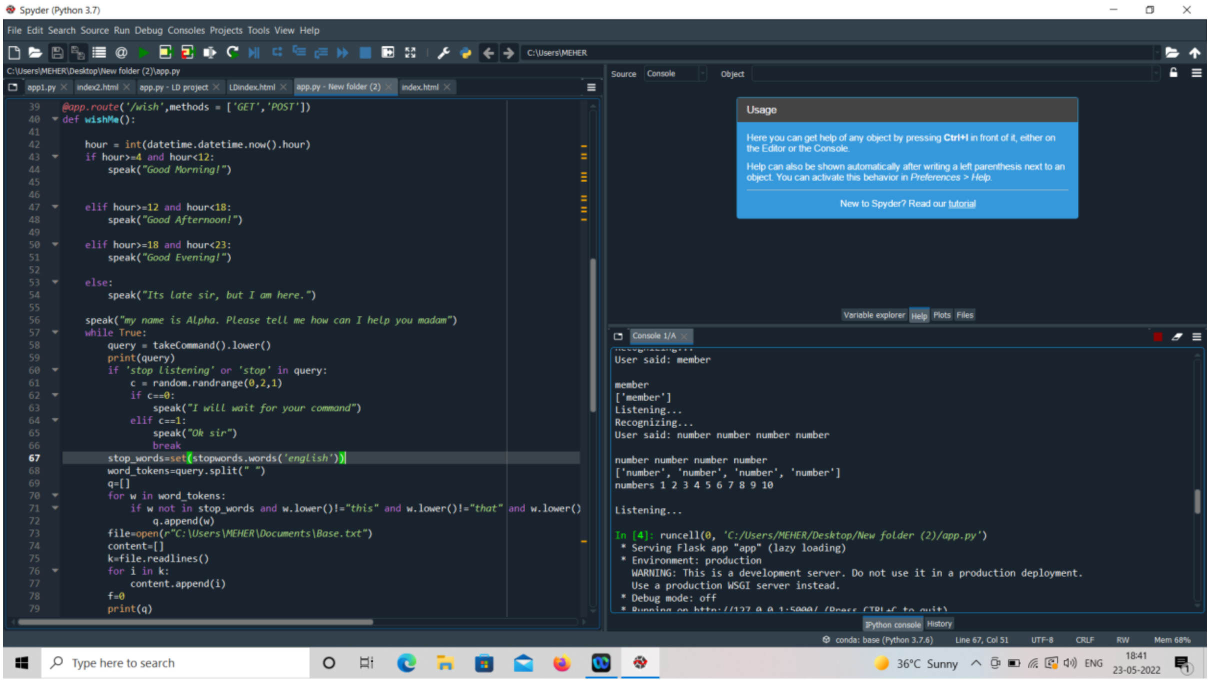


Figure 6: output for numbers

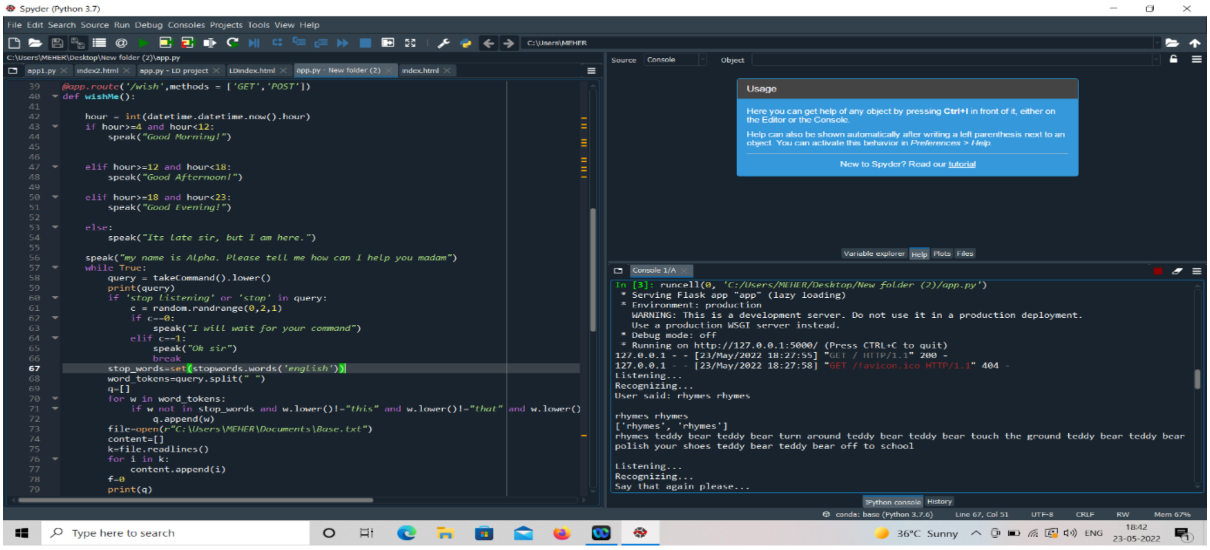


Figure 7: output displaying the local host url

Table 1: Results

Test case	Expected	Responded	Response time	Accuracy
1. tell me numbers	Numbers	Numbers	10 sec	100%
2. tell me alphabets	Alphabets	Alphabets	10 sec	100%
3. sing teddy bear rhyme	Teddy bear rhyme	Teddy bear rhyme	10 sec	100%
4.tell me a fairy story	Fairy tales(stories)	Fairy tales	10 sec	100%
5. tell me 9 table	9 table	9 table	10 sec	100%
6. how many bones are there in our body	Bones count	Bones count	10 sec	100%
7. how many planets are there and their names	Count and names	Count and names	10 sec	100%
8.hahoheehoyyaaa	Say again I can’t understand	Sorry I cloud not get that say that again please	15 sec	100%
9. what is VIBGYOR?	Colors names	Colors names	10 sec	100%
10. tree parts	Names of tree parts	Names of tree parts	10 sec	100%

CONCLUSION

This model efficiently interact with children and helps to grow with a good mental health. We can use this assistant whenever we want and where ever we want. This model help to handle the children in the healthcare and fulfill there requirements through logical interaction. This model help the caretakers and nurses in hospital.

REFERENCES

1. DOUGLAS O'SHAUGHNESSY, SENIOR MEMBER, IEEE, "Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis" proceedings of THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003 Kei Hashimoto1, Junichi Yamagishi2, William Byrne3, Simon King2, Keiichi Tokuda, "An analysis of machine translation and speech synthesis in speech-to-speech translation system" proceedings of 5108978-1-4577-0539
2. "On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant" International Journal of Human Sciences, 13(1), 592-601. doi:10.14687/ijhs.v1311.3549. Harshita Phatnani, Mr. Jyotiprakash Patra and Ankit Sharma' "CHATBOT ASSISTING: SIRI" Proceedings of BITCON-2015 Innovations For National Development National Conference on Research and Development in Computer Science and Applications, E-ISSN2249-8974. Sutar Shekhar, Pophali Sameer, Kamad Neha, Prof. Devkate Laxman, " An Intelligent Voice Assistant Using Android Platform", March 2015, IJARCSMS, ISSN: 232
3. Kei Hashimoto1,junichi yamagishi2,Wlliam Byner3," AN analysis of machine translation and system synthesis in speech to speech translation system"
4. Abdolrahmani, A., Kuber, R., Branham, S. M. (2018). Siri Talks at You: An Empirical Investigation of VoiceActivated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility
5. Bekmyrza, K. (2019). Using Chatbot to Increase Students' engagement in Education Process at High School. https://www.internauka.org/archive2/vestnik/11(61_2)
6. Farzaneh Nasirian, Mohsen Ahmadian, One-ki (Daniel) Lee AI-Based Voice Assistant Systems: Evaluating from the Interaction and Trust Perspectives.
7. Dr. Jaydeep Patil, Atharva Shewale, Ekta Bhushan, Alister Fernandes, Rucha Khartadkar. (2021) A Voice Based Assistant Using Google Dialogflow and Machine Learning. Journal International Journal of Scientific Research in Science and Technology.
8. Martin Loesener Da Silva Viana (2018) A Neural Attention Model for Speech Command Recognition
9. Vinnel Pratap, Awni Hannun, Ronan Collobert (2018) Wav2letter++: The fastest Open-source Speech Recognition System
10. Sercan Arik, Markus Kliegl, Rewon Child, Joel Hestness (2017) Convolution Recurrent Neural Networks for Small-Footprint Keyword Spotting
11. Naveen Suda, Vikas Chandra, Yundong Zhong, Liangzhen Lai (2018) Hello Edge: Keyword Spotting on Microcontrollers
12. Pete Warden (2018) Speech Commands: A Dataset for Limited-Vocabulary Speech Recognition

13. Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio (2016) Neural Machine Translation by Jointly Learning to Align and Translate
14. Zhehuai Chen, Yanmin Qian, Kai Yu (2018) Sequence discriminative training for deep learning based acoustic keyword spotting
15. Ashish Shrivastava; Arnav Kundu; Chandra Dhir; Devang Naik; Oncel Tuzel
Optimize what matters: Training DNN-HMM keyword spotting model using end metric
2021, ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings
https://ieeexplore.ieee.org/document/9414797
16. S Sigtia, R Haynes, H Richards, E Marchi, J Bridle - Interspeech, (2018) Efficient Voice Trigger Detection for Low Resource Hardware.
17. Hochreiter, S., Schmidhuber, J., Nov. 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. http://dx.doi.org/10.1162/neco.1997.9.8.1735
18. Tabibian, S., Akbari, A., Nasersharif, B., 2013. Keyword spotting using an evolutionary-based classifier and discriminative features. Engineering Applications of Artificial Intelligence 26 (7), 1660 – 1670.
19. George Terzopoulos, Maya Satratzemi (2019) Voice Assistants and Artificial Intelligence in Education https://dl.acm.org/doi/10.1145/3351556.3351588
20. M M Schepis , D H Reid, M M Behrmann, K A Sutton "Increasing communicative interactions of young children with autism using a voice output communication aid and naturalistic teaching" https://pubmed.ncbi.nlm.nih.gov/9891394/
21. Teaching a young child with autism to request assistance conditionally: a preliminary study. Reichle J, Dropik PL, Alden-Anderson E, Haley T. Am J Speech Lang Pathol. 2008 Aug;17(3):231-40. doi: 10.1044/1058-0360(2008/022). PMID: 18663108
22. Vladimir Tregubov, Department of Business and Logistics, Yuri Gagarin State Technical University of Saratov, Saratov, Russia "Voice Assistants with Artificial Intelligence for Improving Academic English"
23. S. Popenici and S. Kerr, "Exploring the impact of artificial intelligence on teaching and learning inhigher education," Res. Pract. Technol. Enhanc. Learn., vol. 12, no. 1, p. 22, Dec. 2017, doi:10.1186/s41039-017-0062-8
24. T. Karsenti, "Artificial intelligence in education: The urgent need to prepare teachers for tomorrow'sschools," Form. Prof., vol. 27, no. 1, p. 105, 2019, doi: 10.18162/fp.2019.a166
25. M. J. Timms, "Letting Artificial Intelligence in Education out of the Box: Educational Cobots andSmart Classrooms," Int. J. Artif. Intell. Educ., vol. 26, no. 2, pp. 701 – 712, 2016, doi:10.1007/s40593-016-0095-y.
26. Berdasco, López, Diaz, Quesada, and Guerrero, "User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana," Proceedings, vol. 31, no. 1, p. 51, Nov. 2019, doi: 10.3390/proceedings2019031051
27. C. Pearl, Designing Voice User Interfaces by Cathy Pearl Chapter 1 . Introduction A Brief History ofVUIs. O'Reilly Media, Inc., 2016.
28. Communicative and social consequences of interactions with voice assistants (2020) Sara Aeschlimanna, Marco Bleikera, Michael Wechnerbc, Anja Gampe.
29. Artificial Intelligence-based Voice Assistant(2020) S Subhash; Prajwal N Srivatsa; S Siddesh; A Ullas; B Santhosh
30. Singular Adaptive Multi-Role Intelligent Personal Assistant (SAM-IPA) for Human Computer Interaction(2018) Khawir Mahmood; Tauseef Rana; Abdur Rehman Raza.