iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IV **Month of publication:** April 2021

DOI: https://doi.org/10.22214/ijraset.2021.33663

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

Next Generation AI based Virtual Assistant

Parth Kaushik[1], Shipra Srivastava[2], Adarsh Tripathi[3], Abhinav Sharma[4], Ayush Raj[5]
[1, 3, 4, 5]*B.Tech Student,* [2]*Assistant Professor, Greater Noida Institute of Technology, Greater Noida, India*

Abstract: ***This report discusses ways within which new technology may be controlled to form AN intelligent Virtual Assistant (VA) with attention to user-based data. It will look at examples of intelligent programs with natural language processing that are currently available, with different categories of support, and examine the potential usefulness of one specific piece of software as a VA. This engages the ability to communicate socially through natural language processing, holding (and analyzing) information within the context of the user. It is suggested that new technologies may soon make the idea of virtual assistants a reality. Experiments conducted on this system, combined with user testing, have provided evidence that a basic program with natural language processing algorithms in the form of a VA, with basic natural language processing and the ability to function without the need for another type of human input (or programming), may already be viable. We tend to check this on two variant speakers(masculine & feminine).***

Index Terms: Artificial Intelligence, Virtual assistant, natural language processing, NLU engine, Speech text recognition.

I. INTRODUCTION

The objective of our paper is to demonstrate how An artificial Conversational entity (ACE) may be a programmed to communicates with you. It's a layer on high of, or an entry to a service. Typically it's supercharged by machine learning (the artificial speech communication entity gets smarter the additional you act with it). Or, additionally usually, it's driven mistreatment intelligent rules (i.e. if the person says this, respond with that). The services a synthetic speech communication entity will deliver square measure numerous, necessary life-saving health messages, to ascertain the prognosis or to get a brand new combination of shoes, and anything in between.

The term artificial speech communication entity is similar to text speech communication however is growing quickly through voice communication… "Alexa, what time is it?" (other voice-artificial speech communication entities square measure available!) the unreal speech communication entity will confer with you thru totally different channels; like a Facebook traveler, Siri, WeChat, Telegram, SMS, Slack, Skype, and plenty of others. Customers pay several time mistreatment electronic communication applications (more than they pay on social media) [1]. Therefore, electronic communication applications square measure presently the foremost widespread manner firms deliver artificial speech communication entity experiences to customers.

An Artificial colloquial Entity may be a quiet chatbot that speaks with you helps you to speak over the net helps you to search out the knowledge you required to mistreatment the net. In this, we tend to use python language & its library operates for making our Artificial colloquial Entity for Windows software. it'll not solely give data from the net however it'll conjointly open the appliance put in in your system.

If you say some word to our ACE it'll search it over Wikipedia further as play the most-watched video over YouTube. mistreatment our ACE, you'll be able to even head to a specific website by gap a browser We named our ACE as JAMWANT as a result of some Hindi author afore mentioned("दुनियाकेआरम्भसेजहादुनियाकाअंतहैंसबदेखचूकायहजामवंतहै।".

In this digital era humans are completely dependent on technology. Now after so many great inventions and creations the industry is heading towards smart-bots enabled smart homes, voice assistants, A.I. and Data Mining enabled systems as well. We felt the requirement of a smart assistant chat-bot which will be economic and easy to access and afford for a common use and will let you up to date with the current weather updates and forecasts.

It will provide any information on demand which the user generally search on google, etc. It will solve your basic mathematical problems as well so that you can get rid of your old and outdated calculators [2]. As per requirements and growing industry our chat-bot will work on voice command and will perform a lot more for its users on demand. And also give you information from Wikipedia about a person if available and most watched video of the person.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue IV Apr 2021- Available at www.ijraset.com

II. MAJOR ENTITIES OF AI BASED VIRTUAL ASSISTANT :

The proposed system will have the following entities:

A. NLU Engine

The conversational language used by humans for day to day conversations is not as perfect as the formal language. The input received from the user is in unstructured text format which cannot be understood by the system directly. The unstructured text received from the user is converted to structured format by extracting important words and patterns from the user text using the NLU techniques. NLU techniques enables the system to identify these twerks if the user makes use of them while conversing with the artificial conversation entity, so as to make the user feel that the conversation is taking place between two humans and not between a human and a bot.

NLU systems do not directly understand the meaning of the user sentences. It involves a sequence of processes to derive the actual intent of the sentence. To understand a complete sentence, the NLU system needs to understand each word of that sentence. After knowing the grammatical weightage of each word, all of them are parsed to know the dependency among them. This is the most important step wherein the word with the highest dependency is extracted, from which the intent of the system can be known. It is not possible that the knowledge base would contain the exact sentence that the user has sent. It might contain a sentence with the same intent but with different words used in it.

B. Speech -Text Representation

The discourse sign and each one in all its qualities will be spoken to in 2 distinct areas, the time and the repetition space A discourse signal may be a step by step time differing signal as in, once inspected over a short timeframe (somewhere within the vary of five and a hundred MS), its attributes square measure short-time stationary. this is not matters if we tend to take a goose at a discourse signal below an additional drawn-out time purpose of reading (around time T>0.5 s). Right now, flag attributes square measure non stationary, implying that it changes to mirror the assorted sounds spoken by the utter to own the choice to utilize a discourse flag and decipher its qualities during a legitimate method a portrayal of the discourse signal are liked. Triple State Portrayal: The Triple State Portrayal is one approach to cluster occasions in discourse. The occasions of enthusiasm for the three-state portrayal are:

1) *Silence:* No audio received
2) *Unvoiced:* Vocal cords aren't moving, transfer concerning associate degree non-oscillatory or capricious discourse wave.
3) *Voiced:* Vocal strings square measure strained and moving intermittently, transfer a couple of discourse waves that are semi occasional. Quasi-periodic implies that the discourse wave will be viewed as intermittent over a short timeframe amount (5-100 ms) throughout that it's stationary [2][3].

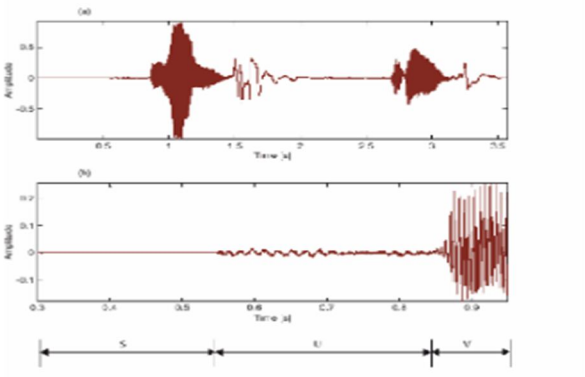


Fig. 1: Triple State Portrayal

The upper plot (a) contains the entire discourse grouping and in the center plot (b) a piece of the upper plot (an) is imitated by zooming a territory of the discourse succession. At the base of Fig.1 the division into a three-state portrayal, according to the various pieces of the center plot is given. The division of the discourse wave forming to all-around characterized states isn't straight forward. In any case, this trouble isn't like a major issue as one can suspect.

©IJRASET: All Rights are Reserved

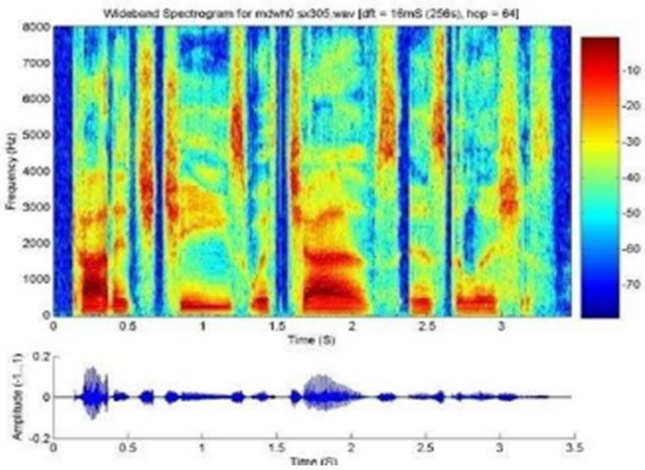


Fig. 2: Spectrogram Using Welch's Method (a) and Speech Amplitude (b)

Here the darkest (dark blue) components represent the components of the speech wave a pair of wherever no speech is produced and therefore the lighter (red) components represent intensity if speech is created. Speech wave is given within the time domain. For the photograph, Welch's technique is employed, which uses averaging changed period grams . Parameters utilized in this technique are block size K= 320, window sort Hamming with sixty-two. 5% overlap leading to blocks of twenty MS with a distance of six.25 MS between blocks. Phonetics& descriptive linguistics The discourse creation starts within the people's mind once the person in question frames a thought which will be delivered and emotional to the audience. within the wake of getting framed planl|theperfect|the best} idea, the individual in question builds Associate in Nursing expression/sentence by choosing Associate in Nursing assortment of restricted unrelated sounds. The essential hypothetic unit for representational process the way to carry linguistics significance to the formed discourse, within the brain, is termed phonemes. Phonemes may be viewed as a technique for a way to talk to the assorted components in an exceeding discourse wave, delivered through the human vocal system and separated into continuant consonant (stationary) or non - continuant consonant components. A speech sound is continuous if the discourse sound is made once the vocal tract is in an exceedingly consistent state. Associate in Nursing In an inverse of this expression, the speech sound is non-continuant once the vocal tract changes its attributes throughout the creation of discourse. as an example, if the region within the vocal tract changes by opening and motion the mouth or moving your tongue in varied states, the speech sound representational process the discourse created is non-continuant Phonemes may be assembled addicted to the properties of either the time wave or repeat attributes and organized in varied sounds delivered by the human vocal tract [4]. The order might likewise be viewed as a division of the areas in Fig three.

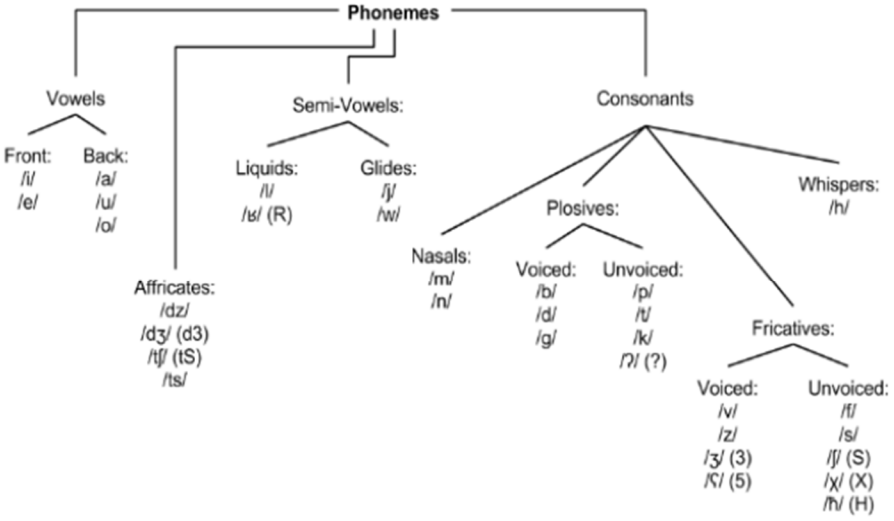


Fig. 3: Phoneme Classification

©IJRASET: All Rights are Reserved

C. Mel Frequency Cepstral Constant

The extraction of the most effective constant portrayal of acoustic signs may be a vital assignment to deliver a superior acknowledgment execution. The proficiency of this stage is important for the following stage since it influences its conduct. MFCC depends on human hearing recognition that can't see frequencies over 1Khz. At the top of the day, in M As appeared in Figure four, MFCC comprises seven procedure advances. every progression has its capability and numerical methodologies as talked regarding quickly within the incidental to a square measure supported an identified form of the human ear's basic transmission capability with repeat [5]. MFCC has 2 forms of the channel that square measure divided directly at low recurrence beneath one thousand rates and exponent separating higher than 1000Hz.

III. PROPOSED ARCHITECTURE

Jamwant has various branches of service but the main feature of Jamwant is it will also play the most-watched video with the search result. Jamwant is designed for our housewives those searches the things with name directly clicking on virtual assistant then they select the video then it plays and video that is shown to them is based on relevance but our virtual assistant directly plays the most-watched video from YouTube. Jamwant responds to basic commands like Open Applications, Close Applications. Operations such as Browsing or Searching for any topic, using Applications that need internet connections like “open site geeksforgeeks.com” will open the site on our web browser.

Steps to start our virtual assistant

A. Install our application on your windows system.
B. After installation click on the icon named Jamwant.
C. Then it will start interacting with you if you are connected over the internet.

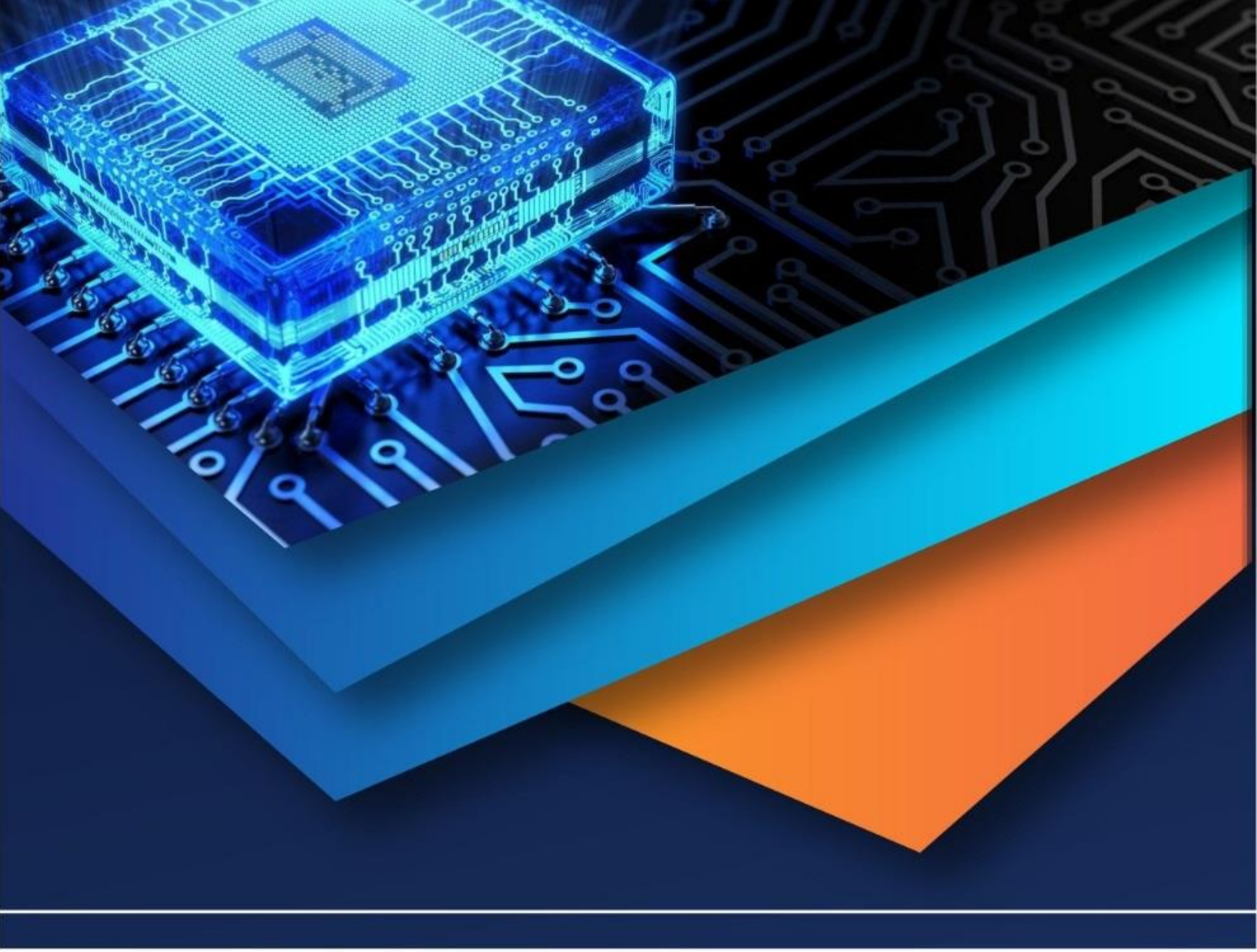
IV. CONCLUSION

A smart virtual assistant will allow the user to access all applications that are installed in your system. It will not only just display google search results it will also play the most-watched video related to your search result. It will answer some of your questions by itself. It will allow you to directly open the site on a browser. It has two speakers one masculine and the other feminine and it will speak in the voice that the user wants. It has two languages that it can speak are Hindi and English. It has various functionalities of windows devices like network connection and managing various applications on just the voice commands. Contains key features like Voice Pattern Detection, Keyword Learning, etc. which helpful for end-user to use various functionalities and services of the windows devices. Hence, Jamwant is language barrier independent which actively responds to user’s voice commands faster than the Online Voice Search applications.

REFERENCES

[1] Christina L. McDowell, Edward Forrst, " The Impact of Artificial Intelligence and Virtual Personal Assistants on Marketing", Encyclopedia of Information Science and Technology, Fourth Edition (10 Volumes), DOI: 10.4018/978- 1-5225-2255-3.
[2] Matthew B Hoy, Ariel F Pomputious." Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants.", MEDICAL REFERENCE SERVICES QUARTERLY , 2018, VOL. 37, NO. 1, 81–88,https://doi.org / 10.1080/ 02763869.2018.1404391
[3] Amrita Sunil Tulshan, Sudhir Namdeorao." Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa:" 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018
[4] Gamal Bohouta, Veton Z Kopuska ." Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home).", IEEE CCWC 2018, The 8th IEEE Annual Computing and Communication Workshop and Conference, University of Nevada, Las Vegas, USA.
[5] Peter Imrie, Peter M. Bednar. "Virtual Personal Assistant." ItAIS 2013, The 10th Conference of the Italian Chapter of AIS, Milan, Italy

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)