iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VI **Month of publication:** June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6442

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: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

Smart Voice Assistant for IOT

Bibi Asma Desai[1], Dr. B. N Veerappa[2]
[1, 2]*Department of Computer Science and Engineering, UBDTCE*

Abstract: IOT has gained the major interest of people in the present days as the innovative achievements have occurred in this area. This results the growth in numbers of IOT devices installed which give rise to the challenge of data acquisition, management and inter-communication among these IOT devices. To make over this, we propose cloud-based IOT platforms which focus on bestowing the higher-priority IOT devices and also peak reduction through solar energy by executing the device supply management algorithm along with the integration of MQTT which is the avant-grade IOT protocol.
Keywords: IOT, MQTT, MATLAB.

I. INTRODUCTION

Now days the improvement of man-made reasoning like AI frameworks that can compose a characteristic human-machine association (through voice, correspondence, signals, facial articulations, and so forth.) are picking up in prevalence. A standout amongst the most examined and mainstream was the course of communication, in view of the comprehension of the machine by the machine of the common human language. It is never again a human figure out how to impart with a machine, yet a machine figures out how to speak with a human, investigating his activities, propensities, conduct and attempting to become his customized aide (personal assistant).

The work on making and improving such customized aides has been continuing for quite a while. These frameworks are always improving and improving, go past close to home PCs and have as of now solidly settled themselves in different cell phones and devices. A standout amongst the most prominent voice collaborators are Siri, Amazon Echo, which reacts to the name of Alex from Amazon, Cortana from Microsoft, Google Assistant from Google, and the as of late seemed clever collaborator from Yandex, under the name "Alice".

Man-made consciousness advances are starting to be effectively utilized in human life, this is encouraged by the appearance what's more, wide spread of the Internet of Things (IoT). Independent gadgets are getting to be more intelligent in their manner to cooperate with both a human and themselves. New limits lead to creation of different frameworks for mix of shrewd things into Social Systems of the Internet of Things. One of the applicable patterns in man-made consciousness is the innovation of perceiving the characteristic language of a human. New bits of knowledge in this subject can prompt new methods for normal human-machine communication, in which the machine would figure out how to comprehend human's language, changing and connecting in it. One of such devices is voice partner, which can be incorporated into numerous other astute frameworks.

Each organization designer of the intelligent aide applies his own particular strategies and methodologies for advancement, which thus influences the final item. One associate can orchestrate discourse all the more qualitatively, another can more precisely and without extra clarifications and rectifications perform assignments, others can play out a smaller scope of assignments, however most precisely and as the client needs. Clearly, there is no all-inclusive assistant who might play out all assignments similarly well. The arrangement of attributes that assistant has depends altogether on which zone the designer has paid more consideration. Since all frameworks depend on AI strategies and use for their creation colossal measures of information gathered from different sources and afterward prepared on them, an significant job is played by the wellspring of this information, be it seek frameworks, different data sources or informal communities. The measure of data from various sources decides the nature of the associate, which can result therefore. In spite of the various ways to deal with learning, various calculations and strategies, the standard of structure such frameworks remains roughly the equivalent. Figure 1 demonstrates the innovations that are utilized to make intelligent frameworks of collaboration with a human by his common language.

The principle advancements are voice actuation, programmed discourse acknowledgment, Teach-To-Speech, voice biometrics, discourse director, normal language understanding and named substance acknowledgment.

©IJRASET: All Rights are Reserved

VOICE TECHNOLOGY	BRAIN TECHNOLOGY
Voice Activation	Voice Biometrics
Automatic Speech Recognition (ASR)	Dialog Management
Teach-To-Speech (TTS)	Natural Language Understanding (NLU)
	Named Entity Recognition (NER)

Figure 1: Innovations for building clever frameworks of collaboration with a human by natural language.

A. Problem Statement

New era is beginning with regards to our relationship with the machines compared to humans i.e., Artificial Intelligence. One of the pertinent trends in Artificial intelligence is recognizing the natural human's language. Voice assistant is prime existing tool among the human machine interaction. The major flow with this existence method is that dependence on Internet and cloud services, which in turn fail to concentrate on solutions to specific problems and lack of private data security. Our challenge is to develop local voice assistant with data privacy. This challenge is taken as “Voice enabled smart IOT”.

B. Proposed System

We consider development of alice voice assistant for understanding the basic arrangement of our proposed system.
"Alice" is an insightful collaborator (intelligent assistant) for cell phones and PCs, which permits to understand normal undertakings of clients, for example, seeking data on the Internet, finding puts on the guide, directing courses, detailing climate figure. For this situation, "Alice" can bolster the discussion, engage the client, and so forth. To do this, "Alice" utilizes the cloud assets of the organization Yandex, to which it alludes by means of the API through the Web. The plan of work is exhibited in Figure 2.

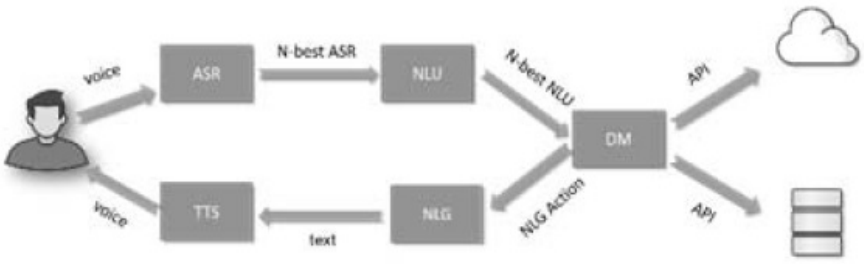


Figure 2: The plan of work of the voice colleague "Alice"

At the principal arrange, enactment happens, for instance, by articulating a key expression. The collaborator continually tunes in around the encompassing sounds, breaks down the nearness of the key expression and, in the event that it is perceived, goes into the dynamic mode. Next, the client says the content, which can disclose to the partner what the client needs to do. The programmed discourse acknowledgment framework transforms the content into N-best theories of what the client said. At that point the characteristic language understanding framework turns the content into N-best alternatives for understanding the client's expression, at that point the discourse motor deciphers and orders these expressions furthermore, figures out what should be done dependent on the data gotten. For instance, contact different administrations for data. In the wake of getting the fundamental information, the framework plays out a procedure of returning data to the client, for example the common language age framework creates content for the client's reaction, at that point the Voice-to-Speech (Text-To-Speech) framework creates sound data dependent on the prepared models, which is declared to the client as a reaction. Notwithstanding a reaction, any move can likewise make place on a cell phone or PC.

One of the significant pieces of the voice right hand from the perspective of the useful is the exchange administrator. There are straightforward contents that can be quickly extricated from the NLU model and duplicated by means of NLG

©IJRASET: All Rights are Reserved

C. Objectives

Artificial Intelligence has gained major role in human life which can be facilitated by dissemination of Internet of Things (IOT). Voice assistant is the existing tool for which understand human language. But it has limitations hence we proposed smart voice assistant and objectives of proposed system as follows:

1) To develop smart voice assistant without internet and cloud services (i.e. local smart voice assistant).
2) To build smart voice assistant for distinct and more specialized personal assistants with narrower tasks.
3) To enable privacy of the private data through local dictionary with less infrastructure.

II. LITERATURE SURVEY

In 2015, Assefi M. [1], performs a test assessment of two prevalent cloud-based discourse acknowledgment frameworks. Cloud-based discourse acknowledgment frameworks upgrade Web surfing, transportation, medicinal services, and so on. Utilizing voice directions enables drivers to remain associated with the Internet by keeping away from traffic dangers. In this way, execution of these kind of utilizations ought to be strong under troublesome system conditions. Client dissatisfaction with system traffic issues can influence the convenience of these applications.

They also assess the exhibition of two famous cloud-based discourse acknowledgment applications, Apple Siri and Google Speech Recognition (GSR) under different system conditions. They assess the deferral of every application under various bundle misfortune and jitter values. Aftereffects of our examination demonstrate that exhibition of cloud based discourse acknowledgment frameworks can be influenced by jitter and parcel misfortune; which are usually happening in WiFi and cell arrange associations. Our outcomes likewise demonstrate that Google Speech Acknowledgment has better execution under various system conditions.

In 2016 Arriany A. A., Musbah M. S [2], provides the idea of smart home, its applications and its framework parts and advances utilized counting the systems administration techniques.

The audit is then expanded to present the voice acknowledgment innovation that empowers the control of any gadget through voice/discourse directions. Besides, they also propose a structure of straightforward model for small scale brilliant home system that is constrained by voice acknowledgment innovation.

The different phases of structure and execution of the proposed model are talked about, and pursued by an essential assessment to the presentation of the model. Assessment results affirm the out-execution of this plan in short-separate calm condition utilizing outer mouthpiece arrangement.

In 2017, Chung H. [3], explains the motivation behind this examination is to look at the capacities and comprehension of virtual individual aides (VPAs) and research which of them gave better comprehension through three programming; Alexa, Siri and Cortana. These menial helpers help individuals and make their life simpler by responding to questions and playing out some advanced activities through voice inquiries. In this trial, he asked each virtual individual colleague fifty-seven inquiries under seven classifications. The after effects of this venture will enable clients to realize that these virtual individual associates are diverse programming, and know which one of them is better. In this way, these outcomes will enable them to choose which gadget they will like to purchase since VPA is one of the principle highlights of these days individual gadgets.

In 2017, López G., Quesada L., Guerrero L. A. [4], introduces an ease of use trial of the most esteemed and globally utilized Speech-based NUI (i.e., Alexa, Siri, Cortana and Google's). An examination of the administrations that everyone gives was additionally performed considering: access to music administrations, motivation, news, climate, schedules and maps or headings, among others. The test was plan by two Human Computer Interaction specialists and executed by eight people. Results demonstrate that despite the fact that there are numerous administrations accessible, there is a ton to do to improve the ease of use of these frameworks. Extraordinarily centred around isolating the customary utilization of PCs (in view of uses that expect parameters to work) and to draw nearer to genuine NUIs.

In 2017, Caranica A. [5], centres around a lot of tests in structure a progression of acoustic and punctuation models for Romanian language, to be utilized in far off discourse acknowledgment situations, for voice driven discourse applications in keen homes or structures, utilizing recently obtained discourse databases in Romanian language, all things considered, conditions, by our exploration gathering.

In 2017, Dempsey P. [6], gives an article which explains about few analysts and there thinking , who were expect the voice-first market to be one of the buyer gadgets stars of 2017. Amazon was first out of the doors with the Echo in 2015, and toward the end of last year Google joined the race with a US-just dispatch of its Home item.

©IJRASET: All Rights are Reserved

III.METHODOLOGY

The proposed framework is with the end goal that it can defeat the disadvantage of the current framework. The venture configuration includes content to discourse. Here whatever the framework gets as info after the direction the yield will get as voice implies discourse. To work with voice acknowledgment, it is prudent to utilize existing frameworks. For instance, the Pocket Sphinx venture. In the created framework, it was picked, as the fundamental methods of voice acknowledgment (recognition).

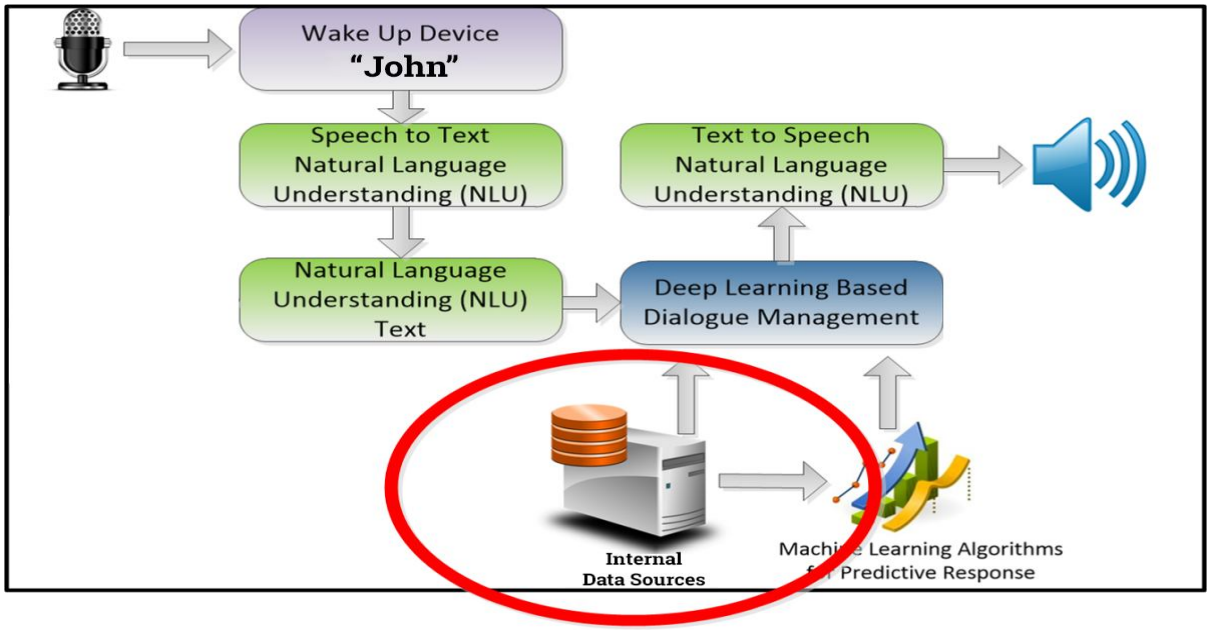


Figure 3: Architecture of the proposed system

PocketSphinx is an apparatus for programmed voice acknowledgment [7], which functions admirably on different low-power installed frameworks, for example, Raspberry Pi, and it is likewise cross-stage, which is the explanation behind picking this system [7].

To create the voice, the most utilized Festival motor was picked, which keeps running on very great attributes of voice peer group. The following stage, on which the scholarly work of the voice associate is done, is the formation of a framework for perceiving the characteristic language of an individual, for example acknowledgment of purpose.

For the framework that can comprehend the client, we utilized preparing scholarly calculations dependent on AI and machine learning techniques. Be that as it may before we show the framework to comprehend the client, we have to play out a few phases of data arrangement which is explained in the next chapter data training.

A. Voice Recognition

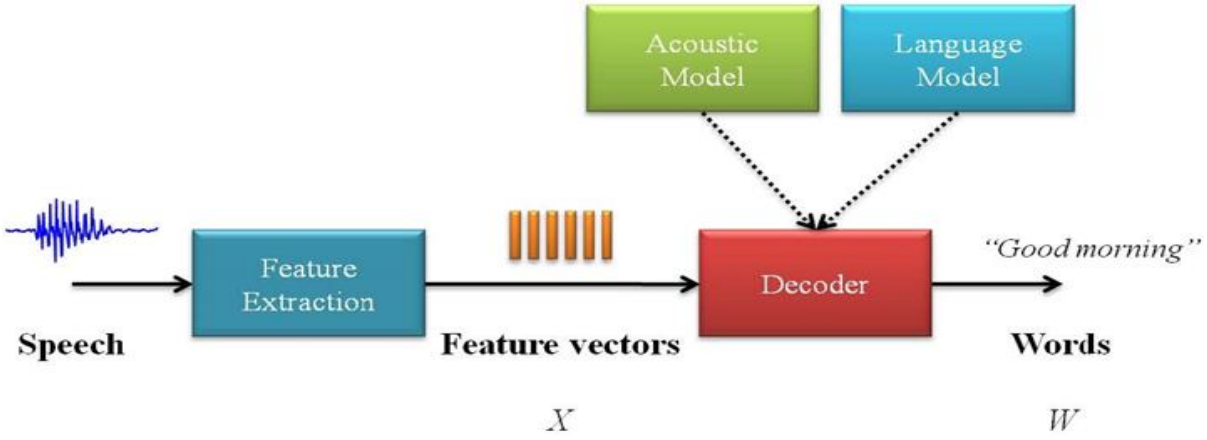


Figure 4: Voice Recognition Architecture

©IJRASET: All Rights are Reserved

Voice is the most well-known sort of individual correspondence and a standout amongst the most energizing examination regions of the signal handling is voice handling. Voice handling is nothing but only learning of language signals and the handling methods of these sign. The sign are generally handled in a digital variant, so voice processing can be seen as a remarkable instance of advanced sign handling which is connected to voice signal.

Voice Recognition is one of the dive examination zones in language (voice) handling, which is otherwise called ASR (automatic speech recognition). Voice recognition innovation enables a PC to focus on individual voice directions and to comprehend singular dialects. Voice recognition is the technique of modifying a given info signal into a progression of words by methods for a calculation that is actualized as a PC program. That is, the voice recognition framework empowers a PC to perceive the words an individual expresses in a mouthpiece or telephone and convert it into coherent content. Voice recognition has various applications, for example, in human services, military, helicopters, communication and different areas and so forth.

Headway in language (voice) innovation was persuaded as individuals needed to create mechanical models that permit the imitating of individual oral declaration capacities. PCs use discourse handling to follow voice directions and different individual dialects.

Figure 4 demonstrates a crucial type of voice recognition framework that mean assorted periods of a plan that pre-processing, speech feature extraction, classification and language model.

The information signal will be changed by pre-processing stage before any data can be extricated at feature extraction stage. The feature extraction stage removes basic vectors required for use in displaying stage subsequent to pre-processing stage. The separated vectors must be robust to clamor (noise) for improved precision.

The language content is perceived by grouping stage by utilizing highlights features and a language layout where the language format contains linguistic structure and semantics identified with the capable language which help the classifiers to distinguish the information explanation feature extraction goes about as an uncommonly significant activity in voice recognition methodology and as it draws out profitable information from test voice it is an imperative piece of research for a long time. The key goal of this strategy is to discover the presentation dimension of various component extraction strategies and afterward choosing one of the techniques among them. It assumes a significant job in exactness of voice recognition.

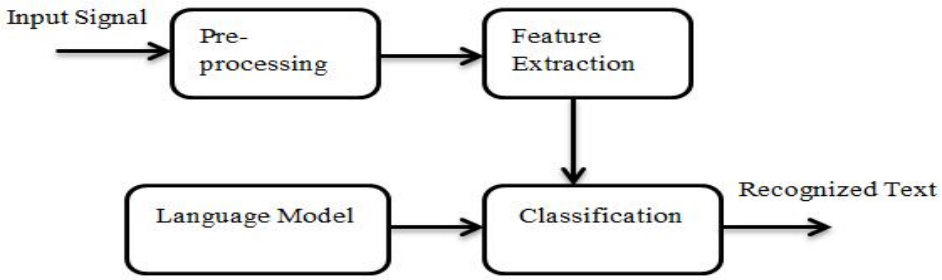


Figure 5: Voice Recognition Stages

To isolate one voice signal from the other feature extraction method assumes a significant job. Since every language has distinctive individual qualities embedded in proclamation, these attributes can be separated from a broad assortment of highlight extraction systems proposed and adequately abused for discourse acknowledgment task.

B. Mel Frequency Cepstral Coefficient (MFCC)

Mel Frequency Cepstral Coefficients (MFCC) is most boundless and significant method used to extract highlighted features. MFCCs utilized in voice recognition depend on recurrence area utilizing the Mel scale which depends on the human ear scale and they are a standout amongst the most acknowledged feature extraction methods. MFCCs which are very much considered to be recurrence area highlights are, all things considered, more exact than time space highlights.

Human Speech as a component of the frequencies isn't direct in nature; along these lines the pitch of an acoustic discourse sign of single recurrence is mapped into a "Mel" scale. In Mel scale, the frequencies dispersing beneath 1 kHz is straight and the frequencies separating over 1 kHz is logarithmic. The Mel frequencies relating to the Hertz frequencies are determined by utilizing condition (1).

$$fmel = 2595 * \log(1 + \frac{f}{700}) \quad (1)$$

©IJRASET: All Rights are Reserved

1) *Pre-Emphasis:* The sound sign are recorded having an inspecting rate of 16 kHz. Each word is put away in independent sound document. This progression incorporates the Pre-accentuation of sign to help the vitality of sign at high frequencies
2) *Framing and Windowing:* The language (voice) signal isn't stationary in nature. So as to make it stationary surrounding is utilized. framing is the subsequent stage after 1st step; in this voice signal is part up into littler casings covered with one another. Subsequent to framing, windowing is utilized to evacuate discontinuities at edges of edges. Hamming Window is the window technique utilized in this exploration
3) *FFT:* Fast Fourier transform is utilized for figuring of the discrete Fourier change (DFT) of sign. This progression is performed to change the sign into recurrence area.
4) *Mel Filter Bank:* The following stage is change from Hertz to Mel Scale, the ranges control is changed into a Mel scale. The Mel channel bank comprise of triangular folded covering channels.
5) *DCT:* The Discrete Cosine Transform (DCT) is utilized in the wake of taking logarithm of yield of the Mel-channel bank.

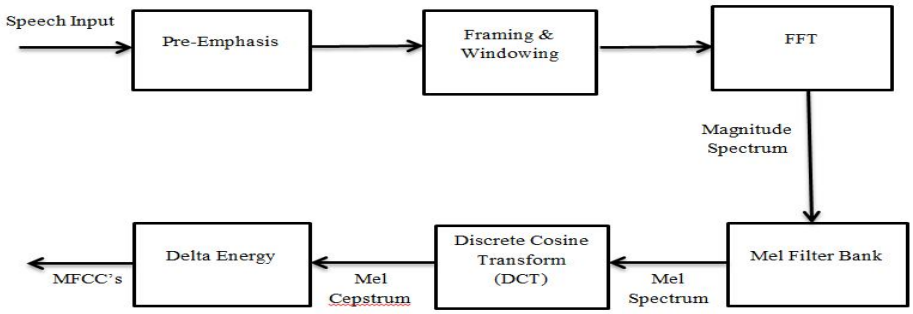


Figure 6: MFCC block diagram

6) *Delta Energy:* In this progression take base 10 Logarithm of yield of past advance. The calculation of Log vitality is basic due to the way that human ear reaction to acoustic discourse signal dimension isn't straight, human ear isn't much touchy to contrast in sufficiency at higher amplitudes. The upside of logarithmic capacity is that it will in general copy conduct of human ear

C. Hardware Implementation

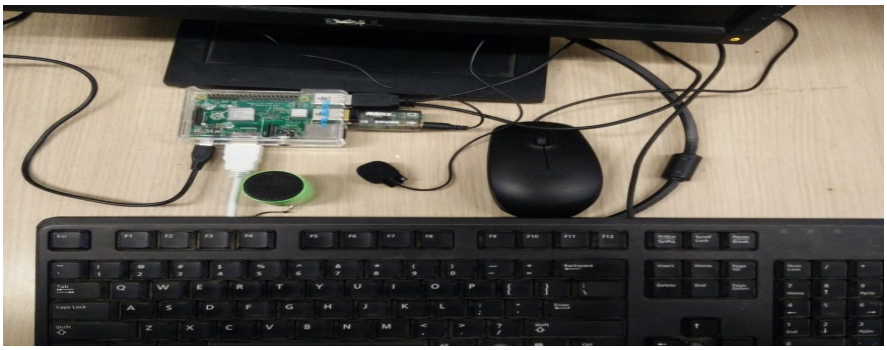


Figure 7: Raspberry Pi to other peripherals

The Raspberry Pi is shown in Figure 7 is connected to various other components such as SD card, LAN, Mouse, Keyboard, Microphone and Speakers. The Microphone is interfaced with Raspberry Pi via USB port, which is the input device for Jasper. The output is taken via 3.5mm audio jack, which is connected to the Speakers. As this is completely offline based project, we'll not be using the network connectivity.

Microphone is utilized to take the audio contribution of the sound. This sound information when further gone through the framework would be hunt down watchwords. These watchwords are basic for the working of the voice directed framework as our modules take a shot at the substance of hunting down watchwords and giving yield by coordinating watchwords.

Keyboard goes about as an information interface mostly for the engineers, giving access to make alters to the program code.

Mouse additionally acts an interface between the framework and the designer and does not have an immediate communication with the end client.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

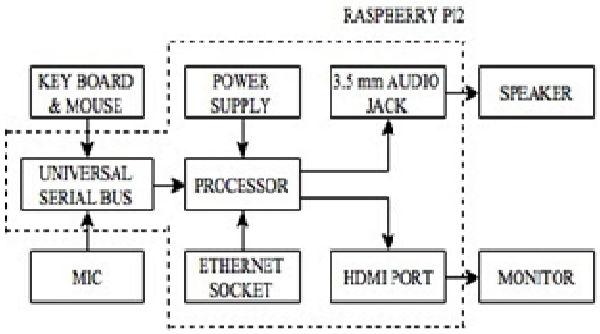


Figure 8: Block Diagram of Raspberry Pi

Raspberry Pi is the core of the voice order framework as it is engaged with each progression of handling information to associating parts together. The Raspbian OS is mounted onto the SD card which is then stacked in the card space to give a working framework.

Screen gives the designer an extra method to look at the code and make any alters assuming any. It isn't required for any kind of correspondence with the end client.

Speakers, when the inquiry set forward by the client has been handled, the content yield of that question is changed over to speech. Presently this speech which is the sound yield is sent to the client utilizing the speakers which are running on sound out.

D. Stream of Events in Voice Command System

Initially, when the client begins the framework, he utilizes a microphone to send in the information. Fundamentally, what it does is that it takes sound contribution from the client and it is nourished to the system to process it further. At that point, that sound info whenever nourished to the speech to content converter, which changes over sound contribution to content yield which is conspicuous by the system and can likewise be handled by it.

Then that content is parsed and looked for watchwords. Our voice order framework is worked around the arrangement of watchwords where it looks the content for watchwords (keywords) to coordinate. What's more, once watchwords are coordinated then it gives the significant yield.

This yield is as of text content. This is then changed over to speech yield utilizing text content to speech converter which includes utilizing an optical character remembrance framework.. This yield is transmitted by means of the speakers which are associated with the sound jack of the raspberry pi as appeared in Figure 4.7.

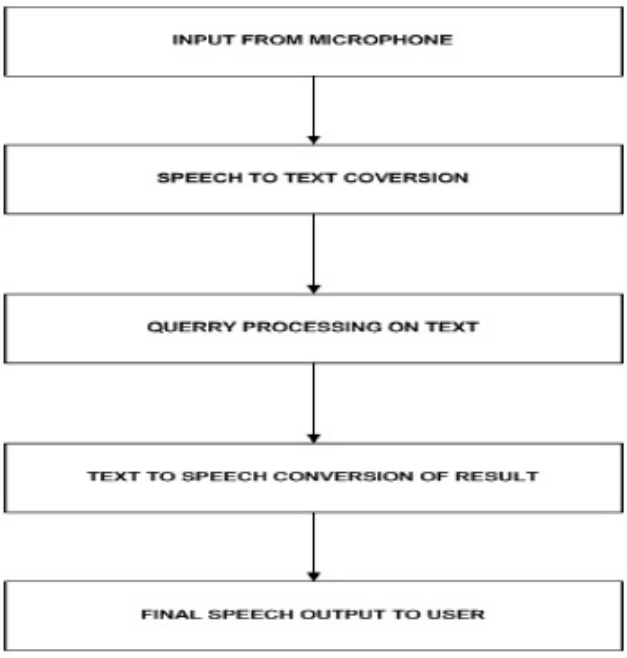


Figure 9: Stream of Events in Proposed System

©IJRASET: All Rights are Reserved

E. Training Data

Since for our situation the yield attributes can be a lot of distinctive classes (that is, n-best speculations of the client's aims, the errand will be a multi-class grouping task where one name can contain marks from various classes. For precedent, as a reaction can return "on; light; washroom", where each piece of the appropriate response has a place with one of a few classes.

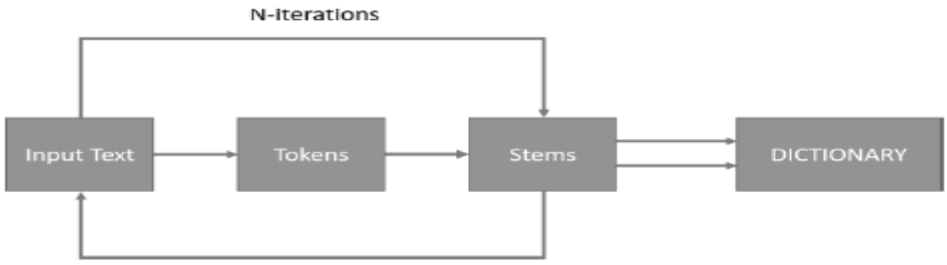


Figure 10: Data Training Process

Gathering the needful information is the most imperative procedure that will enable we to most precisely foresee the outcomes. In Figure 10 demonstrates the way toward getting ready information for training.

In our work, we made a specific framework for dealing with voice command. Thinking about this, the given instances of informational collections will relate to the preparation of the model for such frameworks. For some other errands, the guideline will stay unaltered, as it were the informational collections will change.

To gather information, we have to make a table of equivalent words and distinctive variations of elocution of keywords, as per which the framework will almost certainly fabricate conjectures. Words and equivalent words are demonstrated in the organization in which they are regularly utilized in everyday discourse. A short case of a table of equivalent words is appeared in Figure. 11.

Switch on			
Turn on	Activate	Put	
Switch off			
Turn off	Cut out	Shut off	Disconnect
Light			
Lighting	Illumination	Lamp	Lantern
Split system			
Air conditioning	Cooling	Freeze	
Check			
Verification	Review	Check out	
Mail			
Mailbox	Postbox	Letter-box	

Figure 11: Synonyms Table

After the defining equivalent words, a rundown of answers to be anticipated dependent on the input information is resolved. The anticipated answers can be in any advantageous structure for handling in the application. A case of the appropriate response list is appeared in Figure 12.

On	Off	Light	Split system	Check	Mail

Figure 12: Answers Glossary

©IJRASET: All Rights are Reserved

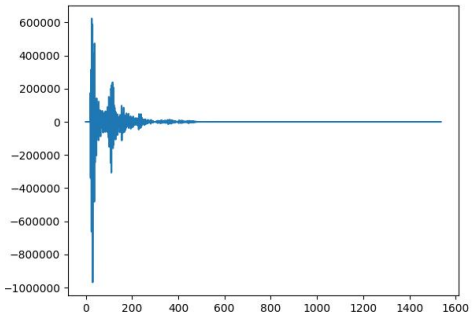
The subsequent stage is to construct a vocabulary from which a teaching data is made. Since the vast majority of the machine learning calculations work on numeric information, we should coordinate each word in the lexicon (dictionary) with a remarkable number inside the word reference nothing but dictionary. It is additionally prudent to decrease the span of the lexicon by using stamping. On huge volumes of information and unique variations of elocution, this technique enables one to lessen the number of various words using magnitude order. This eventually prompts a decrease in information and will improve the nature of acknowledgment. The rundown of answers ought to likewise be set apart with special numeric identifiers.

IV. EXPERIMENTAL RESULTS

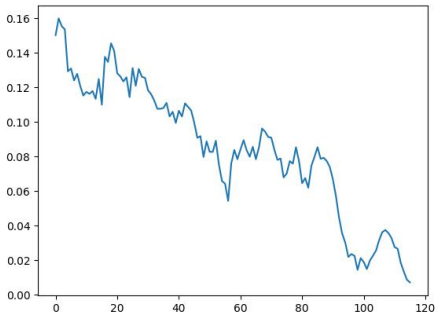
```
INFO: cmn_live.c(126): Update from < 41.00  5.29 -0.12  5.09  2.48 -4.07 -1.37 -1.78 -5.08 -2.05 -6.45 -1.42  1.17 >
INFO: cmn_live.c(138): Update to   < 35.63  8.99 -3.94  7.77 -12.54  9.28  1.57  5.95  5.91  5.62  1.73 -5.65 -7.91 >
INFO: ngram_search_fwdtree.c(1550):     2327 words recognized (21/fr)
INFO: ngram_search_fwdtree.c(1552):   295820 senones evaluated (2714/fr)
INFO: ngram_search_fwdtree.c(1556):   960964 channels searched (8816/fr), 60466 1st, 76222 last
INFO: ngram_search_fwdtree.c(1559):     4597 words for which last channels evaluated (42/fr)
INFO: ngram_search_fwdtree.c(1561):    42827 candidate words for entering last phone (392/fr)
INFO: ngram_search_fwdtree.c(1564): fwdtree 2.67 CPU 2.452 xRT
INFO: ngram_search_fwdtree.c(1567): fwdtree 3.61 wall 3.310 xRT
INFO: ngram_search_fwdflat.c(302): Utterance vocabulary contains 125 words
INFO: ngram_search_fwdflat.c(948):     1682 words recognized (15/fr)
INFO: ngram_search_fwdflat.c(950):    88902 senones evaluated (816/fr)
INFO: ngram_search_fwdflat.c(952):   138141 channels searched (1267/fr)
INFO: ngram_search_fwdflat.c(954):     8475 words searched (77/fr)
INFO: ngram_search_fwdflat.c(957):     6314 word transitions (57/fr)
INFO: ngram_search_fwdflat.c(960): fwdflat 0.54 CPU 0.496 xRT
INFO: ngram_search_fwdflat.c(963): fwdflat 0.57 wall 0.519 xRT
INFO: ngram_search.c(1250): lattice start node <s>.0 end node </s>.67
INFO: ngram_search.c(1276): Eliminated 1 nodes before end node
INFO: ngram_search.c(1381): Lattice has 283 nodes, 1363 links
INFO: ps_lattice.c(1380): Bestpath score: -2820
INFO: ps_lattice.c(1384): Normalizer P(O) = alpha(</s>:67:107) = -211083
INFO: ps_lattice.c(1441): Joint P(O,S) = -230948 P(S|O) = -19865
INFO: ngram_search.c(872): bestpath 0.01 CPU 0.011 xRT
INFO: ngram_search.c(875): bestpath 0.01 wall 0.012 xRT
hello
INFO: continuous.c(275): Ready....
```

Figure 13: Voice recognition of “Hello” keyword using pocket sphinx

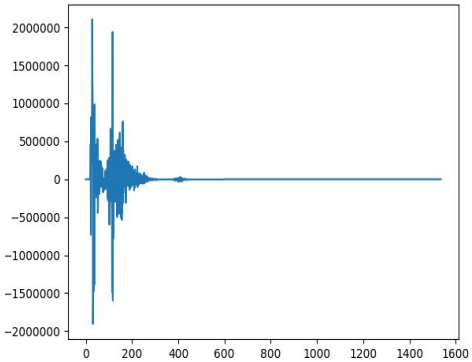
The following figures shows Fast Fourier Transformation (FFT) and Normalized FFT for voice input “HELLO”:



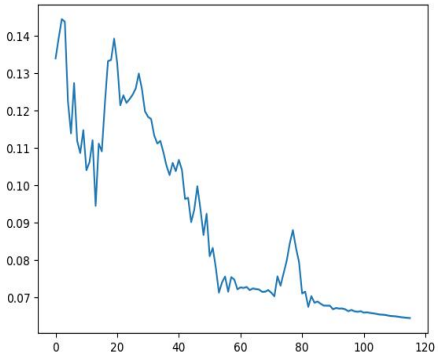
FFT 1



Normalized FFT 1

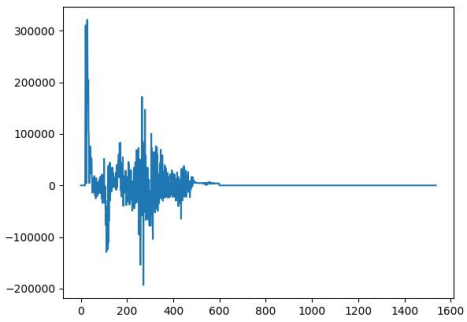


FFT 2

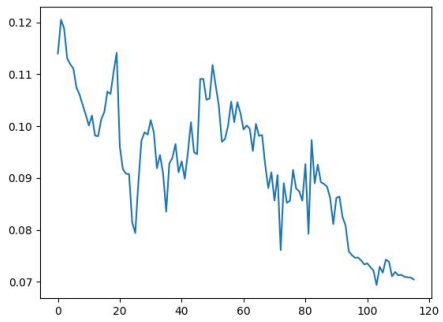


Normalized FFT 2

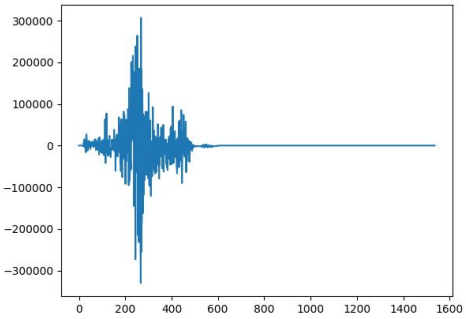
©IJRASET: All Rights are Reserved



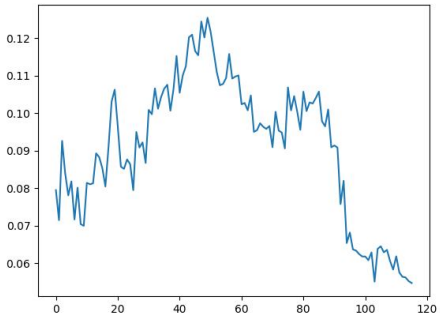
FFT 3



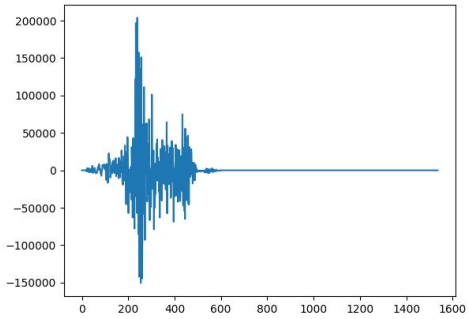
Normalized FFT 3



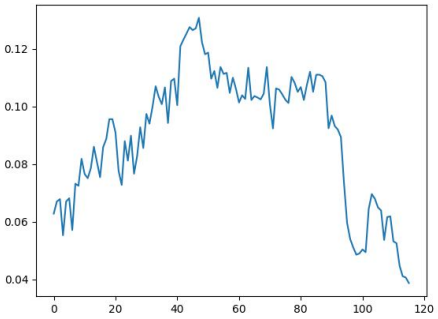
FFT 4



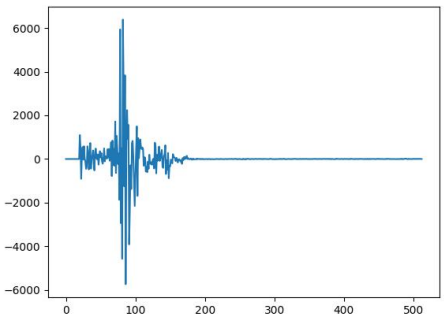
Normalized FFT 4



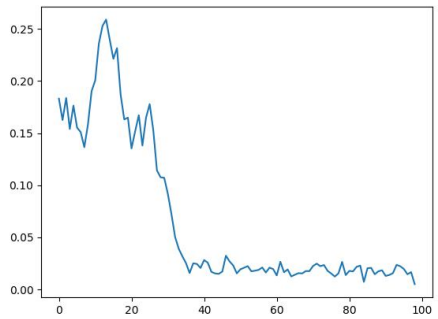
FFT 5



Normalized FFT 5



FFT 6



Normalized FFT 6

©IJRASET: All Rights are Reserved

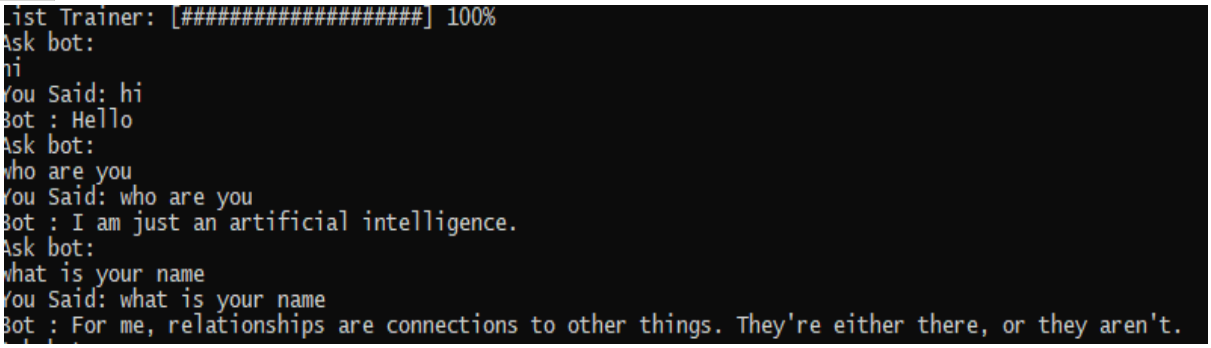


Figure 14: Sample Bot Conversion with User

V. CONCLUSION

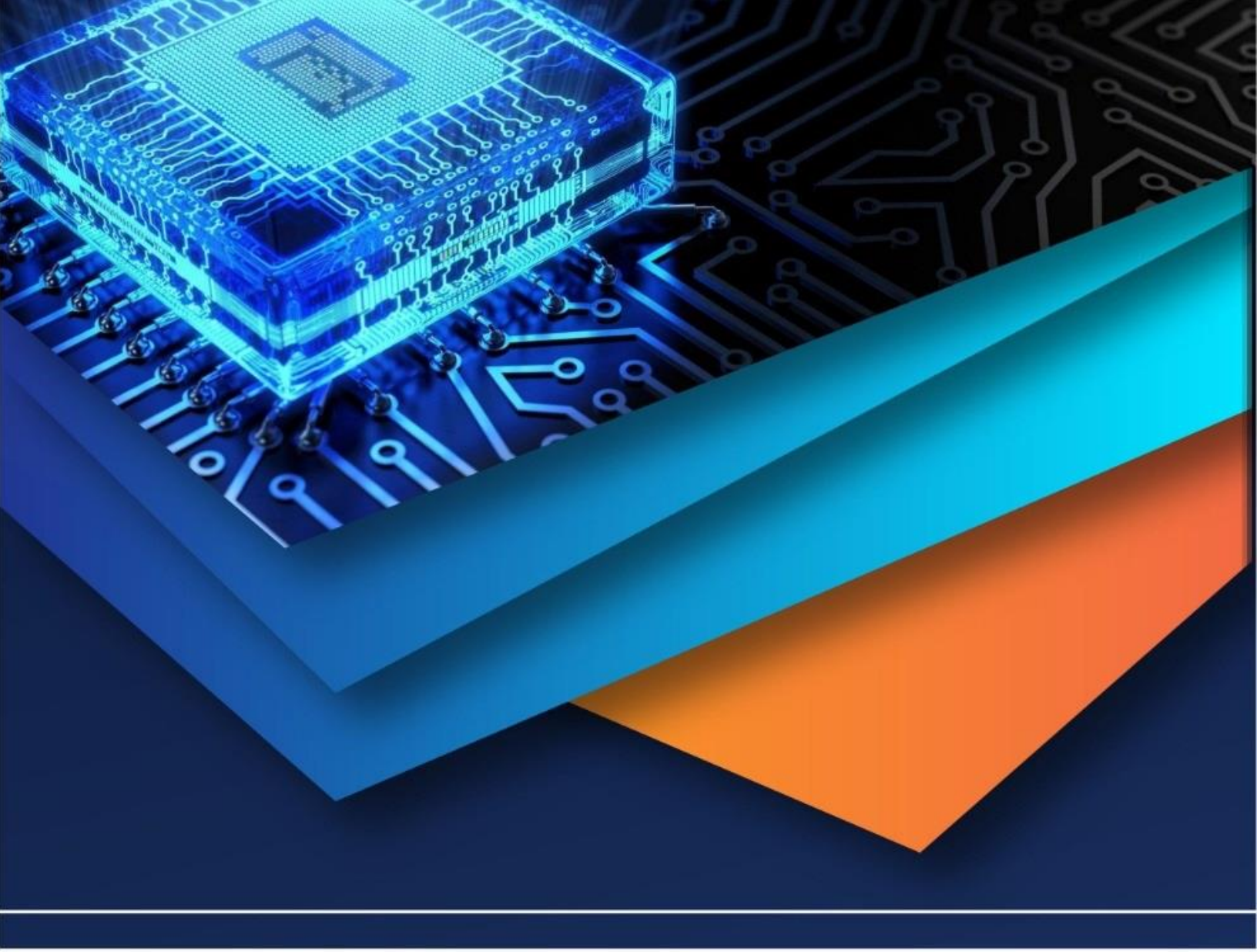
The standards of voice associates which are as of now Spoken to available, were led in the work. The principle deficiencies from significant research were distinguished. A technique for understanding these inadequacies was proposed. In the advance, a voice assistant was fabricated and prepared. Withal, assessment of the calculations' learning capacity for perceiving expectations was performed.

Furthermore, the examine uncovered the creation and the utilization of voice assistant isn't constrained just to cloud administrations. Also, the utilization of nearby frameworks permits to extend the range of assignments in which they can be connected in IoT and IIoT frameworks, keen home frameworks, human services, security and frameworks with an expanded dimension of privacy, where the utilization of cloud advancements can be troublesome.

REFERENCES

[1] Assefi M. et al. An experimental evaluation of apple siri and google speech recognition //Proccedings of the 2015 ISCA SEDE. – 2015.
[2] Arriany A. A., Musbah M. S. Applying voice recognition technology for smart home networks //Engineering & MIS (ICEMIS), International Conference on. – IEEE, 2016. – C. 1-6.
[3] Chung H. et al. Alexa, Can I Trust You? //Computer. – 2017. – Т. 50. – №. 9. – С. 100-104.
[4] López G., Quesada L., Guerrero L. A. Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-Based Natural User Interfaces //International Conference on Applied Human Factors and Ergonomics. – Springer, Cham, 2017. – С. 241-250.
[5] Caranica A. et al. Speech recognition results for voicecontrolled assistive applications //Speech Technology and Human-Computer Dialogue (SpeD), 2017 International Conference on. – IEEE, 2017. – С. 1-8.
[6] Dempsey P. The teardown: Google Home personal assistant //Engineering & Technology. – 2017. – Т. 12. – №. 3. – С. 80-81.
[7] CMU Sphinx Toolkit: http://cmusphinx.sourceforge.net.

©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)