

Intelligent Virtual Assistance Using Raspberry PI

Prof. R. B. Rathod¹ Raut Gauri² BoraKajal³ Budihale Arati⁴ Yash Kumar⁵

¹Professor ^{2,3,4,5}Student

^{1,2,3,4,5}Department of Computer Engineering

^{1,2,3,4,5}PDEA College of Engineering, Manjari(BK), Pune, India

Abstract— The task really comprise of the working of a gadget dependent on usage of a voice order framework as a canny individual aide. The administrations given by the gadget relies upon the information given as voice order by the client and capacity to get to data from an assortment of online sources like climate, reading a clock or getting to online applications to tune in to music. This Voice driven gadget utilizes Raspberry Pi as its primary equipment. Discourse to message motor is utilized to change over the voice order to straightforward content. Inquiry handling is then applied utilizing regular language preparing (NLP) onto this content to decipher the proposed significance of the order given by the client. Subsequent to deciphering the planned significance, text to discourse change is utilized to give suitable yield as discourse. This gadget may give a stage to outwardly weaken to do their everyday undertakings all the more effectively like tuning in to music, checking climate conditions, checking current time or in any event, doing a straightforward numerical estimation. Numerous tests and results were refined and archived.

Keywords: Virtual Personal Assistant, Natural Language Processing, Query Processing, Raspberry Pi, Python Programming

I. INTRODUCTION

A virtual individual right hand can be viewed as a product specialist that comprehends regular language voice orders and finishes errands for the client. Such assignments were prior performed by an individual associate or a secretary that included undertakings like transcription, understanding content or email messages out loud, looking for contacts, booking, settling on telephone decisions and setting updates for arrangements. Yet, in this day and age these drawn-out undertakings that must be overseen by a solitary individual prior have been made simple, viable and more effective by utilizing a gadget. This device has edge over individuals that were recruited for the errands that should have been performed for some basic and clear reasons, which are, it doesn't get drained so it can work productively throughout the day, it just requirements a one-time venture during its buy and needs no compensation along these lines saving the expense of the client, it is smaller and simple to convey anyplace, it can undoubtedly save considerably more information and there's vey less likelihood of information misfortune, it can likewise be utilized for individual use at home like for tuning in to music or setting a clock or setting an alert. Right now most famous individual colleagues are Alexa by Amazon, Siri by Apple and Cortana by Microsoft - by and large utilized for Windows 8.1 and Windows 10.

Individual aides like Alexa and Google Home have inbuilt equipment parts like receivers and speakers. A portion of these aides likewise have inbuilt LED shows. As deciphering sounds takes up a great deal of computational force, gadgets like Alexa and Google Home utilize their own

worker for discourse acknowledgment and undertaking recognizable proof. There is a colossal dataset on their worker which encourages them achieve high exactness and effectiveness.

In the mean time in this gadget equipment parts like mouthpieces and speakers should be appended remotely. This gadget utilizes Google API to change over discourse into text. This content is additionally prepared through NLP. Handling is performed on the actual gadget as there are restricted dataset assets and no worker openness. Give the detail data about the field condition to the client.

The gadget that is utilized as virtual individual collaborator utilizes Raspberry Pi as its significant part. Raspberry Pi can be considered as a little and moderate PC. It has ports on it through which different parts of the gadget like speaker and mic can be associated with it. These segments are then used to take in order from the client as info and afterward give out the outcomes as yields.

The gadget takes in the order from the client through the amplifier being connected to the raspberry pi as info. The client provides its order in his characteristic language to the gadget. The vocal order is changed over to plain content by discourse to message combination. The gadget at that point performs question preparing utilizing regular language handling (NLP) on the plain content got from the order given by the client. During inquiry preparing the gadget attempts to decipher the proposed significance of the guidance given by the client. The watchwords are looked in the sentence and the information that appears to be immaterial to the gadget is disregarded. Gadget sorts out expectations of the client by utilizing these catchphrases from the sentence.

After the order is deciphered gadget essentially plays out the given undertaking that client requested that it do. The exhibition of the gadget fluctuates in various circumstances. Here and there the information that the client provides through his vocal orders isn't taken in or considered by the gadget. One purpose behind this is that the distance between the mouth of the client and the amplifier is huge. Another explanation perhaps that the climate in which the client is providing the order to the gadget is uproarious. Additionally some of the time the gadget may give out some yield that the client was not expecting, this issue may emerge because of the strange complement utilized by the client.

II. LITERATURE SURVEY

A PC fundamentally based methodology for playing out an order through a voice customer interface on a subset of items. The subset is chosen from a fixed of things, each having an article type in any event one taggable field is related with the item type a has a comparing esteem. The arrangement of items is saved in the PC memory. An expression is obtained from the individual and comprises of an order, an article type decision, a tag-peak field determination, and a cost for the taggable control. Receptive to the expression, at any rate one

thing is recovered from the arrangement of devices, the thing of the sort chose through the client and including a cost inside the taggable zone determination that coordinates the taggable field charge acquired from the client the order is done on the thing.

The article incorporates literary substance that is changed over to voice yield. They imagined that sometime PCs will perceive normal language and depend on what we need, while and where we need it, and proactively entire duties for our sake.

Nonetheless, discourse acknowledgment and machine becoming acquainted with have driven forward to be refined, and based records served through bundles and substance suppliers have arisen. We concur with that as PC frameworks end up being more modest and more prominent pervasive [e.g., wearable's and Internet of Things (IoT)]. The recognizer is intended to change a verbal explanation from a person into a substitute strategy for information (e.g., text). A hand held individual partner including a voice-recognizer and a trademark tongue processor is uncovered. This bit of information can be an arrangement for the afternoon, information in the person's logbook or information from the person's location book, Such as a phone number.

The Most notable use of iPhone is "SIRI" which makes the end customer bestow end customer adaptable with voice and it moreover responds to the voice charges of the customer. It is named as Personal Assistant with Voice Recognition Intelligence, which takes the customer commitment to sort of voice or substance and cycle it and returns the yield in various designs like movement to be performed or the thing is coordinated to the end customer. Besides, this proposed structure can change the strategy for interchanges between end customer and the phones. Open Data is at present assembling thought for innovative organization creation, prevalently in the zone of government, bio science, and clever endeavor. In any case, to propel its application more for buyer organizations, a web crawler for Open Data to acknowledge what kind of data is there would be of help.

This paper presents a voice associate which uses Open Data as its learning source. It is featured by change of exactness according to the customer reactions, and acquiring of unregistered data by the customer uphold. We also show an application to help for a field-work and certify its reasonability. The paper gives a graph of the VPA applications, and the ordinary features and future examples. The paper proposes additionally a bound together decision model considering a quantitative evaluation of the meaning of the requesting and the availability of the customer. Virtual Personal Assistant (VPA) is the best in class period of carrier organizations for compact customers. VPA is acknowledged to be the savvy progression of organizations to deal with the consistently extending request by the versatile specialists for convenience and network.

The VPA controls the calls, manages the individual activities though powers the customer to get to his endeavor chairman by methods for voice interfaces, and joins all of the components of Unified Messaging. The Virtual Personal Assistant (VPA) will enable the customer to gainfully deal with growing interest of calls, messages, social events and various activities.

Regardless, a considerable number individuals don't use them reliably. Past examination has inspected the experiences of consistent IPA customers. Using abstract strategies we explore the experience of uncommon customers: people who have endeavored IPAs, at this point pick not to use them reliably. Clearly intermittent customers share a segment of the experiences of progressive customers, for example disappointment at controls on totally sans hands coordinated effort. Basic motivations behind distinction and heretofore unidentified concerns moreover create. Humanness of IPAs began relationships with human partners, looking at their limitations. Specifically, basic concerns rose around security, variation, data permanency and straight imposition. Drawing on these revelations we talk about key challenges, including: illustrating for intrude on capacity; reconsideration of the human comparability; issues of trust and data ownership. Watching out for these troubles may provoke more in all cases IPA use.

As remote helpers advance toward turning out to be me keen and the IVA organic local area of organizations and contraptions reaches out, there's a creating need to grasp the security and assurance threats from this rising development. A couple of late events include critical weaknesses in IVAs. Better definite testing can uncover such weaknesses and brief more dependable structures. It empowers the target customers to associate with PCs and online organizations with a wide group of helpfulness considering distinctive web organizations and web-based media.

There are four standard pieces of the framework; the voice acknowledgment module, the characteristic language handling module, conversational specialist and the substance extraction module. The current screen per customer composing PC programs are not fitting for getting to Internet in context of the base assistance they give for web content and the nonattendance of voice insistence. The Virtual Right hand programming open in the market are not particularly given everything and ill suited to use it also. Some may stand up to give now as well.

This paper presents an ease of use of four Virtual aide voice based and context oriented content (Google colleague, Coratan, Siri, Alexa). Cortana can in like manner read your messages, track your zone, watch your examining history, check your contact list, look out for your date-book, and set up this data together to propose significant information, if you empower it. You can similarly type your requests or sales, if you need to not stand up boisterous. It is just work area based remote helper.

Every business visionary, side prankster and performing various tasks capable out there couldn't want anything more than to have a remote helper right hand to go facing a segment of the dull each day that go with existing in the high level time.

III. PROPOSED WORK

The proposed system is with the end goal that it can defeat from the disadvantages of the current system by making it a standalone personal assistant that can be associated exclusively through the client's voice.

Furthermore, which perform different errands like perusing content from a picture, controlling movement

through voice based indicated directions, and so forth. This system is a model for an assortment of employment.

A. Hardware Components

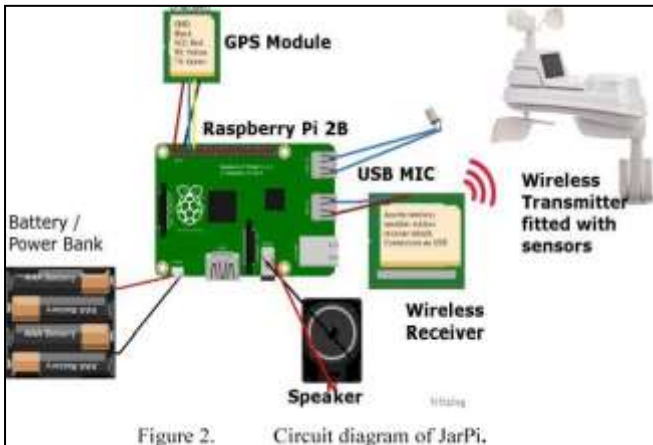


Figure 2: Circuit diagram of JarPi

1) Microphone

The vocal commands given by the user which is used as input is given in through the microphone that is connected to the device. This vocal command is then later converted to simple text and keywords are searched through this text which helps the device to perform its functions and give out the expected results.

2) Raspberry Pi

Raspberry Pi is the major component of the device. It acts as a mini computer. It is indulged in all the activities since the beginning when the user gives the input till the end when the output is presented to the user. It sorts of binds all the components together. All the processing of the data takes place here.

3) Ethernet

The Ethernet cable helps us to provide the internet connection to the device. Internet plays a very important role in the operation of the device as it helps the device to do speech to text conversion, query processing through NLP and text to speech conversion. All these processes take place online that's why the internet connection is very essential.

4) Speakers

Speaker performs the last function in this process. The speaker helps the device to give out the output in the form of speech that is being converted from the text online. The speaker can be connected to the device through an AUX cable.

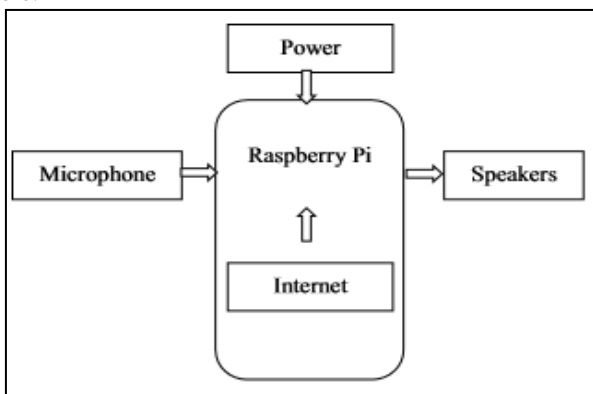


Figure 2:

IV. SOFTWARE IMPLEMENTATION

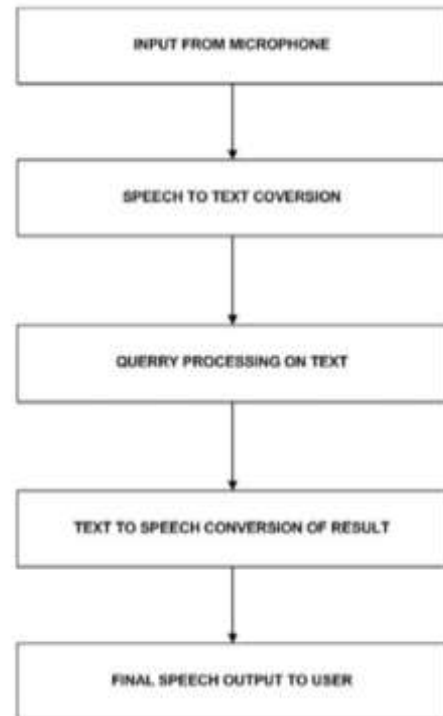


Figure 3: Flow of Events in voice command system

A. Input through Microphone

The user gives in his command verbally to the device through the microphone in the form of input that the device later will process on. But before using the microphone it needs to be configured properly.

B. Speech to Text Conversion

The input that is given by the in the form of a vocal command is first converted to the plain text. This action is performed by using the Google speech recognition API. This API can recognize around 120 languages and variants so that it can support global user base. It's easy to use and very effective.

C. Query Processing

This piece of the entire cycle is the main one. This cycle utilizes normal language handling (NLP) to work. The information that is changed over to the content by the gadget is concentrated in this progression. Entire content is examined, from that point the tokens are distinguished from the content being gotten. By considering the tokens chose by the gadget it attempts to decipher that what could be the activity that the client needs it to perform. When it investigations the proposed activity that client needs it to do it at that point plays out the activity that the client is anticipating that it should perform. Regular Language Toolkit (NLTK) is being utilized here to execute the normal language preparing.

D. Text to Speech Conversion

Once the device interprets the intended meaning of the command given by the user it then performs the action and gives its output in form of text. For the user this text is then converted to the speech. For this conversion of the text to speech python text to speech package (pyttsx) is used.

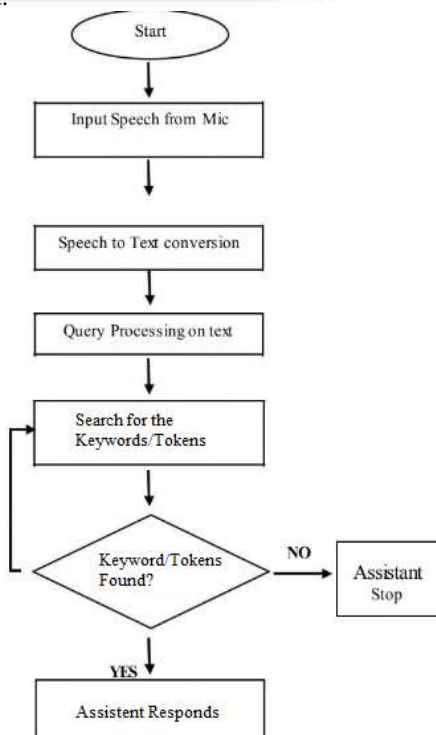
E. Output through Speakers

Once all the above steps are performed then it's time to give the output to the user in the form of speech. This action is achieved by using the speaker. Speaker can be connected to the device by using an AUX cable.

V. WORKING MODEL

Voice acknowledgment framework gives an assortment of administrations through normal language handling utilizing man-made consciousness frameworks prepared through the client's voice orders. The raspberry pi based individual associate includes three crucial modules: Voice Control, Character acknowledgment, and Virtual Assistant.

- 1) Convert the crude content (sentence) into a rundown of words, known as tokens, by parting the sentence by void areas and putting away them in elite "words".
- 2) Remove the accentuations from everything in "words" by supplanting them with nothing.
- 3) Normalize the case by changing over all the words in "words" to bring down case.
- 4) Collect all the prevent words from the English word reference and putting away them in top notch "stopWords".
- 5) Remove the tokens from "words" that are available in the rundown "stopWords".
- 6) Remove the tokens from "words" that have a length under 4.
- 7) Remove the inflectional endings from every token to restore the base or word reference type of a word, known as lemma.
- 8) Find the element that is closest to all the tokens by figuring the comparability of each element, given by a formerly characterized list "highlights", with every token.
- 9) Return the element with most extreme comparability limit.



VI. ALGORITHM

- 1) Set words to the list of words that are split by white spaces.
- 2) Repeat the following steps for each token in words: a. For each character in token repeat : i. If character is a punctuation : Replace punctuation with nothing, i.e. ‘’. ii. Else : Continue
- 3) Change case of token to lower case for each token in words.
- 4) Set stopWords to the set of words that are considered to be stop words in English dictionary.
- 5) Set filteredSentence to an empty list.
- 6) Repeat the following steps for each token in words: a. If token is not in stopWords : Append token to filteredSentence. b. Else : Continue
- 7) Set words to filteredSentence.
- 8) Repeat the following steps for each token in words: a. If length(token) is greater than or equal to 4: Append token to filteredSentence. b. Else : Continue
- 9) Set words to filteredSentence.
- 10) Create an instance of WordNetLemmatizer known as lemmatizer.
- 11) Set lemmas to an empty list.
- 12) Repeat the following steps for each token in words: a. Lemmatize the token (as a noun) and store it in temp. b. Append temp to lemmas.
- 13) Set similarity to an empty list.
- 14) Repeat the following steps for each feature in features : a. Set w1 to a set of all the synonyms of the feature. b. Set sim to an empty list. c. Repeat the following steps for each lemma in lemmas : i. Set w2 to a set of all synonyms of the lemma. ii. Compute the similarity of w1 with w2 and append it to sim. d. Find the maximum from the list sim and append it to similarity. 15) Return the feature having the maximum threshold of similarity.

VII. FUTURE SCOPE

Later on we can build up the gadget and make it a stride further by joining a camera to it and train it to recognize various activities and perform as per those activities. This element possibly utilized for security reasons concerning occasion on the off chance that the gadget recognizes any individual other than the client, it will send a warning on client's mail with a picture of the room by then of time. In the event that the gadget distinguishes that nobody is in the room, it might turn off all the electrical apparatuses like cylinder lights, fans, TV and climate control system. Additionally the gadget can be prepared to respond on a few hand signals like in the event that the client waves his hand, the gadget may advise the current to the client.

VIII. CONCLUSION

Improvement of Virtual Personal Assistant (VPA) is examined in this paper. The vocal orders that the client gives are initial changed over into the content utilizing discourse to message union. Next assignment of the gadget is to break down the content changed over from the discourse. The gadget performs inquiry handling utilizing common language

preparing (NLP). Question preparing encourages the gadget to distinguish the tokens from the content framed. Utilizing these tokens the gadget can decipher what the client needs it to do. The gadget at that point plays out the necessary activity.

REFERENCES

- [1] M. R. and D. Subramaniyan, "Personal Assistant and Intelligent home Assistant via Artificial Intelligence Algorithms-(Raspberry Pi/Pineapple)", *Impact: International Journal of Research in Engineering & Technology (IMPACT: IJRET)*, vol.4, no. 6, pp. 9-13, 2016.
- [2] Ass. Prof. Emad S. Othman, "Voice Controlled Personal Assistant using Raspberry Pi," *International Journal of Scientific & Engineering Research*, vol. 8, no. 11, pp. 1611–1615, Nov. 2017.
- [3] Dahl, George E., et al. "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition." *Audio, Speech, and Language Processing, IEEE Transactions on* 20.1 (2012): 30- 42.
- [4] Chelba, Ciprian, et al. "Large scale language modeling in automatic speech recognition." *arXiv preprint arXiv:1210.8440* (2012).
- [5] Schultz, Tanja, Ngoc Thang Vu, and Tim Schlippe. "Global Phone: A multilingual text & speech database in 20 languages." *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, 2013.*
- [6] Renuka P. Kondekar and Prof. A. O. Mulani, "Raspberry Pi based voice operated Robot," *International Research Journal of Engineering and Technology (IRJET)* Volume: 02 Issue: 12 | December 2017. (<http://www.ijrerd.com/papers/v2-i12/11-IJRERDB597.pdf>)
- [7] Amruta Nikam, Akshata Doddamani, Divya Deshpande and Shrinivas Manjramkar, "Raspberry Pi based obstacle avoiding robot," *International Research Journal of Engineering and Technology (IRJET)* Volume: 04 Issue: 02| Feb-2017.
- [8] Chowdhury Md Mizan, Tridib Chakraborty, and Suparna Karmakar, "Text Recognition using Image Processing," *International Research Journal of Engineering and Technology (IRJET)* Volume: 08 Issue: May-June 2017.
- [9] About Raspberry-Pi 3 model B-plus <https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/>
- [10] Automatic speech and speaker recognition: advanced topics. Editors: Lee, Chin-Hui, Soong, Frank K., Paliwal, Kuldeep (Eds.) Vol. 355. Springer Science & Business Media, 2012. M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [11] Rutuja V. Kukade, Ruchita G. Fengse, Kiran D. Rodge, Siddhi P. Ransing, Vina M. Lomte, "Virtual Personal Assistant for the Blind", *International Journal of Computer Science and Technology (JCST)*, Volume 9, Issue 4, October-December 2018
- [12] Isha S. Dubey, Jyotsna S. Verma, Ms. Arundhati Mehendale, "An Assistive System for Visually Impaired using Raspberry Pi", *International Journal of Engineering Research & Technology (IJERT)*, Volume 8 Issue 05, May-2019
- [13] Narathip Thongpan, & Mahasak Ketcham, "The State of the Art in Development a Lane Detection for Embedded Systems Design, Conference on Advanced Computational Technologies & Creative Media (ICACTCM'2014) Aug. 14-15, 2014
- [14] A. Bar Hillel, R. Lerner, D. Levi, & G. Raz. Recent progress in road and lane detection: a survey. *Machine Vision and Applications*, Feb. 2012, pp. 727–745
- [15] Bohouta, Gamal & Kępuska, Veton. (2018). Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa, and Google Home).
- [16] P. Mane, S. Sonone, N. Gaikwad and J. Ramteke, "Smart personal assistant using machine learning," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, 2017, pp. 368-371.
- [17] O. Portillo-Rodriguez, C. A. Avizzano, A. Chavez-Aguilar, M. Raspolli, S. Marcheschi, and M. Bergamasco, "Haptic Desktop: The Virtual Assistant Designer," 2006 2nd IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications, Beijing, 2006, pp. 1-6.
- [18] V. Kaarthick Raja, A. Kartik, S. Prabhakaran "Voice Assistant for Visually Impaired People" in International Conference on Communication, Computing, and Internet of Things (IC3IoT)
- [19] Maya Rida, Michael Nahas, Milad Ghantous, "iSee: An Android Application for the Assistance of the Visually Impaired", Springer International Publication, 2014