Artificial Intelligence based Personal Assistant

Sulochana Devi
Dept. of Information Technology
Xavier Institute Of Engineering
Mumbai - 400016
bishnoisulochana@gmail.com

Zishan Arif Merchant
Dept. of Information Technology
Xavier Institute Of Engineering
Mumbai - 400016
merchantzishan@gmail.com

Md. Shaqlain Siddiqui
Dept. of Information Technology
Xavier Institute Of Engineering
Mumbai - 400016
ucc.saqlain937@gmail.com

Magdina Lobo
Dept. of Information Technology
Xavier Institute Of Engineering
Mumbai - 400016
lobomagdina161@gmail.com

Abstract—Many people like business men, executives and management persons need personal assistant to help them with their day to day tasks. Personal assistant perform many tasks including replying to emails, managing appointments, scheduling different events etc. Now a days concept of virtual intelligent personal assistant's concept is quite popular. Virtual intelligent personal assistant performs tasks like playing music, dialling phone numbers, searching information on Internet etc. Concept of smart equipments like switching on and off lights and fans etc is also popular. The need of a reliable low cost system has remained an ever desired requirement. This paper combines both concept of personal assistant and control of home appliances. This paper is based on interactive personal assistant capable of processing voice commands given by the user to execute the desired operations like search anything from the internet, fetching stock market information, weather report, basic calculations, opening mailbox, etc. With the help of IOT, proposed system will be able to control home appliances (Smart home concept), thereby making our lives easier and will prove to be a boon for the disabled.

Keywords— Personal Assistant, voice controlled, Natural Language processing.

I. INTRODUCTION

The virtual personal assistant is designed to perform various tasks such as to help in searching answer and respond to various queries based on user input. These systems work based on the concept of Artificial Intelligence, Natural Language Processing (NLP) and Machine Learning (ML). Basically this type of system is an application program capable of understanding natural language and performs tasks. In many ways these virtual personal assistant reduces the need of personal assistant or secretary. Various popular virtual assistant can be listed as Apple's Siri, Amazon's Alexa, Google Now, Microsoft's Cortana etc. These virtual assistants need digital devices with internet connections. Dedicated device assistants are also available in market like Alexa. These virtual assistants are boons in many ways but also cause concerns related to privacy. These devices needs large amount of personal data including emails, user's contacts, search history, location data etc and are always in listening mode in order to provide immediate response.

II. TECHNOLOGY REVIEW

A. Concept of Virtual Personal Assistant

Concept of virtual personal assistant is becoming very popular now and leading companies are launching their devices to assist in day to day life. These devices use speech recognition technique to understand human voice commands. The first tool enabled to perform digital speech recognition was the IBM Shoebox, presented to the general public during the 1962 Seattle World's Fair after its initial market launch in 1961. This early computer, developed almost 20 years before the introduction of the first IBM Personal Computer in 1981, was able to recognize 16 spoken words and the digits 0 to 9.[12] In the 1990s digital speech recognition technology became a feature of the personal computer with Microsoft, IBM, Philips and Lernout & Hauspie fighting for customers. Much later the market launch of the first smartphone IBM Simon in 1994 laid the foundation for smart virtual assistants as we know them today. [13] The first modern digital virtual assistant installed on a smartphone was Siri, which was introduced as a feature of the iPhone 4S on October 4, 2011. [14]

B. Existing Systems

1. Siri: It is Apple's personal assistant, first shipped in 2011 as part of iOS and brought conversational agents into the mainstream. Users can access Siri from their iPhone, iPad, Apple Watch, Apple TV and Carplay in 21 languages. [6]

2. Cortana: It is an intelligent personal assistant created by Microsoft for Windows 10, Windows 10 Mobile, Windows Phone 8.1,

Microsoft Band, Xbox One, iOS, Android, and Windows Mixed Reality. Cortana can set reminders, recognize natural voice without the requirement for keyboard input, and answer questions using information from the Bing search engine. [3]

3. Google Now: The Google Assistant was unveiled during Google's developer conference on May 18, 2016, as part of the unveiling of the Google Home smart speaker and new messaging app Allo; Google CEO Sundar Pichai explained that the Assistant was designed to be a conversational and two-way experience, and "an ambient experience that extends across devices". [15]

C.   Limitations

All the personal assistants present in the markethave some drawbacks. Even the most popular ones like Siri, Cortana, Google Now are platform dependent and only perform device specific tasks. There is no such device that works across platforms and perform a wide variety of tasks without the use of constant internet service.Every day, we see people around us who are unable to do their daily tasks by themselves due to some illness or diseases which have made them bed ridden or in a handicapped position.Also, these systems are of no use when it comes to making your home a smart home. As none of them have found a way by which they can automate appliances without using Bluetooth/WIFI equipped appliances.

BI. PROPOSED SYSTEM

The proposed system is a proof of concept of Artificial Intelligence which uses Natural Language Processing (NLP) and Machine Learning (ML). It is more or less similar to what you get in the name of a personal assistant in the market but with a couple of tweaks which makes it unique. What makes this system unique is platform independence, works without using WIFI/Bluetooth equipped switches/appliances, and doesn't require internet for interaction, controlling electronic appliances, basic calculations. Proposed System has the ability to do take the input commands from voice, and perform the desired task.It is capable of recognizing speech with UN-EN culture and integration with Microsoft speech recognition.It has a feature of Microsoft speech sdk with the help of with it can output speech sentences. This system can control hardware processing unit with the use of Arduino. This system can be programmed to perform various task such as controlling Electronic Home Appliances, control any .exe file of windows, browsers for retrieving anything from the World Wide Web and other task as per requirement.

IV.   LITERATURE REVIEW

1) T. Capes, P. Coles et al(2017) have discussed about Siri On Device Deep Learning-Guided Unit Selection Text-to-Speech System.[1] This paper describes Apple's hybrid unit selection speech synthesis system, which provides the voices for Siri with the requirement of naturalness, personality and expressivity. It has been deployed into hundreds of millions of desktop and mobile devices via iOS and macOS in multiple languages. The system is following the classical unit selection framework with the advantage of using deep learning techniques to boost the performance. In particular, deep and recurrent mixture density networks are used to predict the target and concatenation reference distributions for respective costs during unit selection. This paper presents an overview of the run-time TTS engine and the voice building process. It also describe various techniques that enable on device capability such as preselection optimization, caching for low latency, and unit pruning for low footprint, as well as techniques that improve the naturalness and expressivity of the voice such as the use of long units.

2) Van den Oord et al(2016) have discussed about wave net:a generative model for raw audio. [2] This paper introduces Wave Net, a deep neural network for generating raw audio waveforms. The model is fully probabilistic and autoregressive, with the predictive distribution for each audio sample conditioned on all previous ones; nonetheless we show that it can be efficiently trained on data with tens of thousands of samples per second of audio. When applied to text-to-speech, it yields state-of the-art performance, with human listeners rating it as significantly more natural sounding than the best parametric and concatenative systems for both English and Mandarin. A single Wave Net can capture the characteristics of many different speakers with equal fidelity and can switch between them by conditioning on the speaker identity. When trained to model music, we find that it generates novel and often highly realistic musical fragments. We also show that it can be employed as a discriminative model, returning promising results for phoneme recognition

3) Unnati D. et al. (2015) have discussed about natural language processing using artificial intelligence in [4]. The complete process of the natural language processing system, at many places had made the work easier. Just by our natural language (any language) we can direct the robot, which can do conversation with computer and there is no need of a person to work as a translator for

the conversation between two persons who don't know any common language. Still current program has not reached this level but they may do so very soon. Language technologies can be applied to a wide range of problems in business and administration to produce better, more effective solutions.

4) AvneetPannu (2015) have discussed about Artificial Intelligence and its Application in Different Areas in [5]. The field of artificial intelligence gives the ability to the machines to think analytically, using concepts. Tremendous contribution to the various areas has been made by the Artificial Intelligence techniques from the last 2 decades. Artificial Intelligence will continue to play an increasingly important role in the various fields. This paper is based on the concept of artificial intelligence, areas of artificial intelligence and the artificial intelligence techniques used in the field of Power System Stabilizers (PSS) to maintain system stability and damping of oscillation and provide high quality performance, in the Network Intrusion Detection to protect the network from intruders, in the medical area in the field of medicine, for medical image classification, in the accounting databases, and described how these AI techniques are used in computer games to solve the common problems and to provide features to the games so as to have fun.

5) Dayu Yuan et al (2015) have discussed about Semi-supervised Word Sense Disambiguation with Neural Models. [6] In this paper, two WSD algorithms which combine LSTM neural network language models trained on a large unlabeled text corpus, with labelled data in the form of example sentences and optionally unlabeled data in the form of additional sentences. Using and LSTM language model gave better performance than one based on Word2Vec embedding. The best performance was achieved by our semi-supervised WSD algorithm which builds a graph containing labelled example sentences augmented with a large number of unlabeled sentences from the web and classifies by propagating sense labels through the graph.

6) CiprianChelba et al (2012) have discussed about Large Scale Language Modelling in Automatic Speech Recognition. [7] This paper introduces the large n-gram language models which are a simple yet very effective way of improving the performance of real world ASR systems. Depending on the task, availability and amount of training data used, language model size and amount of work and care put into integrating them in the lattice rescoring step we observe improvements in WER between 6% and 10% relative, for systems

on a wide range of operating points between 17% and 52%-word error rate. I
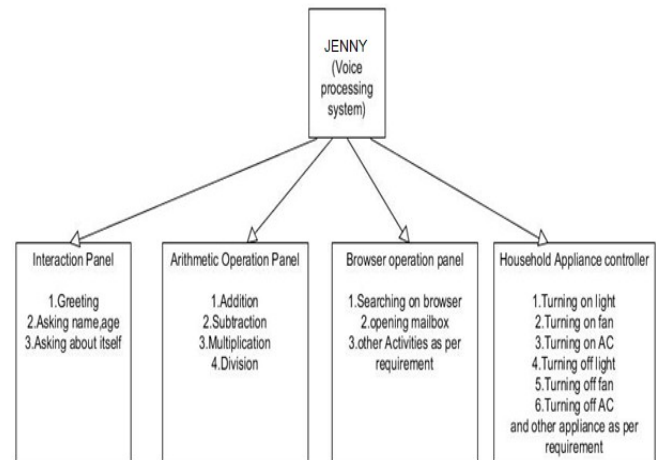
V. MODULE DESCRIPTION

Figure 1: Block diagram of proposed system

The block diagram gives more clear view of proposed system, it provides us with all the information thatsystem will perform. First module is interaction panel in which the user can interact with system and ask various question and will get the answer. Then we have some arithmetic operation panel in which system can perform the basic operation as addition, subtraction, multiplication and division. In browser search panel user can access the information from the internet, for example the user can ask the weather today or the stock exchange etc. Fourth module is the hardware control panel in which the electronic appliances are operated.

V. COMPONENT AND PLAN

Hardware and software requirement are Arduino UNO Board with AT-MEGA 328 Micro Controller, 4 Channel Relay Board , Socket Extension Board, Wireless Speaker with Mic, Microsoft Visual Studio, MicrosoftSpeechPlatformSDK API , MSSpeech SR en US TELE API, MSSpeech TTS en US Helen API, SpeechPlatformRuntime API, Arduino IDE. The following activities arerequired for completion of system. First activity is to create the dictionary of different words. It can be used as a pattern to be match when the voice input is given to system. In this voice input the system will recognize the words and it will check its dictionary and then the output will be produced based on the Grammar that is built. Second activity is building the grammar In this we will frame a sentence. This will create a meaningful statement that is grammatically correct and which is understandable by system. Based on these statementsvoice input will be matched. Next activity is logic development for arithmetic operations, searching anything on internet and controlling the home appliances. Other activity is

connecting and programming the Arduino IDE for controlling the appliances. An Arduino will be connected with laptop from one end and another end connected with relay board. As a command is given for the home appliances to be controlled, Arduino will automatically trip the relay based on the serial data which is received to it via the laptop

VI. CONCLUSION AND FUTURESCOPE

This system will be capable of performing tasks by taking speech commands as input. It processes those input speech commands by using Natural Language Processing (NLP) by using the UN-EN culture, and executesthe desired output. The tasks that our system can perform includes handling electronic appliances like lights, fans, refrigerator, AC, washing machines, etc. These tasks will be performed by using Arduino and Relay boards.

Our system is useful specially for physically handicap and/or bed ridden people as it overcomes the drawbacks of the existing systems of platform dependence, constant need of internet, no possible way to control electric appliances without using WIFI/Bluetooth equipped appliances. Our system is also useful for normal people to make their everyday tasks easier and hence save time.

REFERENCES

[1] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, K. Kavukcuoglu, "Wavenet: A generative model for raw audio,"arXiv preprint arXiv:1609.03499, 2016.

[2] T. Capes, P. Coles, A. Conkie, L. Golipour, A. Hadjitarkhani, Q. Hu, N. Huddleston, M. Hunt, J. Li, M. Neeracher, K. Prahallad, T. Raitio, R. Rasipuram, G. Townsend, B. Williamson, D. Winarsky, Z. Wu, H. Zhang,"Siri On-Device Deep Learning-Guided Unit Selection Text-to-Speech System,"Interspeech, 2017.

[3] New institute explores the future of Cortana, By Roy Zimmermann, Director, Microsoft Research Outreach, February 1, 2018.

[4] UnnatiDhavare, Prof. Umesh Kulkarni, "Natural Language Processing Using Artificial Intelligence', International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 4, Issue 2, March-April 2015

[5]AvneetPannu, "Artificial Intelligence and its Application in Different Areas", International Journal of Engineering and Innovative Technology (IJEIT), Volume 4, Issue 10, April 2015.

[6]Dayu Yuan, Julian Richardson, Ryan Doherty, Colins, Evans, Eric Altendorf, "Semi-supervised Word Sense Disambugation with Neural Model" , Google Mountain View CA, USA,2015

[7] Siri, a virtual personal assistant by Tom Grube. IJRECE VOL. 6 ISSUE2 APR.-JUNE 2018 ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE) INTERNATIONAL JOURNAL OF RESEARCH IN ELECTRONICS AND COMPUTER ENGINEERING A UNIT OF I2OR 201 | Page

[8]CiprianChelha, Dan Bikel, Maria Shugrina, Patrick Nguyen, Shankar Kumar," Large Scale Language Modelling in Automatic Speech Recognition",2012

[9]LapPoonRupertTang, "Integrating Top-down and Bottom-up Approaches in Inductive Logic Programming: Applications in Natural Language Processing and Relational Data Mining", The University of Texas at Austin, December 2003. [10].Michael A Covington "Natural Language Plurals In Logic Programing Queries", Artificial Intelligence Centre, The University of Georgia, Athens, GA 30602 U.S.A, February 1996.

[11] Raymond J. Mooney," Inductive Logic Programming for Natural LanguageProcessing",6th International Workshop S. Muggleton (Ed.), pp. 3-22, Springer Verlag, Berlin 1997.

[12] https://en.wikipedia.org/wiki/Virtual_assistant

[13] "Smartphone: your new personal assistant - Orange Pop". Pop.orange.com. 2016-02-23. Archived from the original on 2017-07-10. Retrieved 2017-12-10.

[14] Darren Murph (2011-10-04). "iPhone 4S hands-on!". Engadget.com. Retrieved 2017-12-10

[15] https://en.wikipedia.org/wiki/Google_Assistant