AIBA

A Personal Assistant

Shaikh Bushra, Khan Bakhteyar, Shaikh Mohd Juned, Dr. Siddiqui Mohd Riyazoddin

IT Department, M. H. Saboo Siddik College of Engineering, Affiliated to Mumbai University, Mumbai.

bushra210596@gmail.com, bakhteyarahmed97@gmail.com, ,shaikhjunaid19965@gmail.com, sriyazoddin4@gmail.com

Abstract—: *Natural language processing gives machines the ability to read and understand human language. This concept of natural language processing is used in our proposed system, AIBA. Our system is able to use natural language processing to process the voice commands given by the user to execute the desired operations. Our system is like a personal assistant and is interactive. Similar type of assistants like Apple's SIRI, Microsoft's Cortana, Google's Google Now already exists but it has drawbacks as they require physical interaction for initiating or starting them up. This drawback is overcome by our system as our system needs voice commands for all the purposes even for initiating or starting it. Other feature of already existing systems is that they perform operations which are device specific and are alsoplatform dependent. We have improvised our system and made it platform independent and also our system will be able to perform operations which will control electronic appliances as well and perform various other tasks as per our requirement. Challenges in natural language processing frequently involve speech recognition, natural language understand, natural language generation (frequently from formal, machine-readable logical forms), connecting language and machine perception, dialog systems, or some combination thereof. Our system has a feature of Microsoft Speech sdk with the help of which it can produce speech sentences as output.*

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

I. INTRODUCTION

According to the father of Artificial Intelligence, John McCarthy, it is "The science and engineering of making intelligent machines, especially intelligent

computer programs". Artificial Intelligence is a way of making a computer, a computer-controlled robot, or a software think intelligently, in the similar manner the intelligent humans think. While exploiting the power of the computer systems, the curiosity of human, lead him to wonder, "Can a machine think and behave like humans do?". Thus, the development of AI started with the intention of creating similar intelligence in machines that we find and regard high in humans. Artificial intelligence(AI) is relevant to any intellectual task where the machine needs to take some decisions or choose the next action based on the current state of the system, in short act intelligently or rationally. In simple words, Artificial Intelligent system works like a human brain, where a machine or software shows intelligence while performing given tasks, such systems are called intelligent systems or expert systems. You can say that these systems can "think" while generating output. The computer would need to possess the following capabilities:

A.   Natural language processing: This unit enables computer to interpret the English language and communicate successfully. [4]

B.   Knowledge representation: This unit is used to store knowledge gathered by the system through input devices.

C.   Automated reasoning: This unit enables to analyze the knowledge stored in the system and makes new inferences to answer questions.

D.   Machine learning: This unit learns new knowledge by taking current input from the environment and adapts to new circumstances thereby enhancing the knowledge base of the system

Our system has the ability to do the natural language processing and take the input from voice, commands are given to the AIBA to perform the task. AIBA is an Artificial Intelligence Based Assistant which thinks to the given input through voice, processes it and gives the required output of various task. AIBA has the capability to control the Electronic appliances, do basic operations, search anything from the internet. It produces speech sentences which is interactive with the user.
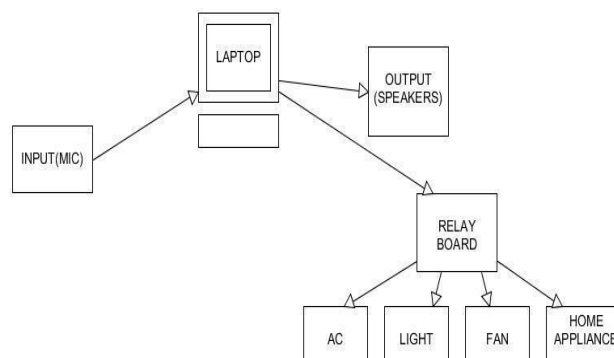

Fig 1. Block Diagram

II. RELATED WORK

1. 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 (e.g. iPhone,

iPad, Mac, etc.) 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. In this paper, we present an overview of the run-time TTS engine and the voice building process. We 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) A. 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) Avneet Pannu (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 labeled 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 labeled example sentences augmented with a large number of unlabeled sentences from the web and classifies by propagating sense labels through the graph.

6) Ciprian Chelba 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.

III.   EXISTING TOOLS

This section describes the existing system in vouge in different part of the world.

A) Siri
Siri, 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]

B) Cortana
Cortana 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]

C) 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".

The problem with the existing tools are that they have physical intervention required in it and they all are device specific and limited to do the functionalities within the device

IV.   PROPOSED SYSTEM

This section proposes an efficient, full proof and easy to use for humans.

A. System overall design

The system proposed by us have the aim to overcome the platform independencies problem and it is not limited   to do task within the device. Our system can work and perform task of controlling   the Electronics appliances. The following are the activities performed by Artificial Intelligence Based Assistant(AIBA).

AIBA is like a personal assistant that is interactive. AIBA is capable of recognizing speech with UN-EN culture and integration with Microsoft speech recognition. AIBA also has a feature of Microsoft speech sdk with the help of with it can output speech sentences. AIBA is a proof of concept to great extent of artificial intelligence. AIBA can control hardware processing unit with the use of Arduino. AIBA can be programmed to perform various task such as controlling Electronic Home Appliances, sending SMS, control any .exe file of windows, browsers for retrieving anything from the world wide web and other task as per requirement
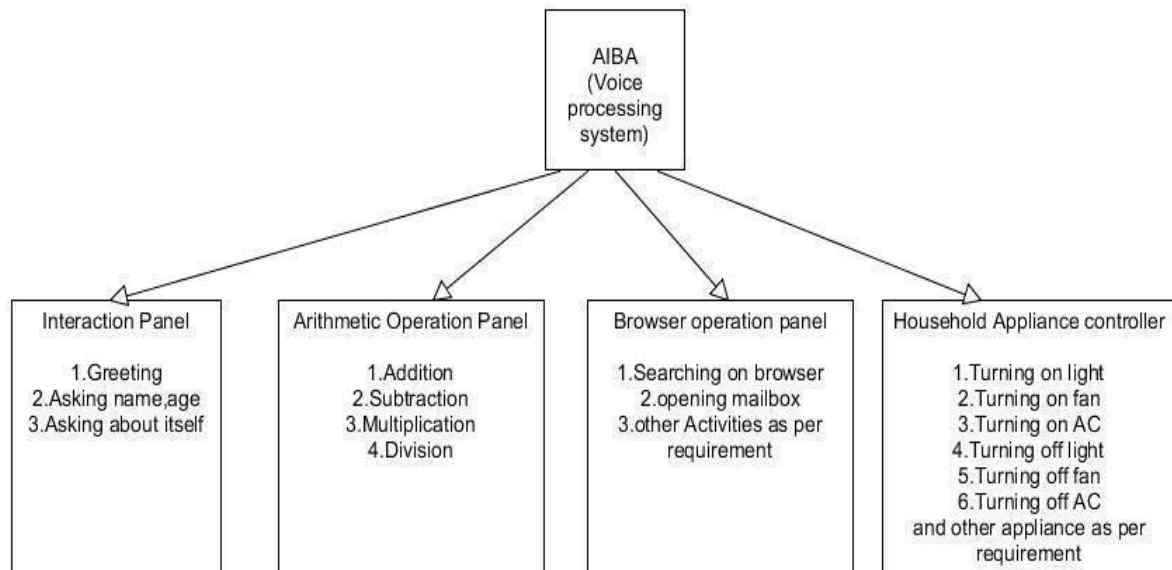

Fig 2. Architectural Diagram

V.   CONCLUSION

Our system, AIBA, is 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 executes the 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 physical intervention to start the system. Our system is also useful for normal people to make their everyday tasks easier and hence save time.

VI.   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]. Unnati   Dhavare,   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]. Avneet Pannu, "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.

[8]. Ciprian Chelha, 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.