INTELLIGENT VOICE ASSISTANT AND SMART SPEAKERS IN EVERYDAY LIFE

Prapulla M S [Student], Guide Asst. Professor Shri Ramu D S

Student, Dept. of Electronics and Communication Engineering, S J C Institute of Technology, Karnataka, India.

ABSTRACT

Virtual assistant is boon for everyone in this new era of 21st century. It has paved way for a new technology where we can ask questions to machine and can interact with IVAs as people do with humans. This new technology attracted almost whole world in many ways like smart phones, laptops, computers etc. Some of the significant VPs are like Siri, Google Assistant, Cortana, and Alexa. Voice recognition, contextual understanding and human interaction are the issues which are not solved yet in this IVAs. So, to solve those issues 100 users participated a survey for this research and shared their experiences. All users' task was to ask questions from the survey to all personal assistants and from their experiences this research paper came up with the actual results. According to that results many services were covered by these assistants but still there are some improvements required in voice recognition, contextual understanding and hand free interaction. After addressing these improvements in IVAs will definitely increased its use is the main goal. Voice control is a major growing feature that change the way people can live. The voice assistant is commonly being used in smartphones and laptops. AI-based Voice assistants are the operating systems that can recognize human voice and respond via integrated voices. This voice assistant will gather the audio from the microphone and then convert that into text, later it is sent through GTTS (Google text to speech). GTTS engine will convert text into audio file in English language, then that audio is played using play sound package of python programming Language.

Key Words: Artificial Intelligence, Big Data, GTTS, Deep Learning, Diagnosis, Machine Learning, *Assistant*

INTRODUCTION

In recent times only in the Virtual Assistants we can experience the major changes, the way user interacts and the experience of user. We are already using them for many tasks like switching on/off lights, playing music through streaming apps like Wynk Music, Spotify etc., This is the new method of interacting with the technical devices makes lexical communication as a new ally to this technology.

The concept of virtual assistants in earlier days is to describe the professionals who provide ancillary services on the web. [1] The job of a voice is defined in three stages: Text to speech; Text to Intention; Intention to action; Voice assistant will be fully developed to improve the current range. [6] Voice assistants are not befuddled with the virtual assistants, which are people, who work casually and can therefore handle all kinds of tasks. Voice Assistants anticipate our every need and it acts, Thanks to AI based Voice Assistants.

AI-based Voice assistants can be useful in many fields such as IT Helpdesk, Home automation, HR related tasks, voice-based search etc., and the voice-based search is going to be the future for next generation people where users are all most dependent on voice assistants for every need. In this proposal we have built the AI- based voice assistant which can do all of these tasks without inconvenience.

Intelligent Virtual Assistant (IVA) is "an application that utilization information, for example, the user's voice. and logical data to give help by noting inquiries in normal dialect, making suggestions and performing activities". Inside the writing the term IVA is utilized conversely with terms, for example, Conversational Agents, Virtual Personal Assistants, Personal Digital Assistants, Voice-Enabled Assistants or Voice Activated Personal Assistants, to give a few illustrations. IVAs join talk affirmation lingo understanding, trade organization, tongue age and talk association to respond to clients request and sales. Voice engaged IVAs like Siri, Google Assistant, Microsoft Cortana and Amazon Alexa are by and large open on cutting edge cell phones, and continuously in homes (e.g. Amazon Echo and Google Home) and automobiles (e.g. Google Assistant blend with Hyundai and KIA). The market for IPAs is foreseen to reach 4.61 billion by the mid 2022s.

LITERATURE SURVEY

In [3] author has shown the primary objective of this is to construct a fully functional voice-based Home automation system that uses Internet of Things, Artificial Intelligence and Natural Language Processing (NLP) to provide a cost- effective, efficient way to work together with home appliances. There are many smart home solutions in the market that aim to automate the basic operations of these home appliances using various technologies such as GSM (Global System for Mobile), NFC (Near-Field Communication) etc. However, most of these systems focus on mimicking the basic operation of the electrical switch. Our project aims at providing a fully automated voice-based solution that our users can rely on, to perform more than just switching

on/off the appliances. The user sends a command through speech to the mobile device, which interprets the message and sends the appropriate command to the specific appliance.

In [4] author has shown how natural language processing (NLP) is a subfield of Artificial Intelligence and getting lot of focus on research and development due to emergence of its applications. The research areas in focus are conversation systems, Language processing, Machine Translation, Deep learning. The researches in these areas lead to development of many tools to build industrial applications. Combining Deep Learning techniques with Natural Language Processing is finding lot of applications in domains such as Healthcare, Finance, Manufacturing, Education, Retail and customer service.
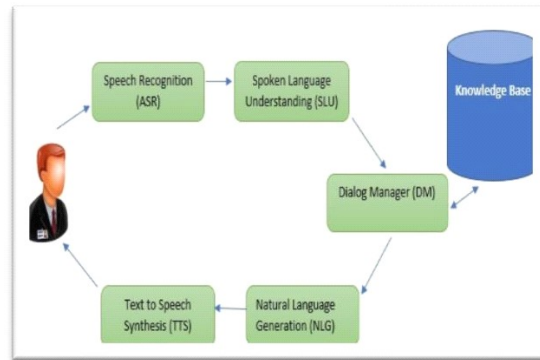
In[5] Voice control is a major growing feature that change the way people can live. The voice assistant is commonly being used in smartphones and laptops. AI-based Voice assistants are the operating systems that can recognize human voice and respond via integrated voices. This voice assistant will gather the audio from the microphone and then convert that into text, later it is sent through GTTS (Google text to speech). GTTS engine will convert text into audio file in English language, then that audio is played using play sound package of python programming Language.

WORKING PRINCIPLE

The dialogue system is one of an active area that many companies use to design and improve their new systems. According to CHM Research, before 2030, millions of us will be using "voice" to interact with machine, and voice- driven services will become part and parcel of smartphones, smart glasses, home hubs, kitchen equipment, TVs, games consoles, thermostats, in-car systems and apparel [8]. There are many techniques used to design the dialogue systems, based on the application and its complexity. On the basis of method used to control dialogue, a dialogue system can be classified in three categories: Finite State (or graph) based systems, Frame based system and Agent based systems [1]. Also, there are many different architectures for dialog systems. Which sets of components are included in a dialog system, and how those components divide up responsibilities differs from system to system. A dialogue system has mainly seven components: Input Decoder, Natural Language Understanding, Dialogue Manager, Domain Specific Component, Response Generator, and Output Renderer [1]. However, there are six main components in the general dialogue systems, which includes the Speech Recognition (ASR), the Spoken Language Understanding (SLU), Dialog Manager (DM), Natural Language Generation (NLG), Text to Speech Synthesis (TTS), and the knowledge base. The following is the structure of the general dialogue system. Voice assistants are all written in programming language, which listens to verbal commands and response according to the user's requests. In this project we have used python programming language to build an AI based voice assistant. How user can say, "Play me a song" are "Open facebook.com", the voice assistants will respond with the results by playing the particular song or by opening Facebook website. The voice assistant waits for a pause to know that user have finished their request then the voice assistant since uses request to its database to search for the request. The request asked by the user gets split into separate commands, so that our voice assistant can able to understand. Once within the commands list, our request is searched and compared with the other requests. The commands list then sends these commands back to the Voice assistant. Once the voice assistant receives those commands, then it knows what to do next. The voice assistant would even ask a question if the request is not clear enough to process it, in other words, to make sure it understands what we would like to receive. If it thinks, it understands enough to process it, the voice assistant will perform the task which the user has asked for.

Figure 1. The Structure of General Dialogue System

TASKS PERFORMED BY THE AI-BASED VOICEASSISTAN

Can remember any person name till the usage session. Voice assistant name can also be changed unlike in other voice assistants. Play/download a song or video from YouTube. When user asked 'can you play/download me a song', 'play movie' the assistant open YouTube and plays the required content for the user or download the requested video/song. Searches anything from google and tells the required content. If asked 'google search' the assistant searches the content asked from the google and opens the required content in browser. Opens the maps and tells the exact location the user asked for. When asked for 'find location' or 'google maps' the assistant ask for the location the user wants and opens the google maps and highlight the location user asked for. Tells the accurate weather of the location the user asks for. When asked for 'current weather in' the assistant tells the exact weather of the desired location of both maximum and minimum in degree Celsius.

Takes a screenshot of the display. When asked for "capture", "capture my screen", "my screen", "screenshot", "take screenshot", the assistant captures the display the user is using and stores it in the path specified. Can able to tell whether the password has been hacked or not. It can translate the words

the user speaks into any language and displays the words of that language which is specified by the user. Can able to shut down or restart the system by just user command etc. These are some features we have added to our AI-based voice assistant as of now and we are working on many more features to embed into this assistant.
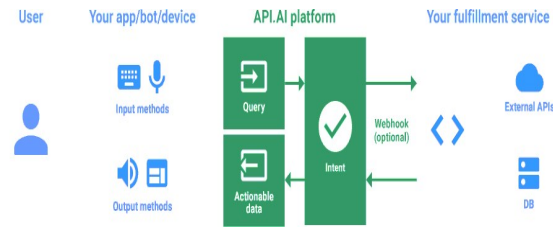

Figure 2: Block Diagram of AI-based voice assistance mode

we can tell how our voice assistant is useful. For e.g., if we want to go to a particular location instead of opening google maps and typing the destination takes a lot of time instead of that with our VA just by command to find the particular location it opens the map and highlight the particular location. In the modern VA's, playing a song on YouTube is just a feature, but they ant download it. For e.g., if you want to download the song "give me some sunshine", in other voice assistants we can't download it but, in our VA, the song is searched in the YouTube database and the Video ID is noted, with this the video is downloaded.
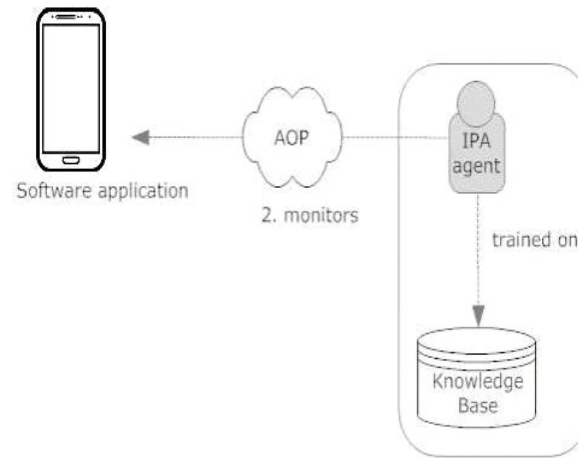

Figure 2. software usage

If we want to get any information, we have to open google and search for it. But in VA by just giving it in the form of command for e.g., "what are super computers?" it collects the best information available on Wikipedia and gives us and if we want to know about what is happening in the world by giving the command "what are the top five news of the day?" it grabs the information and tells us. From this we can gain knowledge. Concerning the aspect of security, if a girl says "I'm in danger" the Voice Assistant extracts the current location of the girl and sends it to their parent's relatives and the nearby police station.

There is a possibility that this task can be handled by personal assistant applications. For helping the user, agent first need to understand him. The training dataset is necessary for training the agent. When user starts using application, the agent is initially in training mode at least for one week. During this period, agent gathers information such as at what time in day the users' phone is in general mode and in silent mode. This information is maintained specifically for each day of week as user may have different schedule on each day of week. The information has many attributes such as phones status, time, day-of-week.

This data is temporarily gets stored in user's smartphone in form of text file and is uploaded on server at suitable time. The suitable time means the time when device is connected to Wi-Fi or mobile network and preferable during night time because data for full day is available for uploading. To reduce the processing load on user's smartphone, the task of machine learning has been implemented on server side. Complex operations such as creating a predictive module using machine learning algorithm is implemented on server.

ADVANTAGES

- There are a number of benefits associated with NLP program. The program is now extensively researched on by teams within the scientific community, those studying academic psychology as well as students of psychotherapy and counseling. The benefits of programming are many, addressing the needs of every segment of society. This methodology continues to help improve and generate highly profitable interactions within the immediate social and business community [6].
- NLP is transforming the way humans and machines interact.
- In business, NLP has played a vital role in strengthening client – salesperson rapport and improving management skills in communication, leadership and employee motivation. In the health care sector, this powerful tool has helped in treatment and recovery of many patients [5].
- Delivers a great customer experience while improving the self-service system's containment rate
- Encourages natural, human-like conversations that create more satisfying self- service interactions with customers
- Automates what touchtone cannot by collecting dynamic data such as names and addresses
- Enables organizations to save agents for more critical tasks.

FUTURE SCOPE

- NLP's future will be redefined as it faces new technological challenges and a push from the market to create more user-friendly systems [8].
- It is also pushing NLP more towards Open Source Development. If the NLP community embraces Open Source Development, it will make NLP systems less proprietary and therefore less expensive. The systems will also be built as easily replaceable components, which take less time to build and more user-friendly [10].
- NLP is a challenging field since it deals with human language, which is extremely diverse and can be spoken in a lot of ways.
- NLP is one of the growing technologies. With constant innovation and research going on in this field, it is only expected to grow in the future. Since this is such an upcoming field, there is a dire need for skilled professionals. If you are interested in working on making computers learn and understand human language, then this is a good time to upskill yourself. NLP offers good prospects and is a high paying field.

CONCLUSION

The required packages of Python programming language have been installed and the code was implemented using PyCharm Integrated development environment (IDE) and the python code we have developed runs in both Python 2.7 and Python 3.x, and below are the few outputs which we have received in our AI-based voice assistant. The system enables the user to get features provided by different applications on a single platform. The application will work and provide profile management automatically without any human intervention. The Reminders in this application i.e. Time, Call and Location based would not let user to miss single important tasks in the user's routine by allowing user to keep track of everything. It would provide Profile management, Reminders and other functionality to assist user in day to day tasks.

REFERENCES

1. Chandhana Surabhi.M, Natural Language Processing Future, International Conference on Optical Imaging Sensor and Security IEEE, Proc IEEE, July 2013.
2. P. Milhorat, S. Schlogl, G. Chollet, J. Boudy, building the next generation of personal digital assistants, 1st International Conference on Advanced Technologies for Signal and Image Processing - ATSIP'2014 March 17-19 2014
3. Mrs. Paul Jasmin Rani, Jason Bakthakumar, Praveen Kumaar.B, Praveen Kumaar.U and Santhosh Kumar, voice-controlled home automation system using natural language processing and internet of things, Third International Conference on Science Technology Engineering & Management IEEE, Proc IEEE, Vol 36, July 2019.
4. Krishna Prakash Kalyanathaya, D. Akila and P. Rajesh, advances in natural language processing – a survey of current research trends development tools and industry application, International Conference on Advances in Signal Processing, Power, Embedded, Soft Computing, Communication and Control Systems, Proc IEEE April 2019.
5. Razieh Adelkhah, Mehrnoush Shamsfard, Niloofar Naderian**,** The Ontology of Natural Language Processing, 2019 5th International Conference on Web Research IEEE, June 2019'
6. Prajyot Mane , Shubham Sonone , Nachiket Gaikwad and Prof. Jyoti Ramteke, Smart Personal Assistant using Machine Learning, International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017).
7. Subhash S, Ullas A, Prajwal N, Srivatsa, Santhosh B, Siddesh S, Artificial Intelligence-based Voice Assistant, Fourth World Conference on Smart Trends in Systems, Security and Sustainability(WorldS4) 2020.
8. Runze Chen, Zhanhong Tian, Hailun Liu, Fang Zhao, Shuai Zhang, Haobo Liu, Construction of a Voice Driven Life Assistant System for Visually Impaired People, International conference on artificial intelligence and big data, Proc IEEE,
9. July 2018
10. Pooja Singh1, Pinki Nayak2, Arpita Datta3, Depanshu Sani4, Garima Raghav5, Rahul Tejpal6, Voice Control Device using Raspberry Pi**,** International Journal of Research in Engineering & Technology IJRET, vol. 4, no. 6, pp. September 2019**.**

11. Nithuna S and Laseena C.A**,** Review on Implementation Techniques of Chatbot, International Conference on Communication and Signal Processing, IEEE, July 2020

12. Prajwal S V, Mamatha G, Ravi P, Universal Semantic Web Assistant based on Sequence to Sequence Model and Natural Language Understanding, International Conference on Technical Advancements in Computers and Communications, Proc IEEE, 2017**.**

13. Piotr Kłosowski**,** Deep Learning for Natural Language Processing and Language Modelling, signal processing algorithms architectures arrangements and applications Spa,

14. Proc IEEE, September 2019

15. Fernando A. Mikic Fonte, Martín LlamasNistal, Juan C. BurguilloRial, and Manuel Caeiro Rodríguez, NLAST: A natural language assistant for students, IEEE Global Engineering Education Conference, Proc IEEE, April 2016.

16. Subhajit Dey**,** Web based real-time home automation and security system**,** International Journal of Electrical and Electronic Engineering & Telecommunications, Volume 4, No. 3, July 2015.