AI BASED VIRTUAL ASSISTANT

Sri Meenakshi Pandey [a, *], S Sindu [a], Ms C A Daphine Desona Clemency [a],

[a] Department of Computer science and engineering, Sathyabama institute of science and technology,

Abstract
Artificial Intelligence has been fast emerging as a noteworthy technology that has the capability to revolutionize the cognitive behaviour of humans by simulating their intelligence for the betterment of the mankind. AI consists of multi-functional technologies which plays a significant role in our everyday lives like home automation where controlling the computer and performing multiple tasks using voice commands to remote monitoring and control activities. This study is aimed at designing an AI based virtual assistant that acts as a human language interface through automation and voice recognition based interaction from human based on Python . The instructions for the Voice Assistant are implemented as per the user requirement .The most successful Speech recognition software's like Alexa, Siri, etc has been the brainchild of AI technology. Speech Recognition API in python converts speech into text thereby sending and receiving the emails without typing, searching the keywords in Google without opening the browser, and carrying out many other tasks like playing music etc., has been made possible through the help of this AI based virtual Assistant software. In the present scenario, innovation in digital technologies has resulted in increased effectiveness and accurateness of several tasks that would have required large amount of human effort and resources. Through utilization of AI in every domain, remarkable transformations have resulted in reduced time and labour. Thus AI based voice assistant software offers highly accurate and efficient solution to minimise human effort and time while performing a task that imitates a human assistant to carry out any particular task. Muti-functional aspects like voice commands, sending emails, reading PDF, sending text on WhatsApp, opening a command prompt or IDE, playing music, performing keyword searches in Wikipedia , giving weather forecast, desktop reminders of your choice etc are some of the major operations that can be performed by the developed AI based virtual assistant which also possess certain basic conversational abilities. ,pyttsx3, Speech Recognition, Date time, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, pyQt etc are some of the tools utilised for the project. A live GUI has been designed for interacting with the AI virtual Assistant as it presents an elegant design framework to carry out the necessary conversation.

Keywords: AI, Python, WhatsApp, Wikipedia, GUI, IDE, Speech Recognition

1. Introduction

Recent times has been witnessing multiple innovative digital technologies based gadgets like smart watches ,fitness bands, multimedia speakers, Bluetooth earphones, smart phones, computer , desktop, TV, etc., which primarily comprise of voice assistants. Almost every digital gadgets arriving to the market are being equipped with virtual assistants that can control the device through speech recognition. With the modern advancement in the form of AI, the conventional speech recognition based computer systems has got transformed into even more sophisticated and efficient. In today's world, these technologies have been playing significant roles in our day-to-day lives, right from utilising them for entertainment, education, or interaction with others. This incredible advancement has lead to design and development of Intelligent Personal Assistants (IPAs), whose main aim has been to work towards making lives easier for end-users [1]. A seamless access to devices through voice commands has been revolutionary in accomplishing human-system interaction. To execute the tasks in a more efficient manner, voice based interaction, conversion of speech to text API has been utilised for thorough understanding the input. Majority of the voice assistants in the modern times answer the users' query through voice commands based on Natural Language Processing (NLP) technology that can be an exact replica of real-time conversations with individuals.NLP mainly focuses on designing methods and ML algorithms for better understanding and generation of languages where users are presented with opportunities to interact with intelligent devices that can exactly replicate another human. The active role played by these virtual assistants in understanding the user's requirements in the form of voice commands and assist them by answering their commands is utilised to design a framework where further advancement is in progress. Popular voice assistants like Google assistant, Amazon's Alexa, Apple's Siri and Microsoft's Cortana are the practical evidences of impressive development and capability of AI. Most of the tech behemoths like Amazon and Apple have their own voice assistants, and the odds are very strong that these devices will become more prominent in our daily lives in the near future too.

A voice assistant begins its operation once activated by the user (Fig 1). Normally certain keywords are used as activation inputs to these voice assistants. For instance, users of Google Assistant use "OK Google", Amazon's Alexa is activated by "Alexa" and Apple users call out "Hey Siri" to activate their respective voice assistants. When a user says something, the voice assistant converts this voice input into actionable data. During the initial stage, any speech obtained from the user is transformed into text followed by means of syntactic and semantic processing of converted text, which refers to the understanding of voice assistant in interpreting the actual meaning of the converted text, by looking at the sentence format, grammar, contextual data and the meaning of individual words. Once the meaning of the acquired information has been understood, the voice assistant explores the query based information from Internet, a cloud platform or an application to answer the question in an appropriate manner, by generating a sentence in the

form of text to convey its answer in response to user's query. The final step thus involves conversion of this text into speech through audio output for the user.

Role of AI in Designing Voice Assistants

Advancements in voice assistants in the modern times are aimed at making the assistants efficient in solving the problems along with guiding the users in arriving at appropriate solutions. Thus making these voice assistants smart can be accomplished through repetitive fine-tuning and refining of parameters by utilising machine learning and deep learning techniques. These techniques that can be seen as a subset of AI as it facilitates the computer systems equipped with voice assistants the capacity to automatically learn and progress themselves through experience and do so without being explicitly programmed by any human. In summary voice assistants make use of voice recognition facility to convert users' speech into audio and then convert them back to speech again. Nevertheless, it is only through the use of AI and especially machine learning that facilitates voice assistants to develop them into much smarter, by operating accurately and efficiently and work towards potentially achieving consumer satisfaction by offering them an enhanced user experience.
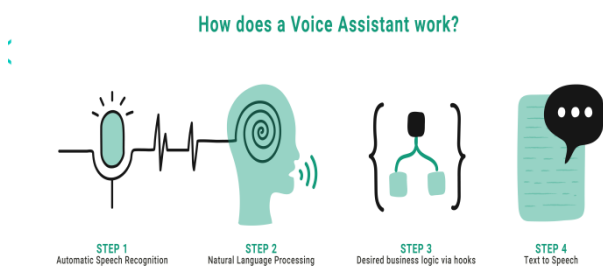
Source: www.Slanglabs.com

Figure 1: Working of Voice assistants

History and Evolution of Intelligent Voice Assistants

Voice control was first employed in the public space via HAL 9000, followed by sentient computer in 2001 for Space Odyssey, which was taken up by Starship Entreprise's and quite recently in Iron Man. The evolution of smart voice assistants in fiction and entertainment industries has been rapid. However, voice assistants were utilised and became to be famous in real life only from 2011, when Apple launched its first smart phone-based assistant, Siri. This section will offer a systematic review of evolution of voice assistants so far designed and its developments.

SIRI

Siri has been acquired by Apple in 2010and is unequivocally considered to be the most well-known and extremely efficient voice assistant that can perform a wide range of operations like sending text messages, scheduling the meetings, makingn phone calls, ,activating the low power mode, enable do not disturb mode (DND) etc., Siri can also answer user questions, send emails, sets up an alarm, makes reservation in restaurants, provides directions to places by its interpretation knowledge of natural languages. Inspite of all the benefits that Siri has been offering,it has its own drawbacks like it can operate only in Apple devices, requires an active internet connection to operate, Siri works well with English commands, but must be spoken clearly to understand, speaking too fast, or with a strong accent, Siri won't be able to understand as there are certain listening problems that reduces Siri's ability to understand user queries, Siri also has trouble with background noise and low-quality audio from headsets, Siri thus needs Wi-Fi if internet connection for its effective operational abilities.

ALEXA

Amazon launched the Amazon Echo in 2014. This smart speaker presented the users with the voice assistant Alexa that has been designed in-house as a key strategy for Amazon to develop its customer base and further increase revenue through facilitating online shopping experience. The main benefits are it's easier to operate process, nonstop music; shopping, timers etc., but the limitations are its mishearing and slower response rates.

Google Assistant

Google's voice assistant is available on smart phones and home devices that was launched in 2016 is not only available on its own products but also offers operation in multiple devices through partnerships with other companies.

Other Voice Assistants

Along with Siri, Alexa and Google Assistant which are still the most widely used smart voice assistants there are other voice assistants like Facebook's M and Microsoft 's Cortana. As people are still continuing to embrace this technology, the evolution of voice assistants is definitely sure to take this to next level.

Our AI based voice assistant has been designed with the following objectives in mind

• To design an effective personal assistant software that uses semantic data sources available on the internet, user generated content and knowledge acquired from knowledge databases.
• To efficiently answer questions posed by users with respect to various domains like business environment, website details, together with an appropriate chat interface.
• To efficiently save the time and efforts by presenting a systematic understanding on several information through detailed research and then making the report terms of our understanding.

• By presenting a rapid voice search mechanism where more time can be saved.

The organization of the paper is as follows: The first section presents a detailed introduction on voice assistant technology along with recent techniques available in market; second section offers a systematic review on various AI based voice assistants and their benefits and limitations. The section 3 is dedicated to our proposed methodology followed by results discussion and analysis in section 4.The study concludes by presenting a summary of our research in conclusion and future enhancements section.

2. Literature Survey

Technological companies like Microsoft's Cortana, Apple's Siri, Amazon's Alexa, Google's Assistant have utlised Natural Language Processing (NLP) to design their voice assistants. These companies have been utilising different approaches to modify their work flow and improve their Virtual Personal Assistants (VPAs) in order to be consistent with the device usage and its complexity. Google has enhanced the functioning of Google Assistant by making use of Deep Neural Networks (DNN) to focus more on dialogue systems .Microsoft employs Microsoft Azure ML Studio with other components to improve the Cortana's language processing system. Amazon makes use of Automatic Speech Recognition (ASR) to transform speech to text, and tongue appreciation to understand the nuances of the text, thereby allows the developer to design voice assistants to enhance consumer experiences and enable practical conversational capabilities.

Majority of the virtual voice assistants have female voice although the user can change the voice and tone as per their need. As voice Assistants allows us to ask about anything, be the weather information or location of places, it also allows us to access translated information in almost over 100 languages. This feature of Google Assistant helps in home automation to control the home from remote places, favorite playlist can be played, and all of this can be done directly from smartphone through hands-free speech recognition process.

Cortana is probably one among the foremost quintessential multi-device, multi-sense surface areas. Cortana is a part of Windows Shell built into Outlook, whose special abilities in scheduling and assigning meetings together with a Bot Framework to build skills needed to engage in conversation with other digital assistants. It also learns about our time to be as useful in offering suitable answers along with completing basic tasks.

Alexa is the voice service of Amazon Echo, Echo Show and Echo Dot which allows customers to form personalized experience by offering capabilities to implement their skills. Companies like Uber, Capital One, Starbucks etc., makes use of Alexa-enabled gadgets to enhance their businesses. Following are some of the common tasks performed by voice assistants like

• Setting Reminders and alarms
• Sending and receiving messages
• Creating calendar entries
• Email briefing
• Scheduling meetings
• Play music
• Entertainment
• Gaming
• Weather forecast
• Voice based home automation
• Multi-language answering abilities
• Location information
• Maps
• Cloud and other online services

The voice assistants discussed above have certain limitations like most of the time is consumed in entering the entries than actual work getting dome and they do not maintain a knowledge database of their own and its understanding comes mostly arises from the information captured from domain models and data models.

3. Proposed System Architecture

Our familiarity with existing voice assistants like Alexa, Siri, Google Assistant, Cortana that uses concept of Natural language processing, and speech recognition. By listening to the command given by the users, the requirements are understood and specific function in performed in an efficient manner. Artificial Intelligence has been used to generate accurate results and reduce the overall effort and time while performing any task. The conventional typing has been reduced completely and this assistant has been designed to imitate a human assistant in facilitating an effective operation in hand. The algorithm used focuses more on the time complexities and reduces time. In order to use virtual voice assistants its mandatory to have accounts like Google for Google assistant, Microsoft account for Cortana etc.,and can be used only with internet connection. Our software is versatile and can be integrated with several devices like, phones, laptops, and speakers etc.

Our proposed smart voice assistant can send emails without typing a word, can search on Google without opening the browser, and perform many other tasks like playing music, opening your favorite IDE etc., with the help of a single voice command. Moreover, its different from other traditional voice assistants which are specific to desktop and requires separate account to use this, this does not require any internet connection while getting the instructions to execute any specific task. The IDE used in this project is PyCharm along with modules and libraries that were used are pyttsx3, Speech Recognition, Datetime, Wikipedia, Smtplib, pywhatkit, pyjokes, pyPDF2, pyautogui, pyQt etc. A GUI has been developed for interacting with the Virtual Assistant whose design and look has been made to enhance the conversation. With further advancement our proposed Assistant can perform any task more effectively than humans. AI is rapidly emerging in every field and decreases human efforts and saves more time and resources. Functionalities of this project include the following

- Sending emails
- Reads PDF
- Sends text on WhatsApp
- Opens command prompt
- Opens our favorite IDE, notepad
- Plays music
- Does Wikipedia searches
- Opens websites like Google, YouTube, etc., in a web browser
- Gives weather forecast
- Desktop reminders can be set of our choice.
- Performs basic conversation.

The system has been designed using the concept of AI and with the help of packages like Python whose libraries and packages have been used to perform the tasks, for example pyPDF2 can be used to read PDF. The data in this project is nothing but user input (Fig 2), whatever the user says, the assistant performs accordingly. The user input comprises of list of tasks that a user wants to carry out specified in human language i.e., English.

Fig 2: Representation of AI Assistant receiving commands from User

AI when used with machines has the capability of thinking like humans. In this project, a computer system has been designed to interact with human using Python which is an emerging language that has been used for scripting the Voice Assistant. The instructions for the assistant are handled as per the user requirement of user and Speech recognition is the Alexa, Siri, etc.
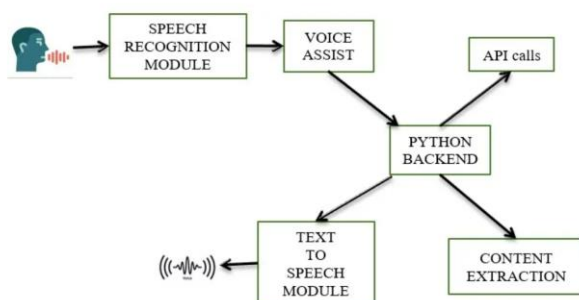

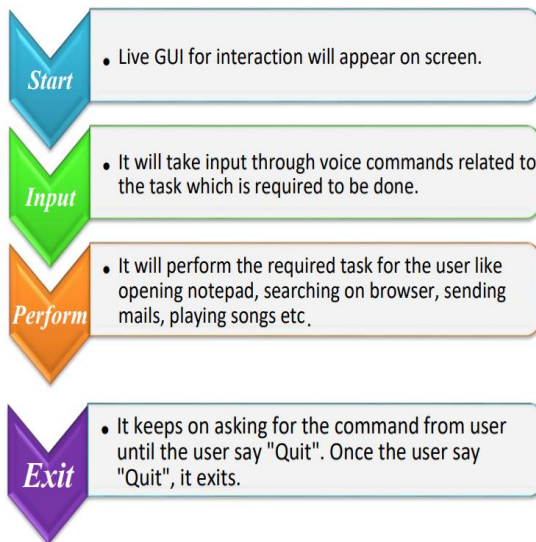**Fig 3: System Architecture for AI Desktop Assistant**

Fig 4: Block diagram of the proposed model

In Python, API known as Speech Recognition is present which facilitates us to convert speech into text thus sending emails and searching on Google can be done without opening the browser, and daily tasks like playing music, can be done with the help of a single voice command. In the current scenario, advancement in technologies like AI is such that performing complex tasks can be done with more effectiveness than us. Our proposed architecture and block diagrams are presented in figures 3 and 4.

4. Results Analysis and Discussion

The desktop assistant is a voice assistant that can perform many daily tasks of desktop like playing music, opening your favorite IDE with the help of a single voice command. Assistant is quite different from other traditional voice assistants in terms that it is specific to desktop and user does not need any account to use this, it does not require any internet connection while obtaining the instructions to perform any specific task. As the first step, install all the necessary packages and libraries. The command used to install the libraries is "pip install" and then import it. The necessary packages included are as follows:

Libraries and Packages

pyttsx3: It is a python library which converts text to speech.
Speech-Recognition: It is a python module which converts speech to text.
pywhatkit: It is python library to send WhatsApp message at a particular time with some additional features.
Date/time: This library provides us the actual date and time.
Wikipedia: It is a python module for searching anything on Wikipedia.
Smtplib: Simple mail transfer protocol that allows us to send mails and to route mails between mail servers.
pyPDF2: It is a python module which can read, split, merge any PDF.
Pyjokes: It is a python library which contains lots of interesting jokes in it.
Webbrowser: It provides interface for displaying web-based documents to users.
Pyautogui: It is a python librariy for graphical user interface.
0S: It represents Operating System related functionality.
sys: It allows operating on the interpreter as it provides access to the variables and functions that usually interact with the interpreter.

Functions

takeCommand(): The function is used to take the command as input through microphone of user and returns the output as string.

wishMe(): This function greets the user according to the time like Good Morning, Good Afternoon and Good Evening.

taskExecution(): This is the function which contains all the necessary task execution definition like sendEmail(), pdf_reader(), news() and many conditions in if condition like "open google", "open notepad", "search on Wikipedia" ,"play music" and "open command prompt" etc.The system testing is done on fully integrated system to check whether the requirements are matching or not. The system testing for desktop assistant focuses on the following four parameters:

Functionality: In this we check the functionality of the system whether the system performs the task which it was intended to do. To check the functionality each function was checked and run, if it is able to execute the required task correctly then the system passes in that particular functionality test. For example, to check whether Assistant can search on Google or not, user said "Open

Google", then Assistant asked, "What should I search on Google?" then user said, "What is Python", Assistant open Google and searched for the required input.

Usability: Usability of a system is checked by measuring the easiness of the software and how user friendly it is for the user to use, how it responses to each query that is being asked by the user. It makes it easier to complete any task as it automatically do it by using the essential module or libraries of Python, in a conversational interaction way. Hence any user when instruct any task to it, they feel like giving task to a human assistant because of the conversational interaction for giving input and getting the desired output in the form of task done. The desktop assistant is reactive which means it know human language very well and understand the context that is provided by the user and gives response in the same way, i.e. human understandable language, English. So user finds its reaction in an informed and smart way. The main application of it can be its multitasking ability. It can ask for continuous instruction one after other until the user "QUIT" it. It asks for the instruction and listen the response that is given by user without needing any trigger phase and then only executes the task.

Security: The security testing mainly focuses on vulnerabilities and risks. As Assistant is a local desktop application, hence there is no risk of data breaching through remote access. The software is dedicated to a specific system so when the user logs in, it will be activated.

Stability: Stability of a system depends upon the output of the system, if the output is bounded and specific to the bounded input then the system is said to be stable. If the system works on all the poles of functionality then it is stable.

Project Management plan

The project has been designed by our team. From installing of all the packages, importing, creating all the necessary functions, designing GUI in PyQT and connecting that live GUI with the backend, was all done by us. We have done all the research before making this project, designed the requirement documents for the requirements and functionalities, wrote synopsis and all the documentation, code and made the project in such a way that it is deliverable at each stage. We have created the front end (.ui file) of the project using PyQt designer, the front end comprises of a live GUI and is connected with the .py file which contains all the classes and packages of the .ui file. The live GUI consists of moving GIFs which makes the front end attractive and user friendly. We have written the complete code in Python language and in PyCharm IDE from where it was very easy to install the packages and libraries, We have created the functions like takeCommand(), wishMe() and taskExecution() which has the following functionalities, like takeCommand() which is used to take the command as input through microphone of user and returns the output as string, wishMe() that greets the user according to the time like Good Morning, Good Afternoon and Good Evening and taskExecution()which contains all the necessary task execution definition like sendEmail(), pdf_reader(), news() and many conditions in if condition like "open Google", "open notepad", "search on Wikipedia" ,"play music" and "open command prompt" etc. While making this project we realized that with the advancement Assistant can perform any task with same effectiveness or can say more effectively than us. By making this project, we realized that the concept of AI in every field is decreasing human effort and saving time. Functionalities of this project include, It can send emails, It can read PDF, It can send text on WhatsApp, It can open command prompt, your favourite IDE, notepad etc., It can play music, It can do Wikipedia searches for you, It can open websites like Google, YouTube, etc., in a web browser, It can give weather forecast, It give desktop reminders of your choice. It can have some basic conversation. At last, we have updated my report and completed it by attaching all the necessary screen captures of inputs and outputs, mentioning the limitations and scope in future of this project.

5. Conclusion and Future enhancements

Using the NLP built-in with synthetic talent we have completed a design of intelligent virtual digital assistant that can manage applications, replies to user queries, and additionally carries out internet searches and converses with the human by interacting intelligently and manipulates the devices, this consists of looking in Wikipedia, opening Google, YouTube, Facebook, MS-Word, MS-Excel, MS-PowerPoint, sending mails, login to your Gmail through voice, lock your PC, restart your PC, shut down your PC, play music, placing alarms, getting climate notifications, etc. Thus, with the aid of examining the current voice assistant systems, we have come up with this proposed machine which is environment friendly and helps us to be more organized. There are lot of enhancements to be developed in the future to address the various user needs like users with various disabilities through recent machine learning algorithms.

REFERENCES

1. Gimpel, H. (2015). Interview with Thomas W. Malone on "Collective Intelligence, Climate Change, and the Future of Work." Business & Information Systems Engineering, 57(4), 275–278. https://doi.org/10.1007/s12599-015-0382-4
2. A. Shawar and E. Atwell, "Chatbots: are they really useful?", LDV Forum, vol. 22, no. 1, (2007)
3. M. Turing, "Computing Machinery and Intelligence", Mind, (1950), pp. 433-460
4. Khanna, "Pandorabots Chatbot Hosting Platform SARANG Bot", (2015) April 19, Internet:http://pandorabots.com/pandora/talk?botid=9f0f09a71e34dcf8/
5. Alice, "Foundation. Free A.L.I.C.E. AIML Set", (2015) March 21, Internet: http://code.google.com/p/aiml-en-us-foundation-alice/
6. Whitby, Author, "Artificial Intelligence", The Rosen Publishing Group, (2009)
7. Xin Lei, Andrew Senior, Alexander Gruenstein and Jeffrey Sorensen "Accurate and Compact Large Vocabulary Speech Recognition on Mobile Devices," in INTERSPEECH. 2013, pp. 662–665, ISCA.

8. Alex Graves, Santiago Fern´andez, Faustino Gomez, and J¨urgen Schmidhuber "Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks," Proceedings of the 23 rd International Conference on Machine Learning, Pittsburgh, PA, 2006.
9. Brian Kingsbury, "Lattice-based Optimization of Sequence Classification Criteria for NeuralNetwork Acoustic Modeling", 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan.
10. Daniel L. Stufflebeam, and Anthony J. Shinkfield, "Evaluation Theory, Models, and Applications", John Wiley & Sons, 2007.
11. David Rybach, Michael Riley, and Chris Alberti, "Direct Construction of Compact Contextdependency Transducers from Data," Proc. of INTERSPEECH, pp. 218–221, 2010. [12] K. Beulen; and H. Ney "Automatic Question Generation for Decision Tree Based State Tying," Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98.
12. Patrick Nguyen, Georg Heigold, Geoffrey Zweig, "Speech Recognition with Flat Direct Models," IEEE Journal of Selected Topics in Signal Processing, Volume: 4, Issue: 6, Dec. 2010.
13. Klaus Beulen, Hermann Ney, "Automatic Question Generation for Decision Tree Based State Tying," Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98.