

Research Article

Implementing a Voice Assistant Using Python and Machine Learning

Ayush Agarwal¹, Abhineet Mishra², Saransh Tiwari³

^{1,2,3}Computer Science and Engineering ,Chandigarh University ,Punjab, India.

I N F O

Corresponding Author:

Agarwal Ayush, Computer Science and Engineering ,Chandigarh University ,Punjab, India.

E-mail Id:

19bcs1270@cuchd.in

Orcid Id:

<https://orcid.org/0009-0002-5950-5903>

How to cite this article:

Agarwal A, Mishra A, Tiwari S. Implementing a Voice Assistant Using Python and Machine Learning. *J Adv Res Intel Sys Robot* 2023; 5(1). 20-24.

Date of Submission: 2023-03-06

Date of Acceptance: 2023-04-02

A B S T R A C T

Artificial has experienced tremendous progress and has a bright future. Natural language processing (NLP) is one area in particular where it has the most potential. Voice assistants can interact with users in natural language thanks to artificial intelligence. They are quite simple to use, it is anticipated that this trend will continue in the rapidly developing digital world. It is reasonable to say that they have cemented their presence in our daily lives, from being in our home speakers to being in our pockets in the form of smartphones. This research article aims to comprehend how voice assistants are becoming increasingly prevalent in our daily lives.

Keywords: Artificial Intelligence , Voice Assistants, Home Speaker, Smart Phones

Introduction

Today, everything we need to live our daily lives depends in some way on this wonderful discovery known as silicon.¹

Artificial intelligence is a new discipline that evolved with the development of the internet. Humans have been fascinated by the idea of building an artificial computer that can perform our fundamental duties for ages, with the emergence of voice assistants like Google Assistant, Apple's Siri, Amazon's Alexa, this dream may very well come true for all of us. Their restrictions are getting smaller by the day, from finding files and documents online to answering basic mathematical tasks.²⁻⁴

Every major tech company in the world is concentrating their efforts and resources on developing a human-like voice assistant that can answer our everyday questions by simply speaking them to our device in recent years due to the growth of the smartphone economy.⁵ The voice assistant (VA) is capable of doing the activities listed below with ease:

- Setting timers or alarms

- Playing music
- Playing videos or any multimedia file
- Making calls and sending messages
- Giving weather information
- Answering queries given by the user
- Controlling other smart devices

Some key elements which differentiates voice assistants from simple programs are:-

- NLP: The ability to understand and process human language
- Machine Learning: The ability to adapt to new things by identifying patterns
- The ability to use stored information and data use it to draw new conclusions

Methodology

The snowball method as outlined by Wohlin (2014) was utilised to find enough and high-quality papers about voice assistant uses. The process consists of the following steps:

- First, run searches in Google Scholar, IEEE Xplore,

Scopus, the ACM Digital Library to compile a starting selection of pertinent documents. The terms “voice assistant”, “smart speaker”, “amazon echo”, “google assistant”, “Alexa” and “Siri” were used

- Use backward and forward snowballing to iterate for the first start set of papers. Forward snowballing refers to finding new papers that cite the paper under consideration, whereas backward snowballing uses the reference list to find new papers to include. New papers that are found in each step using backward and forward snowballing are collected into a pile for the subsequent iteration

Childrens: Several research on youngsters over the age of 14 who use voice assistants have been done. Beirl et al.’s (2019) study looked at how Alexa was used over the course of three weeks. The study’s major goal was to determine how families used Alexa for their kids. Storytelling, games, music were shown to be children’s top interests when it came to voice assistants.⁶

The findings indicated that Alexa was well received and had become a family tradition.

According to the study’s findings, all of the aforementioned contacts promoted social and emotional closeness and increased family cohesiveness. Another significant finding of the study was that families took helper roles to encourage and provide advice for what younger children should say to Alexa when they were struggling to follow the rules of a game or quiz.

Children have also discovered that using voice assistants while playing video games or having fun with their friends can be quite helpful. For instance, a 14-year-old boy once requested Alexa to purchase lunch for him and his pals while they were playing their preferred video game, he claimed that the entire voice assistant food ordering process only took him two to three minutes. The fact that he could now order everything exclusively through this new pathway left a lasting impression on his memory, he no longer needed to use his phone to complete the process.⁷⁻⁸

Elderly People

Savago et al. (2019) look into how elderly adults (65+) use voice assistants.

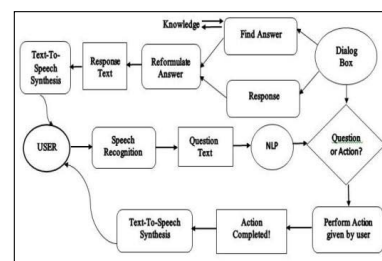
According to survey results, individuals were impressed by the variety of voice assistant possibilities. In addition, because it does not interfere with their process and may be difficult for them to move, as one participant stated, participants enjoyed the concept of being able to complete some activities using only speech. It should be mentioned that voice assistants were mostly utilised as search engines, but they were also useful as translators and even teachers. The authors also emphasise the need for more research on how older individuals utilise voice assistants.⁹

Elderly folks also use it to listen to their older music and perform web searches, which they claim takes three times longer on a phone. When looking for songs, devotional songs take precedence, with family-related song lyrics coming in second. Additionally, it boosts their propensity to ask Siri or Alexa to perform a certain activity, according to researchers.

Adults

Sciuto et al. (2018) studied how households use conversational bots in their daily life and found some interesting results. Specifically, they examined a total of 278,654 voice commands from the logs of 75 Alexa users. Those who have owned an Alexa device for at least six months responded to survey questions about how often Alexa is used in their home. In spite of the fact that no information could be gleaned from the log files about which household member issued each command, 26 of the 75 participants reported having children. Parents who were questioned remembered their kids successfully interacting with Alexa before they had ever done so with smartphones or other technologies.¹⁰

In a study by Song (2019), 433 adult participants participated in an online survey to gauge their attitudes towards voice assistants, their perceived utility, their perceived ease of use. Findings imply that attitudes towards voice assistants and behavioural intentions to utilise this technology are significantly influenced by perceived utility. Additionally, buyers prefer to purchase user-friendly equipment.¹¹



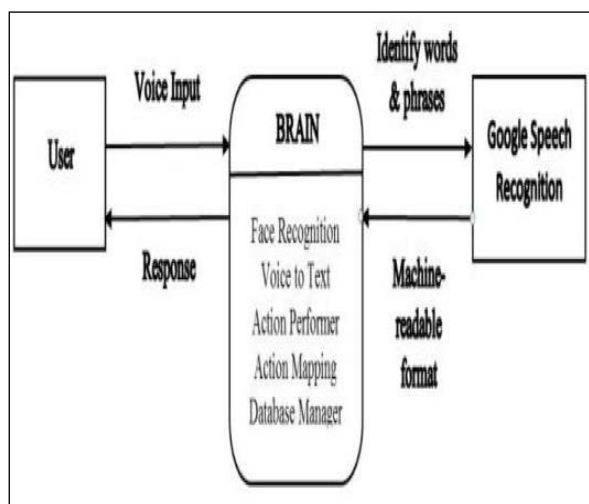


Figure 2.DFD Diagram Level 0

Level 1 DFD

The user can provide input by speaking. The Google Voice API will translate this voice data into text, the voice assistant will carry out the task in accordance with the user's instructions by comparing them to the dialogue box and knowledge base.

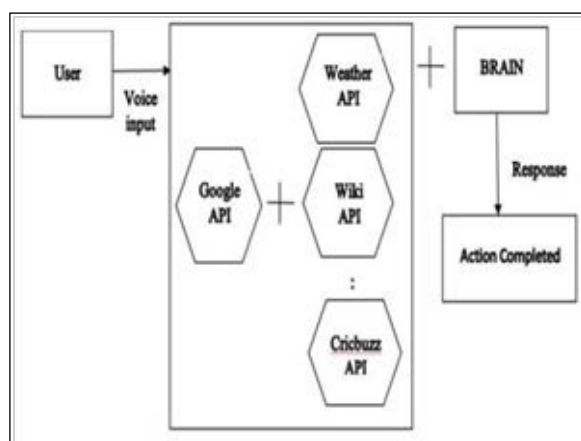


Figure 3.DFD Diagram Level I

Flow Chart

The visual representation of algorithms is a flow chart. Flow charts are represented by various symbols. When the system first starts, it authenticates the authorised user before the voice assistant begins to run in the background and listen for voice commands. Once the user issues a command, the voice assistant analyses it and outputs what is required based on the conditions it has been given. This output is given to speech recognition, which turns the voice into a format that computers can understand. The personal voice assistant subsequently completes the desired task based on the input obtained.¹²

Proposed System Feature

Python offers a sizable standard library that covers topics including operating system interfaces, string manipulation, web services tools, internet protocols. The standard library already contains scripts for many high-use programming tasks, which drastically decreases the amount of code that needs to be developed.

The boost in Python's speed and productivity can be attributed to the language's clear object-oriented architecture, increased process management capabilities, robust text processing and integration capabilities, as well as its own unit testing framework. Python is regarded as a practical choice for creating intricate multi-protocol network systems.

Text in everyday language is translated into speech using a text-to-speech (TTS) system. Concatenating segments of recorded voice that are kept in a database can produce synthesised speech. Speech is delivered as the output. With worldwide hands-free and eyes-free speech control over their mobile devices, this voice assistant can help a lot of consumers. Future voice control technology may take a cue from this device's structure.

Facial Recognition Technique (FRT) has the capacity to eliminate the need for passwords, fingerprint data, even keys.

Security and Privacy Concerns

Although the voice assistant is a huge aid in modern living, there is always a downside. Many independent agencies and experts have highlighted concerns about the potential malicious use of voice assistants in a number of concerts.

Many people have become so reliant on this new technology that they are now asking to perform the most elementary tasks using it.

Many people who do not use the new technology have claimed that it is only a flashy tool for large, multinational corporations to charge their customers more money. They discovered that the use of this technology is not as widespread as was previously thought. They claimed that with the bare minimum of effort, they could have accomplished the same thing on their own.

Voice assistant use among kids was reportedly enjoyable, according to some parents. The majority of them bought the gadgets to be among the first to use and early adopters of the technology. Authors' suggestions include giving kids their own accounts and giving them the option to mute devices via voice commands.

They also advocate for the inclusion of an incognito mode, akin to that of web browsers, to stop corporations from collecting your personal information. McReynolds et al.

(2017) provide indicators while recording to solve privacy concerns for recording devices.

Cho (2019) investigated the modality effects of voice assistants in the context of acquiring health information. 53 undergraduate students from a southeastern university's undergraduate programme used a Google Home device and a smartphone with the Google Assistant.

Participants were told to use voice or text input to ask two different kinds of health inquiries (i.e., less sensitive versus more sensitive health concerns).

However, only when low sensitivity material was involved and when users expressed fewer privacy concerns did the results reveal that voice versus text interactions significantly boosted felt social presence towards the voice assistant.

However, when a person's privacy concerns and the sensitivity of the material were high, text had an equal influence to voice in creating a human-like perception of the voice assistant.

Additionally, research indicates that voice interactions improved the user's perception of their interaction with the voice assistant as a social dialogue, which improved their opinion of it.

This, however, only happened when asking for less sensitive information and when people expressed modest degrees of privacy concerns.

Conclusion

The 21st century has seen a significant change in how we communicate with one another thanks to immersive technologies. Although voice assistants can offer fresh learning opportunities, we also need to be aware of their drawbacks. Since this is a new technology still in its infancy, there hasn't been much research done on it. However, as time goes on, we'll learn more about its benefits and drawbacks. One of the main problems with voice assistants is that only a few languages are supported, which limits the user base. However, several big businesses, such as Google and Amazon, are already making significant progress in resolving this issue by pouring enormous amounts of money into research and development. It is important to teach adults, kids, the elderly about the benefits of this technology and how to utilise it effectively without impairing their mental capacity. Conclusion: For the time being, this new technology has a restricted function, but as the world moves towards greater digitalization, we will see one of the most revolutionary technologies in our lifetimes.

Future Scope

Using the framework voice assistant technology can be updated further. Cyber security concerns are very real these days so it can be combined with the system to give further

boost in security. Technologies such as voice authentication can be used to make sure that the command is given by the right user.

Work and advancements can also be made on how pitch and tone is judged. Accents from different regions can also be upgraded to make it truly a global technology. Further modifications like learning the answers of questions that are yet not known by voice assistant can be implemented.

References

1. Cho E. Hey Google, Can I Ask You Something in Private? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems 2019; 258 ACM.
2. Davie N, Hilber T. Opportunities and Challenges of Using Amazon Echo in Education. *International Association for Development of the Information Society* 2018.
3. Dousay TA, Hall C. Alexa, tell me about using a virtual assistant in the classroom. In EdMedia+ Innovate Learning 2018; 1413–141. Association for the Advancement of Computing in Education (AACE).
4. Ilhan K, Mušić D, Junuz E et al. Scarlet-Artificial Teaching Assistant. In 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO) 2017; 11–14. IEEE.
5. Juniper Research. Amazon Echo & Google Home to Reside in over 50% of US Households by 2022, as Multi-Assistant Devices Take Off 2017. <https://www.juniperresearch.com/press/press-releases/amazon-echo-google-home-to-reside>
6. Kukulska-Hulme, A. Intelligent assistants in language learning: friends or foes?. In the World Conference on Mobile and Contextual Learning 2019; 127–131.
7. López G, Quesada L, Guerrero LA. Alexa vs. Siri vs. Cortana vs. Google Assistant: a comparison of speech-based natural user interfaces. In *International Conference on Applied Human Factors and Ergonomics* 2017; 241–250. Springer, Cham.
8. McLean, G., Osei-Frimpong, K. (2019). Hey Alexa... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, 99, 28–37.
9. Nilsen, J. H., Røyneland, K. (2019). «It knows how to not understand us!» A study on what the concept robustness entails in design of conversational agents for preschool children (Master's thesis).
10. Rzepka, C. (2019). Examining the Use of Voice Assistants: A Value-Focused Thinking Approach. 25th Americas Conference on Information Systems, Cancun, 2019.
11. Trivedi, N. (2018). ProblemPal: Generating Autonomous Practice Content in Real-Time with Voice Commands and Amazon Alexa. In E-Learn: World Conference on E-Learning in Corporate, Government, Health-care, Higher Education (pp. 80– 82). Association for the

Advancement of Computing in Education (AACE).

12. Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering. 2014 38). ACM. DOI: 10.1145/2601248.2601268
13. Wiederhold BK. Alexa, Are You My Mom?" The Role of Artificial Intelligence in Child Develop- ment. *Cyberpsychology, behavior and social networking* 2018; 21(8): 471–472.