

Research Article

A Study on the Usability of Desktop Voice Assistants

Naveen Kumar¹, Bhanu Pratap Singh², Meghna Lekhwla³

^{1,2,3}Chandigarh University Mohali, India.

I N F O

Corresponding Author:

Naveen Kumar, Chandigarh University Mohali, India.

E-mail Id:

19bcs1724@gmail.com

Orcid Id:

<https://orcid.org/0009-0005-2075-2896>

How to cite this article:

Kumar N, Singh BP, Lekhwla M. A Study on the Usability of Desktop Voice Assistants. *J Adv Res Intel Sys Robot* 2023; 5(1): 6-8.

Date of Submission: 2023-03-20

Date of Acceptance: 2023-04-23

A B S T R A C T

A personal voice assistant is software that can carry out tasks and offer various services to the user in accordance with the user's given orders. This is accomplished by a synchronous procedure that starts with the recognition of speech patterns and ends with a synthetic speech response. These assistants enable users to automate a variety of tasks, including media playback, task management, mailing. People are getting increasingly dependent on technology as it advances, the computer is one of the most popular platforms. We all want to make the use of these computers more comfortable, traditional way to give a command to the computer is through keyboard but a more convenient way is to input the command through voice. Giving input through voice is not only beneficial for the normal people but also for those who are visually impaired who are not able to give the input by using a keyboard. For this purpose, there is a need of a voice assistant which can not only take command through voice but also execute the desired instructions and give output either in the form of voice or any other means.

Keywords: Desktop, Software, Management, Keyboard

Introduction

After 2017, the market for virtual assistants has seen a sharp increase in the number of products available. Numerous new features are being added to PCs and mobile phones as a result of technological improvement. We need a method of input that is both quick and dependable if we want to use them in a more enjoyable and easy way. The microphone in our project is utilised to transform acoustic energy into electrical energy when we want to input data into the system. After receiving the input, it is necessary to comprehend the audio signal in order to use the Google API. These voice assistants work as your companion which performs your day by day task with minimum efforts and also help the user to function better by giving daily updates. It was after the recognition of importance of voice

commands in day to day life that we have aimed to develop a personal assistant for desktop which will do every work from playing music to sending an Email.

Features

The following characteristics will be included in the development of the System:

1. When contacted with a specific specified functionality, it wakes up into action after continuously listening inactively.
2. Web browsing based on the user's spoken parameters, followed by the release of the desired output through audio and It will print the output on the screen and print it simultaneously.

Speaking commands cause the computer to perform

additional helpful functions including playing any type of media, viewing weather predictions, setting reminders, shutting down the computer, sending an email, etc.

Project Design

Speech Recognition

Google's Speech Recognition API is the speech recognition module used by the programme, it is imported in Python using the command "import speech_recognition as sr". This module is used to identify the voice that the user provides as input. This is a free API that Google offers and supports. This is a really lightweight API that aids in making our application smaller.

TTS & STT

The speech recognition module initially converts the voice that is provided as input into text. The text is subsequently processed to produce the answer to the user's query. The result of the processed query is converted to voice as the final output in the last phase. STT takes the longest of the two because the system must first listen to the user, various users speak in a variety of voices, some of which are simple to understand and others which are difficult to hear. Our entire execution time is reliant on this phase. Once the speech has been transformed to text, it doesn't take long to execute commands and return the results to the user.

PYTTSX3

Python 2 and Python 3 both support the offline module pyttsx3 which is used for text to speech conversion. Additionally, only this module contains the run and wait capabilities. It affects the time interval between inputs, or how long the system will wait for another input. The Python community offers this free module, which may be installed using the pip command like other modules.

Date Time

The date and time's functionality is supported by the import of the DateTime module. For instance, the user might wish to know the time and date right now or schedule a job for a specific time. In a nutshell, this module allows classes that only conduct operations based on the date and time. This module is crucial, especially for activities where we need to keep track of the time. This tiny module aids in managing the overall size of our programme. The system will lag and respond slowly if the modules are too big or heavy.

Web Browser

This module enables the system to show users web-based data. For instance, the user types "Open Google" when he wants to access any webpage. The user receives a browser with Google open after the input has been processed using the web browser module. The web browser that will be used is the one that is set as the default.

Wikipedia

A Python module called Wikipedia makes it feasible for the virtual assistant to process user requests for information on the website and show them the relevant information. Since this is an online library, an internet connection is required to retrieve the results. The user can manually set the number of lines they want to receive as a result.

OS Module

An operating system-dependent functionality is provided by the OS Module. All of these types of functionalities are available in an OS module if we wish to execute actions on files like reading, writing, or manipulating paths. Any errors, such as invalid names, paths, or arguments that may be correct or incorrect but are simply not accepted by the operating system, cause all of the actions to produce the error "OSError".

SMTPLIB

This module for working with emails and email servers is available in the standard library of Python. The "SMTP client session object," which is defined by the SMTPLIB, is used by the user to send emails. Sendmail(), quit, initialise are the 3 steps. When the optional host and port parameters are supplied, the connect method is used with these arguments during the initialization process.

Innovation in Project

The software's primary function is to carry out the user's instructions when given in either text or speech, depending on the command. The majority of the user's work will be made easier because a whole task can be accomplished with just one command. The virtual assistants seen on Windows are where Jarvis gets its inspiration. Users can communicate with the assistant by speaking commands and carrying out voice searches. This voice assistant will come in handy for anyone who has hand disabilities.

Proposed Architecture

1. The system design consists of
2. Taking the input as speech patterns through microphone.
3. Audio data recognition and conversion into text.
4. Comparing the input with predefined commands
5. Giving the desired output The initial phase includes the data being taken in as speech patterns from the microphone. In the second phase the collected data is worked over and transformed into textual data using NLP.

Result and Discussion

Methodology

1. We made a good plan for this project's work and assigned tasks to each team member.

2. To create a voice assistant that is user-friendly, we are employing Python in this project.
3. It has been created to offer a user-friendly interface for performing a number of operations by using specific, clearly defined instructions.
4. The user can communicate with the assistant by using voice instructions, the assistant will respond.
5. Main role of Desktop Voice Assistant
6. It can send emails on your behalf.
7. It can play music for you.
8. It can do Wikipedia searches for you.
9. It is capable of opening websites like Google, Youtube, etc., in a web browser.
10. It is capable of opening your code editor or IDE with a single voice command.

Voice Assistant Privacy Concerns

Personal assistants require a lot of data and are always listening, thus a user can be concerned about their privacy. Then, practically all of the major companies—Amazon, Apple, etc.—retain and sort this passive data using employees. Concerns have been raised about the kind of information that these employees and contractors were hearing in addition to the finding that the Ai could record our audio chats. Therefore, a privacy policy must exist in the cloud-based voice assistant to safeguard user data.

Conclusion

They have altered the basic genome of the sphere in which they operate. While this has largely created a better world with drastic benefits for communities, which were before kept in dark with reference to technological innovations, they have posed new kind of threats with respect to user's privacy and security. The main agenda and purpose is they taking input from the user in the form of voice. Converting the speech into text to be processed by the assistant. Now, the transformed text is processed to produce the desired outcomes. One or two keywords in the text decide which query will be run. The assistant prompts the user to speak again if the keyword doesn't match any of the coded queries. To provide the user with results, the text-based result is once more transformed to speech.

Acknowledgement

We are thankful to our management "Chandigarh University, Mohali, India" for giving a support to write this.

References

1. Sutar Shekhar, Pophali Sameer, Kamad Neha, Deokate Laxman, intelligent voice assistant using Android platform. *International Journal of Advance research in computer Science and Management Studies*. Volume 3, Issue 3, march 2015
2. Mhamunkar, M. p. v., Bansode, M. k. S., & Naik, L.S. (2013). Android application to get word meaning through voice. *International journal of Advance Research in Computer Engineering & Technology (IJARCET)*. 2(2), pp- 572.
3. Apte, T. V., Ghosalkar, S., pandey, S., & Padhra, S. (2014). Android app for blind using speech technology. *International Journal of Research in Computer and Communication Technology (IJRCCT)*, #(3), 391-394.
4. Anwani, R., Santuramani, U., Raina, D., & RL, P. Vmail: voice Based Email Application. *International Journal of Computer Science and Information Technologies*, Vol. 6(3), 2015
5. Deepak shende, Ria Umahiya, Monika Raghorte, Aishwarya Bhisikar, Anup Bhange. AI based voice assistant using python. *JETIR*, february 2019, vol 6, issue 2.
6. Thakur, N., Hiwrale, A., Selote, S., Shinde, A. and Mahakalkar, N., Artificially Intelligent Chatbot
7. Yu, T. L., Gande, S., & Yu, R. (2015, January). An open-Source Based Speech Recognition Android Application for Helping Handicapped Students Writing Programs. In *Proceedings of the International Conference on Wireless Networks (ICWN)*
8. The Steering Committee of The World Congress in Computer Science, Computer Engineering And Applied Computing (WorldComp).
9. Yannawar, P. (2010). Santosh K. Gaikwad Bharti W. Gawali Pravin Yannawar. A Review on Speech Recognition Technique. *International journal of Computer Application*
10. Brandon Ballinger, Cyril Allauzen, Alexander Gruenstein, Johan Schalkwyk, On-Demand Language Model Interpolation for Mobile Speech Input *INTERSPEECH 2010*, 26-30 September 2010, Makuhari, Chiba, Japan, pp 1812-1815.
11. *IOSR Journal of Engineering* Mar. 2012, Vol. 2(3) pp: 420-423 ISSN: 2250-3021 www.iosrjen.org 420 | "Android Speech to Text Converter for SMS Application" Ms.
12. Anuja Jadhav* Prof. Arvind Patil** Michael Stinson, Sandy Eisenberg, Christy Horn, Judy Larson, Harry Levitt, Ross Stuckless" *REAL-TIME SPEECH-TOTEXT SERVICES.*"