

SPYDER: Intelligent Voice Assistant

Piyush Gupta¹, Sakshi Yadav²

¹U.G. Student-Computer Science Department, Medi-caps University, Indore (M.P), India

²Assistant Professor- Computer Science Department, Medi-caps University, Indore(M.P), India

Abstract - Voice assistants are smart devices that respond to particular voice commands by providing information or completing tasks as asked by the user, using speech recognition, voice synthesis and language processing algorithms. Voice assistants are in extraordinary demands as of late. The fundamental driver for this shift towards voice UIs is changing client requests. There is an expanded in general mindfulness and a more significant level of solace showed explicitly by millennial purchasers. Although requesting a machine to Google search something can seem easy, the technology at the backend is entrancing. This voice assistant will use a microphone to collect voice feedback, which will then be converted to text using pyttsx3 (python text to Speech) module. A keyword check is done after the text has been parsed. Our voice command system works by scanning the text for keywords to imitate. Once the keywords are paired, it will provide appropriate results. The result is stored in the text format. Finally, the Text-To-Speech engine converts text to an English audio file, which is then played back using the Python's play sound package.

Key Words:- Intelligent Voice Assistant, Speech Recognition, Personal Assistant, Speech Synthesis, TTS engine.

1. INTRODUCTION

Digital voice assistants have become one of the most significant advances in user interface and user experience in recent years. We also use them for certain activities such as light switch-off/on and playing music via streaming applications such as Spotify and YouTube Music. This modern way of communicating with technical devices introduces lexical conversation as a new way of this technology. Verbal communication has become the latest interface to technology because of its simple means of communicating with technical systems without the need of any hectic touch

The virtual assistants will be helpful in certain areas such as Home automation, HR-related activities, IT related tasks, voice based search and so on. This virtual Voice search assistant will be future for the next generation, as people have become reliant on voice assistants for all of their needs and daily-routine activities. In this proposal, we have created an Intelligent Voice Assistant that can perform all of these tasks with minimal effort.

Individuals these days need a personal assistant that tunes in to their calls, expects our desires, and can make an important

move when required. This should be possible by utilizing voice assistants that makes our life simpler and can help our day by day life exercises [10]. This sumptuous way of life is currently conceivable in light of voice assistant based on Artificial Intelligence.

As a personal assistant, this project can assist the end-user with day-to-day activities like searching queries on Google, Yahoo (or any search engines), browsing videos on Youtube, knowing the live weather conditions, searching queries on Wikipedia, translating the conversation in any language, Playing Jokes, email exchange, calculating Mathematical expression and get a quick response of each query. The user's voice statement/command is analyzed with the help of machine learning to give optimal solution.

2. LITERATURE SURVEY

Voice Assistants have been the subject of a host of recent studies. We collected a number of papers in order to examine voice assistants and then analyzed the data.

Yash Mittal el al. [1] proposed a study on Smart Home Automation System. This system can be adapted to a user's voice and interpret the voice commands in order to manage their home appliances and devices for various functionalities and their applications. The Voice recognition is processed using an Arduino microcontroller for commands processing and controlling the appliances. Through their automation system prototype they figured out, it can be used to turn existing homes into smart homes at a minimal cost and with ease.

Ravivanshikumar Sangpal [12] et al. came with up a examination on JARVIS which combines with the power of Artificial Intelligence Markup Language with the industry leading Google platform for text-to-speech with male pitch. The use of Artificial Intelligent markup language [AIML] makes Jarvis flexible illustrating its widespread reusability and low maintenance. They also talked about future scope of JARVIS in Artificial Intelligence.

Emad S. Othman [3] implemented Voice Command System as a personal voice assistant that can handle various tasks for end users using Raspberry Pi microcontroller as the main hardware. He outlined the framework configuration of the Voice Command System. Modules with open source systems allows the machine to work at its best in terms of space-time complexity.

Kshama V. Kulhalli [4] claimed that their proposed system can be used with or without Internet access that takes user input in the form of voice or text and analyses it before returning the output in different ways such as an action to be taken or a search result that is dictated to the end user. It is intended to assist blind people, who rely on voice commands. The device is language-independent and responds to user voice commands more quickly than online voice search apps.

Takahiro Kawamura [5] et al. proposed work describes a voice assistant that derives its information from Open Data. It has features such as improved accuracy based on user input and the retrieval of unregistered data by user engagement. They also demonstrate an application to assist its fieldwork and validate its efficacy.

Ankit Pandey [6] et al. presented intelligent voice assistant that can handle calendar meetings, emails exchanges, and operate to-do lists notes. The aim was to create an intelligent personal assistant with Python that allows users to monitor devices with the voice commands, collecting information, and execute desirable tasks on ones desktop system.

Subhash S [7] et al. proposed AI based voice assistant that can commonly used in smart phones and laptops. This AI based virtual assistant will collect audio input from the system's microphone and translate it to text, which will be transmitted via pyttsx3 (Google text to speech) to end user. The author claimed that their proposed work can be helpful in education field, daily life application, home appliances etc.

3. PROPOSED PLAN OF WORK

Based on past works and assessment, the project is accomplished with understanding of an intelligent assistant capable of taking user's command, analyze it and respond to the user by using vocal media [6]. Python libraries and Speech Recognition APIs are used to integrate the personal voice assistant. Python Text To Speech module is used to

translate voice responses.

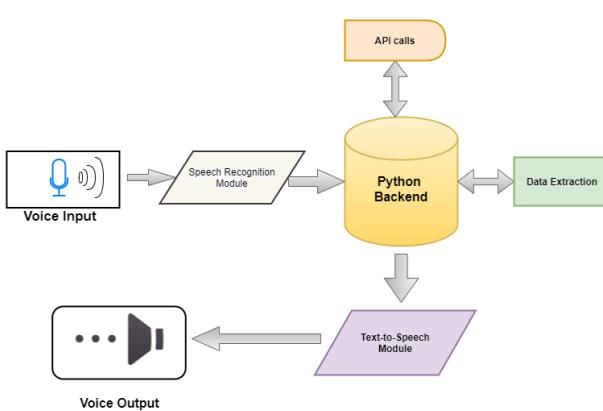


fig -1: major components of our proposed system

The major components are described below-

1. Speech Recognition Modules: As shown in the fig 1.0, the system uses speech recognition module named python-text-to-speech (pyttsx3) module. It is used for converting user voice input to text. The pyttsx3 module in Python has pre-installed speech-to-text (STT) and text-to-speech (TTS) engines. Therefore, it is compatible to use this system without internet access as well [8]. The pyttsx3 module is available in a variety of languages such as English, Hindi, Tamil, French, German etc. It makes this system more interactive with the user so that the task can be performed easily by giving the voice command as input.

2. Python Backend: The python backend receives the user's query in the form of text and determines if the query is an API call or data extraction then result is sent back to python backend, which provides the user with the appropriate output. Python Backend consists of rich libraries that make this system robust. At any point of time, this system is able to react with an appropriate response.

3. API calls: API acronyms for Application Programming Interface. An API is a program interface that enables two programs to communicate with one another. In other words, an API is the intermediary that sends the request to the provider and then returns the response of that request.

4. Data Extraction: Data extraction is the process of extracting organized information from the unorganized machine-readable documents. In most of the cases, this method involves Natural Language Processing (NLP) to interpret human language documents and sentiment analysis [9] so that retrieval of meaningful and qualitative data from the system can be done. In our proposed system, we have extracted the relevant data that has been asked by user. Hence, the use of Natural language processing enhances the performance of our system.

5. Text-to-speech module: The TTS module is used to convert text into voice. The text-to-speech module is extremely beneficial to people who have difficulty in reading. This system uses python text to speech module [pytthsx3] module to implement this functionality. TTS engines can be implemented in a variety of languages, dialects, and advanced vocabularies by using third party libraries.

4. FLOW OF EVENTS OF OUR PROPOSED SYSTEM AND ITS WORKING

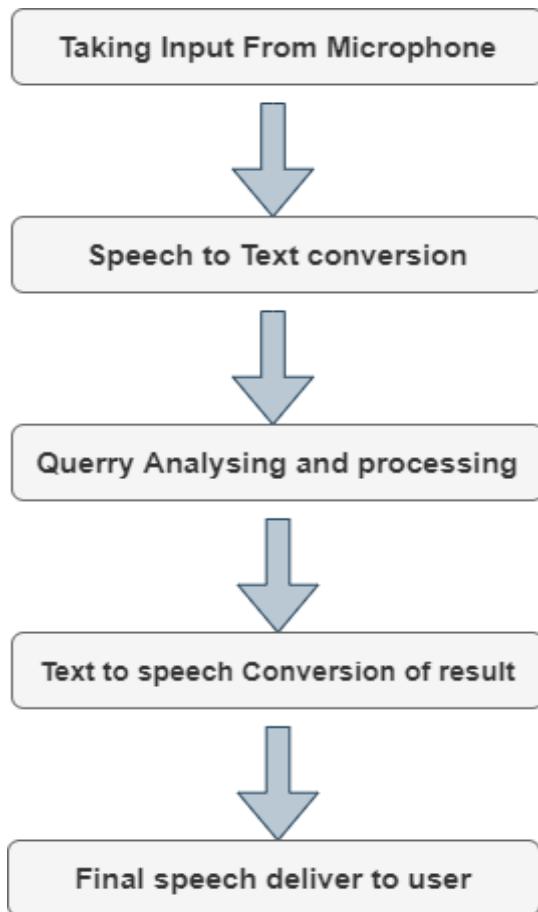


fig -1: flow events of our system

First of all, User sends the input through the system's microphone. The system receives voice input and sends it to a speech-to-text translator that transforms the voice input into the text output. The text is parsed and a keyword search is performed. This proposed system is based on the keywords matching system, which scan the text for keywords to match. Then natural language processing is performed on the text so that meaningful data or keyword can easily be extracted. Once the keywords are paired, it will give the relevant output. The result is in the format of text [3]. The text is then translated to voice using the speech recognition module. Then the voice generated output is to be delivered to the user through speaker.

WORKING

The proposed system's Graphical User Interface was created using the QT designer tool. QT designer is an open source GUI builder that generates UI files rather than code in any programming language. We have used the QT designer GUI platform to make a user-friendly graphical GUI for our proposed framework. Here are some screenshots of Graphical User Interface (GUI) which responds to the query

given as an input.. The Graphical user interface of this system is fully based on verbal communication user interface that makes this system useful for visually disabled people as well [11].

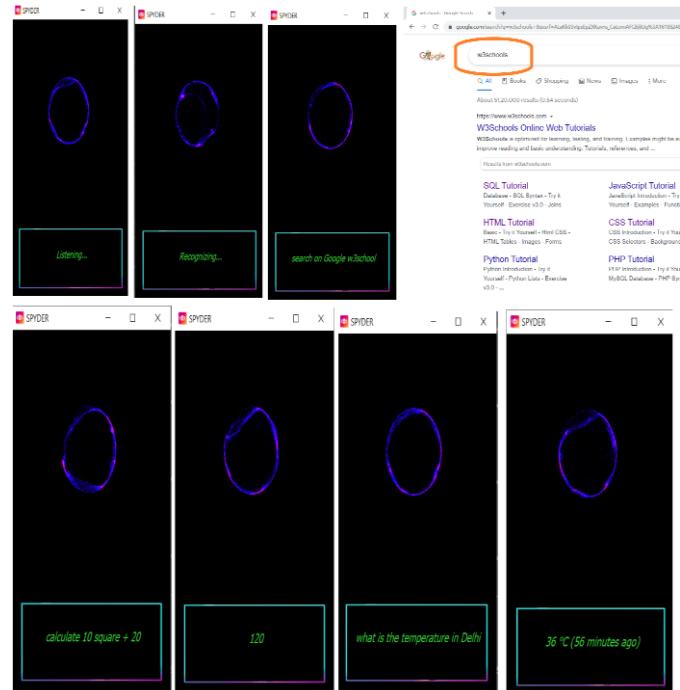


fig -1: ScreenShots of our system

Our project employs a query processing module that functions in a manner close to that of many other query processors. It effectively takes the user's input, searches for relevant output, and then provides the user with a relevant output. Behind the system, we have used python programming language to perform the user's task. This proposed framework utilizes wolframalpha computational intelligence API as the source for integrating query processing in the system. This module has been utilized to retrieve information about famous celebrity, perform numerical estimations, knowing climate condition, depict every general entity and so forth.

This system can be treated as personal assistant that can be perform multiple tasks like searching queries on any search Engine, browsing YouTube videos, translating conversations in any language, Email exchange, playing music, jokes and getting a quick response for any query. The user statements/commands are analyzed with the help of NLP to give an optimal solution. At any point of time, the system is able to answer with an appropriate result.

5. CONCLUSIONS

The Intelligent Voice Assistant system introduced in this paper is a voice-controlled, multi-utilitarian system that can accept voice commands from a particular user and execute the desirable tasks. This framework utilizes wolframAlpha

calculation intelligence API that makes this system, ready to answer any computational question identified with Mathematics, Science and Technology, Society and Culture and so forth. This system uses speech recognition module named pyttsx3 (library in python) that makes this system compatible to use with or without internet access as well. The framework chips away at the normal lines with every intended functionalities that were initially brought up. Moreover, the framework likewise gives sufficient guarantee to the future as it is profoundly adaptable and new functionality can be incorporated with ease without disturbing the basic flow of existing constituents.

REFERENCES

- [1] Mittal, Y., Toshniwal, P., Sharma, S., Singhal, D., Gupta, R., & Mittal, V. K. (2015, December). A voice-controlled multi-funciton smart home automation system. In 2015 Annual IEEE India Conference (INDICON) (pp. 1-6). IEEE.
- [2] Deepak Shinde, Ria Umahia, Monika Raghort, Aishwarya Bhisikar, Anup Bhange. AI Based Voice Assistant Using Python.
- [3] Othman, E. S. (2017). Voice Controlled Personal Assistant Using Raspberry Pi. International Journal of Scientific & Engineering Research, 8(11), 1611-1615.
- [4] Kulhalli, K. V., Sirbi, K., & Patankar, M. A. J. (2017). Personal assistant with voice recognition intelligence. International Journal of Engineering Research and Technology (IJERT).
- [5] Kawamura, T., & Ohsuga, A. (2013). Flower Voice: Virtual Assistant for Open Data. International Journal of Web & Semantic Technology (IJWesT), 4(2)
- [6] Pandey, A., Vashist, V., Tiwari, P., Sikka, S., & Makkar, P. Smart Voice Based Virtual Personal Assistants with Artificial Intelligence.
- [7] Subhash, S., Srivatsa, P. N., Siddesh, S., Ullas, A., & Santhosh, B. (2020, July). Artificial Intelligence-based Voice Assistant. In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4) (pp. 593-596). IEEE.
- [8] <https://www.javatpoint.com/how-to-convert-text-to-speech-in-python>
- [9] <https://www.nature.com/articles/s41598-020-77258-w>
- [10] Yash Agrawal, Mr. Ranjeet Rai, Puneet Kumar Chaubey. BRAIN – THE A.I. (PERSONAL VOICE ASSISTANT).
- [11] Bose, P., Malphak, A., Bansal, U., & Harsola, A. (2017, April). Digital assistant for the blind. In 2017 2nd International Conference for Convergence in Technology (ICCT) (pp. 1250-1253). IEEE.
- [12] Sangpal, R., Gawand, T., Vaykar, S., & Madhavi, N. (2019, July). JARVIS: An interpretation of AIML with integration of gTTS and Python. In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) (Vol. 1, pp. 486-489). IEEE.