International Journal on Emerging Research Areas (IJERA) **ISSN:2230-9993**

DOI:

INTELLIGENT BUDDY

Kevin Roy
Dept. of Computer Science & Eng.
Amal Jyothi College of Engineering
Kottayam, India
kevinroy2023@cs.ajce.in

Lino Shaji
Dept. of Computer Science & Eng.
Amal Jyothi College of Engineering
Kottayam, India
linoshaji2023@cs.ajce.in

Riya G Johnson
Dept. of Computer Science & Eng.
Amal Jyothi College of Engineering
Kottayam, India
riyagjohnson2023@cs.ajce.in

Tince Tomy
Dept. of Computer Science & Eng.
Amal Jyothi College of Engineering
Kottayam, India
tincetomy2023@cs.ajce.in

Jane George
Dept. of Computer Science & Eng.
Amal Jyothi College of Engineering
Kottayam, India
janegeorge@amaljyothi.ac.in

***Abstract*—Software agents known as voice assistants are able to understand human speech and answer with synthetic voices. The most well-known voice assistants, which are built into smartphones or specific home speakers, are Apple's Siri, Amazon's Alexa, Microsoft's Cortana, and Google's Assistant. Users can use voice commands to manage other basic chores like email, to-do lists, and calendars, as well as ask inquiries of personal assistants, operate home automation devices, and playback of media. By holding and analysing information in the context of the user, this engages the capacity for social communication through natural language processing. This research will examine the fundamental operations and typical characteristics of voice assistants in use today. The currently used system operates online and is kept up by a third party. This program will safeguard personal information from others and use the local database, speech recognition and synthesizer.**

Keywords: Speech recognition system, Voice assistants, Machine Learning.

I. INTRODUCTION

A voice assistant is a type of digital assistant that listens for specific voice commands, language processing algorithms, and synthesis to find relevant information or carry out specific tasks as requested by the user. Supported commands, also referred to as "intents", are told by the user, and voice assistants find relevant information by hearing for specific keywords and removing background noise. While voice assistants may be entirely a software system on which all devices can be combined, other assistants are specifically designed for each individual device use, such as the Amazon Alexa. Voice assistants are now incorporated into a number of the gadgets we use on a daily basis, including smartphones, PCs, and highquality speakers. In this work, we hope to create a speech recognition-based assistant that can help users with everyday activities and introduce them to online resources for learning.

II. LITERATURE SURVEY

A Smart Personal Assistant by Shrooti Singh, is a digital Assistant for visually faulty human beings in an indoor environment. The voice assistant system is built with the PyCharm IDLE [1]. Paper focuses on a well-designed algorithm for Natural Language Understanding(NLU) interpreting deep learning by expanding the state networking of voice assistants to accommodate more information and ambiguous responses given by voice assistants such that assistants can easily understand and present the answers.

Desktop Voice Assistant for Visually Impaired [2] by Ankur Sindhu, uses google API to convert input speech into text. The speech is given as an input to google cloud for processing. As an output, the system then receives the resulting text. At backend the python gets the output from speech recognition and after that it identifies whether the command is a system command or a browser command. The output is send back to python backend to give desired output to user.

Voice Assistant using Artificial Intelligence [3] by A ullas, gathers 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. AI-based Voice assistants are the operating systems that can recognize human voice and respond via integrated voices.

SPYDER: Intelligent Voice Assistant by Piyush Gupta [4], uses 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. The proposed system's Graphical User Interface(GUI) 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. Life Assistant for Visually Impaired People using AI [5] by Bhanushali V, uses Natural Language Understanding(NLU) and Voice Recognition. Built-in falling detection algorithm based on gyroscope.

A vision and speech enabled, customizable, virtual assistant for smart environments [6] by Mario Grasso, graph shows expressions and is enabled with speech synthesis and recognition, face detection and face recognition for user identification.

Blind Assistance Device using AI [7] by Shreya Shetty, uses basic algorithm that will be implemented for working are Speech Recognition converts the users voice to text. OCR, Raspberry Pi 4, IPI Camera, Image classification, Text-to-Speech (Google TTS).

Construction of a voice driven life assistant system for visually impared people [8] Fang Zhao, is implemented using falling detection algorithm based on tri-axis accelerometer and object detection algorithm based on Mask R-CNN are used.

Speech based virtual assistant system for visually impared people [9] by P Sumalatha, uses speech based virtual assistant for helping and guiding the visually impaired people for their daily tasks when required. A modular solution is presented for improving the web based accessibilities for visually impaired persons. Uses text to speech interfaces and speech to text interfaces, then user communicates and customizes with the system. This system contains five modules namelyVoice based chatbot, face detection, object recognition, text recognition, and Website access that are currently implemented.

Spectral Voice Conversion For Text to Speech Synthesis [10] by A Kain, uses new spectral conversion algorithm using a locally linear transformation based on Gaussian mixture models whose parameters are trained by joint density estimation. Numerically, it was found to perform roughly equivalent to a previous GMM-based approach, but was more robust for small amounts of training data.

III. Methodology

The Speech Recognition library has many built-in features that will enable the assistant to understand the command given by the user and the response will be sent back to the user in voice, using Text to Speech functions, in the proposed concept effective way of implementing a Personal voice assistant. The algorithms used in the background will translate the user's spoken instruction into text when the assistant hears it. The following features are intended for the proposed system: 1) Using the microphone to capture speech patterns. 2) conversion from audio data recognition. 3) Assessing the input against pre-established commands. 4) Producing the required results. The overall architecture of the system is illustrated in Figure 3.1.

1) Taking the input as speech patterns through microphone: While continuing to listen for commands, the assistant solicits input from the user. The amount of time for listening can be adjusted based on the needs of the user. If the assistant doesn't understand the command correctly, it will keep asking the user to repeat it. Depending on the user's preferences, the voice of this assistant can be changed to be either male or female. To enable the use of speech synthesis and recognition within Windows applications, Microsoft created the Speech Application Programming Interface, or SAPI.

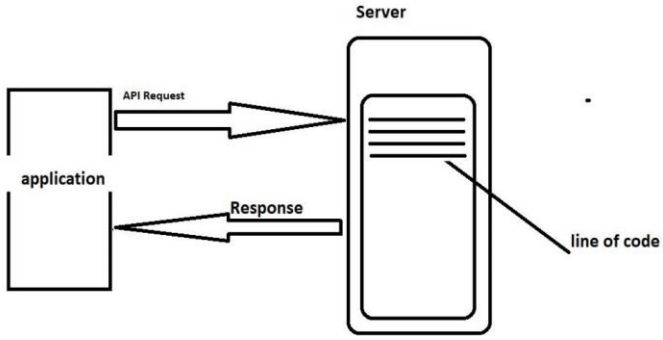

Figure 3.1 Architecture of the system

2) Audio data recognition and conversion into text: Speech recognition, also known as automatic speech recognition (ASR), computer voice recognition, or speech-to-text, is the ability of a program to translate spoken words into written language. Despite the fact that the two terms are occasionally used interchangeably, voice recognition only seeks to identify a specific person's voice, whereas speech recognition focuses on transforming speech from a verbal to a text format. In our project, we'll be using AssemblyAI to recognize speech. The quality and accessibility of the lauded AssemblyAI voice recognition API are unmatched. The AssemblyAI API can be used to understand and transcribe

DOI:

audio and video files using AI models. Since it is free and easy to use, this API was chosen.

3) Comparing the input with predefined commands: Matching spoken words or phrases to a list of predetermined commands is how input is compared to established commands. Natural methods are frequently used for this approaches for natural language processing (NLP), including speech recognition and intent categorization. The system analyzes user input using algorithms, finds the appropriate command, and executes that command. This method enables the voice assistant to comprehend the user's intent and react suitably, giving it a practical and user-friendly tool for carrying out a variety of tasks.

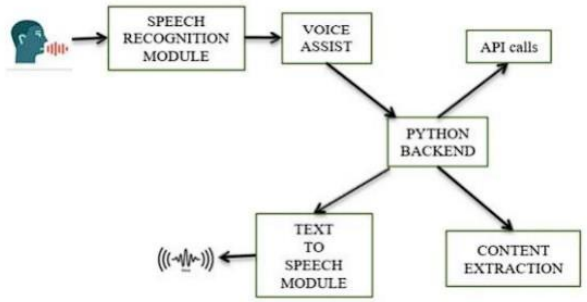

Figure 3.2: Detailed Workflow of the voice assistant

4) Giving the desired output: The voice recognition module's output is obtained by the Python backend, which then determines if the command or the speech output is an API call and context extraction. The output is then returned to the Python backend so that the user can receive the desired output. The desired output of a voice assistant is produced utilizing Text-to-Speech (TTS) technology. A language model like myself texts an answer, which the voice assistant then uses TTS software to turn into speech. Based on the supplied text, the TTS software creates human-like speech that is then output as audio through the device's speaker. Instead of needing to read the voice assistant's comments and information, users may converse with it and hear what it has to say.

IV. Architecture

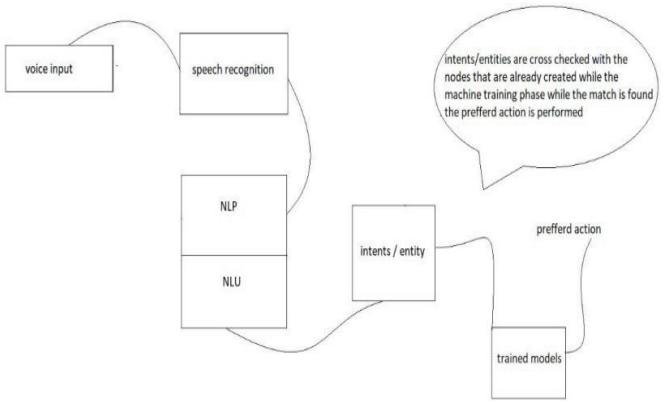

Figure 3.3 Detailed Architecture of the system

After identifying the intents and entity the next step is to perform the intended activity. For performing the intended activity the identified intents/entities are compared with the predefined or the user defined functions. For example, a user defined function "create Todo()", inside this function there will be the program for performing the action such that there will be so many user defined nodes packed up with the functions for perf.orming the activities. The functions are added while the machine training phase encounters with the commands that are not defined machine will be able to notify the admin and the admin will be able to update the database with the required information. The "front end" of a Virtual Personal Assistants(VPA) that interacts with users must be able to comprehend human language. Converting the given natural language input into relevant representations.

Examining the language's various facets. It enables humancomputer connection, is able to comprehend input in the form of sentences delivered via text or speech. Translates human speech into a precise data model of semantic and pragmatic definitions using algorithms.

V. Conclusion

This project provides a thorough explanation of the planning and creation of a Python-based personal assistant with support for static Voice. In comparison to earlier times, the voiceactivated personal assistant of today will be more efficient at saving time and being supportive of people with disabilities. This assistant effectively completes some tasks that the user has assigned. Additionally, this assistant can perform a lot of things with just one voice command, including sending messages to the user's mobile device, automating YouTube, and obtaining data from Google and Wikipedia. With the help of this voice assistant, we were able

DOI:

to automate several services with just one line of code. Most of the user's tasks, such web searching and other things, are made easier by it. It eases the work of the user like searching the web,retrieving the weather forecast details. The aim of the project is to improve the performance of these bots by adding more functions and methods, making it more efficient, reliable, and able to deal with different situations.

REFERENCES

[1] Saunshimath, Nirmala Thakur, Sakshi Singh, Shrooti Singh, Ranavijay .K, Kumuda. (2022). SAHARA -A Smart Personal Assistant. 10.13140/RG.2.2.33808.48642.

[2] Yadav, A., Singh, A., Sharma, A., Sindhu, A., Rastogi, U. (2020). Desktop voice assistant for visually impaired. Int J Recent Technol Eng (IJRTE), 9(2)..

[3] G. Preethi , Abishek. K , Thiruppugal S , Vishwaa D A, 2022, Voice Assistant using Artificial Intelligence, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) Volume 11, Issue 05 (May 2022)

[4] Piyush Gupta, Sakshi Yadav "SPYDER: Intelligent Voice Assistant," International Research Journal of Engineering and Technology (IRJET) Volume 08 Issue 04 (April 2021)

[5] Bhanushali Nayan, V., Kapile Namrata, R., Pawar Shubham, G., Pawase Kalpesh, D., Tambe Prachi, S. LIFE ASSISTANT FOR VISUALLY IMPAIRED PEOPLE USING AI.

[6] Iannizzotto, G., Bello, L. L., Nucita, A., Grasso, G. M. (2018, July). A vision and speech enabled, customizable, virtual assistant for smart environments. In 2018 11th International Conference on Human System Interaction (HSI) (pp. 50-56). IEEE.

[7] Safnaz, Nayana Manju Jogi, Shreya Shetty, Prasad Rathode, Deeksha K R, 2021, Blind Assistance Device using AI, INTERNATIONAL JOUR-NAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) NCCDS – 2021 (Volume 09 – Issue 12),

[8] Chen, R., Tian, Z., Liu, H., Zhao, F., Zhang, S., Liu, H. (2018, May). Construction of a voice driven life assistant system for visually impaired people. In 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 87-92). IEEE.

[9] S. Subhash, P. N. Srivatsa, S. Siddesh, A. Ullas and B. Santhosh, "Artificial Intelligence-based Voice Assistant," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020, pp. 593-596, doi: 10.1109/WorldS450073.2020.9210344.

[10] M. R. Sultan and M. M. Hoque, "ABYS(Always By Your Side): A Virtual Assistant for Visually Impaired Persons," 2019 22nd International Conference on Computer and Information Technology (ICCIT), 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038603.