iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60456

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

AURA-The Privacy Preserving Multifunctional Desktop Voice Assistant

Aditi Dhumal[1], Vishal Yadav[2], Aniket Dhere[3], Navnath Bagal[4]
Computer Department, Savitribai Phule Pune University

Abstract: ***The "AURA" is the project which proposes the development of the versatile desktop voice assistant which mainly focuses on the privacy of the user.The convenience of voice interaction has surged in popularity, but concerns about user privacy remain. Conventional cloud-based voice assistants raise anxieties regarding data collection, storage, and potential misuse. This survey paper explores the burgeoning field of privacy-preserving desktop voice assistants powered by Artificial Intelligence (AI). We delve into the core challenges of on-device speech recognition, natural language understanding, and response generation while maintaining user privacy. We conclude by outlining open research questions and future directions for this evolving field, paving the way for the development of secure and user-centric voice interaction experiences on desktops.***
Keywords: Voice assistant , Natural language processing ,Text to speech , Tokenization , TF/IDF

I. INTRODUCTION

The primary objective of the project is to develop the desktop voice assistant which can perform multiple tasks based on the commands given by the user while concerning the privacy of the user.In today's digital world, voice assistants have become increasingly popular, offering a convenient way to interact with technology through spoken commands. However, these assistants often raise privacy concerns as they require access to user audio data for speech recognition and potentially other functionalities. This is particularly true for desktop environments where users may handle sensitive information.

This survey paper explores the concept of privacy-preserving multifunctional desktop voice assistants. We delve into the current landscape of voice assistants, highlighting their functionalities and the privacy risks associated with their data collection practices. We then focus on the emerging field of privacy-preserving techniques that can be employed to develop secure voice assistants for desktop environments.The survey paper focuses on the desktop voice assistant which can perform multiple tasks and provides a comprehensive overview of existing desktop voice assistants and their functionalities. This survey paper mainly focus on solving all the privacy related issues in the desktop voice assistant.

This survey paper hunt through the flourishing field of privacy- preserving desktop voice assistants powered by AI.By doing a thorough exploration, we examine the rudimentary challenges and the developments in on-device speech recognition, natural language understanding and the response generation in form of speech output. Our goal is to offer valuable insights into the methods used to safeguard user privacy during voice interactions by examining current research and cutting edge strategies.

Furthermore, this article examines and explore3s unresolved research inquiries and forthcoming paths for the advancement of desktop voice assistants that prioritize privacy. By tackling these obstacles and utilizing emerging technologies like federated learning and differential privacy, we envision a future where users can engage in secure and personalized voice interactions on their desktops without jeopardizing their privacy. Ultimately, the objective of this comprehensive study is to contribute to the ongoing conversation regarding privacy-enhancing technologies and promote the creation of inventive solutions that empower users while protecting their confidential data.

II. LITERATURE SURVEY

[1] This paper deals with the Speech Recognition Intelligence System for Desktop Voice Assistant by using AI &IoT, with statistical testing of hypothesis.In the modern era of reckless technology, we are able to carry out tasks that we never could have imagined testing we would be able to prepare for. However, in order to carry out these daydreams, we need a method that makes it simple for us to automate the things we do every day. As a result, we created applications like Voice Assistant that can communicate with us solely through human interaction . A voice assistant can be used by a number of applications, including AI and IoT. It has the ability to alter how users and machines communicate. By using voice commands, the user can access all of the features of this application, which has been designed machines to work with mobile phones.

©IJRASET: All Rights are Reserved

The primary difficulties and drawbacks of various voice assistants will be discussed in this paper. In this paper, we talk about how to make a voice based assistant that doesn't need cloud services, which would help these devices grow in the future.

[2] This paper addresses the primary goal of the trending technology AI is to realize natural human machine dialogue. Various IT-based companies also utilized dialogue networks technology to create various types of Virtual Personal Assistants focused on their products and areas for expanding human-machine contact, such as Alexa, Cortana, Google's Assistant, Siri and so more. Just like the Microsoft voice assistant named 'Cortana', we designed our virtual assistant which performs basic tasks based on the instruction provided to it on the Windows platform using Python. Here, Python is used as a scripting language as it has a large library that is used to perform instructions. Using Python packages, a personalized virtual assistant recognizes and processes the user's voice.

[3] This paper deals with a personal virtual assistant that allows a user to command or ask questions in the same manner that they would do with another human and are even capable of doing some basic tasks like opening apps, doing Wikipedia searches without opening a browser, playing music etc, with just a voice command.

[4] This paper addresses that the Personal Assistants, or conversational interfaces, or chat bots reinvent a new way for individuals to interact with computes. A Personal Virtual Assistant allows a user to simply ask questions in the same manner that they would address a human, and are even capable of doing some basic tasks like opening apps, reading out news, taking notes etc., with just a voice command. Personal Assistants like Google Assistant, Alexa, Siri works by Speech Recognition (Speech-to-text) and Text-to Speech. Keywords: Personal Assistants; chat bots; conversational interfaces; Speech Recognition; Text-to Speech.

[5] 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 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.. Multi-functional aspects like voice commands, sending emails, reading PDF, sending text on Whats-App, 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 utilized 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.

[6] The topic of this essay is the technologies are evolving rapidly day by day with the development of these technologies humans can perform their daily tasks easily. one of the most popular field of technology is artificial intelligence which is growing nowadays. Using artificial intelligence machines can recognize human behavior and on the basis of that it can respond to humans. the desktop voice assistant is the best example of developing artificial intelligence technology. These desktop voice assistants help humans to perform various tasks on the desktop easily just by using their voice. The voice assistant uses speech recognition modules which is useful for recognizing and understanding human input voice and on the basis of user input command it gives the required input queries or performs the given task like opening and closing different applications, can search and send messages on whatsapp without using keyboard or mouse. In this paper artificial intelligence technology is used to create a desktop voice assistant which will be helpful for the visually impaired and the people with disabilities. This desktop voice assistant can also be useful for normal people as it saves time and provides efficiency in doing our day to day tasks.

.

III. PROPOSED SYSTEM

Voice assistant is one of the biggest problem solvers for the users of the desktop .Solution of every problem which can be faced by the user can be found in the fraction of the seconds and only with the voice commands. The voice assistant AURA performs multiple tasks on the voice command given by the user while concerning the privacy of the user.

It is useful for the user to perform multiple tasks only just with the voice command and it is also useful for the impaired persons who are unable to perform the tasks visually so the AURA can help them to perform the tasks just by their voice commands.

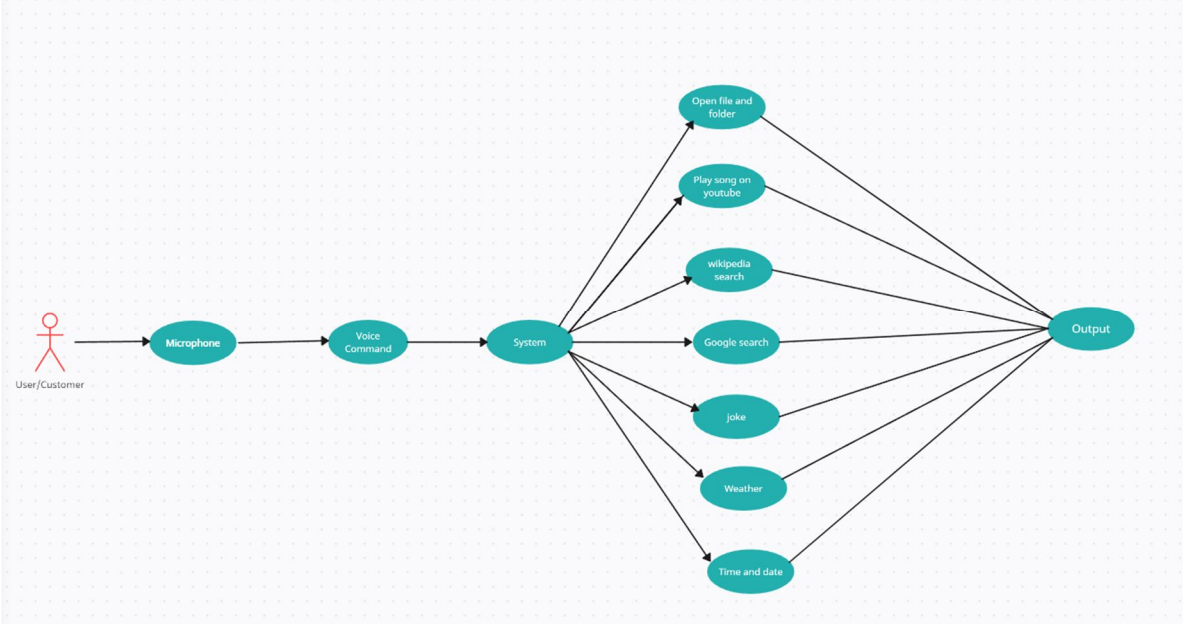


Fig 1: Use Case Diagram

The user can perform multiple tasks using the AURA and some of the tasks are listed as below :

1) *Opening the file or folder:* Every time while opening the file or folder in our system we need to go to that particular location and then open that file or folder but by using the AURA the user can open any file or folder just by giving the name of that file or folder in the voice command.
2) *Playing the video on YouTube:* While users have to play any YouTube video it is mandatory to go to YouTube search for a particular video and then play that video but using AURA the user can play the video by command “play” along with the name of the video . This makes daily life for the user very easy.
3) *Search on Wikipedia:* Wikipedia is the widely used website by many of the user worldwide. It can be used by students , corporate workers , and many more . The AURA helps all the user to search on the Wikipedia only by giving the voice command without opening the Wikipedia website.
4) *Telling some joke:* Now let's be honest, everyone would have had at least one moment in their life where they were so tensed up or had an argument with their close people. So, these moments can be chilled up at least ten percentage with some random joke and the user can get a moment of joy.
5) *Telling the temperature/weather of any location:* Let's start this with a question, why is it important for us to know the weather of the day? or why is it important for us to monitor the weather every day? The answer is pretty simple: it forewarns the users asking about the weather, telling them that "it might rain today so carry an umbrella if you want or weather it is a sunny day.
6) *Searching for what the user asks:* Today in the 20th century, we people often get doubts and we need to clear that doubt as soon as possible else that one doubt will be multiplied and at the end, we'd have n doubts and to clear the doubts searching the question in the internet will give us an answer and clear our doubts and asking that to the assistant will save a lot of time. Other than clearing the doubts, we need to search a lot of questions or topics in the internet to keep up with the trend and we can do this searching just by giving command to our assistant, asking it to search a specific topic/question.
7) *Telling the time and date:* The user can easily see today's date and the current time only by giving the voice command.
8) *Play songs from music folder:* AURA can play the songs which are stored on the folder in your PC just by giving the command and the songs will automatically get played.And the user can get the help with opening the folder without going to that particular folder location.
9) *Technologies used:*

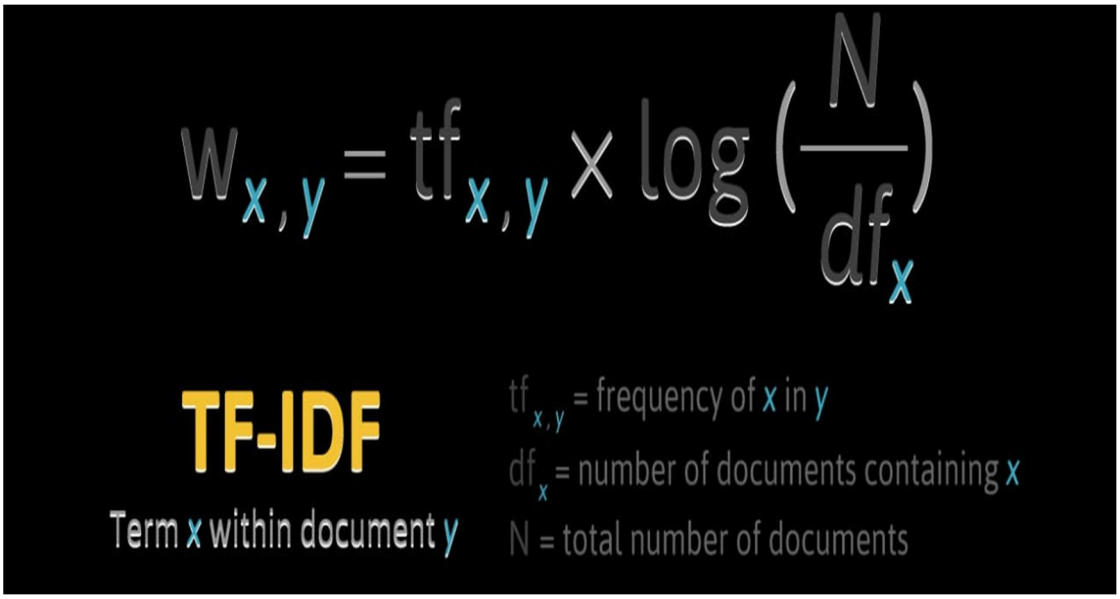
a) *Python :* We have used Python to build the assistant as it supports Object Oriented Programming through which a lot of built-in functions is made keeping it less complicated to build the assistant.

The assistant's query can be modified to suit the user's needs. Speech Recognition is a process of converting the audio into text which can further used by the assistant to find what the user is requesting the assistant to do. The usage of Python is such that it cannot be limited to only one activity. Its growing popularity has allowed it to enter into some of the most popular and complex processes like Artificial Intelligence, Machine Learning (ML), natural language processing, data science etc. Python has a lot of libraries for every need of this project.

b) *Pyttsx3:* Pyttsx3 stands for Python Text to Speech. It is a cross-platform Python wrapper for text-to-speech synthesis. It is a Python package supporting common text-to-speech engines on Mac OS X, Windows, and Linux. It works for both Python2.x and 3. versions. Its main advantage is that it works offline.

c) *NLP and Voice Recognition:* Natural language processing (NLP) techniques are used to process and understand the voice commands the desktop voice assistant receives. This may involve tasks such as speech recognition, language understanding, intent recognition, and context extraction, to accurately interpret the user's commands.

d) *TF/IDF:* Term Frequency - Inverse Document Frequency (TF-IDF) is a widely used statistical method in natural language processing and information retrieval. It measures how important a term is within a document relative to a collection of documents.



10) System Architecture:

The overall system design consists of following phases:

a) Data collection in the form of speech.

b) Voice analysis and conversion to text.

c) Execute Python script.

d) Generating speech from the processed text output.

In the first phase, the data is collected in the form of speech and stored as an input for the next phase for processing. In the second phase, the input voice is continuously processed and converted to text using STT. In the next phase the converted text is analyzed and processed using Python Script to identify the response to be taken against the command. Finally once the response is identified, output is generated from simple text to speech conversion using TTS.

11) Data Flow Sequence:

a) Initialize device: Initialize the device by calling its name.

b) Task Manager: Conversion of Speech-to-Text and Text-to Speech is performed by the task manager.

c) Service Manager: Analysis of commands and matching them with web service and applications.

d) Execute Command: After finding the match for the given command, run the respective python script and give the output.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

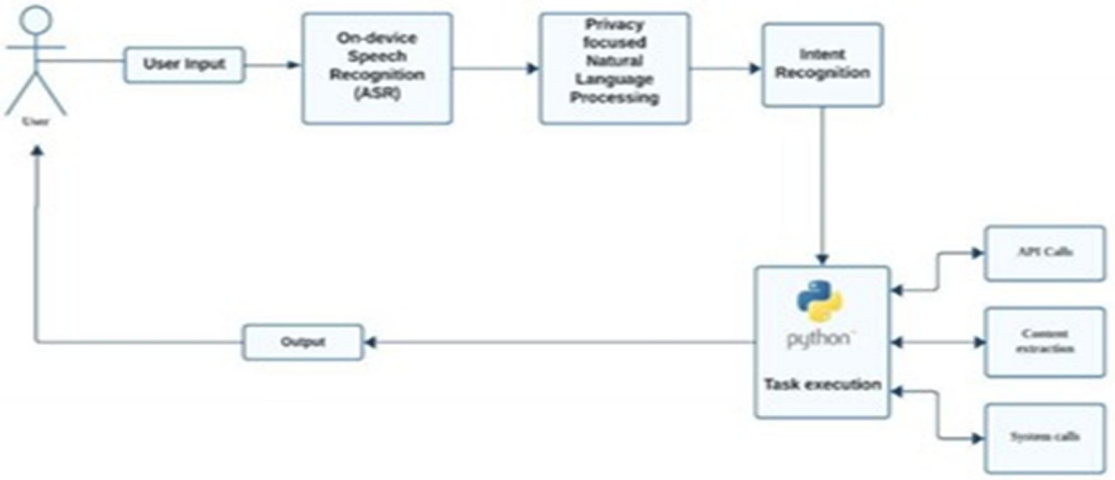


Fig 1: System Architecture

IV. CONCLUSIONS

Through this AURA voice assistant, we have automated various services using a single line command. It eases most of the tasks of the user like searching the web,retrieving weather forecast details, vocabulary help and medical related queries.We aim to make this project a complete server assistant and make it smart enough to act as a replacement for a general server administration. The AURA mainly focus on providing the services while concerning th users privacy .The future plans include integrating Jarvis with mobile using React Native to provide a synchronized experience between the two connected devices.The privacy-preserving multifunctional desktop voice assistant is a significant step forward in the development of voice interaction technologies that prioritize user privacy and security.

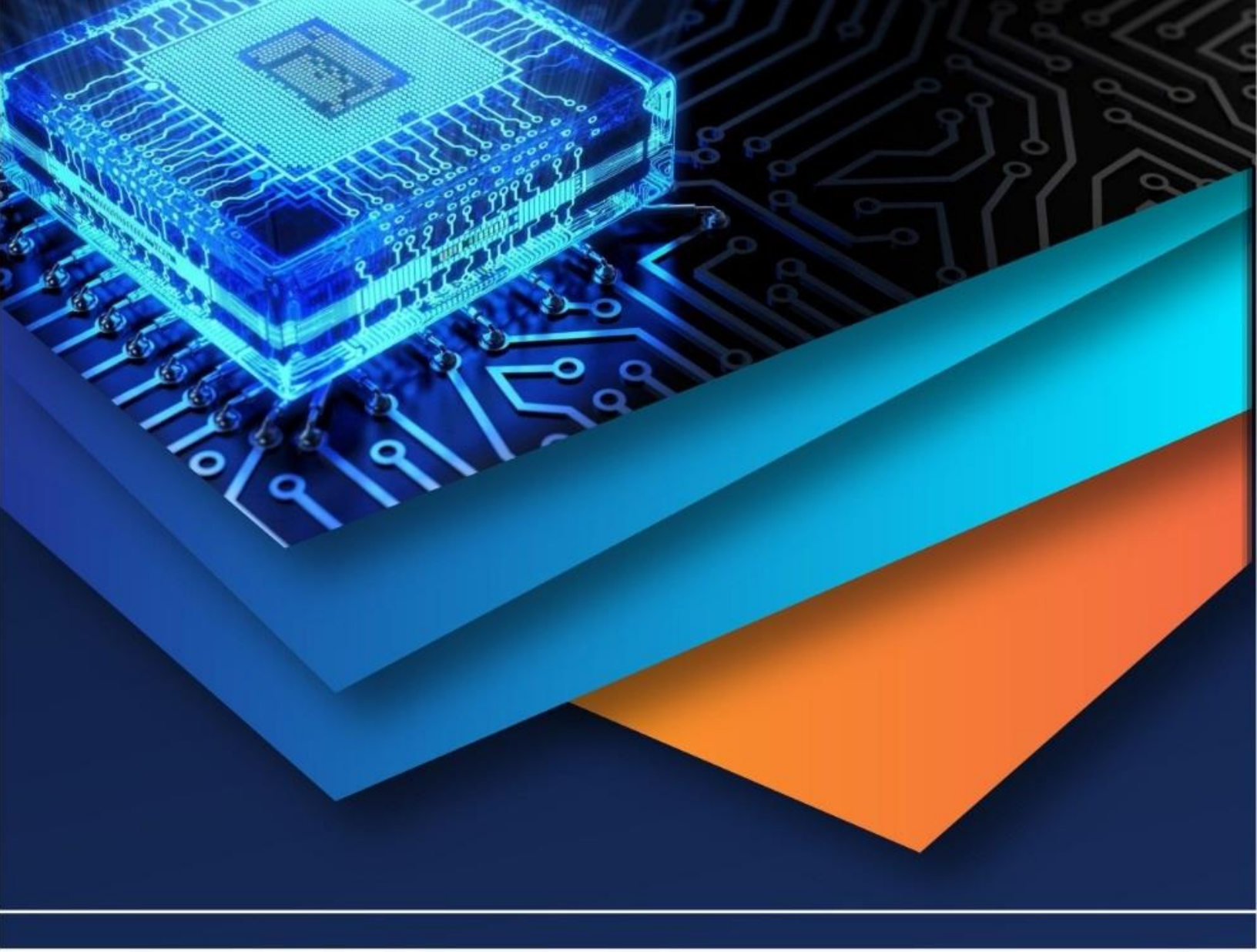
V. FUTURE SCOPE

1) Integration with Smart Devices: extending compatibility to interact with a broader range of smart devices,creating a more connected and streamlined user experience.
2) Enhanced Privacy Features: continuous development and improvement of privacy mechanisms to stay ahead of emerging threats.
3) Advanced AI Capabilities: incorporating cutting-edge AI technologies to improve natural language understanding, context awareness, and personalized user interactions.
4) Ecosystem Expansion: integrating with a wider range of applications and services, making voice assistants an integral part of users' daily tasks and activities.
5) Research and Innovation: staying at forefront of AI research to integrate emerging technologies and maintain a competitive edge in providing state-of-art voice assistant.

REFERENCES

[1] Deller John R., Jr., Hansen John J.L., Proakis John G. ,Discrete-Time Processing of Speech Signals, IEEE Press, ISBN 0-7803-5386-2.
[2] Hayes H. Monson,Statistical Digital Signal Processing and Modeling, John Wiley & Sons Inc. , Toronto, 1996, ISBN 0-471-59431-8.
[3] Proakis John G., Manolakis Dimitris G.,Digital Signal Processing, principles, algorithms, and applications, Third Edition, Prentice Hall , New Jersey, 1996, ISBN 0-13- 394338-9.
[4] Ashish Jain,Hohn Harris,Speaker identification using MFCC and HMM based techniques,university Of Florida,April 25,2004.
[5] Rabiner Lawrence, Juang Bing-Hwang. Fundamentals of Speech Recognition Prentice Hall , New Jersey, 1993, ISBN 0-13-015157-2.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)