# VOICE BASED VIRTUAL ASSISTANT

*A Project Report Submitted By*

# Sriyansh Shivam – 2114110063

# Anupam Sharma – 2114110041

# Raj Harsh Vardhan – 2114110052

**Anurag Singh - 2114110045**

*In partial fulfilment of the requirementFor Project Based Learning of*

# BACHELOR OF TECHNOLOGY

*In*

# COMPUTER SCIENCE AND ENGINEERING *For*

# *Artificial Intelligence*

*Under the guidance of*

***Prof. Priyanka Gurav***



# DEPARTMENT OF COMPUTER SCIENCE AND BUSINESSSYSTEMS

**BHARATI VIDYAPEETH (DEEMED TO BE UNIVERSITY)COLLEGE OF ENGINEERING, PUNE- 43**

2023-24

# BHARATI VIDYAPEETH (DEEMED TO BE UNIVERSITY)COLLEGE OF ENGINEERING, PUNE- 43



**CERTIFICATE**

This is to certify that the Project Based Learning report titled **VOICE BASED VIRTUAL ASSISTANT** submitted by **Raj Harsh Vardhan (2114110052),Anupam Sharma (2114110041), Sriyansh Shivam (2114110063),Anurag Singh(2114110045)** , to the Bharati Vidyapeeth (Deemed to be University), College of Engineering, Pune - 43 for the award of the degree of **BACHELOR OF TECHNOLOGY** in Computer Science and Engineering is a bonafide record of the PBL work done by him/them under my supervision.

Place: Pune Prof. Priyanka Gurav

Date:

# Index

|  |  |  |
| --- | --- | --- |
| **Sr.No.** | **Title** | **Page No.** |
| 1 | Abstract | 4 |
| 2 | Introduction of Topic | 5 |
| 3 | Methodology Used in the Project | 6 |
| 4 | Snapshots of Implementation | 10 |
| 5 | Limitations/  Advantages | 12 |
| 6 | Conclusion | 14 |
| 7 | References | 14 |

**ABSTRACT:**

The project on a voice-based virtual assistant explores the development and implementation of an intelligent digital companion that leverages natural language processing and voice recognition technology to facilitate seamless human-computer interaction. This virtual assistant represents a significant advancement in user-centric AI applications, offering a hands-free and intuitive solution to a wide range of everyday tasks.

The project underscores the potential of voice-based virtual assistants to revolutionize human-computer interaction, catering to a diverse range of users. It is poised to benefit not only individuals seeking a hands-free and efficient way to manage their daily tasks but also those with accessibility needs. Continuous feedback and user engagement play a vital role in shaping the future of this virtual assistant, as it strives to become an indispensable tool in our daily lives.

This project showcases the commitment of our team to innovation and user satisfaction, setting the stage for ongoing developments and the creation of a more intelligent and responsive virtual assistant.

**Introduction**

In an era of rapid technological advancements and the relentless pursuit of simplifying our daily lives, voice-based virtual assistants have emerged as transformative innovations that bridge the gap between humans and machines. These digital companions, powered by artificial intelligence and natural language processing, have revolutionized the way we interact with technology and access information. The concept of voice-based virtual assistants is at the forefront of the digital age, redefining convenience, accessibility, and efficiency in both our personal and professional spheres.

A voice-based virtual assistant, often referred to as a voice assistant, is a software application that enables users to interact with their devices, such as smartphones, smart speakers, or computers, through spoken language. These assistants can understand and respond to voice commands, making it possible to perform a wide range of tasks without the need for manual input. This technology has become an integral part of our daily routines, offering services that range from setting reminders and answering questions to controlling smart home devices and even providing entertainment.

At the core of this innovation lies Natural Language Processing (NLP), a field of artificial intelligence that equips these virtual assistants with the ability to comprehend and interpret human language. Through continuous learning and improvement, voice-based virtual assistants have become increasingly adept at understanding context, nuances, and a variety of accents, making them more accessible and user-friendly.

The prevalence and adoption of voice-based virtual assistants have grown exponentially, with major tech giants offering their own versions, such as Amazon's Alexa, Apple's Siri, Google Assistant, and Microsoft's Cortana, to name a few. These digital companions have not only simplified tasks but have also paved the way for innovative applications, especially in healthcare, automotive technology, and accessibility for individuals with disabilities.

In this digital age, where convenience and efficiency are paramount, the concept of voice-based virtual assistants has transcended being merely a technological novelty. It has become an integral part of our interconnected world, enhancing our daily lives and expanding the possibilities of human-computer interaction. This exploration of voice-based virtual assistants will delve into their evolution, capabilities, challenges, and the transformative impact they have on the way we live, work, and communicate in today's increasingly tech-driven society.

As we journey deeper into the realm of voice-based virtual assistants, we will uncover the complexities and potentials of this technology, shedding light on how it has not only simplified tasks but also altered our very relationship with machines, and consequently, the world around us.

**Methodology:**

Creating a voice-based virtual assistant, like "Jarvis," involves a complex and multi-faceted development process. Here's a general methodology that can be used to create such a virtual assistant:

**Conceptualization and Planning:**

Define the purpose and scope of your virtual assistant (e.g., personal assistant, home automation control, information retrieval, etc.).

Identify the target audience and their needs.

Set specific goals and functionalities the assistant should have.

Plan for any unique features or personalization options.

**Choose Technology Stack:**

Select the programming languages and tools that will be used in development. Common choices include Python, JavaScript, and libraries like TensorFlow, PyTorch, and OpenAI for machine learning and NLP.

**Speech Recognition:**

Implement a speech recognition system to convert spoken language into text. Popular APIs and libraries like Google Speech-to-Text or Mozilla's DeepSpeech can be utilized.

Natural Language Processing (NLP):

Integrate NLP models and algorithms to analyze and understand the text input from the speech recognition system.

Use libraries like spaCy, NLTK, or Hugging Face's Transformers for NLP tasks.

**Voice Synthesis:**

Develop or integrate a text-to-speech (TTS) system to convert the assistant's responses back into human-like speech.

Consider services like Amazon Polly or Google Text-to-Speech.

**Functional Modules:**

Create distinct modules for each functionality (e.g., reminders, weather updates, knowledge retrieval).

Implement algorithms and logic for each module to execute tasks accurately.

**User Interface (UI) Development:**

Design a user-friendly interface for interaction. This may include mobile apps, web applications, or integration with popular smart speakers.

Voice Command Handling:

Develop a voice command handling system that processes user requests, converts them into actionable tasks, and triggers the appropriate module.

**User Authentication and Data Handling:**

Implement user authentication for personalization and data storage.

Ensure secure data handling, including user preferences and previous interactions.

**Machine Learning and Training:**

Train the NLP models with relevant datasets and continually update them to improve the assistant's understanding and responses.

Fine-tune voice recognition to recognize a variety of accents and speaking styles.

**Testing and Debugging:**

Rigorous testing is crucial to ensure the assistant performs accurately.

Test different use cases, edge cases, and corner cases to enhance reliability.

**Privacy and Security:**

Implement robust security measures to protect user data, including voice recordings and personal information.

Adhere to privacy regulations and obtain user consent for data usage.

**User Feedback and Continuous Improvement:**

Encourage user feedback and use it to make necessary improvements and updates to the virtual assistant.

Regularly release updates to enhance the user experience and add new features.

Documentation:

Create user guides and documentation to help users understand how to interact with the virtual assistant effectively.

**Deployment:**

Deploy the virtual assistant on servers or cloud platforms, ensuring scalability and uptime.

**Monitoring and Maintenance:**

Continuously monitor the assistant's performance, troubleshoot issues, and provide regular maintenance and updates.

Creating a voice-based virtual assistant like "Jarvis" is a dynamic process that requires a dedicated development team and ongoing commitment to improvement. The methodology provided serves as a general framework, and the specific technologies and tools used can vary based on project requirements and available resources.

**Implementation**For creating the voice based virtual assistant “**JARVIS”** we have used **PYTHON** programming language.

The assistant can perform tasks like playing songs on YouTube, telling the current time and date, responding to questions about its name and well-being, and providing information about a person from Wikipedia. Let's go through the code line by line:

**import speech recognition as aa**: This line imports the speech\_recognition library and aliases it as aa. This library is used for speech recognition.

**import pyttsx3:** This line imports the pyttsx3 library, which is used for text-to-speech conversion.

**import pywhatkit**: This line imports the pywhatkit library, which allows you to perform various tasks with YouTube, including playing songs.

**import datetime**: This line imports the datetime module for working with date and time.

**import wikipedia**: This line imports the wikipedia library, which can be used to search and retrieve information from Wikipedia.

**listener = aa.Recognizer():** This line initializes a Recognizer object from the speech\_recognition library. This object will be used to capture and recognize speech.

**machine = pyttsx3.init():** This line initializes a text-to-speech engine using pyttsx3. This engine is used to convert text into speech.

**def talk(text):** This function takes a text input as an argument and uses the text-to-speech engine to speak the provided text.

**def input\_instruction():** This function is defined to capture voice input. It uses a with statement to create a context where the microphone is opened, and then it captures and recognizes the spoken instruction using the listener object. The recognized text is converted to lowercase and any occurrence of "jarvis" is removed. The recognized text is then returned.

**def play\_jarvis():** This function is the main logic for the Jarvis assistant. It starts by calling input\_instruction() to capture and recognize the user's voice input.

The code then checks the recognized instruction to see if it matches specific keywords like "play," "time," "date," "how are you," "what is your name," and "who is."

**Code:**

import speech\_recognition as aa

import pyttsx3

import pywhatkit

import datetime

import wikipedia

listener = aa.Recognizer()

machine = pyttsx3.init()

def talk(text):

machine.say(text)

machine.runAndWait()

def input\_instruction():

global instruction

try:

with aa.Microphone() as origin:

print('Listening...')

speech = listener.listen(origin)

instruction = listener.recognize\_google(speech)

instruction = instruction.lower()

if "jarvis" in instruction:

instruction = instruction.replace('jarvis',' ')

print(instruction)

except:

pass

return instruction

def play\_jarvis():

instruction = input\_instruction()

print(instruction)

if "play" in instruction:

song = instruction.replace('play', '')

talk("Playing " + song)

pywhatkit.playonyt(song)

elif 'time' in instruction:

current\_time = datetime.datetime.now().strftime('%I:%M %p')

talk('Current time is ' + current\_time)

elif 'date' in instruction:

current\_date = datetime.datetime.now().strftime('%d/%m/%Y')

talk("Today's date is " + current\_date)

elif 'how are you' in instruction:

talk('I am fine, how about you?')

elif 'what is your name' in instruction:

talk('I am Jarvis, what can I do for you?')

elif 'who is' in instruction:

person = instruction.replace('who is', '').strip()

info = wikipedia.summary(person, 1)

print(info)

talk(info)

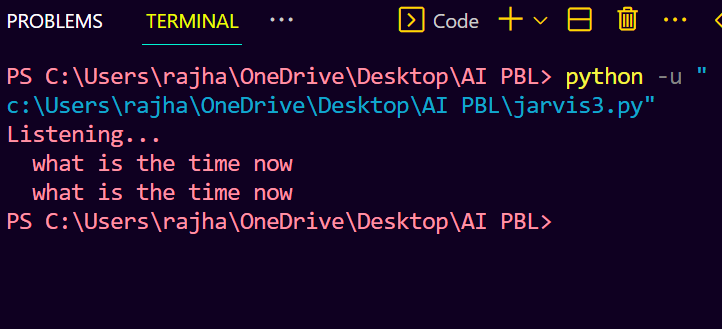
else:

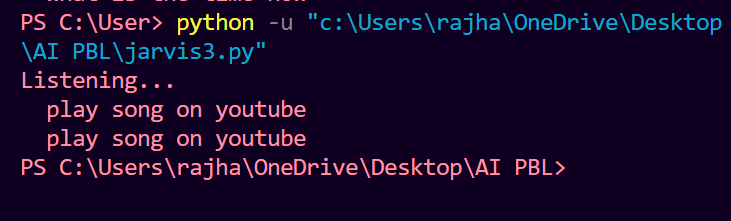
talk('Please repeat')

# while True:

play\_jarvis()

**SNAPSHOTS:**

****

****

**After giving the following command to play song on youtube**

**A screenshot of a video

Description automatically generated**

**LIMITATIONS:**

Voice-based virtual assistants, like the one created in the provided code, have several limitations. Here are some of the common limitations associated with such projects:

Limited Vocabulary and Context Understanding: Voice assistants often struggle to understand and respond to complex or context-dependent queries. They typically work well for predefined commands but may struggle with open-ended questions or unfamiliar words.

Accuracy and Speech Recognition Challenges: Speech recognition technology is not perfect and can misinterpret words, especially in noisy environments or with accents and dialects. This can lead to incorrect responses or misunderstandings.

Limited Language Support: Voice assistants are typically designed for specific languages, and their accuracy can drop significantly when used in languages or dialects they were not trained for.

Privacy and Security Concerns: Voice assistants raise privacy concerns as they continuously listen to and record conversations. Unauthorized access to voice data can lead to security breaches. It's essential to handle voice data with care.

Internet Dependency: Many voice assistants, including the one in the provided code, rely on internet connectivity to function correctly. This means they may not work in areas with poor or no internet access.

Response Time: Some voice assistants may have a noticeable delay between the user's query and the assistant's response. This delay can be frustrating for users who expect immediate responses.

Limited Task Capabilities: Most voice assistants are designed for specific tasks and may not perform well outside their predefined functionalities. They might not handle complex tasks or offer integration with various services.

Lack of Emotional Understanding: Voice assistants lack the ability to understand or respond to the user's emotions. They provide scripted responses and cannot empathize with users.

Lack of Common Sense Reasoning: Voice assistants often lack common-sense reasoning and may provide nonsensical or incorrect answers to certain questions.

Accessibility Challenges: People with speech impairments or disabilities may face difficulties using voice assistants. These technologies are not always designed with inclusivity in mind.

Maintenance and Updates: To keep a voice assistant relevant and functional, it requires constant maintenance, updates, and training. This can be a resource-intensive process.

Inability to Handle Multitasking: Most voice assistants can only handle one request at a time. They may not understand or respond to multiple commands given in a single sentence or handle complex, multi-step instructions.

Misinterpretation of Homophones: Voice recognition systems may have trouble distinguishing between homophones (words that sound the same but have different meanings) in some contexts.

**ADVANTAGES:**

Voice-based virtual assistants offer numerous advantages, making them a popular choice in various applications. Here are some of the key advantages of voice-based virtual assistants:

Hands-Free Interaction: Users can interact with voice assistants without the need to use their hands. This is particularly valuable when driving, cooking, or performing tasks where manual input is inconvenient or unsafe.

Accessibility: Voice assistants can greatly assist people with disabilities, including those with mobility impairments or visual impairments, by providing a convenient and natural way to interact with technology.

Natural Language Interaction: Voice assistants allow users to interact using natural language, making them more intuitive and user-friendly, especially for individuals who may not be tech-savvy.

Convenience and Efficiency: Voice commands are often faster and more efficient than typing on a keyboard or touchscreen. Users can get tasks done quickly, such as setting reminders, sending messages, or checking the weather.

Multi-Tasking: Voice assistants can handle multiple tasks and requests in a single interaction. Users can ask for information, control smart home devices, and perform other tasks in a single conversation.

Smart Home Integration: Voice assistants can control and manage various smart home devices, providing a centralized hub for home automation. Users can control lighting, thermostats, locks, and more using voice commands.

Information Retrieval: Voice assistants can provide quick access to a wide range of information, such as news updates, weather forecasts, sports scores, and general knowledge. This is especially useful for on-the-spot inquiries.

Entertainment and Media: Users can ask voice assistants to play music, movies, podcasts, or audiobooks, making them ideal for entertainment and media consumption.

Productivity: Voice assistants can be used for setting reminders, scheduling appointments, sending emails, and managing to-do lists, increasing personal and professional productivity.

Translation and Language Learning: Voice assistants can assist with language translation and pronunciation, making them valuable tools for travelers and language learners.

Personalization: Many voice assistants can be customized to suit the user's preferences. Users can set their preferred wake word, choose voices, and personalize settings.

Continuous Improvement: Voice assistant technology is continuously improving through software updates and machine learning. This means that they can adapt to user behavior and provide increasingly accurate and personalized responses over time.

Wide Adoption: Voice assistants are integrated into a variety of devices and platforms, including smartphones, smart speakers, and automotive systems, making them widely accessible.

Voice Commerce: Users can make purchases and order products online through voice assistants, creating opportunities for e-commerce and online shopping.

Education and Learning: Voice assistants can help with homework, provide answers to general knowledge questions, and serve as educational tools for learners of all ages.

Remote Control: Voice assistants can serve as universal remote controls for various devices, such as TVs, gaming consoles, and home theater systems.

Third-Party Integrations: Many voice assistants offer third-party app integrations, allowing users to extend their functionality and access a wide range of services and information.

Enhanced Customer Service: In business and customer service applications, voice assistants can provide automated support, streamline customer inquiries, and enhance the customer experience.

**CONCLUSION:**

The development and implementation of a voice-based virtual assistant project has been a significant journey, aimed at enhancing user experiences and simplifying everyday tasks through natural language interaction.

In conclusion, the voice-based virtual assistant project represents a significant step towards creating an intelligent and user-centric digital assistant. Its potential applications are vast, ranging from simplifying daily tasks to supporting individuals with accessibility needs. The project's success has been possible through the dedication of our team, and we are committed to refining and expanding its capabilities to better serve our users in the future.

**REFERENCE:**

[**https://www.techtarget.com/searchcustomerexperience/definition/virtual-assistant-AI-assistant**](https://www.techtarget.com/searchcustomerexperience/definition/virtual-assistant-AI-assistant)

[**https://en.wikipedia.org/wiki/Virtual\_assistant**](https://en.wikipedia.org/wiki/Virtual_assistant)

**https://ieeexplore.ieee.org/document/10141447**