|  |
| --- |
| ***A Mini Project Report on***  **Voice Assistant Using AIML**    **T.E. - I.T Engineering**    **Submitted By**  **Soham Bolla 20104135**  **Sakshi Ahire 20104021**  **Sakshi Gaikwad 20104015**    **Under The Guidance Of**  **Prof. Shital Agarwal**        **DEPARTMENT OF INFORMATION TECNOLOGY**  A.P.SHAH INSTITUTE OF TECHNOLOGY  G.B. Road, Kasarvadavali, Thane (W),  Mumbai-400615 UNIVERSITY OF MUMBAI    Academic year : 2022-23 |

|  |
| --- |
| **CERTIFICATE**      This to certify that the Mini Project report on Voice Assistant System has been submitted by Soham Bolla (20104135), Sakshi Ahire (20104021) and Sakshi Gaikwad (20104015) who are a Bonafede students of A. P. Shah Institute of Technology, Thane, Mumbai, as a partial fulfilment of the requirement for the degree in Information Technology, during the academic year 2022-2023 in the satisfactory manner as per the curriculum laid down by University of Mumbai.      Shital Agarwal  Guide      Dr. Kiran Deshpande Dr. Uttam D.Kolekar  Head Department of Information Technology Principal      External Examiner(s)  1.  2.    Place: A.P.Shah Institute of Technology, Thane Date: |

|  |
| --- |
| **ACKNOWLEDGEMENT**    This project would not have come to fruition without the invaluable help of our guide Shital Agarwal. Expressing gratitude towards our HoD, Dr. Kiran Deshpande, and the Department of Information Technology for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our teacher Ms. Charul Singh who gave us her valuable suggestions and ideas when we needed them. We would also like to thank our peers for their helpful suggestions. |
| **ABSTRACT**  The primary goal of trending technology artificial intelligence (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. Voice assistants are a fantastic advancement in the sector of Artificial Intelligence that can transform people\'s lives in a variety of ways. The voice-based assistant was initially given on cellphones and quickly gained popularity. It was widely acknowledged by all. Previously, voice assistants were largely found in smartphones and laptops, but they are now increasingly available in various home automation setups and smart speakers. Many technologies seem to become wiser in their very own way, allowing them to converse with humans in a simple language. Desktop voice assistants are programs that can identify people\'s speech and answer through an integrated speech system. This paper will outline how different voice assistants work, as well as their primary challenges and limitations. The way of developing a voice-based assistant without requiring cloud services is discussed in this paper, which would promote the future growth of such devices. |
| **TABLE OF CONTENTS**       1. Introduction……………………………………………………………………..1   1.1.Purpose..................................................................................................... 1  1.2.Problem Statement………………………………………………………2  1.3.Objectives................................................................................................. 2  1.4.Scope........................................................................................................ 3   1. Literature Review………………………………………………………………...4 2. Proposed System……….........................................................................................7   3.1. Features and Functionality………………………………………………7   1. Requirements Analysis......................................................................................... 9 2. Project Design.......................................................................................................10   5.1.Use Case diagram……………………………………………………… 10  5.2.DFD (Data Flow Diagram) ……………………………………………..11  5.3.System Architecture……………………………………………………..12   1. Technical specification……………...…………………………………………...13 2. Project Scheduling.................................................................................................15 3. Implementation…………………………………………………………………..16 4. Result and Discussion…………………………………………………………….17 5. Conclusion and FutureScope.................................................................................18 6. References. |

# Chapter 1

**1. Introduction:**

A voice assistant can be a digital assistant that uses human voice, language process algorithms, and synthesis to pay attention to voice commands and come applicable information or perform functions as appealed by the user supported commands, commonly known as intents, tell by the user, voice assistants will come applicable information by hearing for keywords and filtering out the close noise. While voice assistants may be completely a software system primarily built on and ready to combine into all devices, some assistants are sketched individually for every unique device application, like the Amazon Alexa clock.

Now a day, voice assistants are combined into some of the devices we intend to use daily, like cell phones, computers, and good speakers. The main motive of this project is to develop it for physically challenged people. In this project we have developed a static voice assistant using python which will perform operations like copying and paste files from one location to another location, to send a message to the users mobile also order pizza and mobile using voice commands. The mass adoption of AI in users’ everyday lives is additionally refueling the shift towards voice. One of the most popular voice assistants are Siri, from Apple, Amazon Echo, which responds to the name of Alexa from Amazon, Cortana from Microsoft, Google Assistant from Google, and the recently appeared intelligent assistant under the name AIVA.

**1.1. Purpose:**

Virtual assistants are typically cloud-based programs that require internet-connected devices and/or applications to function". The technologies that power virtual assistants require vast amounts of knowledge, powering the platforms, as well as machine learning, language communication processes, and speech recognition arena. There are dedicated devices to provide virtual assistance. The most stylish on the market from Amazon, Google and Microsoft having Alexa, Google Siri and Cortana as AI voice assistants respectively given by each company.

• Provides Instant Access to information.

A voice assistant provides fast answers to questions, resulting in faster completion of projects and work.

* + Manage Timesheet with voice assistant.

Simply tell your assistant what project you’re working on, where you are, how long it should take, and it will tell you exactly how much time you spent on each task.

* + Can be used anywhere, anytime.

75% of smartphone users means 3.1 billon people have access to a voice assistant and almost 2 billion people are using a voice assistant right now.

**1.2. Problem Statement:**

This involves knowing about the speech signal we wish to recognize in order to specify the processing required by the machine. The most striking aspect of speech, which achieves its full impact when ASR is considered, is appalling complexity. The most effective overall picture of speech in relation to the machine and the recognition process is by means of a block diagram such as shown. The connecting arrows stand for ‘entails’, and the circles, with numbers in, are conventional threshold gates, used to compound the entailments. In the top left area of the figure is shown the nature of the speech we wish to recognize. The bottom left area shows the main characteristics of a machine which might be built to recognize speech. The right-hand side of the figure outlines some of the operations that the machine would need to perform.

The main purpose of the software is to perform the tasks of the user at certain commands, provided in either of the ways, speech, or text. It will ease most of the work of the user as a complete task can be done on a single command

**1.3. Objectives:**

The goal of voice assistance is to make people smart and provide them with instant and accurate results. It takes the voice input and converts it into a computer-readable language that can answer their questions. The goal of this service is to provide users with results that they have asked for through the web.

Main objective of building personal assistant software (a virtual assistant) is using semantic data sources available on the web, user generated content and providing knowledge from knowledge databases. The main purpose of an intelligent virtual assistant is to answer questions that users may have.

* + It has been designed to provide a user-friendly interface.
  + Users can interact with the assistant through voice commands.
  + It's bringing AI and machine learning together to recognize our voice and do what we ask it
  + The Objective of AI Virtual assistant in today’s world is to replace secretaries and personal Assistant.
  + Requires less consumption of time in writing text.

**1.4. Scopes:**

Modern voice assistants can do almost anything you might be able to think of. You can use them to play music, answer your questions, make phone calls, send messages, get a live location, get a temperature. By 2030, speech recognition will feature truly multilingual models, rich standardized output objects, and be available to all and at scale. Humans and machines will collaborate seamlessly, allowing machines to learn new words and speech styles organically.

It will further develop the work process, consumer loyalty, and deals and develop your suggestions to your customers. Artificial intelligence-fueled Virtual Assistants can likewise make customized Emails - email automation for every one of your customers and take notes of significant central issues during a gathering.

# Chapter 2

**Literature Review:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sr no. | Title | Author(s) | Year | algorithm | limitation | result |
| 1. | Simple AI-  Assistant. | Rutuja V. Kukade | 2018 | OCR (Optical character recognition) | Permitting only through voice command, can  also install input text | Implemented model perform tasks by taking human voice commands |
| 2. | Virtual assistant system. | Shivangee Kushwah | 2019 | CART, XGBoost,  SVM | Implementation of voice  command only verifying voice system. | The project’s attributed to the use of a stacked model, which was able to classify the voice as either male or female. |
| 3. | AI- Speech Recognition system. | Rijwan Khan | 2020 | SVM and Random  Forest, Artificial Intelligence for  Speech  Recognition | Restrict the mobility, does not build the model based on growing generation | The project is to create a system that will allow people with limited mobility to grow along with society. |
| 4. | JARVIS-  Clone system for desktop | Harsh D. Shah | 2021 | Automatic Speech Recognition(ASR | Email sending hampers sending unstructured format emails and grammer | This project allows users to send email using  their voice  commands and the address book of the person |

**Paper 1:**

1. Nil Goksel, CanbekMehmet ,EminMutlu, On the track of Artificial Intelligence: Learning with Intelligent Personal Assistant, proceedings of International Journal of Human Sciences, 2016. Keerthana S, Meghana H, Priyanka K, Sahana V. Rao, Ashwini B Smart Home Using Internet of Things , proceedings of Perspectives in Communication , Embedded systems and signal processing, 2017. Sutar Shekhar, P. Sameer, Kamad Neha, Prof. Devkate Laxman, An Intelligent Voice Assistant Using Android Platform, IJARCSMS, ISSN: 232-7782, 2017. Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K, An Intelligent Chatbot using Natural Language Processing, International Journal of Engineering Research , Vol.6 , pp.281-286, 2017. Luis Javier Rodriguez-Fuentes, Mikel PeÃ±agarikano, AparoVarona, GermÃ¡n Bordel,

GTTS-EHU Systems for the Albayzin 2018 Search on Speech Evaluation, proceedings of

IberSPEECH, Barcelona, Spain, 2018. Ravivanshikumar ,Sangpal,Tanvee ,Gawand,SahilVaykar, JARVIS: An interpretation of AIML with integration of gTTS and Python, proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kanpur, 2019. Luis Javier RodriguezFuentes, Mikel Peagarikano, Iavaroni, German Bordel, GTTS-EHU Systems for the Albayzin 2018 Search on Speech Evaluation, proceedings of IberSPEECH, Barcelona, Spain, 2018.

**Paper 2:**

1. Shaughnessy, IEEE, Interacting with Computers by Voice: Automatic Speech Recognition and Synthesis proceedings of the IEEE, vol. 91, no. 9, September 2003. [2] Patrick Nguyen, Georg Heigold, Geoffrey Zweig, Speech Recognition with Flat Direct Models, IEEE Journal of Selected Topics in Signal Processing, 2010. Mackworth (2019-2020), Python code for voice assistant: Foundations of Computational Agents- David L. Poole and Alan K. Mackworth.

## Chapter 3

**Proposed System:**

In this proposed concept effective way of implementing a Personal voice assistant, Speech Recognition library has many in-built functions, that will let the assistant understand the command given by user and the response will be sent back to user in voice, with Text to Speech functions. When assistant captures the voice command given by user, the under lying algorithms will convert the voice into text.

Proposed Architecture

The system design consists of

* Taking the input as speech patterns through microphone.
* Audio data recognition and conversion into text.
* Comparing the input with predefined commands.
* 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. In the next step, this resulting string the data is manipulated through Python Script to finalize the required output process. In the last phase, the produced output is presented either in the form of text or converted from text to speech using TTS.

3.1 Features and Functionality:

* It keeps listening continuously in inaction and wakes up into action when called with a particular predetermined functionality.
* Browsing through the web based on the individuals’ spoken parameters and then issuing a desired output through audio and at the same time it will print the output on the screen.
* It's software that carries out everyday tasks via voice command. It's bringing AI and machine learning together to recognize our voice and do what we ask it.
* Voice assistant software can be found on smart speakers, smartwatches, mobile phones, tablets, and other devices.

Getting Current News: about his/her motherland, about world, about technologies, about sports or about entertainment of the industry and much more, the user can easily get the news just by giving voice input to assistant to open news so it will open new tab and it can also fetch the data from the websites and return it to the console and read out for user without any labor.

Weather Forecast, through this feature users can see the weather forecast for any location. In addition, the temperature and humidity of Kelvin will return the weather.

Open Applications like , YouTube, google search engine , launching websites , system applications with the help of web browser python library and os for opening system applications(like, code editor, notepad,chrome,etc.) Open Applications like , YouTube, google search engine , launching websites , system applications with the help of web browser python library and os for opening system applications(like, code editor, notepad,chrome,etc.)

WhatsApp Messages, the application works by taking mobile number of the receiver or the name of the receiver, message to send , time when to send as a query. As a result , voice assistant will send the message and inform you. This is done with the help of pywhatkit python library. And the history of messages will be saved in pywhatkit database file .

Checking Internet Speed, the application is done with the help of speed test python library by which assistant will check and return the result on the console.

Sending Mails, this feature allows users to send an email to someone whose contacts include an email address. It then sends the successful execution of the task back to the user.

## Chapter 4

**Requirement Analysis:**

Speech recognition, also known as automatic speech recognition (ASR), computer speech recognition, or speech to-text, is a capability which enables a program to process human speech into a written format. While it’s commonly confused with voice recognition, speech recognition focuses on the translation of speech from a verbal format to a text one whereas voice recognition just seeks to identify an individual user’s voice Speech recognition processing Speech recognition is the process of converting human sound signals into words or instructions.

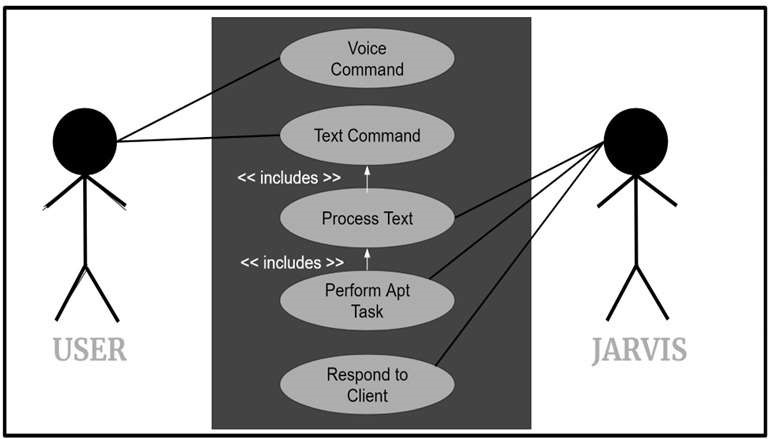
The research of speech recognition involves many subject areas such as computer technology, artificial intelligence, digital signal processing, pattern recognition, acoustics, linguistics, and cognitive science. It is a multidisciplinary comprehensive research field. Different research areas have emerged based on research tasks under different constraints.

## Chapter 5

**Project Design:**

The first step that needs to take place is to record the spoken words and convert it to text, followed by ascertaining that they is able to extract the intent using its artificial intelligence algorithms. The next logical step is to confirm that the Virtual Assistant is able to respond based on the intent deduced in the prior step. Some responses need to execute system commands, others need to get information from third party Application Programming Interface (API) (like weather, and other applications) or changing some values on the Internet of things (IoT) devices.

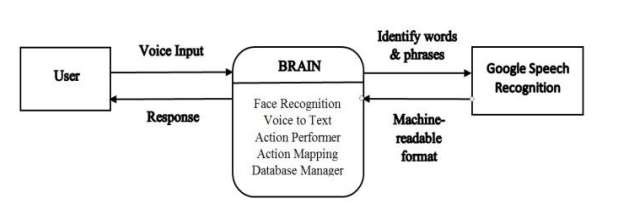
**5.1. Use Case Diagram**



**5.2. DFD (Data Flow Diagram):**

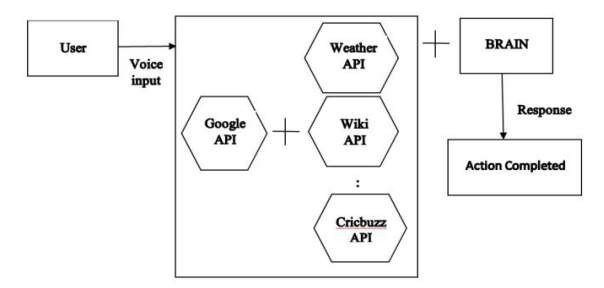
Level 0 DFD:

The user gives the input in the form of voice; this voice command is recognized by the application. Then it will check whether it is the authorized user, then the action is performed as per the command given by the user. Command given is compared as a form of action and question and response with the dialog box or search through the knowledge base.

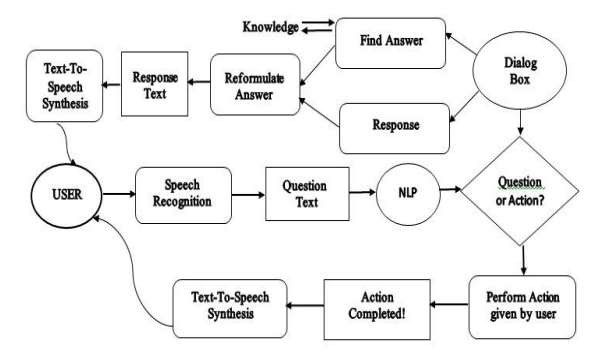


Level 1 DFD:

Input is given by the user in the form of voice. Google Voice API will convert this voice data into text form and then the action is performed by the voice assistant according to the command given by the user by comparing with the dialog box and knowledge base.



**5.3. System Architecture**



**Chapter 6**

**Technical specification:**

GUI: Python (Visual Studio IDE)

The device first converts speech input from the user into text using python module speech recognition algorithm. From voice input taken from users, we can obtain texts from specialized corpora arranged on the research center's computerized network server, that are briefly keep in the computer system before being transferred to python's module for recognizing speech. Then, the central processor accepts the similar text and feeds it.

Backend: Python Backend

The backend of python parses the speech recognition module's response to examine whether the speech or command outcome is either a System Calls, Send Mail, API Call, or Context Extraction. The data is then sent out towards the backend of the python server to furnish the user with the relevant results.

OS: Windows

Windows is a graphical operating system developed by Microsoft. It allows users to view and store files, run the software, play games, watch videos, and provides a way to connect to the internet. It was released for both home computing and professional works.

**Chapter 7**

**Project Scheduling:**

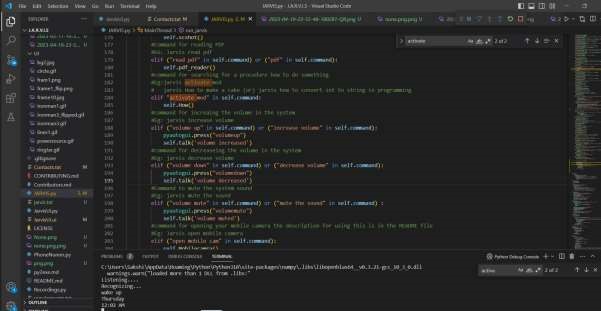
|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. No** | **Group Member** | **Time duration** | **Work to be done** |
| **1.** | Soham Bolla | 2nd half of January | Prepared the abstrat and dissuss the problem Statement with the guide*(includes the abstrat technical data and problem state.)* |
| 2nd Week of February | Designing GUI and which algorithm  to be used, objectives and function.*(NPl,ASR)* |
| **2.** | Sakshi Ahire | 3rd Week of February | Designing the GUI Model based on  the properties and funtion disscued.*(GUI Interface)* |
| 4th Week of February | Implementing 1st Module*(GUI Working testing and project model implementation.)* |
| **3.** | Sakshi Gaikwad | 2nd Week of March | Implementing 2nd  Module/Funtions*(Implemented all the funtions and features for the Assisant connection the model with the GUI)* |
| End of march | Implementing 3rd  Module/Funtionalities*(Implementstion*  *of the fetures and disscued the changes with respect to the model with the guide)* |

**Chapter 8**

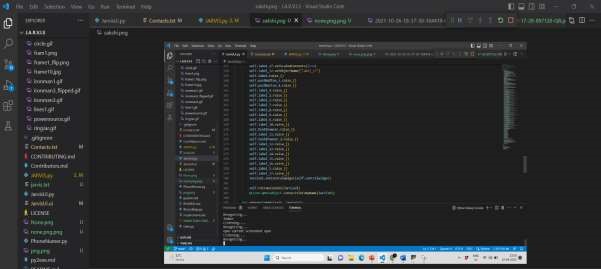
**Screenshot 1: (GUI)**



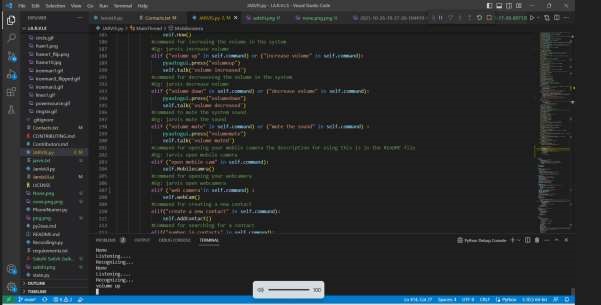
**Screenshot 2:(Waking up Jarvis):**



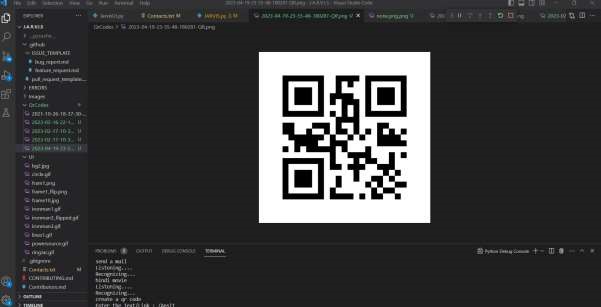
**Screenshot 3: (Take Screenshot):**



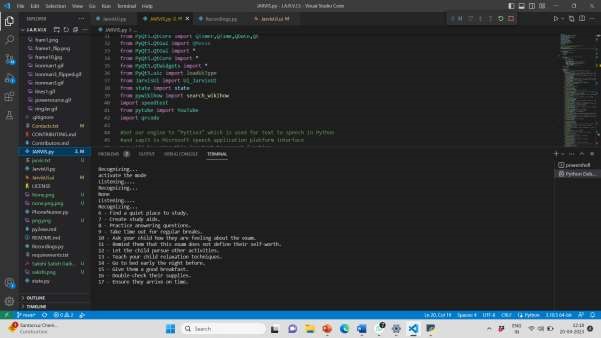
**Screenshot 4:(Volume up):**



**Screenshot 5:(Generate QR code):**



**Screenshot 6:(Activate Mode):**



**Chapter 9**

**Result and Discussion:**

AIML (Artificial Intelligence Markup Language) is a programming language that allows developers to create conversational AI applications, including voice assistant systems. In this system, users interact with a virtual assistant through voice commands , and the AI-powered system responds to the users in a natural language.

There are several benefits of using AIML for developing voice assistant systems. Firstly, it is a well-established and widely used programming language for developing conversational AI applications. Secondly, AIML has a relatively low learning curve, making it easier for developers to create and modify chatbots or voice assistants using AIML. Lastly, AIML allows developers to create custom templates and rules that enable the chatbot or voice assistant to recognize and respond to a wide range of user inputs.

The results of implementing a voice assistant system using AIML can be impressive. The system can understand and interpret users' natural language inputs accurately, and provide relevant responses in a conversational tone. Additionally, voice assistant systems can be integrated with other applications or services, such as calendars, weather forecasts, and news, to provide a more comprehensive and personalized user experience.

However, there are also some limitations to using AIML for developing voice assistant systems. Firstly, AIML-based systems may not be as sophisticated as other conversational AI frameworks, such as neural network-based chatbots or voice assistants. Secondly, AIML relies heavily on rule-based programming, which may limit the system's ability to understand more complex or ambiguous user inputs. Lastly, AIML-based systems may struggle to handle multiturn conversations, where the user's response is dependent on the previous interaction with the system.

**Chapter 10**

**Conclusion and Future Scope:**

In conclusion, the voice assistant system using AIML is a promising technology that enables users to interact with virtual assistants in a conversational and natural way. The system has proven to be effective in accurately interpreting and responding to user inputs, providing personalized and comprehensive information to the users.

The future plans include integrating Assistant with mobile using React Native to provide a synchronized experience between the two connected devices. Further, in the long run. APSIT is planned to feature auto deployment supporting elastic beanstalk, backup files, and all operations which a general Server Administrator does. The functionality would be seamless enough to replace the Server Administrator with voice assistant.

The future scope of voice assistant systems using AIML is vast. As the demand for conversational AI technology grows, developers can expect to see improvements in the capabilities of the AIML programming language. This can include enhancements in natural language processing, machine learning algorithms, and speech recognition technology.

Moreover, voice assistant systems can become more personalized by integrating with other technologies such as Internet of Things (IoT) devices, social media, and other online platforms. As these integrations become more seamless, voice assistants can help users accomplish more tasks and provide a more personalized experience.

Finally, the future of voice assistants may involve the development of multi-modal interfaces that can combine voice, text, and touch inputs to create a more immersive and natural interaction with the user.

**Chapter 11**

**References:**

1. [https://www.analyticsvidhya.com/blog/2020/09/ai-virtual-assistant-usingp](https://www.analyticsvidhya.com/blog/2020/09/ai-virtual-assistant-using)ython/
2. <https://pythongeeks.org/python-voice-assistant-project>
3. <https://codewithharry.com/videos/python-tutorials-for-absolute-beginners-120>
4. <https://www.geeksforgeeks.org/personal-voice-assistant-in-python>