Smart Assistant using Machine Learning

Ravita Prajapati[1], Abhishek Singh[2], Anirudh Agrahari[3], Pooja Singh[4], Abhishek Shahi[5]

[1,2,3,4]B.Tech Student, [5]Asst. Professor

[1,2,3,4,5]Deportment of Computer Science and Engineering, Buddha Institute of Technology GIDA Gorakhpur, Uttar Pradesh, India.

Abstract - **A smart assistant is a virtual assistance that executes all of a user's tasks.For instance, computer-based apps. For example, consider a personal assistant who is always nearby and keeps an eye on the user's every move. Give him sound advice on his profession. In everything, there is help. A single user working in his lab can then make him feel as if there is someone else in his lab who is assisting him with his task. Providing him with guidance. He is warning the user before he makes a mistake. Users who are alone should be removed. Use its abilities and gradually develop them. This is a difficult concept to grasp, but it can be accomplished with the help of the Python programming language and machine learning. It's incredible in the digital life of 2021, where computer programmers are using machine learning to make everything simple. No one could have imagined this even 70 years ago. Machine learning, on the other hand, is now widely used. This is incredible. No one could have imagined this even 70 years ago. Machine learning concepts, on the other hand, are now being applied. The implementation is in progress. Day by day, things are becoming simpler. In the statistical context, it is an application of artificial intelligence in which available data is processed or assisted in the processing of statistical data using algorithms.**

I. INTRODUCTION

We often remark that no one helps us without being selfish, and this is somewhat true. But there is only one thing in this world that will aid you without being selfish, and that will help you without ceasing. Yes, we are talking about machines that aid us without being selfish if they merely require electric power. You don't have to do anything for the computer. Simply issue the command, and the outcome will appear in front of you. Between the 1940s and the 1950s, computers could only do what humans programmed them to do. The difference between man and machine was that a computer could not learn from its own actions.

By repeating the same task, a person can achieve something different. However, the machine lacked the ability to learn on its own. It has to be taught how to work by the coder.

However, technology has progressed significantly in recent years. Now the machine is capable of making its own decisions. A computer game capable of defeating the world champion was developed in 1950.

In this software, the computer game player assisted in taking the optimal action in a short amount of time, as a result of which many computer games such as Chess, Ludo, and others exist today.

Gairy Kasparow, the World Champion, was defeated by Deep Blue Computer.

A smart assistant is a computer application that acts as a personal assistant. You don't need a person for anything if you have a smart assistant programme, for example. If we are in a library, for example, we may ask our smart assistant which type of book to read and what the advantages of reading that book are. Aside from that, let our programme read the book, and if you get bored, it can figure out what you're saying. If there is any additional vital task that we have overlooked while studying, we will be reminded of it.

Smartphone computers and internet access have become indispensable in today's business world.

A virtual assistant is a self-employed individual who works from home on his computer and phone to give online services to a company, its customers, administrative, technical, creative, and other areas that may be accessed over the internet.

II. BIRTH OF VIRTUAL ASSISTANT

In the 1990s, digital speech recognition technology became a feature of personal computers, competing with IBM Philips and Learn Out. In 1994, IBM Simon launched the first smartphone on the market, laying the groundwork for today's sophisticated virtual assistants.

Dragon's natural speaking software cloud recognised and transcribed natural human speech into a paper at a rate of words per minute without stop in 1997. It is still possible to download a version of naturally speaking. Today, it's still in use. Many doctors in the United States and the United Kingdom, for example, utilise their medical records solely to document their patient's condition.

Colloquis publicly introduced smarterchild in 2001 on platforms such as AIM and MSN Massenger because playing fully text-based smarterchild games allowed users to check the weather, see information, and interact with each other to some level.

Shiri, the first smartphone-based digital virtual assistant, debuted on October 4, 2011 as a feature of the iPhone 4s.

After purchasing Siri Inc. in 2010, a spin-off of SRI International, a research institute funded by DARPA and the US Department of Defense, Apple Inc. developed Siri.[10] Its goal was to make it easier to do things like send text messages, make phone calls, check the weather, and set an alarm. It has evolved over time to offer restaurant recommendations, internet searches, and driving directions.

Alexa was introduced alongside the Echo in November of 2014.

Amazon launched a service in April 2017 that allows users to create conversational interfaces for any sort of virtual assistant or interface.

III. LITERATURE AND SURVEY

Speech recognition-based learning system.

Lavin Jalan's study examines how English is a globally acceptable language with a variety of accents spoken in different places of the world.

The study uses an isolated character dictionary based on speech recognition to deliver precise meaning of spoken words with great accuracy.

It can be considered as a one-of-a-kind approach to a variety of human-computer interaction domains.

The acoustic phonetic method, Pattern Recognition method, and artificial intelligence method are used to create a real-time speech-based learning/training system distributed between client and server that incorporates speech recognition and linguistic processing for recognising a spoken question and providing an answer to the end user in a learning or training environment implemented on the internet.

There are two fundamental processes in pattern recognition.

Pattern training and pattern comparison are both involved. The stochastic approach and the template approach phonetic are two methods for recognising patterns.

The core of this strategy is the acoustic phonetic approach and the process of locating sounds.

Artificial intelligence is the third decoding approach. It can be thought of as a hybrid of pattern recognition and acoustic phonetic approaches.

Designing recognition algorithms, demonstrating speech units, and selecting relevant units are all part of this process.

Artificial intelligence is the most efficient form of speech recognition among all approaches.

Voice recognition intelligence in a personal assistant

Dr. KshamaKulhalli [et.al] offered a work in which they proposed Nowadays, mobile technology is well-known for its user experience, as it is incredibly simple to access applications and services regardless of your geographic location.

Google has also built a voice search application for Android phones, but it relies heavily on an internet connection, which is one of its primary drawbacks.

Context is used by a personal assistant system to intelligently make contacts and provision an address book. When the user does not provide a preference, the system analyses historical data to estimate whether calls will be successful.

C. [3]Digital assistant for the blind

The approach offered by Prince Bose [et.al] creates a speech-based intelligent system for the blind, allowing them to use technology to express words and then hear the computer recite them, as well as command a computer using their voice, rather than having to look at the screen and keyboard[5].

This research demonstrated the interaction between a user guide and an intelligent assistant.

We reasoned that measuring user feedback through user research would be a suitable way to demonstrate the usefulness of the suggestion.

We will also add a multi-modality-based suggestion system to this methodology.

D. [4] Speech synthesis for generation of artificial personality

RoshanJadhav [et al.] proposed work that produced a learning model for a voice-based intelligent system that is secure.

Voice recognition has the potential to be an effective solution to a variety of security issues, including telephone-based crime, fraud, and many more.

In the banking and e-commerce industries, biometric speech recognition software is widely employed. [1] Voice synthesis is a method of artificially producing speech for text-to-speech conversion.

Speech synthesis is demonstrated in this work, with the development of intonation patterns, referred to as rules, as the central component.

These rules are organised in a frequency structure that includes a group, a vowel group, and a consonant group.

*1) Presentation Tier:*The presentation layer is the application's highest level. It is a layer that consumers can directly access. Login, home screen, modules (calculator, file writing, read document, Gmail, command mode) are the GUI interfaces used. It is a layer that allows users to access the following screen, which is the home screen, by entering their username and password. The view structure of modules completed in the project system is displayed on this home screen. The modules are selected, and then the next steps are taken. Each module goes over each one in depth. The arithmetic operations to be performed are contained in the Calculator module. Save and edit options are included in the file authoring programme, and specific sentences can be dictated and saved.

Composing and sending messages are available with Gmail. A command module contains commands such as open terminal, browser, and so forth.... All of them are part of the presentation tier, and the escape button returns the user to the home screen. When each action is accomplished, the filter module is well viewed since the adults' words must be dismissed in order to speak accurate words.

*2) Data tier:*This is the application's second-highest layer. The data tier contains data persistence technologies such as the SQL Apache database. As we all know, SQL is used to interface with databases, and SQL statements are used to execute operations like updating and retrieving data from them. The user data tables will hold the user's data as well as any information entered while logging in to the system. The storage of data is a specific goal of programming languages.

*3) Application tier:*To access the Smart Assistant System via the GUI, the user must first log in. A username and password will be given to the user upon login. The database will save this information. The application layer is the third tier, and it is here that the database system stores the results of the system's actions. The username, the file-writing parameters, readable documents, and the composing mail.

IV. EXISTING SYSTEM

Many systems exist to operate machines using voice commands, such as Cortana on Windows, Jarvis on Linux, and others, but they all rely on the quality of hardware to predict the correct words; otherwise, they may interpret words incorrectly.

Cortana isn't able to distinguish between symbolic and alphanumeric letters. Cortana also lacks a functionality that allows the administrative user to modify his or her password, userId, or username. Cortana instructs the user on how to change their password or username but does not change it for them. Another disadvantage of Cortana is that it requires precise sentences and words in order to do a task. Only the requested task will be completed if the end user speaks those particular words or sentences; otherwise, an error notice will appear on the screen.

V. PROPOSED SYSTEM

Because we know that mobile devices have high-quality hardware such as microphones and speakers, as well as a plethora of libraries for mobile OS to perform speech synthesis, we are using an Android device as an intermediate for predicting words. We also use webservices to deliver commands to the computer over HTTP and perform actions.

Speech recognition is a technology that allows a computer or other machine to understand and compute human speech. This is a significant AI application that creates two-way communication via instant messages or voice

commands. There are a variety of systems that can convert human speech to text. We can use the text we get from the voice to text engine as an assistant. It will make use cases easier to understand, and the words can also be utilised as commands. Our systems follow the same approach of transferring text from speech to a text engine and processing it as commands. The use of voice commands to control the computer is possible with voice assistant systems.

Open cmd, type in a notepad, send or receive emails, check the weather, create reminders, and so forth. Our project entails developing an AI-based application that uses voice instructions to communicate between humans and our computer. In today's world, this technology has become increasingly attractive. Market-available software is machine-specific, posing system compatibility difficulties that will be addressed in our system. The project is carried out in accordance with the phases of the software development lifecycle. We utilised an agile approach. Every project should be handled differently, according to the Agile paradigm, and existing methods should be adjusted to best suit.

The project's specifications. The agile method is based on iterative planning, which makes it simple to adjust when requirements change. The technique is iterative, with each iteration delivering a viable software build. In terms of features, each build is incremental. Using AI and speech recognition algorithms, man's dream to construct a machine that looks like him and can converse with humans in natural language has become a reality. The more technological advancements are made, the more sectors in which they can be used. These programmers may be beneficial or even detrimental to the domain of individuals who use them. Project organization is a method for arranging and making decisions concerning the project's implementation.

It includes the structuring, organizing, and configuring. It creates software for physically challenged people to make operating a computer easier for them and to take use of emerging technology. Speech Recognition is organized in Windows 10 to enable users control a system using their voice. The features allow you to open apps, programs, and _les, which lets you to enter in a document by hand. All of this is made possible by the use of a microphone, either built-in or external.

VI. METHODOLOGY USED

Android Speech to text engine–

This is the fundamental feature of Android OS, which allows us to transform speech to text immediately via the Voice Recognizer thread.

PHP web server

This web server is used to transfer commands from the mobile device to the machine where the action is to be performed.

Virtual Key set:

This library's main purpose is to handle activities depending on commands received from a mobile device.

Login:

The first step in using this system is to log in. As a result, the user must provide his or her username and password to access the system. After entering a valid login and password, the system will proceed to the next step. Incorrect data

can lead to incorrect outcomes. The right login is demonstrated with a speech output that says "Welcome to smart assistant, how may I serve you," demonstrating the text to speech module. After the login page is completed, the valid one will be taken to the home screen, which will display several modules.

File Writing Module:

We use the vinterface database in this module to communicate with files. The vcommand table is made up of timestamps. A command with a long timestamp is regarded as recent. Whatever we say is recorded in the vcommand table, and the most recent command is presented with the help of a buffered reader. It's vital in File writing mode since it receives speech input and shows it on the editor screen. When you use the save command, the file is saved for future use.

Calculator Module:

The calculator uses voice recognition to perform arithmetic computations. Addition, multiplication, division, and subtraction are among the operations enabled by this system. The square root operation is also available. The clear function is used to clear the previous result. The result is displayed using the Result function. Calculator operation: The user speaks the first number, followed by the operation he or she wishes to execute, and finally the second number. The output is displayed on the screen when the user speaks result. The outcome is removed when you speak clear.

Reader module:

Another aspect of the Smart Assistant System is this module. Because this module is not only an absolute user requirement, but it is also a requirement of today's generation for all age groups, from children to the elderly. Now it's Button's turn to 'Read out.' If the file isn't selected, the path is empty, or the file isn't in.txt format, the Option Pane displays a notice that says "PLEASE SELECT A TEXT FILE." Because we're using the Lexar class, which only performs lexical analysis on normal text. For a time, we need lexical analysis to stop the speaker at a full stop or comma.

Other than '.txt' format files, formatted text, which can include graphics, is not processed by the Lexar class. We return to the previous UI of the Welcome () class when we close the window again. The convertor window will open when the command "Document to Audio" is issued. Everything in this window is identical to "Read Out Document," with the exception that "Read Out Document" simply reads the text file for the current instance, whereas "Document to Audio" stores the file in ".wav" format, allowing the user to listen to it anywhere, at any time. It is a programme that converts text files to audio files. This feature can be used to turn text books into audible books.

After using the Editor to store a task in a text file, the user can convert that text file to an audible file that he or she can listen to at any time. This programme works in the same way as 'Read out Document.' Simply use the File Dialog class's open file function to choose the file. In addition, the function showSaveDialog() of the Class JFileChooser object allows us to obtain the user's audio file saving location.

Command Mode:

The command processor is used to launch many OS utilities. However, the user must first specify which command they wish to use to open the utility. The following is the user interface for the command processor and adding new commands to the system:

This module uses the same package imports as the Editor, such as InputStreamReader, BufferReader, URL, URLConnection, and swing. The first figure depicts the Smart Assistant System's command prompt, which requires the user to issue commands through speech.

We've included built-in commands for several typical uses. When you use the 'close' command to close the Command Screen, it will self-dispose and launch the same Home Screen from which it came.

Mail:

One of the modules seen on the system's home screen is Mail Composing. a) Composing and b) Sending capabilities are part of the Gmail module's functionality. The user composes the necessary information to be sent to the person via email. The user must transmit, and the first composing function has been completed. This is accomplished through vocal input, and the data to be sent is preserved. The sending function is the second feature.

This method is called when the user has finished constructing the data, however it is not required. The sender's Id is chosen, and the word SEND is spoken as input. As a result, the email is sent to the specified sender. And the sender's device displays "SUCCESSFULLY SENT" as a notification.

VII. CONCLUSION

As a result, we've read numerous articles in order to design software for the physically challenged that is based on voice recognition algorithms. This included researching all aspects, including technologies, market trends based on AI applications, home appliances using voice recognition systems, and areas still under development in these fields, in order to find a way to provide an efficient and accurate system to the user so that they can control their device using speech rather than typing and staring at a screen. Through this effort, we have reached to the conclusion that man's deepest yearning to construct a computer that looks like him and can connect with others can become a reality.

Speech recognition has emerged as the most desired technology for today's human needs. People can use speech-based systems to send emails, read weather reports, get daily news, dial numbers, and do other things with their gadgets. We investigated all of the domains where speech-based technology are used extensively in this project. All technological aspects of this subject have been investigated, and we have determined that our system would be useful for physically challenged persons who can communicate with and control gadgets via a speech recognition system.

The software available on the market based on these technologies has a number of downsides, including the ability to interact with Bluetooth devices and the fact that it may be purchased, therefore these drawbacks were investigated and addressed in our system. Our system swiftly composes emails, searches for a specific _le, writes and saves documents, reads aloud a selected _le and stores it in audio format, and we've also developed a speech calculator that can handle a wide range of tasks. As a result, we have completed the successful development of our project.

VIII. ACKNOWLEDGMENT

REFERENCES

[1] Yahya S H Khraisat" speaker independent voice recognition calculator" Department of Electrical and Electronics Engineering Contemporary engineering science vol 5,2012 no -3, 119-125
[2] Li Deng, Fellow, Xiano Li, "Machine Learning Paradigms For Speech Recognition "IEEE TRANSACTION ON AUDIO AND LANGUAGE PROCESSING, VOL 21 MAY 2013, ISBN: 1558-7916
[3] K Naga Abhishek Reddy, "A Comparative study on Speech Recognition Approaches and models" IJCTT-Volume 43 Number 2-Jan 2017.

[4] Khaled M. Alhawati, "Advances In Artificial Intelligence" International Journal of computer, Vol: 9, No: 6, 2015
[5] Mathew P. Aylett; Alessandro Vinciarelli; MirjamWesters; "Speech synthesis for the generation of artificial personality "IEEE TRANSACTIONS ON AFFECTIVE COMPUTING 2017 ISSN: 1949-3045
[6] Md .Almin; SyedaZinathAmanl "Design of Intelligent Home Assistant", 2016 7th International Conference on Intelligent Systems, Modelling and Simulations.
[7] Dr. KshamaKulhalli; AbhijeetJ.Patinas, "Personal Assistant Using Voice Recognition Intelligence, "International Journal of Engineering Research and Technology", ISSN: 0974-3154 Volume 10 number 1(2017)
[8] Prince Bose; ApurvaMalpthak; UtkarshBansal, "Digital Assistant for the Blind" 2nd International conference for Convergence in Technology (I2CT) 2017.