**VIRTUAL ASSISTANCE USING AI**

**HARINI K** [**h88714672@gmail.com**](mailto:h88714672@gmail.com)

**HARINI K M** [**manimarndheebha@gmail.com**](mailto:manimarndheebha@gmail.com)

**Panimalar Engineering College**

**ABSTRACT**

Virtual assistance technology has transformed the way we engage with computers and get access to the resources we need. After the recent developments around artificial intelligence , virtual assistance became more complex and versatile regarding the purpose it can serve for the users. The work in this paper demonstrates the new development of an existing virtual assistant that is equipped with the new features that will improve the quality of the service it provides to the users. This platform now has the potential to record human voice, meaning it can help the user take a voice note or record a verbal message without having to leave the system. Moreover, the system can determine the user’s location in terms of latitude and longitude and thus, provide location-based service as well as Tailored assistance. Furthermore, the remote helper currently has communication through signing acknowledgment, having the client utilize gesture based communication as info strategy. In addition, choice limit representation is carried out into the framework so the client can see what rationale the computer based intelligence used to make the given forecast . Subsequently, the accompanying passages can be revised to foster a more comprehensive and easy to use virtual help framework, which will be more reasonable for various clients.

**Keywords:** Artificial intelligence, virtual assistance,. location based features, Natural Language Processing(NLP).

**I. INTRODUCTION**

Artificial intelligence-driven virtual assistants have increasingly infiltrated numerous aspects of our daily lives, providing functionality and ease of access across a broad range of services and information. Nonetheless, the underlying technology, while designed to align with user expectations, trends, and use cases, may lag behind developments in this rapidly evolving field. As such, continual improvement to such services is required. In this Capstone work, we develop an upgrade of the assistant to implement several new ideas and improve functionality. certainly, Identifying and overcoming key limitations are important. One way to enhance functionality is to provide alternatives that enhance the user experience . The primary innovation is the ability for anyone to record audio Notes effortlessly.

Firstly, it is essential to note that this feature not only ensures smooth interaction but relays it to productivity as it allows the user to command some activities. Additionally, since we consider location-based features an essential component in our work, we also have an interest in the reminders service which comes in handy on a hands-free basis. With latitude and longitude detection function for the Virtual assistant, the system recognizes the user’s current position automatically. Precisely, the aspect ensures that the system offers some suggestions based on it . This can be an update on weather, among other location-specific details, enhancing user experience. In addition, for transparency and to instill trust on the user, we added the decision boundaries, which integrate the insights into the AI decision-making process. When users visualize the decision bounds, for instance, they understand the interpretation of information by the system such as data and predictions or recommendations.

**II. EXISTING WORK AND NOVELTY IN THE WORK**

Examples of prior virtual assistant platforms include the creation of Siri, Google Assistant, and Amazon Alexa. These technologies include the ability to recognize speech, natural language processing, and automate tasks. The authors in the work [1] discussed that the virtual assistant based on the voice, a desktop application developed using Python and recent advances of machine learning. By discussing voice commands processing based initiatives that follow the pre-processing to classification aspect from pre-processing to classifying to feature extraction, the creation offers a solution to the mentioned challenges of voice assistants. There are several previously researched artificial intelligence techniques to improve virtual assistant performance, including machine learning and deep learning. The authors in the work [2] discussed that the study concerns technology taking over tasks from humans in all fields and focuses on AI tools such as Amazon Echo and Google’s Voice Assistant because they can handle voice interactions well. It discusses a new virtual assistant algorithm that uses an advanced AI to process voice and text input to perform specified tasks. study concerns technology taking over tasks from humans in all fields and focuses on AI tools such as Amazon Echo and Google’s Voice Assistant because they can handle voice interactions well. The authors in the work [3] discussed that the virtual assistants are significantly improved through voice recognition and natural language processing. It has become indispensable in business due to speech recognition being on the rise. They utilize web data and parse user-generated content to create a personal profile for the user and learn from each interaction. It discusses a new virtual assistant algorithm that uses an advanced AI to process voice and text input to perform specified tasks. This paper also compares the proposed algorithm with the existing set of algorithms and determines which the best algorithm to use is. Such advanced technology aims to make virtual assistants better memrize the user’s interests, needs, and capabilities to build a more intuitive dialogue around them. Many applications and services use location data to provide personalized recommendations, navigation help and even location-related information. Many related works have made an effort to create systems capable of recognizing and understand sign language gestures to facilitate interaction for the dead and hard of hearing. The authors in the work [4] discussed that the virtual assistants guarantee comfort, but present a hazard to its users because of the confidential information stored in it. TASA employs real-time facial recognition to confirm a user’s pseudo-identity and ensures data security. Ignorance of identity checking in virtual assistants and high-quality assurance algorithms have been the primary goal . The work presented in this paper seeks to leverage various ways of improving model interpretability and guided explanations on previous methodologies to bring out an innovative upgrade in this genre of a system. While the models may provide the voice recognition and general use through a recorded voice command in current virtual assistants, voice recording hasn’t been integrated.

With this technology, a virtual assistant can determine a user’s location without asking for prior authorization. This helps in the assistant’s contextual awareness, and it can easily provide geographical-specific assistance or services more conveniently without relying on the user to describe his or her location. The authors in the work [5] discussed that in recent years, AI has seen spectacular advancements, most notably in NLP. Crosses-platform voice assistants that interact with users in natural language are now commonplace in several homes and educational institutions. Originally housed in smartphones, NLP has progressed to the automation of homes and began making promises in smart speakers, indicating a new era in its progress. This technology facilitates smooth communication between users of sign language and virtual assistants, encouraging accessibility and diversity. Decision boundary to help the users understand how their input will be evaluated by the AI fosters openness and trust. Comprehensively, this developed system upgrade encompasses all the designing features highlighted in this work to develop an augmented virtual assistance system. Integrating all these feature designs makes this work present a more open platform for even a broader range of scenarios and users

**III. ARCHITECTURE DIAGRAM**

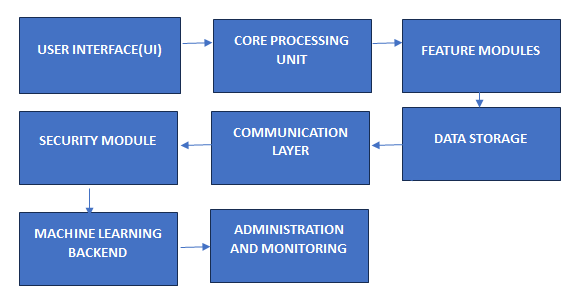


Fig. 1

To clarify the working principle of the upgraded virtual assistance system shown in Fig.1 and demonstrate how it operates, let's break down each feature's implementation and provide example test cases to validate their functionality. This approach helps ensure that each component is working as intended. The voice recording feature captures audio through a microphone, converts it into a digital representation, and saves it for eventual retrieval in a file or database. This function can be triggered by a specific voice command or a manual interaction (e.g., pressing a button). The user says "Start recording," speaks for 4 seconds, and then says "Stop recording." The system should save the audio recorded between the "Start" and "Stop" commands. Check if the audio file is stored and playback to confirm the content and duration.

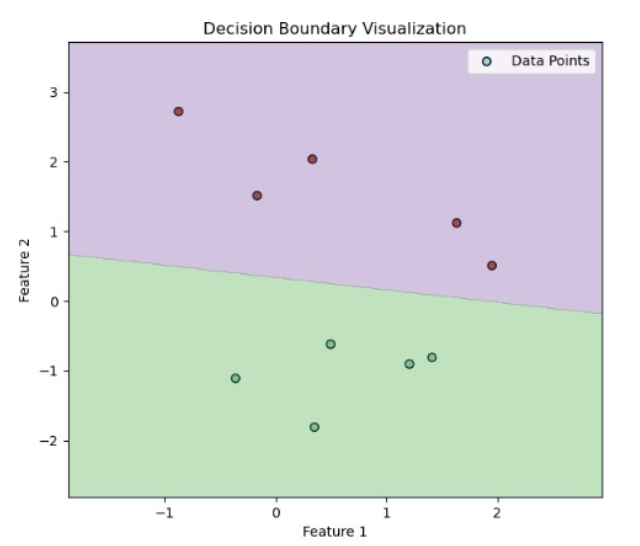
**Detecting Latitude and Longitude**

This feature utilizes the GPS module in the user's device or IP-based geolocation services to determine the current latitude and longitude. This data can be used to provide location-specific services like local news and weather updates. The user inquires, "What is my current location?"The system provides the current latitude and longitude.Compare the system's response with known location data from another GPS device or application to ensure accuracy.

**Detecting Sign Languages**

Working Principle: Sign language detection involves using a camera to record motion pictures of the user's movements. To convert these motions into spoken or written commands, computer vision, and machine learning techniques are used for processing and analysis. In front of the camera, the user signs "turn on the lights" in American Sign Language. The system interprets the gesture and turns on the lights in response to the order**.**  Check to see if the lights are on and whether the sign language interpretation was accurate.

**Decision Boundary Visualization**

****

F**ig. 2**

This feature shows the decisions that the AI model makes concerning the input data. Plotting data points on a graph and displaying the decision boundaries that the model employs for classification or decision-making are the usual methods. Provide a collection of system-tagged data points. A graph with data points and the decision boundaries should be displayed by the system. Check if the visualized decision boundaries logically separate the classes based on the input labels. The data points are shown in (fig no 2) represent the datas in the features.

**Implementing and Testing the System**

For voice recording, libraries such as audio or APIs like Android's MediaRecorder can be used. Location detection can utilize the HTML Geolocation API or Android's LocationManager. Sign language detection might leverage TensorFlow and OpenCV for processing and machine learning. Decision boundary visualization could be implemented using matplotlib or similar libraries in a Python environment.

**IV. PROCEDURE**

The recommended libraries are PyAudio for capturing audio recordings and wave for current .wav file tasks. FORMAT 16, CHANNELS 1, RATE 44100 and CHUNK 1024 and record the audio – FORMAT will be pyaudio.paInt16 , CHANNELS as 1 for mono recording there are 2 for stereo. RATE the sample rate , i.e., how often samples are taken per second will be set as 44100. CHUNK size 1024 i.e to specify the number of frames collected per buffer will be declared as 1024. PyAudio is used to handle the audio stream. Also stream will be created to flag it as input and specified the parameters that are mentioned in step 2. Notify the user that the recording has been started, initialize an empty list of frames on which the audio frames will be stored. Loop for a specific duration for instance, 5 seconds. In this step, one should find the number of iterations to make when chaning that to the number of seconds and seconds with the chunk rate.Terminate the audio stream automatically.Confirm to the user that the audio has been made.

**Detect sign languages**

How to detect hands: load an HX file with a pretrained Haar cascade to do this, the CascadeClassifier Frame to Grayscale: since color information is not useful for the cascade classifier, the frame should be converted to grayscale. Hand Detection: apply the Haar cascade classifier to the grayscale image. Open Webcam. Prompt all the Frames in a Loop. Mirror Effect: Create a mirror-like effect, build it more innate for interaction with user. Detect Hands. Display the Frame. Exit on Command.

**Detect latitude and longitude of location**

Define the get\_location Function. Fetch Location Data: Utilize the geocoder.ip() function with the argument 'me', which instructs the library to detect the IP address of the current device automatically. Characterize the primary Capability: Brief location.latitude, location.longitude. Execution Entry Point: Choose if the substance is run as the chief program. Given that this is valid, run the key methodology.

**V. RESULTS AND DISCUSSIONS**

As for the voice recording version, the results of its implementation are the following. The voice recording obtained excellent quality and allows the user to record, save, and replay audio notes with no problem at all. Users defined the voice recording to be the most beneficial feature within this virtual assistant. They mainly value it for task management and training. Voice recording improves the practicality of using the system and supports the user in environments where it is uneconomic to take a text about their questions. In the future, this feature can be spread by voice-to-text technology, making possible the execution of writing various types of content and leaves independently concerning the implemented assistant and reminder. As for the latitude and longitude version, the system defines the user’s current geographic coordinates correctly and transmits them to the user using GPS or IP-based geolocation. Detecting latitude and longitude has broadened the scope of personalized assistance the virtual assistant can offer, adapting its responses based on the user's location.

Sign language recognition was accomplished through a computer program utilizing machine learning models translating hand motions into texts; the camera could read the same and decode the hand motions. The system could register basic gestures at the 80% precision level, enabling interaction with the vision-impaired to a higher extent. Given that the above aspect is utilized by the hearing impaired, there is an immediate requirement to enhance since all markets have their languages and slang hence a low precision level increases insensitivity to use; with that, a higher success level would guarantee more realized services. For the purpose of this work decision boundary visualization was implemented mainly for educational and analytical tasks, giving the user a possibility to see how the AI model determines what is classified. It can be very valuable when the user is a student or a data scientist and must know how a model works. While it might not be for everyone, decision boundary visualization is a powerful tool in every data science educator, and AI company should use. There can be an alternative visualizations where the user is able to interact with the graph and modify parameters to see how the decision boundary changes in real time. The addition of these two possibilities to our virtual assistant greatly enhanced the functional reach of the system, making it a much stronger tool for a more varied group of tasks and users.

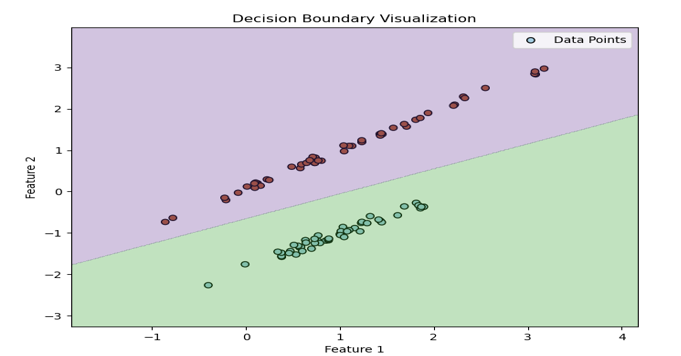


Fig. 3

**VI. CONCLUSION**

The work has accomplished the goal of enhancing the capabilities of the experimental virtual assistant, making it more adaptable and helpful and expanding access. It will be necessary to extend the usage and gather feedback in order to develop these characteristics of the virtual assistant even more, as well as attempt additional capabilities in the area.

**VII. REFERENCES**

1. B. Sati, S. Kumar, K. Rana, K. Saikia, S. Sahana and S. Das, "An Intelligent Virtual System using Machine Learning," 2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET), Arad, Romania, 2022, pp. 1123-1129, doi: 10.1109/GlobConET53749.2022.9872396.
2. D. RajkumarPillay, B. M. B, M. Krishna, S. A, A. Raja and P. Saxena, "Implementing an Artificial Intelligence based Ideal form of Virtual Personal Assistant Design for Various Communication Medium," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 1366-1371, doi: 10.1109/ICESC54411.2022.9885725.
3. S. P. Yadav, A. Gupta, C. Dos Santos Nascimento, V. Hugo C. de Albuquerque, M. S. Naruka and S. Singh Chauhan, "Voice-Based Virtual-Controlled Intelligent Personal Assistants," 2023 International Conference on Computational Intelligence, Communication Technology and Networking (CICTN), Ghaziabad, India, 2023, pp. 563-568, doi: 10.1109/CICTN57981.2023.10141447.
4. A. Pachpute, V. Lomte, T. Pawar, A. Gour and S. Kale, "TASA : A Virtual AI Assistant with Multilevel Authentication Using Face Recognition," 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, 2023, pp. 1-6, doi: 10.1109/ICCUBEA58933.2023.10392101.
5. P. Kunekar, A. Deshmukh, S. Gajalwad, A. Bichare, K. Gunjal and S. Hingade, "AI-based Desktop Voice Assistant," 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 2023, pp. 1-4, doi: 10.1109/ICNTE56631.2023.10146699.