JARVIS: Voice Assistant with Smart Home Automation

Rishabh S. Shah
Department of Artificial Intelligence and Data Science
SVKM's Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
rishabh.shah033@djsce.edu.in

Alan George
Department of Artificial Intelligence and Data Science
SVKM's Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
alan.george093@djsce.edu.in

Yukta V. Saraf
Department of Artificial Intelligence and Data Science
SVKM's Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
yukta.saraf028 @djsce.edu.in

Rujuta A. Jariwala
Department of Artificial Intelligence and Data Science
SVKM's Dwarkadas J. Sanghvi College of Engineering
Mumbai, India
rujuta.jariwala015 @djsce.edu.in

***Abstract*— This paper explores the integration of virtual assistant technology with home automation systems. The voice assistant JARVIS leverages natural language processing to understand user requests and perform tasks such as information retrieval from Wikipedia, open YouTube, and manage reminders and alarms. By simplifying daily routines and enhancing accessibility, JARVIS has the potential to revolutionize how we interact with technology in our homes. We use artificial intelligence (AI) and natural language processing (NLP) methods like Hidden Markov Models (HMM) and Mel-Frequency Cepstral Coefficients (MFCC) for speech recognition. Our setup involves using Raspberry Pi 3B+ to link with IoT devices and sending signals over Bluetooth and Wi Fi. We have prioritized the security of the user by ensuring that the data does not get stored in the cache memory. To maintain functionality when disconnected from the internet, the system will explore alternative communication methods for signals between system components. This research takes a thorough look at how we can blend virtual assistant tech into home automation systems to make living spaces intelligent.**

Keywords*— *Virtual Assistant, Smart Home Automation, Artificial Intelligence, Natural Language Processing, Raspberry Pi 3B+, Speech Recognition, Hidden Markov Model (HMM), Mel-Frequency Cepstral Coefficients (MFCC).

I. INTRODUCTION

Home automation is the technology that allows users to automate tasks and control the devices in their homes. This includes anything from remotely turning on lights, or adjusting the thermostat. The goal is to create a more convenient and efficient living environment for the user. Internet of Things (IoT) are the interconnected devices that collect and share data allowing researchers to make daily life easier through automation that streamlines daily tasks. Both of these concepts are integrated to form a Smart Home Automation system which makes it possible for users to control various appliances in their homes at once.

Virtual home assistants (VHAs) like Alexa and Google Home offer convenience for basic voice-based tasks like turning ON/OFF lights, operating smart plugs to control appliances, setting alarms, creating shopping lists etc. However, they often fall short and address only a few sets of challenges such as basic automation. This research is focused on the potential of virtual assistants and aims to transform them from simple assistants into intelligent tools that make our life more efficient and convenient.

Current virtual assistants excel at handling straightforward commands like adjusting lights or playing music. However, their limitations in natural language processing limit their ability to understand complex user requests. Specific requests require a deeper understanding of user context and intent, which is where this research comes in.

We propose an approach that uses Large Language Models (LLMs), which are powerful AI models pre-trained on huge amounts of text data, which gives them an exceptional grasp of the English language. By incorporating LLMs into our virtual assistants, we aim to enhance their NLP capabilities. This will allow them to understand complex commands and enable our virtual assistants to understand the true meaning behind our words.

We focus on how our virtual assistants can learn from our routines and adjust the environment based on our needs which enables them to anticipate our needs based on past behaviour such as dimming the lights as you settle in for bed, etc.

The virtual assistant will be able to tailor its responses based on the situation where real-time monitoring is necessary, by offering responses which are aware of the user's context, by analysing the tone of voice and language patterns.
Security and user privacy are also extremely valued in our increasingly connected world. This research prioritizes user data protection by ensuring voice commands are not stored on the device itself.
To validate our proposed system, we will construct a real-world smart home system using Raspberry Pi 3B+, which allows us to develop and test a personalized system that seamlessly integrates into daily routines.

By empowering our assistants with advanced NLP capabilities and prioritizing user data privacy, we pave the way for a future where these become intelligent companions, working in the background to make our lives easier and more comfortable.

II. RELATED WORKS

Recently there has been a lot of advancements in the fields Recently there has been a lot of advancements in the fields of Artificial Intelligence and Home Automation, leading to development in studies, helping researchers in carrying out tasks, making the ideal benchmark for real-life use case scenarios of home automation, which in the last few years were not comprehensible.

Some of the recent studies, as well as the related works of various researchers have set the benchmark/base for individual tasks, functions occurring in the realm of home automations.

One recent study[2],[7] explored a secure and intelligent home automation system using Internet of Things (IoT) devices. One area of focus is secure communication. This might involve using Bluetooth for short-range communication within the home, combined with an internet-based verification process to ensure only authorized devices can connect. Another key concept is a decentralized approach. This means devices can communicate and manage some functions directly with each other, without relying solely on a central hub. This can improve both reliability and security. Additionally, some studies explore using machine learning algorithms to categorize device states. For instance, the system might learn to identify when a light is on or off based on sensor data and then automate actions accordingly.

In another recent study[1],[6] , a prototype for a home automation system built around a Raspberry Pi minicomputer was presented. The proposed system used various sensors to collect environmental data and offers remote appliance control and automation functionalities. The key benefits of the automation system described in the paper were the remote monitoring and control which helped avoid situations where there would be a wastage of energy on forgotten appliances. The system also allowed scheduling the use of appliances where the system could be programmed to automatically switch ON/OFF appliances at particular times such as turning the lights off as soon as you left home etc. By using sensor data, the proposed system adjusted the appliance's usage of energy which optimized energy consumption

There were also some challenges regarding the proposed system which included the security of the web application used for controlling and monitoring the system, integration with real world appliances outside the LEDs used for testing purposes to check the real world efficiency of the proposed system and the compatibility with various brands and appliances and a user friendly UI/UX interface while also keeping in mind the overall cost of the system including the hardware and software needed to be weighed against the benefits of the system.

III. PROPOSED METHODOLOGY

In the proposed methodology, the user's audio is taken through a microphone connected to the breadboard. The audio processing is an essential part of the system architecture which enables seamless communication between the user and the smart home environment, simulating the experience of talking to an assistant and not just an AI/Robot.

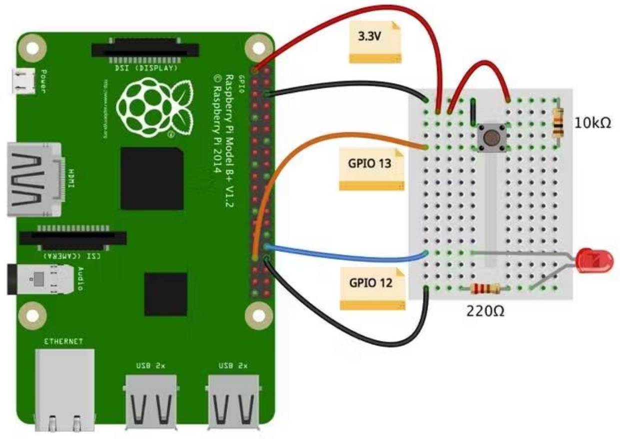

Fig 1. Circuit diagram for Raspberry Pi GPIO pin layout

The preprocessing step consists of removal of background noise, thus enhancing the accuracy of speech recognition. This enhances the quality of the audio by isolating and filtering out unwanted noise components present in the background while the user is having a conversation with the assistant.

In the phase of Automatic Speech Recognition, the captured audio signals are converted into text format.

The voice assistant is programmed to activate upon detecting the predefined wake work 'JARVIS' in the incoming audio stream. This trigger starts the interaction between the user and the assistant.

The text extracted from the Automatic Speech Recognition (ASR) component undergoes further analysis to extract meaning and context. Our system aims to decode the semantic meaning of the text output, discerning the user's intents and preferences.

Our system then generates and synthesizes a response based on the context of the conversation history. This is crucial to ensure a coherent and contextually appropriate response.

To provide additional functionalities and resources beyond its core capabilities, the system makes use of external services and APIs. Such integrations allow for extended features such as weather updates, integration with third-party smart devices and platforms.

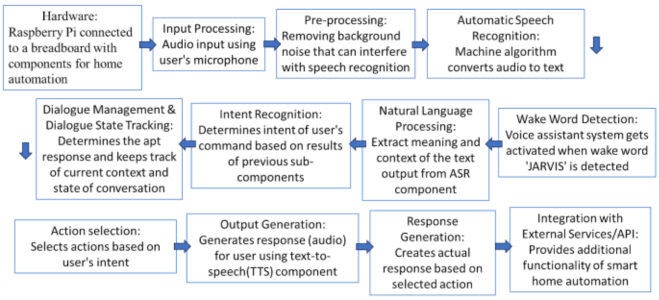

Fig 2. Proposed system architecture

IV. Results and Discussions

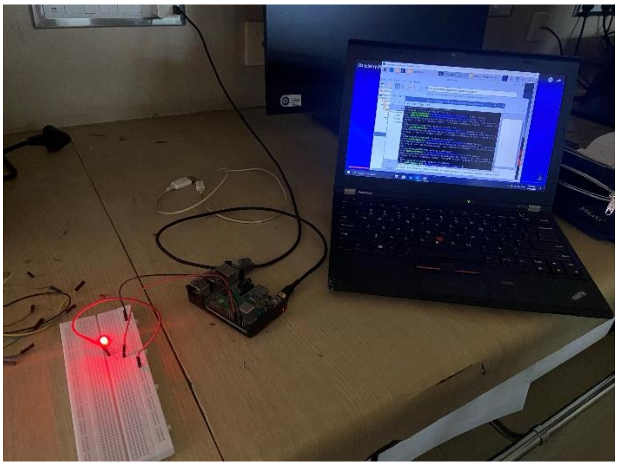

Fig 3.1

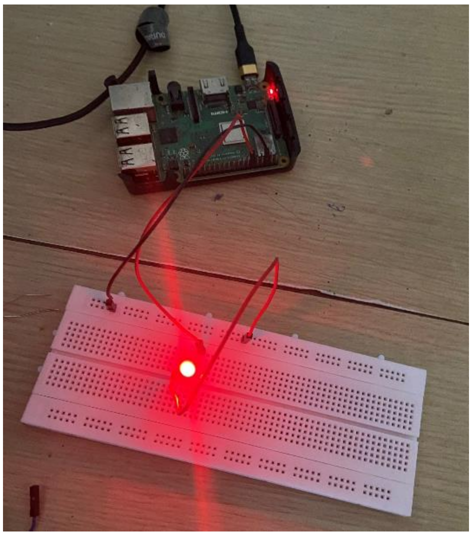

Fig 3.2

Fig.3: Switched ON/OFF LED light via internet

Pseudocode for MFCC:

```
function MFCC(audio_signal, sample_rate, window_size, hop_size, num_filters)
    # Pre-processing
    Preprocess(audio_signal)
    # Frame blocking
    frames = SplitSignalIntoFrames(audio_signal, window_size, hop_size)
    # Feature extraction for each frame
    mfcc_features = []
    for frame in frames:
      # Compute features
      mel_cepstral_coefficients = ExtractMFCCFeatures(frame, sample_rate, num_filters)
      mfcc_features.append(mel_cepstral_coefficients)
    return mfcc_features
# Helper functions (placeholders for specific functionalities)
function Preprocess(signal):
    # Implement pre-emphasis filtering here
function SplitSignalIntoFrames(signal, window_size, hop_size):
    # Split the signal into overlapping frames here
function ExtractMFCCFeatures(frame, sample_rate, num_filters):
    # This function combines power spectrum calculation, mel filterbank application,
    # logarithm, DCT, and selection of desired coefficients
    # Implement these steps here
```

The features performed by JARVIS are:

1. The voice assistant JARVIS recognizes and responds to natural language voice commands.
2. JARVIS performs standard voice assistant functions such as using search engines to access sites like Google, YouTube, and others
3. Smart lights/fans/other appliances can be turned ON/OFF via voice activation
4. Answer your doubts easily
5. The system has security features to ensure user privacy and data protection
6. Search on Wikipedia for you
7. Basic calculations can be performed
8. Setting up an Alarm clock and reminders
9. Allows exiting from the prompted programme
10. One can have a basic conversation with JARVIS
11. Only trusted users have access to the voice assistant.
12. It can provide recipes based on ingredients in your pantry
13. Provides personalized recommendations based on user preferences.

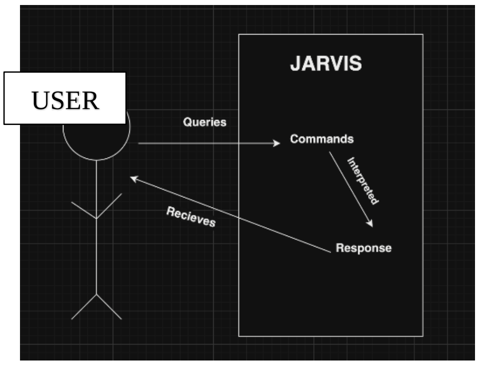

Fig. 3 Use case diagram

V. CONCLUSION AND FUTURE SCOPE

This paper explores the potential of voice assistants within the realm of smart homes and as a transformative technology shaping how we interact with the world around us. The voice assistant Jarvis offers a user-friendly way to adjust lights, fans, and appliances, creating a convenient and comfortable environment for the user.

However, creating a truly seamless experience requires overcoming certain problems including privacy concerns regarding data collection and security vulnerabilities. Technical limitations in compatibility across various smart home devices necessitate further development.

To build a smarter future, robust security protocols, anonymized user data, and user control over data collection are crucial for trust. Standardized protocols and open-source platforms could improve compatibility between smart home devices and voice assistants, while advancements in speech recognition algorithms will ensure understanding of diverse languages and accents. Offline functionality through local processing can further reduce reliance on a stable internet connection.

These advancements pave the way for evolutionary innovations beyond the smart home. Voice-controlled smart homes of the future could utilize AR overlays activated by voice commands, allowing for virtual furniture placement and real-time appliance repair instructions. Voice assistants could connect with smart city infrastructure, report issues. Reminding users to take medication, or even connecting them with healthcare professionals in case of emergencies. Finally, voice assistants can become powerful tools for increased productivity and accessibility for users with physical limitations, allowing them to manage calendars, schedule appointments, or dictate documents using just their voice. By addressing current limitations and embracing these future trends, voice assistants have the potential to transform our living spaces, revolutionize how we interact with technology, and enhance our lives in countless ways.

There could be an option of local inference where the voice assistant can process basic tasks and respond to user queries even without a stable internet connection. This could offer offline responses but it comes with a trade-off, the pre-trained model's knowledge and capabilities will be limited to the data it was trained on, potentially hindering its ability to respond to new information. Future research can explore techniques for enabling local inference.

JARVIS faces a latency issue i.e. time lag between the user asking the question and the response generated, which could be improved over time by faster processing and smarter algorithms.

VI. REFERENCES

[1] N. Valov and I. Valova, "Home Automation System with Raspberry Pi," 2020 7th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE), Ruse, Bulgaria, 2020, pp. 1-5, doi: 10.1109/EEAE49144.2020.9278998.

[2] R. B. A. Tonoy, Mahmudunnabi, H. M. Zilany and R. M. Rahman, "A Smart and Intelligent Home Automation System," 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2021, pp. 0446-0452, doi: 10.1109/UEMCON53757.2021.9666614.

[3] Y. Wang, C. Chen and W. Zhang, "Design of Smart Home Control System Based on Internet of Things," 2023 IEEE 6th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2023, pp. 1680-1683, doi: 10.1109/ITNEC56291.2023.10082440. keywords: {Wireless communication;Wireless sensor networks;TV;Microcontrollers;Web pages;Smart homes;Control systems;smart home;Internet of things;microcontroller;wireless communication}

[4] P. S. Nagendra Reddy, K. T. Kumar Reddy, P. A. Kumar Reddy, G. N. Kodanda Ramaiah and S. N. Kishor, "An IoT based home automation using android application," 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, India, 2016, pp. 285-290, doi: 10.1109/SCOPES.2016.7955836. keywords: {Pins;Androids;Humanoid robots;Temperature sensors;Home appliances;Rain;Arduino;IoT;Android;Wi-Fi;Home Automation},

[5] Iqbal, Salman, et al. "Internet-of-Things based Smart Home Automation System using Android Phone." Sir Syed University Research Journal of Engineering & Technology (SSURJET) 11.2 (2021).

[6] P. Dubey, H. Chourasia and S. Ghode, "A Novel Approach to Design Home Automation using IoT Applications," 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 2023, pp. 801-806, doi: 10.1109/InCACCT57535.2023.10141754. keywords: {Home appliances;Fans;Home automation;Wheelchairs;Operating systems;Lighting;Programming;sensor;microcontroller;Internet of Things (IoT);relay;blynk app;Raspberry pi etc},

[7] Tonoy, Rahat & Mahmudunnabi, & Zilany, Hasan & Rahman, Mohammad. (2021). A Smart and Intelligent Home Automation System. 0446-0452. 10.1109/UEMCON53757.2021.9666614.

[8] Y. Wang, S. Rawat and F. Metze, "Semi-automatic audio semantic concept discovery for multimedia retrieval," 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 2014, pp. 1375-1379, doi: 10.1109/ICASSP.2014.6853822. keywords: {Semantics;Acoustics;Vocabulary;Videos;Multimedia communication;Streaming media;Speech;audio semantic concept discovery;semiautomatic;multimedia retrieval},

[9] I. Feki, A. Ben Ammar and A. M. Alimi, "Query sound-by-example video retrieval framework," 2014 14th International Conference on Hybrid Intelligent Systems, Kuwait, Kuwait, 2014, pp. 297-302, doi: 10.1109/HIS.2014.7086165. keywords: {Speech;Support vector machines;Indexing;Music;Silicon;Hidden Markov models;Accuracy;audio concept;sound indexing;query sound;video retrieval},

[10] R. Chengalvarayan and Li Deng, "HMM-based speech recognition using state-dependent, discriminatively derived transforms on mel-warped DFT features," in IEEE Transactions on Speech and Audio Processing, vol. 5, no. 3, pp. 243-256, May 1997, doi: 10.1109/89.568731. keywords: {Hidden Markov models;Speech recognition;Discrete Fourier transforms;Feature extraction;Psychoacoustic models;Discrete cosine transforms;Cepstral analysis;Data mining;Speech processing;Filter bank}