

F.R.I.D.A.Y – Voice Assistant

Bhavya Jain¹, Arpan Saini², Mr. Kuwar Pratap Singh³

^{1,2}Department of Computer Science & Engineering, SRM Institute of Science and Technology, Ghaziabad

³Assistant Professor, Department of Computer Science & Engineering, SRM Institute of Science and Technology, Ghaziabad

Abstract : A voice assistant is a software that interprets human voice and responds via speech synthesis to come up with a utility through a specific application. Some of the most famous voice assistants are Amazon's Alexa, Apple's Siri, Google's Assistant and Microsoft's Cortana, they are implanted in smart devices. Users can get the answers to their queries by asking their questions to the assistant. Voice assistant also help us to control home automation devices, and manage other basic tasks such as call, message, music apps and many more things. Using this assistant, users would be able to control our YouTube, could solve their arithmetic questions, and also could be able to search any information on the internet. Voice assistants ease our daily life as it does different functions fastly and also reduces time and make searching more smoother. It can also help visually impaired, physically challenged people and illiterate people as they can speak up their questions directly instead of typing.

Key Words: F.R.I.D.A.Y, Voice Assistant System, LBPH, Speech Recognition, Web Scrapping, Face Recognition and Identification.

1. INTRODUCTION

In today's world automation has become a part of our life in performing simple tasks like switching on/off the appliances to the complex one's like self-driving cars. Previously, human voices were used only for communication purpose but now it can be used to interact with various technical devices for performing various day to day activities to the complex one's. One of such devices is Voice Assistant like Apple's Siri, Microsoft's Cortana, Amazon's Alexa, Google's Assistant which are used to do different tasks like set a timer, email to a specific person, making calls, messaging, fetching weather information, browsing on the web, playing songs on YouTube, Spotify etc. This digital Voice search assistant could be destiny for the subsequent generation, as people have turn out to be reliant on voice assistants for all of their needs and daily-habitual activities.

In this project, we have created a voice assistant named F.R.I.D.A.Y (Face Recognition and Identification with Digital Assistant of Yours). The main reason for creating this project is to cover up the security issues in other available voice assistants by using the face recognition and identification system due to which only the authorized person could use the assistant and it makes the assistant very personalized.

This project works by taking input in the form of voice or text and if input is in the voice format, then it is converted into text using Speech Recognition. This text is used to perform different actions like searching on browsers, playing videos on YouTube, translating text into different languages, performing system activities like creating and editing files of various types like word file or power point file.

An extraordinary aspect of speech reputation is to facilitate for humans with purposeful incapacity or different types of handicap. To make their day-by-day chores easier, voice manipulate will be helpful. With their voice they might perform the pc or laptop gadgets shutdown or perform a few different packages of pc. This ends in the dialogue approximately intelligent laptop wherein those operations may be made to be had for the not unusual place guy in addition to for handicapped

2. LITERATURE SURVEY

Speech recognition has an extended record with numerous waves of predominant innovations. Speech popularity for dictation, search, and voice instructions has emerge as a preferred characteristic on smartphones and wearable devices. Design of a compact large vocabulary speech popularity gadget which can run efficiently on cellular devices, as it should be and with low latency.

Voice Assistants were the issue of a number of recent studies. We accrued some of papers so as to have a look at voice assistants after which analyzed the data.

Ravivanshi Kumar Sangpal [1] et al. got here with up a exam on JARVIS which mixes with the electricity of Artificial Intelligence Markup Language with the industry main Google platform for text-to-speech with male pitch. The use of Artificial Intelligent markup language [AIML] makes Jarvis bendy illustrating its extensive reusability and coffee maintenance. They additionally talked about future scope of JARVIS in Artificial Intelligence.

Emad S. Othman [2] implemented Voice Command System as a private voice assistant that may manage numerous responsibilities for cease customers the use of Raspberry Pi microcontroller because the main hardware. He mentioned the framework configuration of the Voice Command System. Modules with open supply systems permits the device to paintings at its best in phrases of space-time complexity.

Yash Mittal et al. [3] proposed a examine on Smart Home Automation System. This gadget may be tailored to a user’s voice and interpret the voice instructions that allows you to manage their domestic home equipment and gadgets for diverse functionalities and their applications. The Voice popularity is processed the use of an Arduino microcontroller for instructions processing and controlling the home equipment. Through their automation gadget prototype, they figured out, it may be used to turn current houses into clever houses at a minimum fee and with ease.

Ankit Pandey [4] et al. supplied smart voice assistant which can cope with calendar meetings, emails exchanges, and function to-do lists notes. The intention turned into to create a smart private assistant with Python that lets in customers to monitor gadgets with the voice commands, gathering information, and execute suited obligations on one’s laptop system.

Subhash S [5] et al. proposed AI based voice assistant that can generally utilized in clever telephones and laptops. This AI based digital assistant will accumulate audio enter from the system’s microphone and translate into textual content, which will be transmitted thru pyttsx3 (Google textual content to speech) to cease user. The author claimed that their proposed work may be useful in education field, day by day lifestyles application, domestic home equipment etc.

3. THE PROPOSED MODEL

The Facial Recognition to the remote helper gives more noteworthy security to the framework by the Haar Cascade and LBPH calculation. The face is perceived by the camera and it is taken care of through OpenCV in which it distinguishes the countenances and articles put before the camera and the photos taken is stood out from the pictures in the in the framework by Haar Cascade Algorithm. After that either audio or text is taken from the user as the command to voice assistant and it is then converted into string, to perform this action we have created two functions each for taking input via audio or text. For audio we have named a function recording which Microphone() and Recognizer() are used to receive and recognize the audio provided by user which are methods of speech recognition module. After that, it identifies the specified keyword in the command which the user give to it and according to which, it either calls some method or perform some action. After that, the user receives the output in the form of audio from the voice assistant, to perform this action we have created a method named speak() in which the object of pyttsx3 module calls the method say() to give an audio as the output. It also calls the method runAndWait() to wait for the user to give next command

FACE RECOGNITION:

Create a database in which all the training images are stored and every one of the pictures are caught by the camera are put away in information base. The face acknowledgment is

done by utilizing two algorithm i.e. Haar Cascade Algorithm and LBPH Algorithm. In the Haar Cascade calculation depends on Course classifiers which comprises of haar highlights. The course classifiers are the connection of a bunch of powerless classifiers used to make a solid classifier. These classifiers recognize the human face reliant upon the hugest features like eyes, eyebrows and lips. Haar highlights are identified dependent on the solid calculation, in which we apportion a pixel force to each and every pixel identified with grayscale values inside the extent of 0 to 255 where 0 talks to the white concealing and 255 addresses the dark shading.

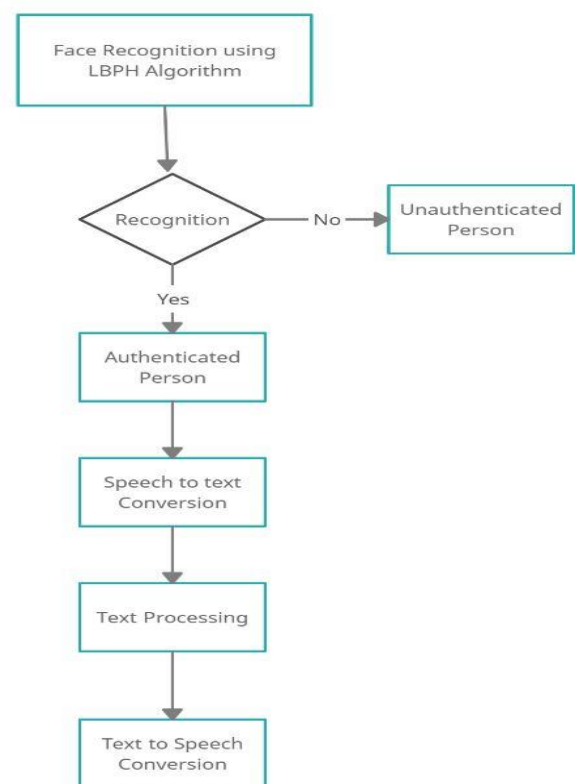


Fig 1: Block Diagram of the Proposed Model

In this first camera catches the photograph of the client and afterward it prepares the models for the pictures put away in the information base with separate to the highlights of the individual. In the identification segment it distinguishes the face dependent on the Haar XML record which has a few pre-characterized highlights of face. In the wake of recognizing the face, it will begin separating the highlights of the face. By considering the highlights the pictures are ordered in like manner. These grouped pictures are contrasted and the prepared models. On the off chance that it matches the client can get to the virtual help. Camera catches the photograph of the client and afterward it prepares the models for the pictures put away in the information base with separate to the highlights of the individual. In the identification segment it distinguishes the face dependent on the Haar XML record which has a few pre-characterized highlights of face. In the

wake of recognizing the face, it will begin separating the highlights of the face. By considering the highlights the pictures are ordered in like manner. These grouped pictures are contrasted and the prepared models. On the off chance that it matches the client can get to the virtual help.

FACE IDENTIFICATION:

Face Identification using LBPH (Local Binary Patterns Histogram) algorithm. Lightening conditions does not affect this algorithm and in this there is a 3x3 window which is to be moved across an image. In each move the pixel present at the centre of the image is compared with the surrounded pixels. If the value of the intensity of the light is lower or equal to the intensity of light at central pixel then the surrounding pixel is given 1 as a value else the surrounding pixel is given 0. Then, 0/1 values are to be read in a clockwise order under the 3x3 window. This will generate a binary pattern which is local to a particular area of the picture. In the end, there will be one histogram for every face which is available in the training dataset. This means, if 100 images are available in the training dataset then the LBPH algorithm will make 100 histograms after the training of the model and will use these for later recognition of the face. Now when a new image is feed to the recognizer then the recognizer will generate a histogram for that new image using the above method. Now this histogram is to be compared with all the histograms which are available after the training of the model. Now finally, if the algorithm finds the best match of the histogram then it returns the person label associated with that best match.

Voice Assistant	Voice Response	Security	Face Recognition	Online Dependency
Siri	YES	NO	NO	YES
Alexa	YES	NO	NO	YES
Cortana	YES	NO	NO	YES
Google	YES	NO	NO	YES
Our Voice Assistant	YES	YES	YES	YES

Fig 2: Comparison of our voice assistant with others available in the market.

4. IMPLEMENTATION

1. The face recognition and identification is the first step in this project and it is done using Haar Cascade classifier and LBPH algorithm.
2. If the user is successfully recognized by the voice assistant then it allows the person to enter and to provide commands to the voice assistant else it does not allow the person to access the voice assistant.
3. Now the commands given by the authenticated person either through text or speech is converted into the string using speech to text conversion which is a module in Python.
4. Now in this string some specific keywords are to be checked and if that keyword is found in the Python script then an action specific to that keyword would be performed.

5. Finally, the result is again converted back to the speech format from the text format using pyttsx3 module present in the Python and is returned back to the user.

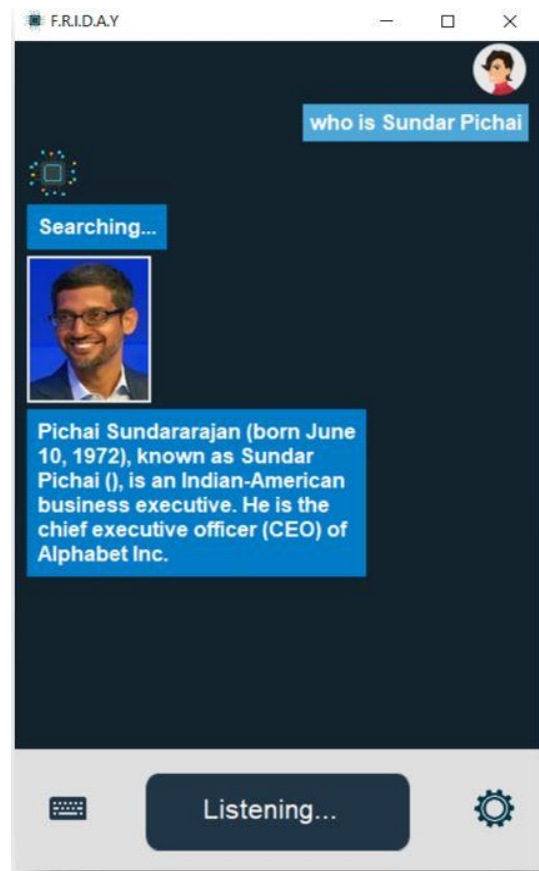


Fig 3: Web search result for Sundar Pichai by F.R.I.D.A.Y

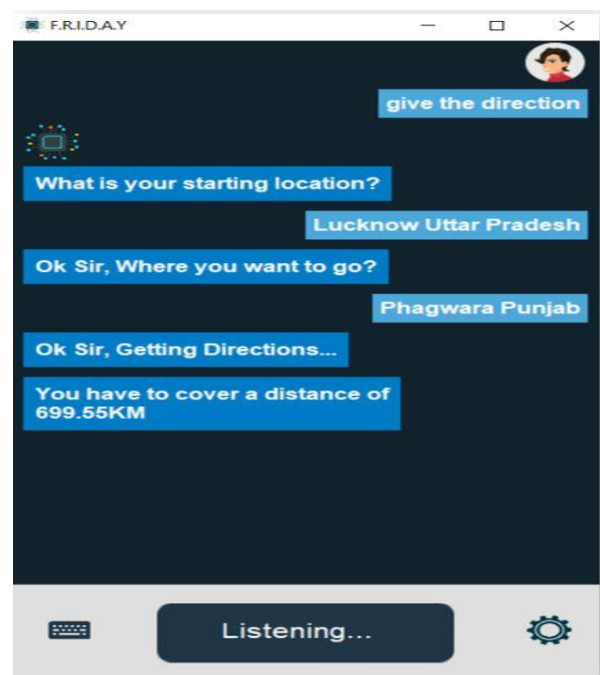


Fig 4: Distance estimated by F.R.I.D.A.Y using Google Maps

5. CONCLUSION

F.R.I.D.A.Y. can make day to day tasks easier to perform. We can send emails, search queries in Google, search for videos on YouTube, retrieving images, live weather conditions, word meanings, and telling the user about the scheduled events and tasks. It helps to make a regular user, a power user. The project is made in such a way that it can tackle every user query efficiently with security by face recognition. By working on the project we broadened our horizon of knowledge. This project is just the first step towards learning advanced technologies.

6. REFERENCES

- [1] Sangpal, R., Gawand, T., Vaykar, S., & Madhavi, N. (2019, July). JARVIS: An interpretation of AIML with integration of gTTS and Python. In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT) (Vol. 1, pp. 486-489). IEEE.
- [2] Othman, E. S. (2017). Voice Controlled Personal Assistant Using Raspberry Pi. *International Journal of Scientific & Engineering Research*, 8(11), 1611-1615.
- [3] Mittal, Y., Toshniwal, P., Sharma, S., Singhal, D., Gupta, R., & Mittal, V. K. (2015, December). A voice controlled multifunctional smart home automation system. In 2015 Annual IEEE India Conference (INDICON) (pp. 1-6). IEEE.
- [4] Pandey, A., Vashist, V., Tiwari, P., Sikka, S., & Makkar, P. Smart Voice Based Virtual Personal Assistants with Artificial Intelligence.
- [5] Subhash, S., Srivatsa, P. N., Siddesh, S., Ullas, A., & Santhosh, B. (2020, July). Artificial Intelligence-based Voice Assistant. In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4) (pp. 593-596). IEEE.
- [6] <https://www.javatpoint.com/how-to-convert-text-to-speech-in-python>