IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Jarvis Voice Recognition Using ML

[1] Mohammad Ateeq, [2] Vineet K Joshi, [3]D Naga Praneeth, [4]Dr. K. C. Sreedhar

[4]Associate Professor, Department of Computer Science and Engineering, SNIST, Hyderabad-501301, India

[1,2,3]B. Tech Scholars, Department of Computer Science and Engineering, SNIST, Hyderabad-501301, India

Abstract:- Technology field has made a lot about progress recently, & it keeps becoming better. We could only perform specific jobs in beginning because there were only computer systems available, but new technologies have since been created, including machine learning, artificial intelligence, deep learning, & others. In recent years, artificial intelligence (AI) has advanced significantly, & its capabilities are constantly growing. Natural language processing (NLP) is an area where artificial intelligence is used. Natural language processing (NLP) enables users to converse in their native tongue with voice assistants & other computer systems. To address challenges people face while interacting with machines, a variety about voice assistants have been created & are continuously being enhanced for better performance. Various modules, including OS, Pysttx3, speech recognition, date/time, Wikipedia, smtplib, & many others, will be used. Python is being used to create a voice assistant that enables users to complete a variety about tasks without using keyboard. This article's goal is to study intelligent behaviour about voice assistants & how they might be applied to daily tasks, learning, & other objectives.

I.INTRODUCTION

Virtual reality, augmented reality, voice interaction, & Internet of Things (IoT) are examples about future-oriented technologies that are reshaping digital experiences & altering how people interact with world. One critical jump in human-machine communication made conceivable by advancement about man-made brainpower is voice control. Artificial intelligence, machine learning, & neural networks, among other technologies, In modern era, we can teach our machines to do their jobs on their own or to think like humans. Virtual assistants allow us to communicate with our robots as well. Voice assistants like Apple's Siri, Google's Assistant, Microsoft's Cortana, & Amazon's Alexa have experienced a meteoric rise in popularity recently as a result about widespread use about smartphones. A variety about services are offered by voice assistants, which make use about technologies like natural language processing (NLP), speech synthesis, & voice recognition to enable users to carry out tasks on a computer by simply speaking commands into device. Additionally, using a voice assistant lets you complete a task without having to repeatedly type commands. Virtual assistants are very helpful for older people, people with special needs or impairments, & young children because they make it easier for them to interact with machines & give them ability to multitask.

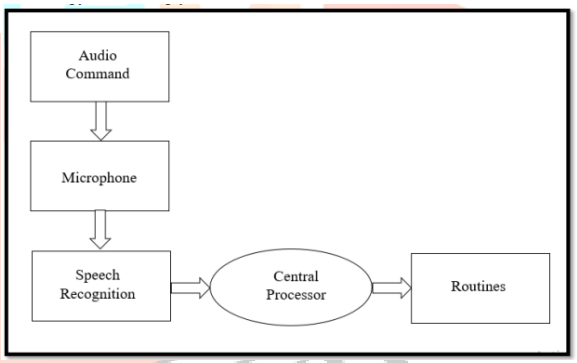

Fig.1: Example figure

A virtual assistant is used to operate devices like laptops & PCs at your discretion. A virtual assistant is an application programme that can carry out activities for users by understanding natural language & voice commands. A virtual assistant is used to carry out common tasks like showing date & time, managing emails, open applications, etc. Virtual assistants are now highly helpful to people. ability to run PCs or laptops using solely voice commands simplifies life for people. It takes less time to use a virtual assistant. By using a virtual assistant, we can save time & help with other projects. majority about time, virtual assistants are cloud-based programmes that need internet-connected gadgets. freedom to only hire a virtual assistant for services they require. Start with fundamentals about Python to create a virtual assistant for your computer. Virtual assistants are focused on their tasks. capacity about virtual assistants to comprehend & carry out instructions. Virtual assistants are pieces about software that carry out tasks given by clients & comprehend spoken & written instructions. Virtual helpers can understand human speech & reply with synthesised voices. There are a number about voice assistants available on market, including Alexa, a smart speaker made using a Raspberry Pi, Siri for Apple TV

remote, Google Assistant for smartphones, & Cortona for Windows 10.

II.LITERATURE SURVEY

A number about academic studies & commercial businesses have investigated application about machine learning for voice recognition in personal assistants like Jarvis. Examples about similar work are shown below:

1. Google Assistant: For speech recognition, Google's voice assistant uses machine learning methods including DNNs & RNNs. It is connected with a variety about Google services & can recognise speech in more than 100 different languages.

2. Amazon Alexa: Using machine learning algorithms, Amazon's voice assistant can understand natural language orders & carry out user requests. It has ability to play music, control smart home appliances, & deliver information on a variety about subjects.

3. Microsoft Cortana: Using machine learning techniques, Microsoft's voice assistant can recognise natural language instructions & execute them, such as sending emails & setting reminders. It is connected with other Microsoft services as well as Windows operating system.

4. Apple Siri: Apple's voice assistant employs machine learning algorithms to recognise human requests & commands in natural language. It is integrated with iOS operating system & has ability to play music, control smart home appliances, & deliver information on a variety about subjects.

The promise about machine learning for voice recognition in personal assistants like Jarvis is shown overall by these instances. Additionally, they emphasise significance about employing big & varied datasets to train machine learning models as well as necessity about ongoing improvement through user feedback & system review.

III.METHODOLOGY

The Jarvis voice assistant system doesn't specify precise machine learning models or methods that it uses for voice detection. However, there are some standard methods & procedures that are frequently employed in speech recognition systems & might be applied to Jarvis:

1. Hidden Markov Models (HMMs): Speech recognition systems frequently employ HMMs, statistical models. They use a collection about hidden states that can be seen through voice signal to describe probability distribution about speech sounds.

2. Deep Neural Networks (DNNs): A subset about machine learning method, DNNs are particularly effective in speech recognition tasks. It is possible to train them to spot patterns in frequency domain about voice sounds & map these patterns to particular phrases. Convolutional Neural Networks (CNNs) are yet another class about machine learning technique that can be applied to speech detection. They work particularly effectively for examining speech signal temporal patterns.

4. Recurrent Neural Networks (RNNs): RNNs are a special class about neural network that excel at processing sequential input, including speech signals. They can be applied to model temporal dependencies between various speech signal components.

These are just a few about methods that a machine learning-based voice recognition system like Jarvis can conceivably employ. It is important to keep in mind that precise implementation details & methods employed in Jarvis are probably proprietary & not available to general public.

PROPOSED WORK:

Machine learning is being used to recognise Jarvis' voice. For such a system, following are some possible steps:

1. Data Gathering: Gather a sizable collection about audio recordings about speech in a range about dialects, accents, & settings. This dataset ought to be varied & reflective about intended audience.

2. Data preprocessing: Before voice recognition, audio recordings should be cleaned up to remove noise & other distortions. Techniques including filtering, normalisation, & feature extraction may be used in this step.

3. Preparing Training Data: Segmenting audio recordings into single words or phrases & labelling them with their associated text transcriptions can help you prepare training data. Using automated voice recognition software or manual transcription are both options for this step.

4. Model Construction: Create a machine learning model that maps acoustic characteristics about speech to text transcriptions utilising strategies like HMMs, DNNs, CNNs, or RNNs. Using labelled training data created in step 3, this model can be trained.

5. Model Assessment: Utilise a distinct dataset about audio recordings & related text transcriptions to assess model's performance. Measuring metrics like word error rate (WER), sentence error rate (SER), & recognition accuracy can be a part about this stage.

6. Model Deployment: To give consumers real-time voice recognition, deploy trained model in Jarvis voice recognition system. This step may entail accelerating & streamlining model for target hardware platform.

7. Continuous Improvement: Track system performance over time & gather user feedback to pinpoint areas that could use improvement. training set about data can be updated using this feedback, & machine learning model

can be improved over time.

Of course, particulars about a suggested machine learning system for Jarvis voice recognition would rely on a number about criteria, including intended user base, desired level about accuracy, & available hardware & software resources.

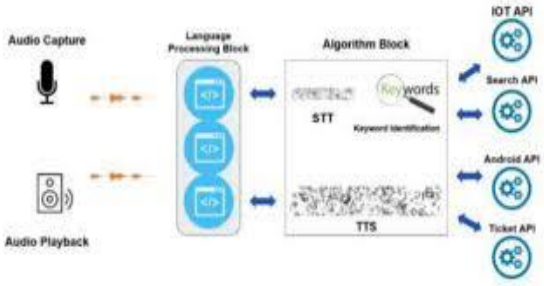

Fig.2: System architecture

The user's spoken input is transmitted to system via an audio input device, such as a headset or a microphone. Preprocessing about audio input entails filtering, noise reduction, & segmentation into small frames. From preprocessed audio frames, system extracts pertinent features including Mel Frequency Cepstral Coefficients (MFCCs), which are frequently employed in speech recognition. A machine learning model called Language Model, which calculates likelihood about a string about words, is applied to data from Acoustic Model. Language Model predicts subsequent word in sequence using Natural Language Processing (NLP) methods like n-grams or recurrent neural networks (RNNs). An intent recognition component analyses Language Model's output to ascertain purpose about user's command. Keyword detection approaches or intent categorization using machine learning models like Support Vector Machines (SVMs) or Decision Trees may be used for intent recognition. Based on outcomes about intent recognition, Dialogue Management component chooses appropriate response to user's demand. To improve system's response over time, this component may employ a rule-based system or machine learning methods like reinforcement learning. A Text-to-Speech (TTS) system receives response from Dialogue Management component & produces an audio output about system's response. An audio output device, such as a speaker or a headset, plays audio output back to user.

IV.IMPLEMENTATION

A. Speech Recognition

The system makes use about Google's online speech recognition system to convert speech input into text. contribution about discourse data community's PC network server has a specific corpus about texts that clients can get to. These messages are taken from amplifier & briefly kept in framework prior to being sent to Google cloud for discourse acknowledgment. corresponding text that was just received is then fed to central processor.

B. Python Backend

The voice acknowledgment module's result is parsed by Python backend to learn whether an order or discourse yield is a Programming interface Call, Setting Extraction, or Framework Call. To give client ideal outcomes, result is accordingly sent back to Python backend.

C. API Calls

This term is abbreviated as API, or Application Programming Interface. communication between two apps is made possible by a software interface known as an API. To put it another way, an API is intermediary that sends your request to provider & returns result.

D. Context Extraction

The process about extracting structured data from unstructured or semi-structured machine-readable resources is known as context extraction (CE). Natural language processing, or NLP, is typically used in this activity to look at documents written in human languages. Test results for context extraction are provided by recent advancements in multimedia document processing, such as automatic annotation & content extraction from photos, audio, & video.

E. System Calls

When a piece about software asks kernel about operating system on which it is running for a service, this is called a system call. creation & execution about new cycles, equipment related administrations (such getting to a hard drive), & collaboration with fundamental piece capabilities like interaction booking are instances about this. Framework calls give an interaction's admittance to working framework.

F. Text-To-Speech

Text-to-discourse (TTS) alludes to capacity about PCs to peruse text out loud. A phonemic representation about written text is converted into waveforms that can be used by a TTS Engine to produce sound. Third parties offer TTS engines in a wide range about languages, dialects, & specialized vocabulary.

V.RESULTS

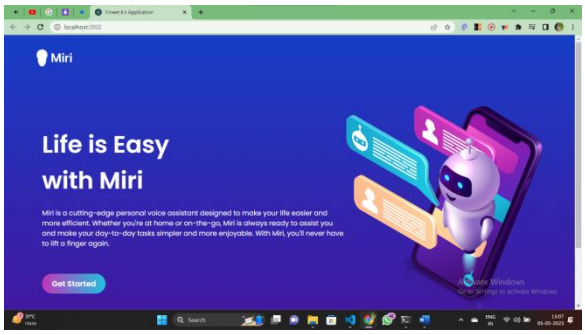

Fig.3: Cover page

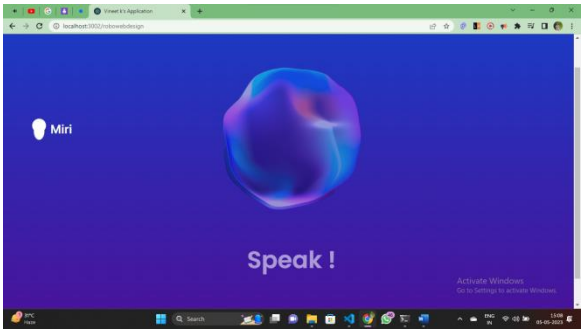

Fig.4: Here we can speak to our application

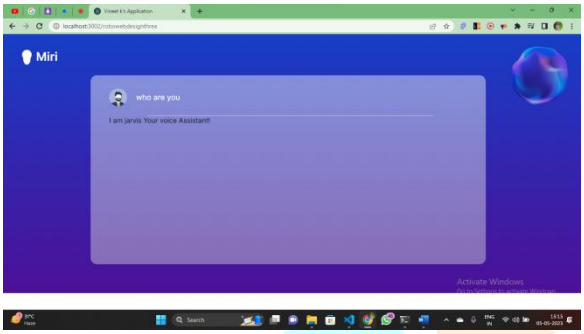

Fig.5: Here we can ask

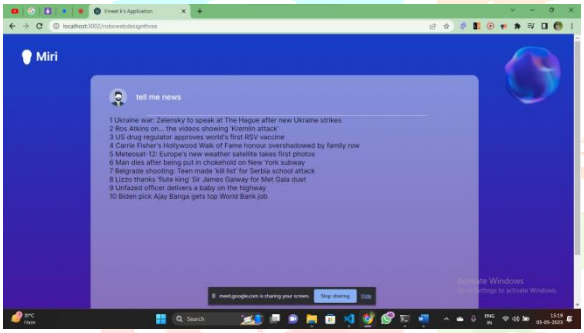

Fig.6: Result

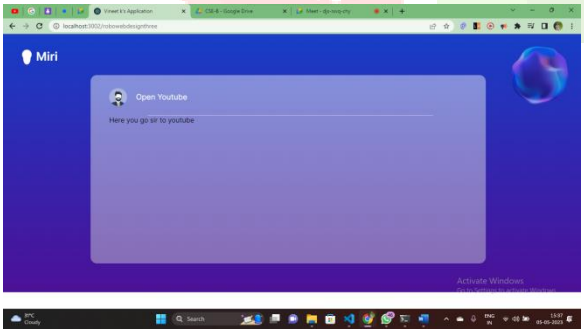

Fig.7: Here we can open Youtube

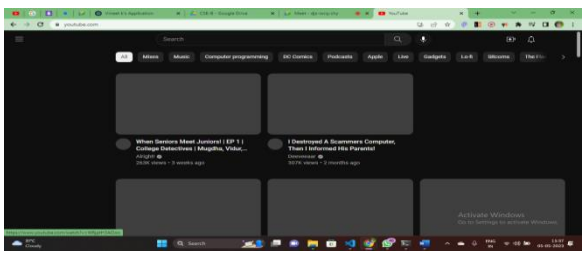

Fig.8: Visual about youtube

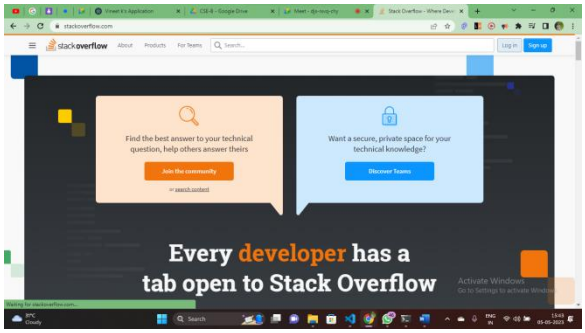

Fig.9: Home page about stack overflow

VI.CONCLUSION

In conclusion, using Python to create a voice assistant like Jarvis may be a pleasant experience that presents many chances for growth & learning. A reliable & knowledgeable voice assistant that can carry out a variety about activities & offer users helpful guidance can be built with aid about numerous Python tools & frameworks. A Jarvis voice assistant can comprehend human orders, offer information, & even handle difficult tasks with ease because to Natural Language Processing (NLP), speech recognition, & machine learning. functionality & capabilities about voice assistant can also be improved by integrating a number about APIs & services. Although creating a voice assistant like Jarvis may need considerable technological know-how, finished product can be a useful tool for both individuals & organisations. possibilities for voice assistants like Jarvis are only limited by our imagination as artificial intelligence develops.

VII.FUTURE SCOPE

The Python-based Jarvis voice assistant has a very promising future because there are many methods to develop & advance technology. following are some possible areas for growth: 1. Improved Natural Language Processing: A voice assistant's capacity to comprehend & respond to natural language commands is one about its most crucial features. Future developments in NLP will make voice assistants like Jarvis more intuitive & user-friendly by enabling them to comprehend subtleties about human speech. 2. Integration with Additional APIs & Services: Jarvis can now integrate with a variety about APIs & services to offer users useful data & support. However, when more APIs & services become accessible, Jarvis will have a greater chance about offering users even more benefit. 3. Personalization: Voice assistants like Jarvis will be able to provide users increasingly more customised experiences as AI technology develops. This can entail figuring out consumer preferences, predicting needs, or even coming up with a distinct personality. 4. Multi-Lingual Support: With internet & business both becoming more globally oriented, there is an increased demand for voice assistants that can converse in a variety about languages. Future versions about Jarvis might be built to support a variety about languages, making it a useful resource for people all across world. Overall, Python-based Jarvis voice assistant has a promising future

& offers limitless room for development. We may anticipate even more sophisticated voice assistants as technology develops, which will change how we interact with technology & outside world.

VIII.REFERENCES

[1] Elizabeth Sucupira Furtado, Virgilio Almedia & Vasco Furtado, "Personal Digital Assistants: Need about Governance"

[2] Zecheng Zhan, Virgilio Almedia, & Meina Song, "Table-To-Dialog: Building Dialog Assistants To Chat With People On Behalf about You"

[3] Yusuf Ugurlu, Murat Karabulut, Islam Mayda "A Smart Virtual Assistant Answering Questions About COVID-19" Mathangi Sri "NLP In Virtual Assistants"

[4] Anxo Pérez, Paula Lopez-Otero, Javier Parapar. "Designing An Open-Source Virtual Assistant"

[5] C K Gomathy & V Geetha. Article: A Real Time Analysis about Service Based Using Mobile Phone Controlled Vehicle Using DTMF For Accident Prevention. International Journal about Computer Applications 138(2):11-13, March 2016. Published By Foundation about Computer Science (FCS), NY, USA,ISSN No: 0975-8887

[6] Dr.C.K.Gomathy , K. Bindhu Sravya , P. Swetha , S.Chandrika Article: A Location Based Value Prediction For Quality about Web Service, Published By International Journal about Advanced Engineering Research & Science (IJAERS), Vol-3, Issue-4 , April- 2016] ISSN: 2349-6495

[7] C.K.Gomathy & Dr.S.Rajalakshmi.(2014),"A Business Intelligence Network Design For Service Oriented Architecture", International Journal about Engineering Trends & Technology (IJETT) ,Volume IX, Issue III, March 2014, P.No:151-154, ISSN:2231-5381.

[8] "VIRTUAL PERSONAL ASSISTANT (VPA) FOR MOBILE USERS"

[9] D. SOMESHWAR, DHARMIK BHANUSHALI, SWATI NADKARNI, "IMPLEMENTATION about VIRTUAL ASSISTANT WITH SIGN LANGUAGE USING DEEP LEARNING & TENSORFLOW"

[10] C.K.Gomathy.(2010),"Cloud Computing: Business Management For Effective Service Oriented Architecture" International Journal about Power Control Signal & Computation (IJPCSC), Volume 1, Issue IV, Oct - Dec 2010, P.No:22-27, ISSN: 0976-268X.

[11] Dr.C K Gomathy, Article: A Semantic Quality about Web Service Information Retrieval Techniques Using Bin Rank, International Journal about Scientific Research In Computer Science Engineering & Information Technology (IJSRCSEIT) Volume 3 | Issue 1 | ISSN : 2456-3307, P.No:1563-1578, February-2018

[12] Dr.C K Gomathy, Article: A Scheme about ADHOC Communication Using Mobile Device Networks, International Journal about Emerging Technologies & Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:320-326, Nov-2018.

[13] Dr.C K Gomathy, Article: Supply Chain-Impact about Importance & Technology In Software Release Management, International Journal about Scientific Research In Computer Science Engineering & Information Technology (IJSRCSEIT) Volume 3 | Issue 6 | ISSN : 2456-3307, P.No:1-4, July-2018

[14] Hemalatha. C.Kand N. Ahmed Nisar (2011)., Explored Teachers' Commitment In Self Financing Engineering Colleges, International Journal about Enterprise Innovation Management Studies (IJEIMS), Vol2. No2. July-Dec 2011 ISSN: 0976-2698 Retrieved From Www.Ijcns.Com

[15] Dr.C K Gomathy, Article: Efficient Automatic Water Control Level Management Using Ultrasonic Sensor, International Journal about Computer Applications (0975 – 8887) Volume 176 – No. 39, July 2020.