PERSONALIZED INTELLIGENT VOICE ASSISTANT ANNA FOR SPEECH RECOGNITIONUSING NATURAL LANGUAGE PROCESSING

N.Sreevidya, Sushma Sai Rekulapally, Prathani Akshitha, Bhavani Medi

Assistant Professor, Department of IT, Sreenidhi Institute of Science & Technology (A), Hyderabad, INDIA

IV B.Tech Students, Dept. of IT, Sreenidhi Institute of Science & Technology (A), Hyderabad

Abstract: As the use of artificial intelligence increases, this technology is more advanced than ever. Artificial intelligence is integrated into the components of life. What was unthinkable 10 years ago is now a casual act, with fluent conversations with computers. This was made possible by concepts of machine learning and speech recognition. That is, you can only run and complete user-designed quests. Plus, get started with audiobooks, SIM tracking, and Google Meet right away. Even if we combine voice technology with human-like conversations, There is no need to talk to our devices and have a conversation to complete our task of choice. The real focus should be on unspoken commands in completing the user's task as quickly as possible. Voice UI is getting popular as it makes the work easy and when it is integrated with a device it will become a good user product. This is why we developed a personalized user-designated model that fulfills all the features and provides a good experience.

Keywords: Artificial intelligence, speaking, proceeding, LANGUAGE

INTRODUCTION

Artificial intelligence is becoming more advanced than ever as more people utilize it. Artificial intelligence has become ingrained in the fabric of life. What seemed unimaginable ten years ago is now a commonplace occurrence, with machines conversing fluently. Machine learning and speech recognition principles make this possible. Before it can determine your request, your assistance must wait for you to complete speaking. Then it switches to a spoken response. After that, you must wait for the assistant to finish speaking before proceeding to the next stage. When the voice assistant makes a mistake or error in understanding the commands or the task assigned by the user we will have to wait until the voice assistant performs the misunderstood task completely before correcting it or interrupting it while the voice assistant is fulfilling the wrong task which results in terrible model.

[1] IVAs are supported by specific foundations that evolved swiftly and have recently become the focus of extensive research. Whatever the case may be, the research focused on understanding the users' experiences with IVAs is rather limited. There have been ideas for the selection and exploitation of this from our point of view IVAs.

For example, despite their widespread and widely advanced thinking mobile phones Individuals are more likely to use IVAs on a sporadic basis or not at all. [2]According to the study, 98 percent of iPhone users had used Siri before. However, just 30% of people used it. 70% of people use it regularly, with the remaining 30% using it only sometimes. Instead, you usually end up re-entering your original command. When you look at smart speakers and smartphones, it's not surprising that I'll bite with a tiny change in how you phrase that command. User-directed requests are the most common use cases for voice assistants. Users of smart speakers are asked to play music, get the weather forecast, or set a timer or an alarm.
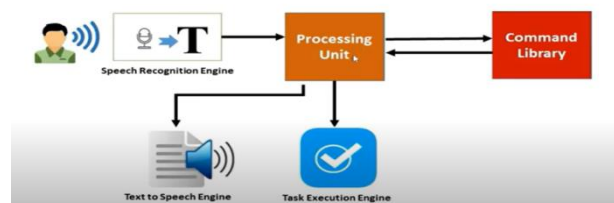

Fig 1 Architectural flow for voice assistant

LITERATURESURVEY

[3]Advancements are taking place a lot now in home improvement in Natural language processing and are resulting in development of highly performing networking models since the beginning of 21 st century. [4]. With applications starting from eGovernment [5], to accessibility support for humans with disabilities [6], [7], to straightforward convenience, IVAs were there in more than 41% of homes within the United States of America itself, generating a projected revenue of up to $19 billion [8]. In fact, the common Google search interest for the term has quadrupled since 2013, after having suffered minor spikes in 2010 and 2012, respectively [9]. These spikes roughly coincide with the release of Apple's Siri as a standalone app in 2010 and its subsequent bundling into iOS in 2011.

The increase in global interest has further corresponds with the delivery of Microsoft Cortana in 2013, Amazon's Alexa in 2014, and Google's Assistant in 2016. IVAs (Intelligent Voice Assistant) operate through an agent/client software on a suitable device, typically provided by the same vendor as the IVA software. Voice commands and instructions are captured through internal microphones of the device and sent to a cloud natural language processing platform, which converts the voice recording to machine-interpretable data, performs linguistic analysis, and then retrieves additional information from other services or sensors connected to the IVA,depending on the voice command entered by the user.

IVA functionality is mainly dependent on proprietary cloud infrastructure; the devices themselves are but a physical front-end for the software features. To activate, these devices perpetually listen for a "wake-word," (e.g., "Hey Siri," "OK, Google," or just the name of the product).

One natural concern becomes what else these devices hear and what conversations and noises in the household are recorded generally. Recently, not solely the IT press [10], [11], the industrial vendors themselves [12], [13], and academia [14], [15] have stated how IVAs will record entire conversations of the people in the room. citing that up to 80% of IVA users are concerned about possible breaches of their privacy. This concern is widespread, as law enforcement agencies discourage users of certain models of smart TVs with bundled IVA functionality from discussing confidential information in the vicinity of the device[16]. This is significantly regarding in jurisdictions, wherever interacting with IVAs implies that the users legally relinquish their right to sufficient privacy (as is the case, for instance, due to the third-party doctrine exception to the 4th Amendment of the United States of America, see [17]). IVAs do not only pose privacy issues for the user through their innate functionality, but also present physical attack vectors. For instance, lasers may be used to initiate the devices and permit injecting simulated voice commands into the device's internal mike [18]. Many researches were done to crack open the device using voice directions into the voice assistant by using ultrasonic waves that also helps in answering the calls on the cellphone easily [19].When the hardware attacks take place on the cloud where the device information is stored.[3]. Whereas past studies, merchant investigations, and IT press looks to concentrate on hardware attacks, the results from a user perspective appear to be relegated to associate implicit circumstance. Therefore,in this article, we focus importance on the user perspective.

METHODOLOGY

Language or technology used
Python language was used to write the code. Python provides a wide variety of libraries for scientific and computational usage. The model is implemented in PyCharm environment.

Speech Recognition module
This module deals with converting speech into text i.e., the commands given the user is identified and then converted into text and then the machine forms a response in order to perform a task. The machine takes the commands through microphone. Pyttx3 is also a module that converts audio to text.

Pyjokes
Here in this module, we can easily access the module and will be able to create different type of jokes for the programmers. This is a python library that can be imported directly in the command line .We can use this module to return jokes from a certain category and inthe required languages.

Pyautogui

Here, this module is a library that is helpful in keyboard and mouse control .It is useful when the environment or machine wants to in interact with other applications. It implements automation that means performing more than one task at a time.

Selenium Web
The selenium package is employed to automatize web browser interaction from Python. Various browsers/drivers like Firefox, Chrome, Internet Exploreraresupported as well as the Remote protocol. Supported Python Versions Python 3.7+ Installing If you have pip on your system, you can simply install or upgrade the Python bindings: pip install -U selenium

pyttsx3
It's a text to speech conversion library in python. This package supports text to speech engines on Mac os x, Windows and on Linux.

Wikipedia
Wikipedia is a multilingual online encyclopedia created and maintained by a community of volunteer editors utilizing a wiki-based editing system as an open cooperation effort. We'll look at how to utilize Python's Wikipedia module to get a range of data from the Wikipedia website in this post.
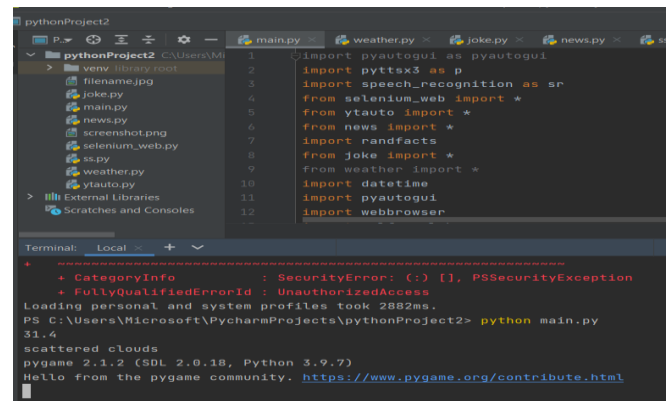

Fig.2. importingmodules

RESULTS AND ANALYSIS

The model is tested, and the voice assistant reacts to the voice commands given by the user accordingly.Itis used to check the functionalities,errors, and execution of the model. The analysis is later made, and desired functionalities and software requirements are added accordingly. It is helpful in ensuring the model meets its requirements and meets the user needs without fail.
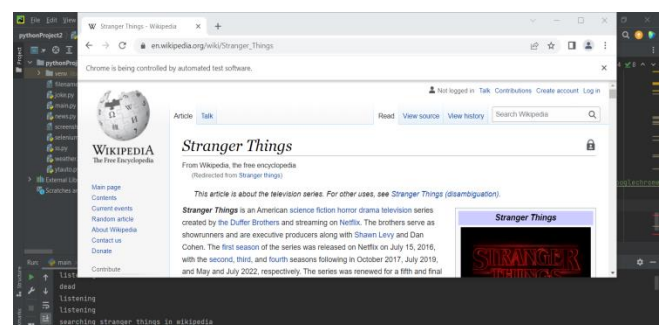

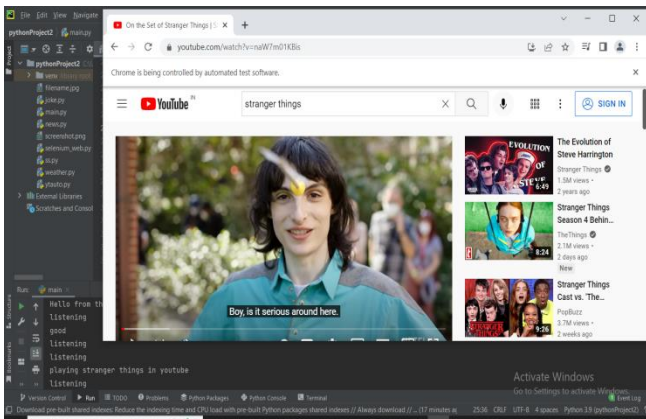Fig. 3. Searching Wikipedia output
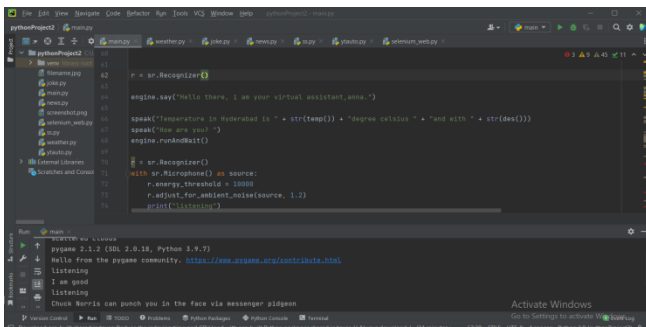
Fig. 4. Play video output

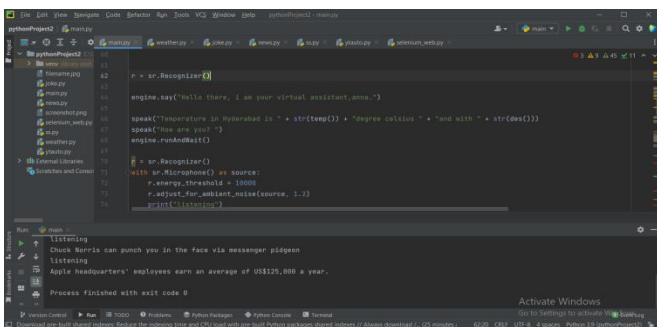Fig.5.   Jokes output
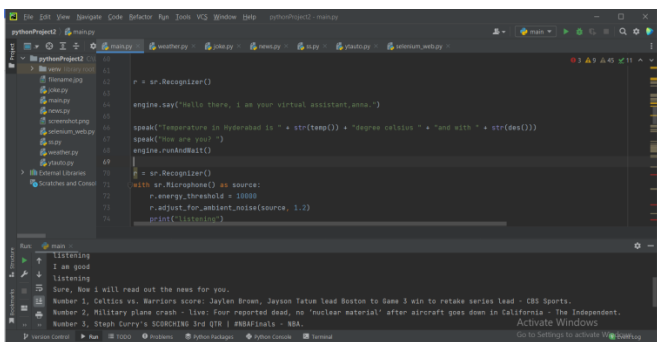

Fig. 6   Random facts output
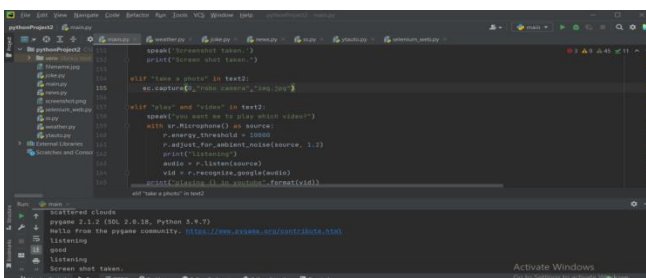

Fig. 7. News output

Fig. 8. Screenshot output

CONCLUSION

Improving the voice of colleagues is ceaseless interaction. You can add a lot more elements as time goes on. Voice aides are before long turning into a piece of organizations, enabling their efficiency, and teaming up their endeavors. Be that as it may, the speed is hazardously delayed with the voice being at its outset. It is yet to track down the necessary speed increase to turn into an essential portion of business organizing and our lives. Notwithstanding, extraordinary personalities have taken their advantage in releasing their true capacity and the fate of voice associates looks brilliant.

REFERENCES

[1] https://towardsdatascience.com/building-an-intelligent-voice-assistant-from-scratch-3d5749f4af07

[2] https://www.askpython.com/python-modules/speech-recognition

[3] J. S. Edu, J. M. Such, and G. Suarez-Tangil, "Smart Home Personal Assistants: A Security and Privacy Review," arXiv preprint arXiv:1903.05593, 2019.

[4] Citius Minds. The Evolution of Smart Homes. Available: https://www.citiusminds.com/blog/the-evolution-ofsmart-homes/. Accessed on Sept. 10, 2020.

[5] E. C. Paraiso and J.-P. Barthès, "A Voice-Enabled Assistant in Multi-Agent System for e-Government Service,"Proceedings of the Intl. Symposium and School on Advanced Distributed Systems, pp. 495–503, 2005. Available: https://doi.org/10.1007/11533962_45

[6] M. Jefferson, "Usability of Automatic Speech Recognition Systems for Individuals with Speech Disorders: Past, Present, Future, and a Proposed Model," University of Minnesota Digital Conservancy, 2019. Available: http://hdl.handle.net/11299/202757

[7] N. Friedman, A. Cuadra, R. Patel, S. Azenkot, J. Stein, and W. Ju, "Voice Assistant Strategies and Opportunities for People with Tetraplegia," Proceedings of the 21st. Intl. ACM SIGACCESS Conf. on Computers and Accessibility, pp. 575–577, 2019. Available: https://doi.org/10.1145/3308561.3354605

[8] S. Perez, Smart Speakers Hit Critical Mass in 2018. Available: https://techcrunch.com/2018/12/28/smartspeakers-hit-critical-mass-in-2018/. Accessed on Sept. 10, 2020.28

[9] Google Trends 2020. Worldwide relative interest in the search term "voice assistant." Available: https://trends.google.com/trends/explore?date=all&q=voice%20assistant. Accessed on Sept. 10, 2020.

[10] S. Perez, 41% of Voice Assistant Users Have Concerns About Trust and Privacy, Report Finds, 2019. Available: https://techcrunch.com/2018/12/28/smart-speakers-hit-critical-mass-in-2018/. Accessed on Sept. 10, 2020.

[11] T.Ammari, J. Kaye, J. Y. Tsai, and F. Bentley, "Music, Search, and IoT: How People (Really) Use Voice Assistants," ACM Transactions on Computer-Human Interaction, vol. 26, no. 3, pp. 1–28, 2019. Available:https://doi.org/10.1145/3311956

[12] Microsoft, Inc. (2019). Voice Report. From Answers to Action: Customer Adoption of Voice Technology and Digital Assistants. White Paper. Available: https://about.ads.microsoft.com/en-us/insights/2019-voice-report.Accessed on Sept. 10, 2020.

[13] C. Fisher, Amazon Enlists 30 Companies to Improve How Voice Assistants Works Together. Web resource, 2019. Available: https://www.engadget.com/2019-09-24-amazon-voice-interoperability-initiative.html.Accessed on Sept. 10, 2020.

[14] S. Lafia, J. Xiao, T. Hervey, and W. Kuhn, "Talk of the Town: Discovering Open Public Data via Voice Assistants," Proceedings of the 14th Intl. Conf on Spatial Information Theory, pp. 10:1–10:7, 2019.

[15] Khairunisa Sharif, Bastian Tenbergen, "Smart Home Voice Assistants: A Literature Survey of User Privacy and Security Vulnerabilities", University of New York at Oswego, Available: https://csimq-journals.rtu.lv/article/view/csimq.2020-24.02DOI: 10.7250/csimq.2020-24.02

[16] A. Holmes, The FBI just issued a warning about the risks of owning a smart TV — here are its suggestions for protecting your privacy, 2019. Available: https://www.businessinsider.com/smart-tv-security-fbi-warning-2019-12. Accessed on Sept. 10, 2020.

[17] M. Ford and W. Palmer, "Alexa, are you listening to me? An analysis of Alexa voice service network traffic,"Personal and Ubiquitous Computing, vol. 23, no. 1, pp. 67–79, 2019. Available:https://doi.org/10.1007/s00779-018-1174-x

[18] L.Armasu, L. Alexa, Google Assistant, Siri Vulnerable to Laser Beam Hacking, 2019. Available:https://www.tomshardware.com/news/light-commands-laser-beam-hack-alexa-google-assistant-siri. Accessed on Sept. 10, 2020.

[19] W. Yan, K. Liu, Q. Zhou, H. Gui, and N. Zhang, "SurfingAttack: Interactive Hidden Attack on Voice Assistants Using Ultrasonic Guided Waves," Proceedings of the Network and Distributed Systems Security Symposium, 2020. Available: https://doi.org/10.14722/ndss.2020.24068