Original Article

Check for updates

Available online at
https://jiicet.gnt.com.pk/index.php/jiicet/

Journal of Innovative Intelligent Computing and Emerging Technologies (JIICET)

Vol. 1, No. 1, 2025

Rule-based AI for Virtual Voice Assistants

Umair Aziz[1], Meerab Mujahid[2],Muhammad Tahir[3]

[1]Department of Computer Science, Air University Multan Campus, 6000, Multan, Pakistan; umairaziz3000@gmail.com (F.A.L)
[2]Department of Computer Science, Air University Multan Campus, 6000, Multan, Pakistan; meerabmujahid8@gmail.com (F.B.L)
[3]Department of Computer Science, Comsats University Islamabad Vehari Campus, 61100, Multan, Pakistan; muhammadtahir1007@gmail.com (F.C.L)

* Correspondence: umairaziz3000@gmail.com

Article History

Received 01 November 2024
Revised 21 December 2024
Accepted 11 March 2025
Available Online 21 March 2025

Keywords:
Voice Virtual Assistant (VVA)
Rule based AI
Voice Commands Control
Voice Assistant
Virtual Voice

Abstract

Virtual Voice Assistant (VVA) is a service-based application that can understand and execute voice commands, or act as an intermediary for the systems. With the development of deep connections between humans and artificial intelligence voice assistants (VA), the relationship between humans and machines has been enhanced. This study purposed the rule-based framework for the Virtual Voice Assistant (VVA) that captured the user's voice and performed an SST operation on it to extract the text. Furthermore applied NLP operations to preprocess the raw data and pass it to the rule-based dictionary for the appropriate action. This system is especially for people with differently-abled (blind, and handicap) and performs their tasks on the system.

1. Introduction

Nowadays, AI-based Virtual Voice Assistant (VVA) application-based services are widely used in the industry and the way of human-computer interaction is gradually changed due to this. It is no longer limited to the voice interaction of one question and one answer but also performed numerous actions according to voice commands and also define the next generation of human-computer interaction [1] so that the interaction between humans and machines is more natural. Different researchers developed and established a complete "end-to-end" human-computer interaction-related technology stack to maintain the international cutting-edge technology level. But the intelligent virtual assistant is still in the development stage [2].

The core of the development of intelligent virtual voice assistants lies in the following three pillars

1. Contextual awareness at the interaction layer: Voice intelligent assistants need to be able to recognize a large number of unstructured speech and respond based on their AI network. The network depth of the interaction layer also determines the integrity of the intelligent assistant's functions, especially in the deep-level interaction based on existing interactions [3].

2. The construction of the knowledge base: The scale and type of the knowledge base determine the scenes that the intelligent assistant can respond to and the granularity of decision-making responses in each scene. The large number of classified data sets it provides will serve as an important input for the learning of the emotional interaction layer [4].

3. Integration of application capabilities: The function of the intelligent assistant is mainly for user interaction, but the further realization of the function needs to be realized through other applications. This

requires the intelligent assistant to be well integrated with other first-party applications or other third-party applications to ensure the actual scenario-based application of the intelligent assistant [5].

In the future of scientific and technological life, the importance of voice commands based action in human-computer interaction greatly changed the world especially for visually impaired people and reduced the threshold for them to interact with machines while under the needs of the digital transformation and upgrading of enterprises, it has also become a new way to help impaired people for their success interaction with the help of AI and made the interactive platform [6].

This article is divided into five major sections. The first section covers the introduction and importance of this work, the second section is a literature review that covers the information about the adopted approaches and findings of previous work which have been done by different research in the same field. The third section holds the operational information about the adopted methodology that is based on several steps to achieve the objective of this study. In the fourth section, we will cover the evaluation approaches and GUI of the system. In the last section, we will discuss the conclusion of this work and future recommendations.

2. Literature Review

PDA with ad-hoc can facilitate the people to visit any historical sites naturally rather than in the traditional way with audio and video guides. This Chabot provides ease of guidance to the users at any place. To provide ease of guidance is the basic goal of the author. This paper also enlightened the advantages of this Chabot the users can do the natural interaction with this Chabot due to the speech recognition technologies. The other advantage of this project, it is also usable on different devices like mobile phones. This Chat robot is based on the multi-model which provides the users' ease in visiting heritage sites [7]. The technique of the laparoscopic hysterectomy by using the computer-enhanced surgical robot. This is the computer-enhanced robot that is used in the surgical system. This is the first report of using the computer-enhanced surgical robot that performs hysterectomy in human beings. The results of this surgical robot were satisfactory because the patients of ages from 22 to 77 years and with a maximum weight of 100 tolerated and were satisfied with this entire procedure. The development of this technology will increase the usage of this application in gynecology because it is well tolerated and reasonable [8]. An intelligent web-based voice Chabot. The major goal of this paper is to present the development of the chat voice robot, which provides an ease for the customer to communicate the server from any place with web service. The java language was used to develop this paper. The basic function of this web-based bot makes the possible; the user can interact with all types of clients from any platform. This paper has also enlightened the advantage of this Chabot; if the user asks a question to the bot which is not understood by the third party and the response will be archived [9]. The mobile robot assistant. This paper is provided with this robot functioned for communication, interaction. They faced many challenges during the development of the mobile assistant robot but dependability was one of them. The main advantages of this project enable to supervise the implementation of a task by robot and the picture transferred to the control panel. This mobile robot assistant is helpful in the museum as well as exposition areas. This mobile assistant can move like humans and detect hurdles by the laser scanner, it can also directly interact with the visitor [10]. Intelligent Chabot Tool to assist high school students in learning general knowledge subjects. Recently Chabot's are only used in banking and other smart business applications to assist the clients a run time. The authors purpose the solution that is based on the intelligent chat agent which helps the high school students in the general knowledge subjects' preparations. It takes the technology of intelligent chat agents in the education field to make the learning process smart, easy, and interactive so that the learning becomes fun. It is a web-based tutoring system that will be available 24/7. To answering the quires of the users it uses natural language processing techniques. The data set on which this system is trained is based on the Large Questions And Answers Of The Domain. It Also Interacts with the user but for limited quires. They also design the mobile application and compare the performance with the other consisting of intelligent chat systems [11]. The voice recognized the bot with image processing that can speak like a human. This robot is based on human-robot interference and human-computer interference. This project is divided into four parts. In the first part, the design and communication system has been discussed by the author. To recognize the voice, they used the grammar for speech which enables the robot to recognize what the user said. In the second part, they used different technologies for speech processing. The basic purpose of speech processing is to provide a suitable interface for new users. In the third part, the author discussed the analysis of speech. The result of speech processing with different parameters with the image processing system and various microcontrollers is used to implement the speech processing. This is the mobile robot that can recognize the voice and also have the capability to detect the image [12]. The basic purpose of this project is to provide ease and improve the living

standard of the older and the disabled persons of the society. This is a mobile home care system that can fulfill the task of the home environment. This system encourages people to live an independent life and gives the person's help in daily life. The main system of this bot is that it can communicate with the user and help the user to move from bed to chair. Furthermore, it is easy to operate for all users [13]. Human-like interaction with an intelligent assistant. Artificial intelligence has been developed day by day. The implementation of artificial intelligence in human life has shown efforts and is somewhat successful to decline the gap between human life and computer life. The presence of the human-like attributes with the intelligent computer is the basic purpose of this paper. The results indicate that the user interacts with the intelligent assistant by giving instructions. Intelligent assistants carried out these instructions' human-like instead of belief and human behavior [14].

3. Methodology

The adopted methodology is based on different procedures to achieve the objective of this study as can be seen in Fig. 1.

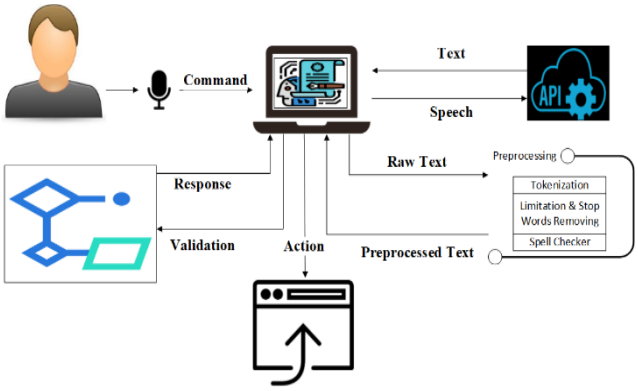

Figure 1. Methodology of Rule-based AI for Virtual Voice Assistants

In the above-mentioned methodology, as the audio command is given to the system, then need to be performed operations.

A. *Speech to Text Operation*

In the STT operation, we performed speech-to-text conversion using Google API. The Primary input for Google API is speech (command) and the primary output is Text.

B. *Preprocessing of Raw Text*

The extracted raw text is noisy and unclean. Therefore, we need to perform different operations to eliminate these anomalies from the textual data. There is the following operation are performed:

- Tokenization: Using the NLTK library we performed word-level tokenization. Word-level tagging is to use spaces and punctuation to divide a piece of text into many words. The smallest dimension of word-level tagging is words.

- Limitation and stop words removing: NLTK is a natural language toolkit for text preprocessing and it has a list of stop words in 16 different languages by using this we filtered the textual data and remove all stop words.

- Spell checker: Use built in function of NLTK to analysis the spell of text and make the necessary correction as per requirement to make text richer to further processing.

C. *Rule-based Dictionary*

Rule-based dictionary plays a vital role in this work, it takes the preprocessed text as a primary input, and then initiate the matching the procedure which is based on "From left to right, gradually remove the words on the right (bottom) for a new round of matching", and in the last, the appropriate response will return to the system.

The last core step of this purposed system is to performed action according to response and in this way, the system achieved the key objective of this work which was initiated through the voice command.

IV. VIRTUAL VOICE ASSISTANTS' EVALUATION

This section describes the evaluation methods for Rule-based AI for Virtual Voice Assistants which were adopted by different researchers to monitor the overall performance. The standard evaluation has not been established and might be out of date quickly as the technology evolves. However, proper evaluation is vital for VVA(Virtual Voice Assistant) development, and we categorized evaluation approaches into two main categories:

A. *Content Evaluation*

The first category is content evaluation, where the evaluation focuses on the response context from the system for appropriate action. The content evaluation heavily involves responses generated by a computer or a machine. Content evaluation is beneficial to researchers in the field of natural language processing and generation because it allows the evaluation to be conducted quickly and without the expensive manual work of human judges. For evaluation of content, we used BLEU (bilingual evaluation understudy) algorithm as follow,

$$\text{BLEU} = \text{BP} \cdot \exp\left(\sum_{n=1}^{N} w_n \log p_n\right)$$

BP(Brevity Penalty)

- *N: No. of n-grams(unigram, bigram, 3-gram or 4-gram)*

- *w_n: Weight for each modified precision,*

by default the value of N is 4,

so,

w_n is 1/4=0.25

- *P_n: Modified precision*

BLEU =0.906978545415645

4. User References Satisfaction

The second category, user satisfaction, is a popular method of evaluation because this evaluation is done according to the satisfaction of users (human) and at this rate, the satisfaction is usually in form of a Likert scale. We consider 100 individuals with 5 points (Strongly Agree, Agree, Neither Agree Nor Disagree, Disagree, Strongly Disagree) to measure the satisfaction of the user as shown in Figure 2. According to this, we found the maximum number of users are strongly satisfied with Virtual Voice Assistants (User-Friendly Environment, Achieve your goals and Recommend to others)

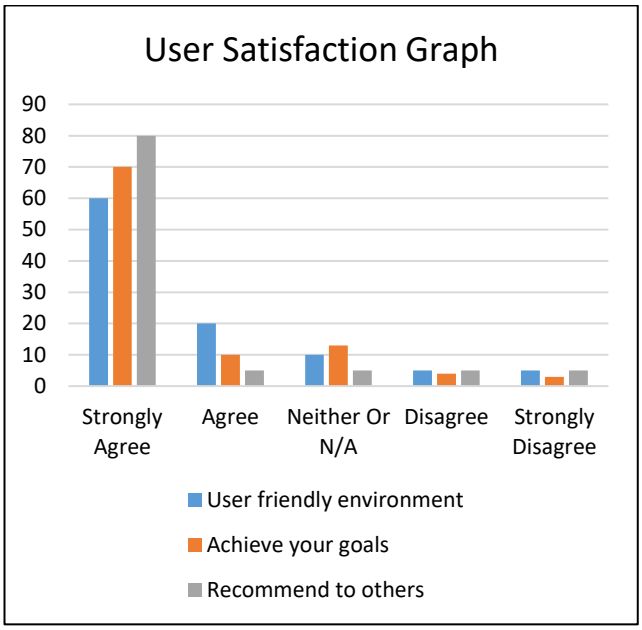

Figure 2. User Satisfaction Graph

Figure 3 shows the main GUI (Graphical User Interface) of Virtual Voice Assistants. User Through the mic entered their commands and after that VVA will invoke action call to perform desired operations.

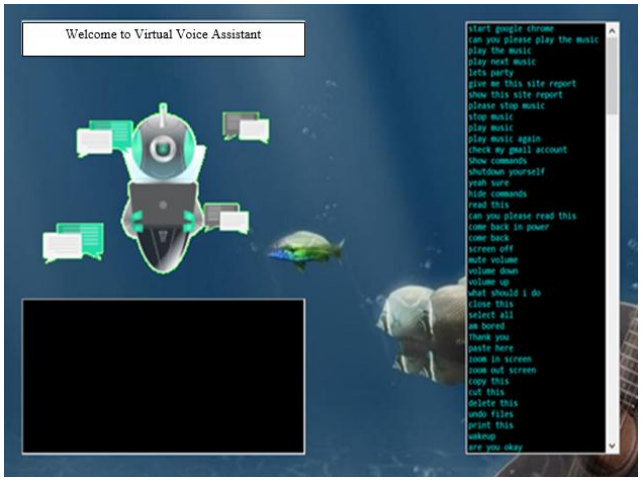

Figure3. GUI Level-0 (VVA)

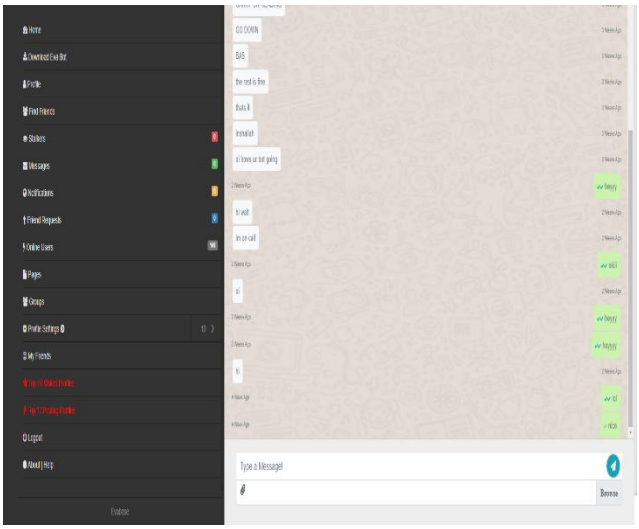

Figure 4. Interact with Social Media Sites (WhatsApp)

Figure 4 shows the messages which bot fetches from the WhatApp alerts and will give you a reminder to give a reply to the sender and User through the mic entered their reply and after that VVA will invoke an action call to perform to send the reply.

For Example, Senders send the messages:

- Message#1: Hello User, How are you doing?
- Message#2:Good Morning User2, Have a nice day

The user through the mic can enter their reply like Reply

- Message#1 Input Hello User2, I am doing great what about you?

And after that VVA will invoke an action call to perform to send the reply.

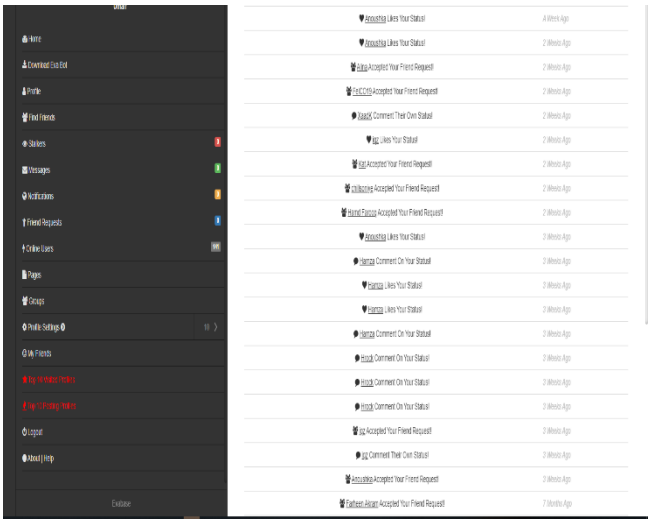

Figure 5. Interact with Social Media Sites (Gmail)

Figure 5 shows the mails from the user's E-mail address and notifications from the social network. Bot fetches the daily mails and notifications and alerts the user about the sender details and sender message.

For Example, Senders send the mail:

- Mail#1:Hello User, Meeting at 10 PM!
- Notification#2:User2 Liked Your Picture!

and User through the mic can enter their reply.

4. Conclusion And Future Work

In recent years there has been a significant improvement in the development and use of AI-based Agents with substantial benefits in many domains. In this work, We developed the Rule-based AI for Virtual Voice Assistant (VVA) application-based services especially for people with differently-abled (blind, and handicap) to established interaction with systems and perform different tasks. For the evaluation of rule-based VVA we adopted two different strategies: Content Evaluation and User satisfaction. For the content evaluation, we used the BLEU score which is 0.9069785454156454, nearby 1 that means the VVA is a trustworthiness system with respect to appropriate translation. While we used the Likert scale to monitor the satisfaction of users about VVA and we found an interesting fact that 60% 70% 80% are strongly agreed, 20% 10%5% agrees with User-Friendly Environment, achieve your goals and recommend to others respectively. For future work, we need to extend this rule-based VVA with a hybrid approach and add a pitch classification feature and allow only authorized users to interact with the system.

References

1. Gaouar, L., Benamar, A., Le Goaer, O., & Biennier, F. (2018). HCIDL: Human-computer interface description language for multi-target, multimodal, plastic user interfaces. Future Computing and Informatics Journal, 3(1), 110-130. doi:10.1016/j.fcij.2018.02.001
2. Mihret, E. (2020). Robotics and Artificial Intelligence. International Journal of Artificial Intelligence and Machine Learning, 10(2), 57-78. doi:10.4018/ijaiml.2020070104
3. Graham, C., & Jones, N. (2016). Intelligent Virtual Assistant's Impact on Technical Proficiency within Virtual Teams. International Journal of Virtual and Personal Learning Environments, 6(1), 41-61. doi:10.4018/ijvple.2016010104
4. Abbaschian, B., Sierra-Sosa, D., & Elmaghraby, A. (2021). Deep Learning Techniques for Speech Emotion Recognition, from Databases to Models. Sensors, 21(4), 1249. doi:10.3390/s21041249
5. Berdasco, López, Diaz, Quesada, & Guerrero. (2019). User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana. Proceedings, 31(1), 51. doi:10.3390/proceedings2019031051
6. Smith, P., & Smith, L. (2020). Artificial intelligence and disability: too much promise, yet too little substance? AI and Ethics, 1(1), 81-86. doi:10.1007/s43681-020-00004-5
7. Napolitano, R., Scherer, G., & Glisic, B. (2018). Virtual tours and informational modeling for conservation of cultural heritage sites. Journal of Cultural Heritage, 29, 123-129. doi:10.1016/j.culher.2017.08.007
8. Diaz-Arrastia, C., Jurnalov, C., Gomez, G., & Townsend, C. (2002). Laparoscopic hysterectomy using a computer-enhanced surgical robot. Surgical Endoscopy, 16(9), 1271-1273. doi:10.1007/s00464-002-8523-5
9. Kadali, B., Prasad, N., Kudav, P., & Deshpande, M. (2020). Home Automation Using Chatbot and Voice Assistant. ITM Web of Conferences, 32, 01002. doi:10.1051/itmconf/20203201002
10. Bodiwala, K., Shah, S., Patel, Y., Prajapati, P., Marolia, B., & Kalyankar, G. (2017). Simultaneous Estimation of Ofloxacin, Clotrimazole, and Lignocaine Hydrochloride in Their Combined Ear-Drop Formulation by Two Spectrophotometric Methods. Journal of AOAC INTERNATIONAL, 100(1), 38-44. doi:10.5740/jaoacint.16-0229
11. IEEE Robotics and Automation Magazine Volume 11, Number 2, June 2004. (2004). IEEE Robotics & Automation Magazine, 11(2), 01-01. doi:10.1109/mra.2004.1310927
12. Graf, B., Hans, M., & Schraft, R. (2004). Care-O-bot II—Development of a Next Generation Robotic Home Assistant. Autonomous Robots, 16(2), 193-205. doi:10.1023/b:auro.0000016865.35796.e9
13. Vanichvasin, P. (2021). Chatbot Development as a Digital Learning Tool to Increase Students' Research Knowledge. International Education Studies, 14(2), 44. doi:10.5539/ies.v14n2p44
14. Thomson, R. (2011). Computer Assisted Pronunciation Training: Targeting Second Language Vowel Perception Improves Pronunciation. CALICO Journal, 28(3), 744-765. doi:10.11139/cj.28.3.744-765