GJR Publication
OPEN ACCESS

Global Journal of Research in Engineering & Computer Sciences
ISSN: 2583-2727 (Online)
Volume 04 | Issue 02 | March-April | 2024
Journal homepage: https://gjrpublication.com/gjrecs/

Review Article

Multi-Lingual AI Chatbot for Web Optimization

*Vaibhavi[1], Prof. Athmaranjan. K[2], Chaya.K[3], Kishan Kumar[4], Arpith.K[5]

[1,2,3,4,5,6]Information Science and Engineering, Srinivas Institute of Technology Mangalore, Karnataka, India

DOI: 10.5281/zenodo.10842079 **Submission Date:** 11 Feb. 2024 | **Published Date:** 20 March 2024

***Corresponding author:** Vaibhavi
Information Science and Engineering, Srinivas Institute of Technology Mangalore, Karnataka, India

Abstract

Official websites often suffer from navigation challenges that delay users from finding crucial information effectively. This project proposes the implementation of innovative technologies to transform the navigation experience. By utilizing AI-driven assistance, the project aims to deliver a user-centric, intuitive browsing experience. This initiative seeks to enhance official websites' accessibility and user satisfaction, ultimately improving their effectiveness in serving the public.

Keywords: *AI-based chatbot, Website navigation, Multimodal interaction, Natural language processing, Machine learning, User intent, Personalization, Accessibility, User engagement, Continuous improvement.*

I. INTRODUCTION

In an era dominated by digital interactions and online experiences, the ease and efficiency with which users navigate websites play a vital role in shaping their overall satisfaction. The exponential growth of information available on the internet has led to a pressing need for innovative solutions that enhance user experience, ensuring seamless access to desired content and services. Recognizing this imperative, our project endeavors to introduce a groundbreaking advancement in web optimization through the development of a Bilingual AI Chatbot. Navigating the vast landscape of the internet can be a challenging task for users, often resulting in frustration and diminished satisfaction.

II. LITERATURE SURVEY

A. AN INTELLIGENT WEB-BASED VOICE CHATBOT [1]

The paper introduces the design and development of an intelligent voice recognition chat bot with a technology demonstrator supporting a web service. Using a black box approach for communication, the system allows clients from any platform to interact with the server through a generated interface, enhancing accessibility. The artificial brain incorporated in the web-based bot generates personalized responses, and a third-party expert system processes queries not understood, improving the bot's intelligence over time. While the main advantage lies in developing an intelligent web-based voice recognition chat bot, potential limitations include the system's dependency on the lifespan of the third- party expert system, posing risks to the technology demonstrator's longevity. The paper acknowledges the need for improved convenience by introducing voice-based interactions but does not extensively address ethical and privacy concerns associated with voice recognition and data use, which could be considered a limitation.

B. SMART COLLEGE CHATBOT USING ML AND PYTHON [2]

This document provides a comprehensive overview of a smart college chatbot developed using machine learning (ML) and Python. The chatbot, designed to enhance user engagement with the college website, employs AI and ML technologies such as WordNet, natural language processing (NLP) and various algorithms for effective response generation. While the paper successfully details the system's architecture, user interaction and administrator management, a notable drawback is the absence of a discussion on potential limitations and challenges associated with the implementation. A more thorough exploration of drawbacks, including the need for continuous training, potential limitations in understanding complex queries, and privacy considerations, would contribute to a more balanced understanding of the practical implications of the proposed chatbot system in an educational setting.

C. CONVERSATIONAL AI POWERED CHATBOT FOR DELIVERING TELE-HEALTH AFTER COVID-19 [3]

The paper, "Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele Health after COVID-19," introduces a novel approach to providing primary healthcare education and advice to chronic patients in India through a multilingual conversational bot named "Aapka Chikitsak." The chatbot leverages natural language processing (NLP) and voice user interface (VUI) to offer information, preventive measures, home remedies, and interactive counseling sessions, with the aim of increasing healthcare access, especially in rural areas, and addressing challenges posed by the COVID-19 pandemics. The document also discusses future enhancements, including anti-depression support and live connectivity with doctors. Potential disadvantages of the paper could include scalability challenges for the chatbot, ensuring accuracy and reliability of medical information, and the need for ongoing maintenance to keep the chatbot's knowledge base current. Ethical and privacy concerns related to AI in healthcare, such as data security and patient confidentiality, may also require more thorough exploration. Addressing these considerations is crucial for the successful development and deployment of AI powered healthcare solutions.

D. VOICE ASSISTANT INTEGRATED WITH CHAT GPT [4]

The document explores the integration of voice assistant technology with ChatGPT, an AI-driven chatbot, and its applications across various domains. It underscores the growing role of AI in daily life, especially in virtual assistants and smart speakers.

The focus is on ChatGPT's development, detailing its features, capabilities, and limitations, with use cases ranging from answering questions to writing code and providing medical advice. It delves into the potential of voice assistants in transforming customer service, particularly in banking, addressing tasks like client authentication, request classification, and account management. The document emphasizes the significance of this integration in digitalization, customer service, and gaming. Furthermore, it discusses the limitations and ethical considerations of AI- driven chatbots, emphasizing the need for solutions to address issues like misinformation and scams. The conclusion highlights the importance of AI-enabled chatbots in providing personalized customer service through machine learning and sentiment analysis. Combining voice assistant capabilities with ChatGPT enhances customer service, providing an interactive and personalized experience. This integration can automate processes, improving efficiency in areas like client authentication and request classification. Moreover, in gaming, this approach has been used to teach new players game mechanics, offering a unique and immersive experience.

E. SIRI -THE INTELLIGENT PERSONAL ASSISTANT [5]

Siri operates through a voice-controlled natural language interface, utilizing sequential inference and contextual awareness to assist IOS users with personal tasks. Like many other key technological features in Apple's IOS products, Siri's development can be traced back to federal funding and research. The program is an artificial intelligence system that incorporates machine learning, natural language processing, and a web search algorithm (Roush 2010). The process of how Siri works can be broken down into several phases. Firstly, when a user issues a command, their device captures their voice, converts it into an audio file (which is then translated into binary code), and sends it to Apple servers. These servers, located in the cloud, process the information instead of the user's mobile device. This approach has two major advantages for Apple: it offloads much of the work to powerful computers, rather than taxing the limited resources of the mobile device, and it allows Apple to use the data it collects to continuously improve the service. The next phase involves understanding the meaning of the user's command, which relies on natural language processing. This field deals with the challenge of interpreting the various ways in which people can express the same concept using endless combinations of words. Once the meaning is understood, Siri transforms it into actionable instructions. To improve its functionality, Siri could adopt certain features from its competitors. For example, it could be enhanced to complete tasks on behalf of the user, similar to Google Now, or it could learn about the user over time and use this information to assist with tasks. Additionally, there is ongoing research into using smartphone-captured fundus images for efficient and rapid clinical screening, particularly to address the lack of ophthalmologists and limited accessibility to retinal image capture devices. This research highlights the potential of smartphone technology in improving healthcare accessibility and efficiency.

F. ANALYSIS OF CNN BASED SPEECH RECOGNITION SYSTEM USING RAW SPEECH AS INPUT [6]

Cutting-edge automatic speech recognition (ASR) systems typically handle the relationship between acoustic speech signals and phones in two distinct stages, optimized independently. Initially, the speech signal undergoes transformation into features, usually involving a phase of dimensionality reduction and information selection based on task-specific knowledge. These phases are carefully crafted, leading to advanced features such as Mel frequency cepstral coefficients (MFCCs) or perceptual linear prediction cepstral features. Following this, the likelihood of sub-word units, such as phonemes, is estimated using generative or discriminative models In terms of experimental setup, the TIMIT acoustic-phonetic corpus serves as a resource, containing 3,696 training utterances (sampled at 16kHz) from 462 speakers, excluding the SA sentences. The cross-validation set comprises 400 utterances from 50 speakers. In tasks related to

connected word recognition on the Aurora2 corpus, a CNN-based system is utilized to compute the posterior probabilities of word states.The input features consist of raw data, representing a window of the temporal speech signal, with normalization to ensure zero mean and unit variance. The system's performance is compared with a standard HMM/A*NN system using cepstral features. An Artificial Neural Network (ANN) with a single hidden layer is trained, taking MFCC features with several frames of preceding and following context as input, without the need for pre-training the network.The network's hyperparameters, including the input window size, kernel width of the first convolution, kernel width, shift of Wep of max pooling layers, and hidden layer width, are fine-tuned through early-stopping on the validation set. An analysis of filters suggests that features learned between the first two convolution layers of the CNN tend to model the spectral envelope of sub-segmental speech signals. Despite being susceptible to noise, these features demonstrate greater robustness compared to MFCC features (without normalization). To enhance the CNN-based system's robustness, potential strategies include filtering the speech signal using the Wiener filter technique in the Aurora Advanced Front End before inputting it into the CNN.

G. OVERVIEW OF THE SPEECH RECOGNITION TECHNOLOGY [7]

Speech recognition refers to the process by which machines interpret and understand spoken language, responding appropriately to commands or statements. This technology focuses on analyzing the human voice, allowing machines to automatically identify and comprehend spoken language through the processing of speech signals and pattern recognition. Hidden Markov Model (HMM) is a statistical model that has been extensively used since the 1970s and 1980s for analyzing and modeling acoustic signals. By the 1990s, HMM had become a core technology in computerized speech recognition and multi-user detection in mobile communications. Artificial Neural Network (ANN) technology, inspired by biological nervous systems, involves numerous simple processing units connected in parallel to create a complex information processing system. In voice recognition, ANN technology is applied in various ways, including reducing modeling units to improve overall system recognition rates, studying acoustic and auditory models, and incorporating context information to reduce the impact of voice changes on speech signals. Hybrid network models, such as HMM + NN, integrate knowledge sources like phonemes, vocabulary, syntax, and word meaning to enhance the understanding of voice recognition research and improve system performance. Despite advancements, speech recognition systems still face numerous challenges and areas for improvement to achieve widespread use. However, it is anticipated that with continued progress in voice recognition technology, these systems will become more sophisticated and widely applied. Various types of speech recognition systems are expected to emerge in the market, prompting people to adapt their speech patterns to accommodate different recognition systems. While creating a speech recognition system comparable to human abilities remains a significant challenge, progress is ongoing to enhance these systems step by step.

H. KANNADA SPEECH SEGMENTATION AND RECOGNITION FOR SPEECH TO TEXT CONVERSION [8]

Numerous technologies have emerged and disappeared over time. However, each new technology has the potential to significantly change how individuals interact and work with computers. These newer technologies often aim to simplify existing ones. Speech, as the primary mode of human-human interaction, also provides a means to communicate with computers, thereby addressing various human-computer interaction challenges. With speech being a signal that reflects acoustic power and is represented as an electrical pulse, it can be modulated using amplitude modulation.Speech or voice is considered a pressure wave, which is then converted into numerical values for digital processing. The initial step involves speech acquisition, where speech samples are recorded in real-time for each speaker and stored in computer memory. Another method involves using a Graphical User Interface (GUI), which provides a graphical display containing controls allowing users to interact with the system. Feature extraction techniques such as pitch period, endpoint detection, Vector Quantization LBG (Linde, Buzo, and Gray) algorithm, and Mel Frequency Cepstrum Coefficient (MFCC) are used. Template Matching, specifically using the Euclidean Distance, compares an unknown speaker's voice represented as characteristic vectors with a codebook database stored in computer memory.Once these processes are executed, the result of the template matching step is displayed as a text character on the screen. The developed GUI can perform Kannada speech segmentation, recognition, and speech-to- text conversion. The push buttons and displays are programmed to execute the aforementioned steps, with the record push button recording speech, performing endpoint detection, and saving the speech signal in the required folder in wave format.

REFERENCES

1. Du Preez, S. J., Lall, M., & Sinha, S. (2009, May). An intelligent web-based voice chat bot. In IEEE EUROCON 2009 (pp. 386-391). IEEE.
2. Koundinya, H., Palakurthi, A. K., Putnala, V., & Kumar, A. (2020, July). Smart college chatbot using ML and python. In 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-5). IEEE.
3. Bharti, U., Bajaj, D., Batra, H., sLalit, S., Lalit, S., & Gangwani, A. (2020, June). Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after covid-19. In 2020 5th international conference on communication and electronics systems (ICCES) (pp. 870-875). IEEE.

4. Yadava, T., & Jayanna, H. S. (2019, December). Automatic Isolated Kannada Speech Recognition System under Degraded Conditions. In 2019 4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT) (pp. 146-150). IEEE.
5. Reehal, S. (2016). Siri-The Intelligent Personal Assistant. International Journal of Advanced Research in Computer Engineering and Technology, 5(6), 2021-2024.
6. Moncy, A. M., Athira, M., Jasmin, H., & Rajan, R. (2020, December). Automatic speech recognition in Malayalam using DNN-based acoustic modelling. In 2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS) (pp. 170-174). IEEE.
7. Bhat, S., Kalaiah, M. K., & Shastri, U. (2021). Development and validation of tulu sentence lists to test speech recognition threshold in noise. Journal of Indian Speech Language & Hearing Association, 35(2), 50-56.

CITATION
Vaibhavi, Athmaranjan. K, Chaya.K, Kishan K., & Arpith.K. (2024). Multi-Lingual AI Chatbot for Web Optimization. In Global Journal of Research in Engineering & Computer Sciences (Vol. 4, Number 2, pp. 31–34). https://doi.org/10.5281/zenodo.10842079