



RESEARCH ARTICLE

NLP-Based Voice Assistant Usage on Consumer Shopping

Karthik Baburaj, Navaneeth kattil Madathil, Roshini Barkur*

Abstract

In the digital age, convenience, and accessibility have become paramount considerations for users at a personal level, particularly for those with diverse needs. The objective of this study is to introduce a Voice-Based Food Ordering Application designed to revolutionize the food ordering experience for tech-savvy individuals, with a special focus on customers with disabilities, including those who are differently abled and visually impaired. Leveraging cutting-edge technologies, such as Natural Language Processing (NLP), Speech Recognition, Google APIs, Text-to-Speech (TTS), and speech-to-text (STT), the application enables users to seamlessly navigate menu options, place orders, and confirm transactions using simple voice commands. The intuitive, user-friendly interface ensures that the app is accessible to all users, promoting independence and ease of use. In addition, the application prioritizes user security by implementing robust data protection measures and safeguarding personal information throughout the ordering process. By offering a straightforward and modern approach to food ordering, this app not only enhances convenience but also bridges the gap between technology and inclusivity, empowering users to interact with digital systems effortlessly. This innovation represents a significant step forward in making everyday tasks more accessible and efficient for everyone, regardless of their physical ability.

Keywords: Voice-Based Food Ordering Application, Speech Recognition, Text-to-Speech, Speech-to-Text, Voice enabled, Natural Language Processing.

Introduction

Traditional food-ordering applications face difficulties for users who must or prefer to interface with technology using voice commands. To solve these problems, this study aims to develop a voice-activated food ordering application. This system offers an easy-to-use and accessible voice-driven interface for food ordering, thereby enhancing social inclusion and accessibility. The difficulties faced by people who had trouble adopting touch-based interfaces were the basis of this idea. This study has developed an application that utilizes voice commands to enhance the food ordering process, addressing the need for an improved approach. The goal was to create a user-friendly system that would

be accessible to everyone. This application or software will help users navigate through available menus and categories as well as place orders.

Additionally, users can review past orders using voice commands. Users can specify the number of quantities and finalize the order. This makes the app particularly useful for visually impaired individuals or for those who are differently abled and prefer a hands-free user experience. Despite having many of these advantages, it also has some disadvantages, such as speech recognition accuracy that will or may be affected by accents or speech clarity and background disturbances, as well as navigating the numerous menus only by means of voice. Privacy concerns may also arise when users speak orders aloud in public. To address these issues, this study introduced Natural Language Processing (NLP) algorithms that help improve speech recognition to the next level by making it more precise in capturing words and enhancing accuracy while also facilitating easy navigation through available menus.

The system helps visually impaired users order food (as they would in the real world), allows busy professionals to multitask, and even makes life easier for older people, who might find traditional interfaces cumbersome to use. Previous systems have attempted to use voice technology with food ordering. The Intelligent Voice Food Ordering System (IVFOS) utilizes speech recognition and semantic understanding to simplify food ordering, making it as easy as

Department of Computer Application, Krupanidhi College of Management, Bengaluru, Karnataka, India.

***Corresponding Author:** Roshini B, Department of Computer Application, Krupanidhi College of Management, Bengaluru, Karnataka, India., E-Mail: roshini.krupanidhi@gmail.com

How to cite this article: Baburaj, K., Madathil N.K., Barkur R. (2025). NLP-Based Voice Assistant Usage on Consumer Shopping. The Scientific Temper, 16(spl-2):46-50.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.spl-2.08

Source of support: Nil

Conflict of interest: None.

talking about it. Additionally, it provides recommendations on foods that customers may like based on their past orders (Wang *et al.*, 2020; Xu *et al.*, 2020). A few other systems have been built for blind users to read and send emails via speech recognition, demonstrating how voice interfaces could be an approach to compensate for accessibility (Belekar *et al.*, 2020; Badigar *et al.*, 2018; Yadav *et al.*, 2020). Similarly, one of the prototypes created with Google Assistant demonstrated a simplified food ordering process, allowing users to place orders by voice by simply stating what they wanted, and it makes a call to an Application Programming Interface (API) to do so (Höijer & Jansson). A prototype of a voice dialogue system for fast food orders showed that an ordering dialog based on a combination of dialogue subsystems and voice interfaces could be successful (López-Cózar *et al.*, 1997; López-Cózar *et al.*, 1998).

The current study combines an algorithm to ensure efficiency by implementing speech-to-text (STT) in conjunction with text-to-speech (TTS) and natural language processing (NLP), which interprets the user's intent, and STT converts the text responses back into spoken language. This voice-enabled food ordering application will transform the traditional food ordering process, making it more accessible and convenient. Implementing these advanced methods can address the limitations of existing systems and provide a seamless and hands-free ordering system experience. This research streamlines the food ordering process compared to traditional methods, enhancing the quality of life for individuals with accessibility needs and promoting a more inclusive society.

Literature Review

In a few studies, online food ordering is an easy process; however, selecting desired foods can be difficult, and there are only a few systems that use both speech recognition and semantic understanding to engage with users. Based on these technologies, the study proposes an IVFOS on the Iflytek AIUI platform to simplify the ordering process. They also suggest products based on customers past orders for a personalized shopping experience (Wang *et al.*, 2020; Xu *et al.*, 2020). One of the proposed works provides an email system for visually impaired people who can follow vocal instructions in composing and reading messages without an intermediary. This system utilizes speech recognition, allowing users to operate the software without relying on keyboard shortcuts, which will be greatly appreciated by individuals with various disabilities or those who are illiterate. Currently, it is desktop-based, but it will be voice-operated and mobile-ready in the future. Future changes will include increased security during logins to protect users (Belekar *et al.*, 2020; Badigar *et al.*, 2018; Yadav *et al.*, 2020). Another study investigated how voice assistants can be integrated into a food-ordering system. To enable this behavior, the researcher created a prototype for speaking

to Google Assistant to place an order. The spoken order is passed through the Dialog flow to detect keywords, and then it goes through the Amazon Web Services (AWS) API Gateway. The user places an order, and the API Gateway forwards it to a lambda function to process the order. Lambda's reply also includes the order quantity and its price, which is sent back to the end user. They also verified the order using Zeldio API receipt. The prototype functions with a few minor bugs and meets the project objectives (Höijer and Jansson).

A study by Lopez-Cozar *et al.* provided an initial voice dialogue system for fast-food orders and questions. There are two components to this: a rule-based expert system, known as a dialogue subsystem, and a voice interface. The dialogue subsystem builds on natural language processing using dialogue from an authentic restaurant. This mechanism is explained by considering knowledge representation, grammar, and structure. The conclusion and future work are also included (López-Cózar *et al.*, 1997; López-Cózar *et al.*, 1998). The World Wide Web has revolutionized many aspects of our lives, particularly in the way users communicate, which is now faster and easier than ever before. This study deals with the methods used for STT and TTS conversion in an interactive voice-based email system. The aim is to discover the most efficient STT and TTS techniques. This review found that Hidden Markov Models (HMM) are best for use in both STT and TTS. HMM + Artificial Neural Network (ANN) for STT, and HMM for (recognizer model) (Nagdewani & Jain, 2020; Dongmei, 2021). Voice-controlled systems have enabled a better way to communicate with the computer or other related devices than the traditional methods in this paper and have introduced or proposed classification or showcases a prototype E-commerce app using IBM Watson's speech recognition; using these technologies, it will benefit in different domains (Kandhari *et al.*, 2018; Duttaroy *et al.*, 2022; Sadi *et al.*, 2022; Kumar *et al.*, 2024). The Internet has become an important factor in our daily needs using Internet email and has been the most popular way to communicate, but it will be challenging for some differently abled people like visually impaired cases; therefore, this research proposes a voice-controlled email system that aims to provide simple and easy email use not only for visually impaired but also for other disabilities (Tiwari *et al.*, 2020; Indalkar *et al.*, 2022; Naidu *et al.*, 2024; Golait *et al.*, 2023).

Materials and Methods

Description of NLP

NLP technology is being used by NLP-based websites to grasp and interpret in human language. It uses textual analysis to respond to the questions and do tasks such as generating content, translating among languages, and summing articles, and these websites involve computer interactions with more effective meanings of the words.

Natural, while using complex technologies to read and understand. They are frequently found in virtual assistants, chatbots, and creative content applications, which will help humans evolve with modern technology.

Algorithms Used

NLP is a modern technology that enables computers to fully understand, analyze, and generate human-like speech. This enables the execution of work information retrieval, text summarization, and natural interactions. STT is a method that transforms spoken words into printed text. By using this modern technology, users can directly interact or speak on laptops or other devices, such as mobile phones, and it will convert their voice into text. TTS is also a modern technology that will help to convert the written text into spoken words; it will also provide audio for the electronic version reading aloud and the transportation infrastructure and by enabling the devices with permission for information to read aloud.

Algorithm applied to work

NLP, exceptional STT, and TTS assist in providing a voice-enabled food ordering application with an engaging experience. NLP will process and interpret spoken commands such as "History," "Order," and "Logout," protecting the application to determine the user's intended purpose and handle the conversion of spoken words to text, allowing it to capture and analyze the user inputs accurately. Users can communicate with it via vocal commands and interests, similar to speaking to humans. TTS takes the conversation to a higher level by making the text reply to the application its voice counterpart. The app functions as audible feedback, providing users with choices and confirmation, allowing them to hear their responses. Collectively, these combined technologies enable users to engage seamlessly in the application via voice. This makes processes such as navigating menu items, selecting what food you want, choosing quantities, and ordering all seem easy for wait staff. This comprehensive strategy not only offers expediency but also ensures a more natural and humane way of placing orders.

System Architecture

In this system architecture, an intelligent voice user interface is used, followed by people speaking, which is the voice input (Figure 1). This input is further passed on to feature extraction, and important information or keywords are extracted from these spoken phrases. The extracted features are then compared with the corresponding data in the database to extract relevant points on the intention of the user. The pre-processed spoken input is then converted to textual information using the STT algorithm, which enables us to understand and process what users request. Similarly, the TTS algorithm converts these responses into spoken words for accurate and seamless user feedback or

instruction delivery. After that, the system will proceed to a decision-making step, which will assess the information processed and how the data was input through handwritten recognition, in order to approve or decline this request based on all aspects previously described. Finally, the User Environment confirms the selected feature or action back to the user via aural feedback, achieving a full voice cycle interaction with a high success rate.

Pseudocode

Step 1: User Login and Main Menu

User Logs Using Voice Recognition. The application verified the credentials.

Application verbally lists options: "History," "Order," "Logout."

Step 2: View Order History

If the user selects "History."

The application retrieves and verbally lists previous orders. The application verbally provided feedback on the last ordered item.

The user provides verbal feedback. The application confirms this feedback.

Step 3: Place New Order

If the user selects "Order"

The application verbally listed the available food categories. The user selects a category by using voice commands.

The application verbally listed the food items in the selected category.

The application asks the user if they want to select a food item. The user confirms or declines.

If confirmed

The application asks for a quantity using voice commands. The user specifies a quantity.

The application confirms this selection.

The application asks whether the user wants to continue, delete the last item, or finish.

Step 4: Order Management

If the user chooses to continue,

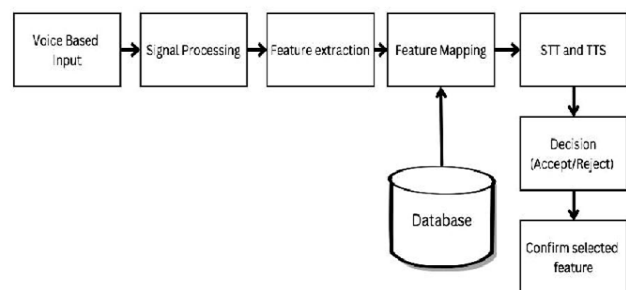


Figure 1: System Architecture Diagram

The application returns to food categories (step 3). If the user chooses to delete the last item,

The application removes the last added item. The application confirms this deletion.

The application returns to food categories (step 3). If the user chooses to finish,

The application confirms this order. The application places the order.

The application verbally confirms order placement and thanks the user.

Step 5: Logout

If the user selects "Logout," the application logs the user out. Session ends.

Results and Discussion

The voice-enabled food ordering application successfully implemented NLP, STT, and TTS technologies to create an intuitive and accessible user interface. The system effectively registered users and authenticated logins using voice commands (Figure 2). Upon successful login, users could navigate the main menu options ("History," "Order," "Logout") entirely through voice interactions (Figure 3).

When selecting the "Order" option, the application verbally presented available food categories (Figure 4).

Users could browse categories, select specific items, and specify quantities using voice commands. The system provided verbal confirmation at each step and allowed users to continue ordering, delete items, or finalize their order (Figure 5).

The "History" option enabled users to review past orders through voice interactions. Additionally, the system allowed users to provide verbal feedback on their most recent order, demonstrating the application's capability for two-way voice communication. By leveraging NLP, STT, and TTS technologies, the application successfully interpreted user commands and provided clear verbal responses. This approach significantly enhanced accessibility, particularly for visually impaired users or those preferring hands-free interactions.

The implemented system architecture (Figure 1) proved effective in processing voice inputs, extracting relevant features, converting speech to text and vice versa, and making appropriate decisions based on user intent. The seamless integration of voice commands throughout the ordering process, from browsing menus to confirming orders, resulted in an intuitive and user-friendly experience. The application's



Figure 2: Login page



Figure 4: Food Category

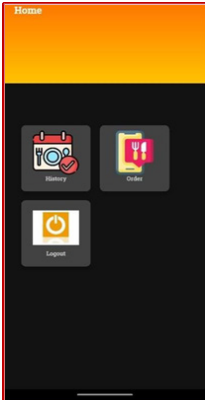


Figure 3: Home page

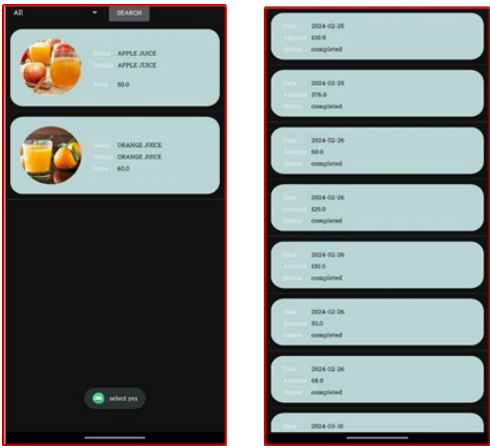


Figure 5: Selected food category

ability to provide verbal feedback and confirmations enhanced user confidence in the system. While the current implementation demonstrates the viability of voice-enabled food ordering, there is potential for further enhancements. Future work could focus on implementing personalized recommendations based on order history, integrating voice-based payment options, and refining the system's ability to handle diverse accents and background noise.

This voice-enabled food ordering application successfully combines NLP, STT, and TTS technologies to create an accessible and efficient ordering process. The system's ability to interpret natural language commands and provide verbal feedback represents a significant step forward in making digital services more inclusive and user-friendly.

Conclusions

The new voice-enabled food-ordering app is a natural intersection of NLP, STT, and TTS technologies, providing users with an easier and more intuitive way of communicating with the platform. Users can easily scroll through existing options by simply talking, navigating previous orders, adding items, and providing feedback. The method allows for greater accessibility and indicates that it is possible to use voice as a user interface along with resources such as dialogs to break down complex experiences. Working with advanced speech recognition and synthesis technologies, on the other hand, allows for a prompt interpretation of user commands as well as a clear communication of responses, thereby securing an uninterrupted and error-less process for the current user. This application describes how technologies can be used daily, making it easier and more interactive with the technology.

Possible future work for voice-enabled food ordering applications could involve implementing personalized recommendations based on the user's past orders. We can also implement machine learning algorithms so that they can provide better analyses of taste preferences and dietary restrictions. Additionally, implementing a voice-based payment option would help users to order seamlessly. Integrating these features will improve the adaptability of the system as a result of technological advancement.

Acknowledgments

The authors appreciate the support and guidance provided by the Krupanidhi College of Management in facilitating this work.

Conflict of Interest

The authors declare no conflicts of interest.

References

Badigar, M., Dias, N., Dias, J., & Pinto, M. (2018). Voice based Email application for visually impaired. *International Journal of Science Technology & Engineering (IJSTE)*, 4(12), 166-170.

Belekar, A., Sunka, S., Bhawar, N., & Bagade, S. (2020). Voice based

Email for the Visually Impaired. *International Journal of Computer Applications*, 175(16), 8-12.

Dongmei, L. (2021). Design of English text-to-speech conversion algorithm based on machine learning. *Journal of Intelligent & Fuzzy Systems*, 40(2), 2433-2444.

Duttaroy, N., Angre, A., Powar, S., Patil, P., & Hegde, G. (2022). Voice Controlled E-commerce Web App. *International Research Journal of Engineering and Technology (IRJET)*, 1(9).

Golait, S. S., Chalkhure, S., Balamwar, S., Shinde, R., Turkar, R., & Gaikwad, S. (2023). Review on Voice Based Email Assistant for Visually Blind People. *International Journal of Innovations in Engineering and Science*, 8(2), 20-25.

Höijer, D., & Jansson, H. (2021). *Voice-controlled order system*. Halmstad University

Indalkar, O., Ghorpade, S., Sonone, A., & Pathak, K. (2022). Review on voice based email system for blind people. *Sustainable Challenges and Smart Practices in Engineering, Technology & Management*, 140.

Kandhari, M. S., Zulkemine, F., & Isah, H. (2018, November). A voice controlled e-commerce web application. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, (pp. 118-124). IEEE. doi: 10.1109/IEMCON.2018.8614771.

Kumar, G., Matharu, M. K., Joshi, G., Chauhan, J., & Kumar, A. (2024). Voice Assisted E-Commerce Website. *Grenze International Journal of Engineering & Technology (GIJET)*, 10.

López-Cózar, R., García-Teodoro, P., Verdejo, J. E. D., & Rubio, A. J. (1997). A voice activated dialogue system for fast-food restaurant applications. In *Eurospeech 5th European Conference on Speech Communication and Technology* (pp. 1783-1786).

López-Cózar, R., Rubio, A. J., García, M. P. A., & Luna, J. C. S. (1998, May). A spoken dialogue system based on dialogue corpus analysis. In *LREC* (pp. 55-58).

Nagdewani, S., & Jain, A. (2020). A review on methods for speech-to-text and text-to-speech conversion. *International Research Journal of Engineering and Technology (IRJET)*, 7(05), 4459-4464.

Naidu, P. R., Jahnavi, D. S., Jalpa, C. J., Edwin, L. J., & Mahathi, D. N. (2024, June). Smart mail: Voice based email system for blind. In *AIP Conference Proceedings* (Vol. 3122, No. 1). AIP Publishing.

Sadi, M. T. A. H., Kadir, M. I., Rahman, M. S., & Khan, M. M. (2022, March). Development of a Voice Controlled Web based E-Commerce. In *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1-8). IEEE.

Tiwari, P. A., Zodawan, P., Nimkar, H. P., Rotke, T., Wanjari, P. G., & Samarth, U. (2020, February). A review on voice based email system for blind. In *2020 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, (pp. 435-438). IEEE. doi: 10.1109/ICICT48043.2020.9112539

Wang, X., Song, P., & Li, L. (2020, December). Research and application of intelligent voice food ordering system. In *2020 8th International Conference on Orange Technology (ICOT)* (pp. 1-3). IEEE.

Xu, Z., Zhao, X., Tian, Z., Zhuang, Z., & Wang, J. (2020, April). The design of nutrition ordering system without queuing in the college canteen. In *2020 International Conference on Computer Information and Big Data Applications (CIBDA)* (pp. 381-386). IEEE.

Yadav, A., Singh, A., Sharma, A., Sindhu, A., & Rastogi, U. (2020). Desktop voice assistant for visually impaired. *Int J Recent Technol Eng (IJRTE)*, 9(2).