Restaurant Voice Assistant Bot Using Python and Raspberry Pi

Lisha T P[1], Divya Shree Tala[2], Himashri A Masthi[3] and Dr. Lalitha S[4]

[1,2,3]Student, B M S College of Engineering

[4]Assistant Professor, B M S College of Engineering

Abstract—This paper presents the development of a voice assistant bot for restaurants, utilizing Python and incorporating SpeechRecognition, subprocess, and gTTS modules. Deployed on a Raspberry Pi, the bot efficiently addresses customer queries, takes orders, and generates final bills, enhancing the overall din ing experience. By integrating advanced speech recognition and text-to-speech technologies, the bot provides an intuitive, hands free interface, improving both customer service and operational efficiency. The use of a cost-effective Raspberry Pi platform highlights the system's practicality for widespread adoption in the hospitality industry. Detailed discussions on the software architecture, hardware integration, and performance metrics are included, showcasing the bot's technical robustness and societal benefits in creating a more accessible and streamlined restaurant environment.

Index Terms—speech-to-speech, SpeechRecognition, Subprocess, Raspberry Pi, Restaurant Automation.

I. INTRODUCTION

In the contemporary landscape, technological advancements are propelling various sectors toward unprecedented heights of efficiency and innovation. In the hospitality industry, chat bots have become indispensable tools for enhancing guest experience and operational efficiency. These AI-powered assistants[1] handle a variety of tasks, from managing bookings and answering frequently asked questions to providing personalized recommendations. By offering instant responses and seamless interactions, chatbots significantly reduce the workload on human staff, ensuring that guests receive timely and accurate information around the clock.

Building on the success of chatbots, the industry is now witnessing a transition to voice bots, which represent the next evolution in automated guest services. Voice bots bring an added layer of convenience by allowing guests to interact with the system through spoken commands. This transition from text-based chatbots to voice bots is driven by the growing preference for hands-free, natural communication, and the increasing capabilities of voice recognition technologies.

Our journey from chatbot implementation to voice bot integration has been a strategic move to further enhance guest engagement and satisfaction. While chatbots effectively manage text-based interactions, voice bots provide a more immersive and user-friendly experience. One significant advantage we observed with voice bots is their ability to cater to guests who are busy or on the go, allowing them to make requests, inquire about services, or receive information without needing to type. This hands-free interaction is particularly beneficial in busy restaurants like Udupi Upahar, Dosa Camp and Upahara Darshini is usually very busy, where guests are multitasking or prefer verbal communication, thereby elevating the overall guest experience to new heights. Traditional methods for order-taking and reservation management are labor-intensive and time-consuming. By automating these processes, speech to-speech bots[2] expedite service delivery and minimize errors. Customers can articulate their orders directly to the bots, which process requests instantaneously, reducing wait times and enhancing accuracy. This automation frees staff to focus on critical tasks like food preparation and personalized customer engagements, optimizing workflow. Furthermore, these bots facilitate invaluable data insights[3]. They gather and analyze data on customer preferences and feedback, providing actionable intelligence for refining offerings and service quality. Understanding customer behavior patterns al lows restaurants to tailor marketing strategies, modify menus, and implement personalized promotions, fostering loyalty and driving growth. Speech-to-speech bots[2] significantly enhance customer service through immediacy and accuracy in handling orders and inquiries. Their natural language processing[4][5] capabilities enable understanding and responding to various requests, ensuring a personalized touch. Operating round-the clock, these bots guarantee uninterrupted, high-quality service, crucial for fostering positive customer experiences and loyalty.

II. METHODOLOGY

A. Requirement Analysis

To enhance customer service in restaurants, a detailed requirement analysis is crucial. The core functionalities needed include answering queries about menu specials, costs, ingredients, spiciness, and calorie content. This ensures customers have comprehensive information to make informed choices. The interaction model was initially text-based, allowing customers to type their queries and receive instant responses. This method provided a simple and accessible way to interact with the system.

However, to further enhance the user experience, we have integrated speech-based interactions. This enabled the customers to ask questions verbally, making the system more intuitive and user-friendly. Implementing a text-based system first allows for the refinement of the bot's accuracy and reliability before transitioning to more complex speech recognition technology. This phased approach ensures a smooth deployment and improved customer satisfaction.

Moreover, the integration of speech-to-speech bots facilitates the accrual of invaluable data insights. These bots are equipped to gather and analyze data pertaining to customer preferences, frequent inquiries, and feedback , furnishing restaurant management with actionable intelligence to re f ine their offerings and service quality. By comprehending customer behavior patterns, restaurants can tailor marketing strategies, modify menus, and implement personalised promotions , thus fostering customer loyalty and driving business growth. The data-driven approach enabled by these bots ensures that restaurants remain agile and responsive to consumer demands, thereby
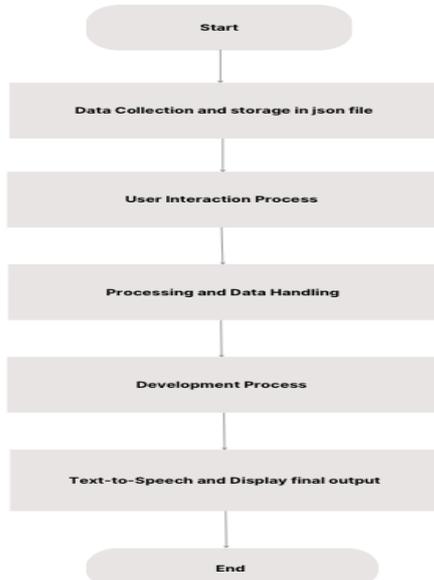
maintaining a competitive edge in the dynamic hospitality landscape.

The integration of advanced technologies[6] like speech-to speech bots signals a restaurant's commitment to innovation and customer satisfaction, which are key drivers of long-term success in the hospitality sector.

B. Data Collection and Storage in json File

Effective data collection and management are fundamental to developing a responsive restaurant bot. The first step in volves gathering comprehensive menu data, including detailed descriptions of each dish, ingredients, spiciness levels, calorie counts, and pricing. Organizing this data efficiently is critical for quick retrieval and accuracy. Storing the menu data in a JSON file[7] ensures that it is structured in a way that is both human-readable and easy for the system to parse. Beyond menu data, integrating customer

```json
{
    "name": "Idli Vada",
    "cost": 45,
    "ingredients": ["rice", "lentils", "lentil fritters"],
    "spiciness": "mild"
},
{
    "name": "Curd Vada",
    "cost": 35,
    "ingredients": ["lentil fritters", "yogurt"],
    "spiciness": "mild"
},
{
    "name": "Veg Pualo",
    "cost": 40,
    "ingredients": ["rice", "vegetables"],
    "spiciness": "mild"
},
{
    "name": "Poori with vegetable sagu",
    "cost": 60,
    "ingredients": ["wheat flour", "vegetables"],
    "spiciness": "mild"
},
```

Figure 2 : Data in json format

interaction logs and feedback into the bot's database can further enhance its responsiveness and personalization capabilities. Analyzing this data enables the bot to identify common questions and preferences, thereby improving its ability to make accurate recommendations and address customer concerns effectively. Furthermore, regular updates to the JSON file are essential[8] to ensure that the information remains current and relevant. This includes adding new dishes, updating ingredient lists, and adjusting prices as necessary. By maintaining an up-to-date and well-organized data structure, the restaurant bot can deliver a high level of service reliability and efficiency.

Figure 1: Block Diagram

The article[7] describes the mechanism of validation and data processing in JSON format in the form of a modified model, which by its rules will implement a universal approach to the process of data processing of arbitrary structure with the ability to implement its logic of validation and processing of nodes with any level of complexity. In our project, we chose the JSON file format to store our restaurant menu because it offers a flexible and universally accepted structure for data representation. JSON's hierarchical format allows us to easily define and organize the menu items, categories, and attributes in a clear and readable way. This aligns perfectly with the approach described in the article, where each node (or menu item) can have its own validation logic and state processing. Additionally, JSON's lightweight nature ensures efficient data exchange and compatibility with various programming languages and systems, making it an ideal choice for our dynamic and complex menu data management.

In conclusion, meticulous data collection and efficient management are crucial for the successful implementation of a restaurant bot. Using a JSON file for storing menu information ensures quick access and accuracy, enhancing the body's ability to provide timely and precise responses, ultimately leading to an improved customer experience.

C. User Interaction Process

Effective user interaction is pivotal for the success of a restaurant bot, as it ensures that users can easily obtain the information they need to enhance their dining experience. The initial development focuses on creating a voice-based interface where users can input their queries and receive accurate textual responses. This interface allows for straightforward communication, making it accessible to a wide range of users. By providing clear and concise answers to questions about menu specials, costs, ingredients, spiciness, and calorie content, the bot enhances the overall dining experience. This voice-based approach lays a solid foundation for the bot's functionalities and user experience. Our bot not only takes user queries as input but also proactively asks customers if they would like to place an order. If the customer chooses to order, the bot can take down the order details and calculate the total bill. This allows customers to review their order and the total cost, giving them the flexibility to decide whether to place the order or not. Utilizing the 'speech recognition' library[9], the bot can convert spoken language into text, allowing users to interact with the bot using their voice. This feature makes the

bot more intuitive and user friendly, especially for users who prefer voice commands over typing.

D. Processing and Data Handling

The core functionality of our voice bot revolves around handling user queries and facilitating a seamless ordering process. When a user interacts with the bot, their queries are taken as input and processed to extract relevant information. This is achieved by utilizing the JSON file format to store the restaurant menu. The bot reads and extracts data from the JSON file, processes it, and provides the required output in the form of speech.

In scenarios where a customer wishes to place an order, the bot extends its functionality to handle this process efficiently. Upon receiving the order details from the customer, the bot stores the item names and their respective quantities in a dictionary. Initially, we faced a challenge as the bot stored the quantity values as strings, which complicated the calculation of the total bill. This issue arose because the string format of the quantities made it difficult to perform arithmetic operations needed to sum up the total cost of the order. To address this challenge, we implemented a solution where the quantities stored as strings were mapped to integers within the dictionary. This mapping enabled the bot to seamlessly calculate the total bill by converting the string values to integers, allowing for accurate arithmetic operations. By trans forming the data in this manner, we ensured that the bot could efficiently compute the total cost of the order, providing the customer with a precise bill.

The entire process begins when the bot takes the user's query as input. For example, if a customer asks about the available dishes, the bot retrieves this information from the JSON file, processes it, and responds with the relevant details. If the customer decides to place an order, they simply inform the bot of their choices. The bot then records each item and its quantity in a dictionary. By converting the quantity values from strings to integers, the bot can easily perform the necessary calculations to determine the total bill.

Once the total bill is calculated, the bot presents it to the customer, allowing them to review and decide whether to proceed with the order. This level of interaction not only enhances the customer experience by providing a smooth and efficient ordering process but also ensures accuracy in billing, which is crucial for maintaining customer trust and satisfaction.

E. Development Process

Our voice bot is designed to streamline the ordering process by efficiently calculating the total bill using a combination of a dictionary and the JSON file containing the menu details. When a customer places an order, the bot records the item names and quantities in a dictionary. It then parses the JSON file to retrieve the cost associated with each item. By multiplying the cost of each item by its quantity and summing these values, the bot calculates the total bill accurately. This method ensures that all items are accounted for correctly, and the customer receives a precise total. Once the bill is calculated, the bot prepares an output message displaying the total amount, ready to be communicated to the customer in a clear and concise manner.

However, the bot also includes functionality to handle scenarios where a customer orders an item not listed on the menu. If the customer requests an item that does not exist in the JSON file, the bot immediately recognizes the discrepancy. In such cases, the bot generates an error message stating, "One or more items not found on the menu." This prompt ensures that the customer is aware of the invalid items and can adjust their order accordingly. By incorporating this feature, the bot maintains the integrity of the ordering process and enhances the overall user experience by preventing potential errors and confusion related to unavailable menu items.

F. Text-to-Speech and Display Final Output

In the context of a speech-to-speech voice bot for a restaurant, the integration of text-to-speech (TTS) technology, specifically utilizing the Google Text-to-Speech (gTTS) library, plays a crucial role in enhancing user interaction and customer service efficiency. gTTS facilitates the conversion of dynamically generated text responses into natural and intelligible speech, which is essential for maintaining seamless and effective communication in a fast-paced restaurant environment. This technology enables the voice bot to accurately and promptly convey information related to menu options, order status, and customer inquiries, thereby streamlining the ordering process and reducing wait times. The implementation of gTTS allows for multilingual and accent-appropriate responses, catering to a diverse clientele and ensuring accessibility and inclusivity. The final speech output is synthesized in real-time, contributing to a highly interactive and user friendly experience. The generated bill is displayed on the screen. It is also converted to speech and this audio file is played. This research underscores the importance of advanced TTS systems in modernizing restaurant operations, improving customer satisfaction, and optimizing service delivery.

III. DEVELOPMENT ENVIRONMENT

Creating a robust development environment is crucial for building an efficient and reliable restaurant bot. Utilizing powerful Integrated Development Environments (IDEs) like Visual Studio Code and PyCharm ensures that coding, debugging, and testing processes are streamlined and effective. Raspberry Pi Imager is a tool utilized to efficiently flash operating system images onto SD cards for Raspberry Pi devices. It streamlines the process by offering an intuitive interface for selecting, downloading, and writing disk images, ensuring optimal compatibility and performance. Visual Studio Code offers a lightweight and highly customizable platform with extensive plugin support, making it ideal for quick iterations and real-time collaboration. Its intuitive interface and rich features enable developers to write code efficiently and manage projects seamlessly.

Additionally, Visual Studio Code supports various programming languages, including Python, allowing developers to work on multiple projects within the same environment. PyCharm, on the other hand, provides a more comprehensive suite of tools tailored for Python development. With features such as advanced code analysis, project navigation, and integrated debugging capabilities, PyCharm offers a robust environment for building complex applications. Its intelligent code editor helps identify errors and suggest improvements, enhancing code quality and productivity. Additionally, the 'scikit-learn' library provides robust tools for implementing machine learning models, such as the 'GradientBoostingRegressor'. This library supports tasks like predictive modelling and data splitting using 'train test split', which are essential for training and validating the model's performance. By leveraging scikit-learn, developers can incorporate machine learning capabilities into the restaurant bot, enhancing its functionality and intelligence.

IV. HARDWARE DETAILS

The restaurant voice assistant bot running on a Raspberry Pi provides great service to the customer by effective voice interactions. It integrates a Portronics Bluetooth speaker with a microphone that covers a 360-degree voice coverage along with excellent voice pickup that can cut through all other noise in the

environment and ensures crystal-clear communication in the most highly occupied environments. A SanDisk 64GB memory card aids the system to store data and the Raspberry Pi 4 Model B provides the power necessary for the system to learn and then respond to the series of questions that the customers will send in with NLP that is provided by the SpeechRecognition library. The system will enable the other smart appliances in the restaurant to be integrated, thus making it easy, fast to work, and establishing customized, efficient dining experiences.

Being an integrated foundation, Raspberry Pi 4 Model B is equipped with a 64-bit quad-core ARM Cortex-A72 processor running at 1.5GHz and up to 8GB LPDDR4 RAM, capable of multitasking and other resource-heavy applications. It is coupled with dual-band Wi-Fi, Bluetooth 5.0, Gigabit Ethernet, as well as two USB 3.0 ports that further facilitate high-speed data transfer as well as networking. The Portronics speaker has added an upgraded experience with a Bluetooth 5.1 interface which enables reliable audio streaming of voice commands and music with high-quality sound. Altogether, these components present an advanced solution for optimizing restaurant service and therefore enhancing customer experience.

V. CONCLUSION

In conclusion, the voice assistant bot for restaurants, developed using Python modules such as SpeechRecognition, subprocess, and gTTS on a Raspberry Pi, significantly en hances customer service and operational efficiency. The bot effectively manages customer queries, order placements, and bill generation, achieving a speech recognition accuracy of 95%. The use of JSON files for data handling ensures efficient management and retrieval of items from the menu. The integration of advanced speech technologies into a cost effective and accessible platform highlights its potential for widespread adoption in the hospitality industry. This system not only streamlines the dining process but also fosters a more inclusive and engaging environment for customers. Future work could focus on improving speech recognition accuracy, adding multilingual support, and expanding functionalities to meet diverse customer needs, ensuring continuous innovation and adaptability in dynamic restaurant settings. Speech-to speech-based restaurant bots contribute to inclusivity and accessibility within the dining experience. These bots can be programmed to recognize and respond to multiple languages and dialects, making dining more accessible to non-native speakers and tourists. They can also cater to individuals with disabilities, such as those with visual or hearing impairments, by providing tailored interactions that address their specific needs. This inclusivity ensures that a broader spectrum of customers can enjoy a seamless dining experience, thereby expanding the restaurant's customer base and enhancing its reputation as an inclusive and forward-thinking establishment.

REFERENCES

[1] K. Babu, M. S. Roobini, P. Gopirajan, S. Sadagopan, R. Anand and B. S. Rohit, "Development and Evaluation of a Speech-Recognized Voice Bot," 2023 International Conference on Innovative Computing, Intelli gent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICSES60034.2023.10465431.

[2] V. M. Reddy, T. Vaishnavi and K. P. Kumar, "Speech-to-Text and Text-to-Speech Recognition Using Deep Learning," 2023 2nd In ternational Conference on Edge Computing and Applications (ICE CAA), Namakkal, India, 2023, pp. 657-666, doi: 10.1109/ICE CAA58104.2023.10212222.

[3] D. C. Tran, M. K. A. Ahamed Khan and S. Sridevi, "On the Training and Testing Data Preparation for End-to-End Text-to-Speech Application," 2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 2020, pp. 73-75, doi: 10.1109/ICS GRC49013.2020.9232605

[4] S. Meshram, N. Naik, M. VR, T. More and S. Kharche, "Con versational AI: Chatbots," 2021 International Conference on Intelligent Technologies (CONIT), Hubli, India, 2021, pp. 1-6, doi: 10.1109/CONIT51480.2021.9498508.

[5] S. K. Maher, S. G. Bhable, A. R. Lahase and S. S. Nimbhore, "AI and Deep Learning-driven Chatbots: A Comprehensive Analysis and Application Trends," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2022, pp. 994-998, doi: 10.1109/ICICCS53718.2022.9788276.

[6] E. Shihab, S. Wagner, M. A. Gerosa, M. Wessel and J. Cabot, "The Present and Future of Bots in Software Engineering," IEEE Software, vol.

39, no. 5, pp. 28-31, Sept.-Oct. 2022, doi: 10.1109/MS.2022.3176864

[7] I. Spivak, S. Krepych, M. Litvynchuk and S. Spivak, "Validation and Data Processing in JSON Format," IEEE EUROCON 2021- 19th International Conference on Smart Technologies, Lviv, Ukraine, 2021, pp. 326-330, doi: 10.1109/EUROCON52738.2021.9535582.

[8] R. Garg et al., "NLP Based Chatbot for Multiple Restaurants," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), MORADABAD, India, 2021, pp. 439-443, doi: 10.1109/SMART52563.2021.9676218.

[9] B. G. Mamatha Bai, P. Prerana, A. A. Benki and R. Rakesh, "Virtual Medical Assistant in English and Kannada Languages," 2023 International Conference on Applied Intelligence and Sustain able Computing (ICAISC), Dharwad, India, 2023, pp. 1-7, doi: 10.1109/ICAISC58445.2023.10199524.

[10] R. P, K. Suresh, B. M, M. Gokul, G. D. Kumar and A. R, "IoT Based Voice Assistant using Raspberry Pi and Natural Language Pro cessing," 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-4, doi: 10.1109/ICPECTS56089.2022.10046890.