Department of Computer Engineering
K. K. Wagh Institute of Engineering Education and Research
Hirabai Haridas Vidyanagari, Amrut Dham, Panchavati, Nashik-422003

Mini-Project Report

TE Computer **Semester: VI**

Academic Year: 2022-23 Programme: Computer Engineering
Course: LP-II Course Code: 310258
Class: TE-A Date of Report: 18/05/2023

Title of Mini-Project: AI Voice assistant integrated with ChatGPT.

1. Objectives:

- Develop an AI voice assistant that integrates with ChatGPT to provide conversational interactions.
- Enhance user experience by allowing natural language voice commands and responses.
- Enable the voice assistant to perform various tasks such as answering questions, providing information, setting reminders, and controlling smart devices.
- Improve the system's accuracy and efficiency through continuous learning and updates.
- Create a user-friendly interface for seamless voice interactions

2. Problem Statement:

The objective of this project is to develop an AI voice assistant that can understand and respond to voice commands in natural language. Integrating it with ChatGPT will enable more sophisticated conversational interactions, providing users with a versatile and intelligent voice-based assistant. The system should be able to handle various tasks such as answering questions, providing recommendations, and performing actions based on user requests.

3. Problem Definition:

The problem addressed in this project is the need for an AI voice assistant integrated with ChatGPT. Traditional voice assistants often lack the ability to engage in natural and dynamic conversations, limiting their usefulness and user experience. Additionally, they may struggle with understanding complex queries or providing contextually relevant responses. This project aims to overcome these limitations by integrating a voice assistant with ChatGPT, a powerful language generation model.

The specific problems to be addressed include:

1. Limited conversational capabilities: Existing voice assistants often provide predefined responses or struggle to engage in meaningful conversations beyond basic commands. This limits their ability to understand and respond to user queries effectively.

2. Lack of context awareness: Voice assistants may fail to maintain context across multiple user inputs, leading to disjointed conversations and inaccurate responses. This can be frustrating for users and hinder the usefulness of the voice assistant.

3. Inability to handle complex queries: Traditional voice assistants may struggle to understand and respond to complex queries or requests that require deeper understanding or contextual knowledge. This limits their ability to provide accurate and relevant information to users.

4. Limited language generation capabilities: While voice assistants can understand speech, they may lack sophisticated language generation capabilities, resulting in less engaging and coherent responses. Users may feel disconnected or unsatisfied with the voice assistant's interactions.

By integrating ChatGPT, a state-of-the-art language model, with the voice assistant, these problems can be addressed. ChatGPT's advanced language generation capabilities and contextual understanding can enhance the voice assistant's ability to engage in natural conversations, provide accurate responses, and generate contextually relevant information. This integration aims to create an AI voice assistant that can understand and respond to a wide range of user queries and provide a more intuitive and satisfying conversational experience.

4. Software/Hardware Requirements:

- ChatGPT: The AI language model should be implemented using an API or library that provides access to the ChatGPT model.
- Speech Recognition: A speech recognition system is required to convert user voice inputs into text for processing.
- Text-to-Speech (TTS): A TTS engine is needed to convert the system's responses into voice output.
- Microphone: A microphone is required to capture user voice inputs.
- Speaker: A speaker is needed to play the voice output to the user.
- Processing Unit: Sufficient computational resources are necessary to run the AI voice assistant system smoothly.

5. Theory:

The AI voice assistant integrated with ChatGPT combines several technologies to enable voice-based interactions and intelligent conversations:

1. Natural Language Understanding (NLU): NLU is a key component of the voice assistant, responsible for extracting meaning and intent from user queries. It involves techniques such as intent classification, entity recognition, and sentiment analysis. NLU allows the voice assistant to understand user commands and identify the specific actions or information being requested.

2. Dialog Management: Dialog management is crucial for maintaining context and handling multi-turn conversations. It involves keeping track of the conversation history, managing user intents and responses, and ensuring smooth transitions between user queries and the voice assistant's replies. Dialog management enables a more natural and interactive conversation flow.

3. Speech Recognition: Speech recognition technology is employed to convert user voice inputs into text. It involves acoustic and language models that analyze audio signals and transcribe them into written form. Accurate speech recognition is vital for the voice assistant to understand and process user commands effectively.

4. Contextual Understanding: Contextual understanding refers to the ability of the voice assistant to interpret user queries in the context of the ongoing conversation. It involves analyzing the conversation history, user preferences, and previous interactions to generate contextually relevant responses. Contextual understanding enhances the voice assistant's ability to provide accurate and personalized information.

5. Language Generation: Language generation plays a crucial role in the voice assistant's ability to generate coherent and contextually appropriate responses. ChatGPT, as a powerful language model, uses techniques such as deep learning and natural language processing to generate human-like text based on user queries and context. The integration of ChatGPT allows the voice assistant to leverage its language generation capabilities, resulting in more engaging and informative responses.

6. User Feedback and Adaptation: The voice assistant can benefit from user feedback and adapt to individual preferences. By incorporating a feedback loop, the system can learn from user interactions, refine its responses, and improve over time. User feedback can be used to train and update the underlying models, ensuring the voice assistant continues to enhance its performance and cater to user needs.

7. Privacy and Data Security: As an AI voice assistant, privacy and data security are paramount. It is essential to handle user data responsibly, ensuring that personal information is protected and used only for the intended purposes. Implementing secure data storage, encryption, and compliance with privacy regulations are crucial aspects of developing a trustworthy voice assistant system.

8. Integration with External Services: The voice assistant can be integrated with external services and APIs to expand its capabilities. This includes integrating with weather services, mapping services, music streaming platforms, and other relevant APIs to provide comprehensive and diverse functionality to users.

By understanding and implementing these theoretical aspects, the AI voice assistant integrated with ChatGPT can offer a sophisticated and intelligent conversational experience, allowing users to interact naturally and receive accurate and contextually relevant information and assistance.

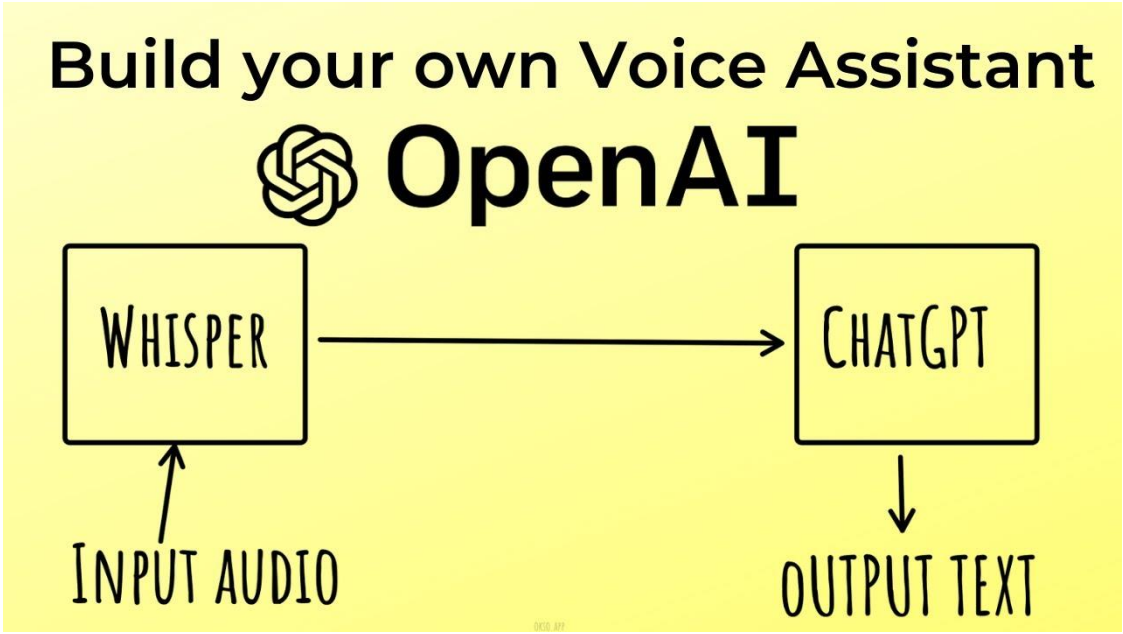


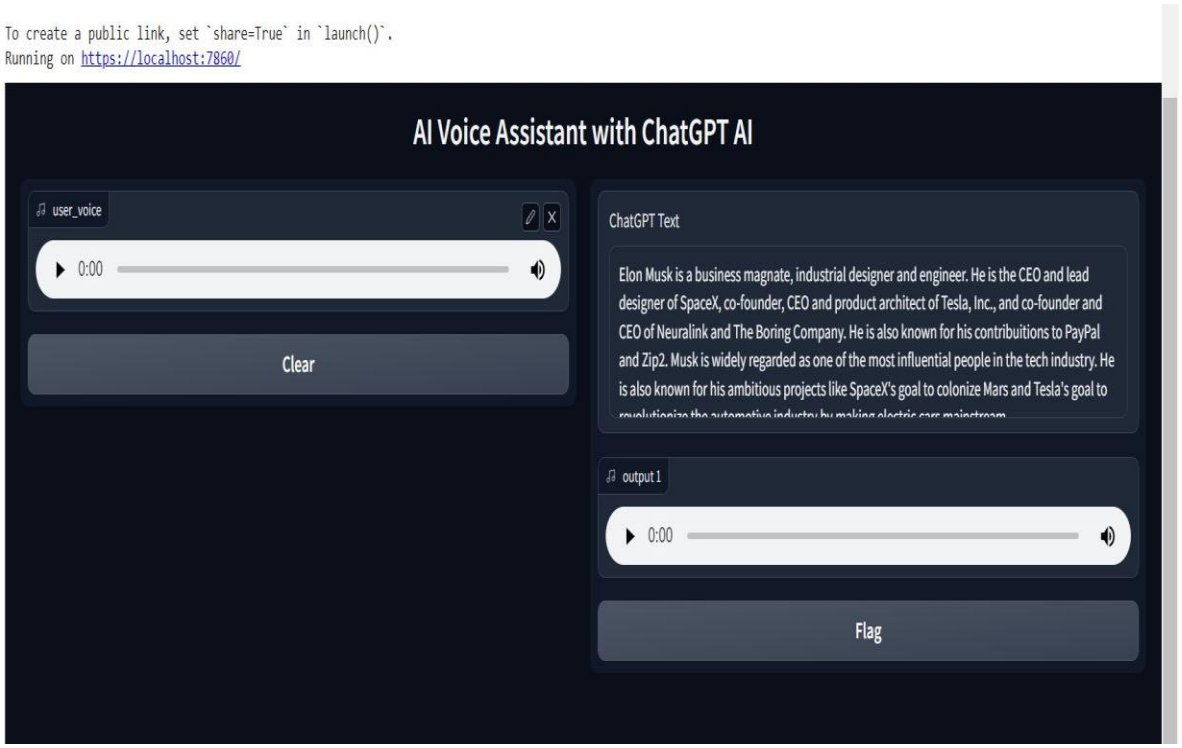
Fig 1.1: Flow diagram of System.

6. Outputs of the Mini-Project:

1. Code:

```
[ ]  def voice_chat(user_voice):

        messages = [
        {"role": "system", "content": "You are a kind helpful assistant."},
        ]


        user_message = model.transcribe(user_voice)["text"]

        #reply = user_message

        messages.append(
            {"role": "user", "content": user_message},
        )

        print(messages)

        chat = openai.ChatCompletion.create(
            model="gpt-3.5-turbo", messages=messages
        )

        reply = chat.choices[0].message.content

        messages.append({"role": "assistant", "content": reply})

        tts.tts_to_file(text=reply, file_path="output.wav",)

        return(reply, 'output.wav')
```

2. Execution:



7. Test Cases:

Test Case ID	KKW_001	**Test Case Description**	AI voice assistant integrated with ChatGPT.		
Created By	Atharva	**Reviewed By**	Prof. D. M. Kanade	**Version**	1.1

QA Tester's Log	AI voice assistant integrated with ChatGPT is being worked accurately in version 1.1

Tester's Name	Atharva	**Date Tested**	19-05-2023	**Test Case (Pass/Fail/Not Executed)**	Pass

S #	Prerequisites:
1	Knowledge of ChatGPT working
2	Software testing skills
3	TTS, whisper and chatGPT API's- Turbo 3.5

S #	Test Data
1	Userid = admin
2	Pass = admin
3	

Test Scenario	Verify asking different question and test if it is converting speech to text and text to speech properly.

Step #	Step Details	Expected Results	Actual Results	Pass / Fail / Not executed / Suspended
1	What is the capital of France?	Paris	As Expected,	Pass
2	Who wrote 'Harry Potter'?	Rowling	As Expected,	Pass
3	What is the square root of 64?	8	As Expected,	Pass
4	Play me a song?	Lyrics must be provided.	Unexpected Json error.	Fail
5	What is the color of the sky?	Blue	As Expected,	Pass
6	Who is the current president of the United States?	Biden	As Expected,	Pass
7	What is the chemical symbol for gold?	Au	As Expected,	Pass
8	What is the largest planet in our solar system?	Jupiter	As Expected,	Pass
9	What is the capital of Australia?	Answer should be correct.	As Expected,	Pass
10	Check if the system properly notifies participants of any changes or cancellations to an event.	Canberra	As Expected,	Pass
11	Who painted the Mona Lisa?	Da Vinci	As Expected,	Pass
12	What is the symbol for hydrogen?	H	As Expected,	Pass

8. Applications of Mini Project:

- Personal Assistant: Users can rely on the voice assistant to manage their daily tasks, such as setting reminders, scheduling appointments, and sending messages. For example, a user can say, "Remind me to pick up groceries at 6 PM," and the voice assistant will set a reminder accordingly.

- Information Retrieval: The voice assistant can provide quick access to information on various topics. Users can ask questions like, "What's the latest news?" or "Who won the Oscar for Best Actor this year?" and receive up-to-date and relevant answers.

- Language Translation: The voice assistant can assist with language translation. Users can ask for translations of phrases or sentences in different languages. For instance, a user may say, "Translate 'Where is the nearest restaurant?' to Spanish," and the voice assistant will provide the translated phrase.

- Entertainment Recommendations: Users can seek entertainment recommendations from the voice assistant. For example, they can ask, "What are some popular movies to watch?" or "Can you recommend a good book?" and receive personalized recommendations based on their preferences.

- Smart Home Control: The voice assistant can integrate with smart home devices to control various functions. Users can issue voice commands to turn on/off lights, adjust room temperature, play music, or even lock/unlock doors. For example, a user may say, "Turn off the living room lights" or "Set the thermostat to 72 degrees."

- General Knowledge: The voice assistant can answer general knowledge questions, such as definitions, historical facts, or trivia. Users can ask questions like, "What is the capital of Australia?" or "Who painted the Mona Lisa?" and get informative responses.

- Personalized Recommendations: Over time, the voice assistant can learn user preferences and provide personalized recommendations. For instance, it can suggest movies based on a user's previous watching habits or recommend restaurants based on their culinary preferences.

- Conversational Interactions: The voice assistant can engage in natural and dynamic conversations with users, providing responses and engaging in back-and-forth exchanges on various topics. Users can discuss current events, ask for opinions, or simply engage in friendly banter.

  These real-life examples showcase the versatility and usefulness of an AI voice assistant integrated with ChatGPT, demonstrating how it can assist users in multiple aspects of their daily lives and provide a seamless and intuitive conversational experience.

9. Conclusion:

In conclusion, the development of an AI voice assistant integrated with ChatGPT offers significant benefits in providing an enhanced and natural conversational experience. By leveraging speech recognition, natural language understanding, dialog management, and ChatGPT's language generation capabilities, the voice assistant can understand and respond to user voice commands effectively.

The integration with ChatGPT allows for context-aware responses, enabling more engaging and intelligent conversations with the voice assistant. Users can interact with the system using natural language voice queries, enabling a more intuitive and seamless user experience.

Through the implementation of various functionalities such as information retrieval, task execution, and device control, the AI voice assistant becomes a versatile and helpful companion for users. Whether it's answering questions, setting reminders, providing recommendations, or performing actions based on user instructions, the voice assistant aims to assist users in their day-to-day tasks and enhance their productivity. The continuous learning aspect of the system ensures that it evolves and improves over time. By leveraging user feedback and

interactions, the voice assistant can adapt to individual preferences, learn from user habits, and refine its responses. This iterative learning process allows for a personalized and tailored experience for each user.

However, it is important to note that the performance and accuracy of the voice assistant depend on the underlying AI models, the quality of speech recognition, and the integration with ChatGPT. Ongoing updates and improvements to these components are crucial to ensure the voice assistant's effectiveness and reliability.

In summary, the development of an AI voice assistant integrated with ChatGPT opens up new possibilities for natural language voice interactions. By combining advanced technologies, the voice assistant aims to provide users with a conversational and intelligent experience, making tasks more convenient and enjoyable. The continuous learning and refinement of the system contribute to its ability to adapt and cater to the unique needs of individual users.

10.Course Outcomes Achieved (COs):

- CO3: Design and develop an expert system

11. Name of Group Members:

Roll No.	Name of Students	Exam Seat No.
71	Atharva Thakkar	
72	Aman Vohra	
56	Yash Patil	
63	Pushkaraj Salunke	

Date: 20-05-2023 Evaluated by: Dated Signature of Guide: _____

Name of Guide: Prof. D. M. Kanade