AI VOICE ASSISTANT

A MINI PROJECT REPORT

Submitted by

Aadarsh Joshi [RA2011026010061]
Sarvesh Ahuja [RA2011026010095]
Anindya Shankar Dasgupta [RA2011026010120]

Under the guidance of
Dr. M. S. Abirami
(Assistant Professor, CINTEL)

In partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING
With specialization in AI and Machine Learning

SCHOOL OF COMPUTING

COLLEGE OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR - 603203

APRIL 2023

SRM
INSTITUTE OF SCIENCE & TECHNOLOGY
Deemed to be University u/s 3 of UGC Act, 1956

COLLEGE OF ENGINEERING & TECHNOLOGY
SRM INSTITUTE OF SCIENCE & TECHNOLOGY
S.R.M. NAGAR, KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that Mini project report titled **"AI VOICE ASSISTANT"** is the bona fide work of

Aadarsh Joshi [RA2011026010061]
Sarvesh Ahuja [RA2011026010095]
Anindya Shankar Dasgupta [RA2011026010120]

who carried out the minor project under my supervision. Certified further, that to the best of my knowledge, the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

4/5/23

Signature

Dr. M. S. Abirami
Assistant Professor
Department of Computational Intelligence

Signature

Dr. R. Annie Uthra
Head of Department
Department of Computational Intelligence

ABSTRACT

This project aims to implement an AI-based chatbot for therapy using the ChatGPT API. The chatbot should be able to **understand natural language** input and provide appropriate **feedback** and guidance to users seeking therapy or counseling. The project presents several challenges, including understanding the **nuances** of human conversation, tailoring responses to individual user needs and preferences, and ensuring user privacy and confidentiality.

However, it also presents opportunities to provide **accessible** and **affordable** mental health support to a wider audience, reducing the stigma surrounding mental health and making it easier for individuals to seek support when they need it.

INDEX

INTRODUCTION

What is AI?

AI stands for Artificial Intelligence, which refers to the development of computer systems that can perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. AI involves the use of algorithms and statistical models to analyze and process data, and to recognize patterns and make predictions.

AI systems can be categorized into various types, including rule-based systems, machine learning, natural language processing, and computer vision, among others. AI has numerous applications in various industries, such as healthcare, finance, transportation, and manufacturing, among others.

How does AI work?

In order to make predictions or decisions, AI uses algorithms and statistical models to analyze and learn from data. The following steps are typically included in the procedure:

1. *Collection of data*: Gathering important information from different sources.
2. *Preprocessing of data*: Data must be cleaned and transformed so that it can be analyzed correctly.
3. *Extracting attributes*: using the data to find relevant features or variables for the model.
4. *Building models*: constructing a mathematical model that is capable of learning from the data and making decisions or predictions.
5. *Training*: analyzing the data with the model and adjusting its parameters to reduce errors and improve accuracy
6. *Testing*: evaluating the model's performance on new data to determine its ability to generalize beyond the training data.
7. *Deployment*: putting the model into action in the real world to make predictions or decisions.

What is Chatbot and how does it work?

A chatbot is a computer program that pretends to talk to real people, especially over the internet. It utilizes regular language handling (NLP) and AI (ML) calculations to comprehend and answer client questions or orders.

Using a pre-defined set of rules or machine learning models, chatbots generate an appropriate response after analyzing user input and determining the message's intent. Some chatbots utilize a standard based approach where explicit reactions are customized for explicit catchphrases or expressions. Others use ML algorithms to improve their responses over time and learn from previous conversations.

Chatbots can be incorporated into a variety of platforms, including websites, messaging applications, and virtual assistants, making it simpler for users to gain access to information or services without the assistance of a human. They are frequently utilized in the healthcare, e-commerce, and customer service industries to provide assistance, respond to questions, or automate routine tasks.

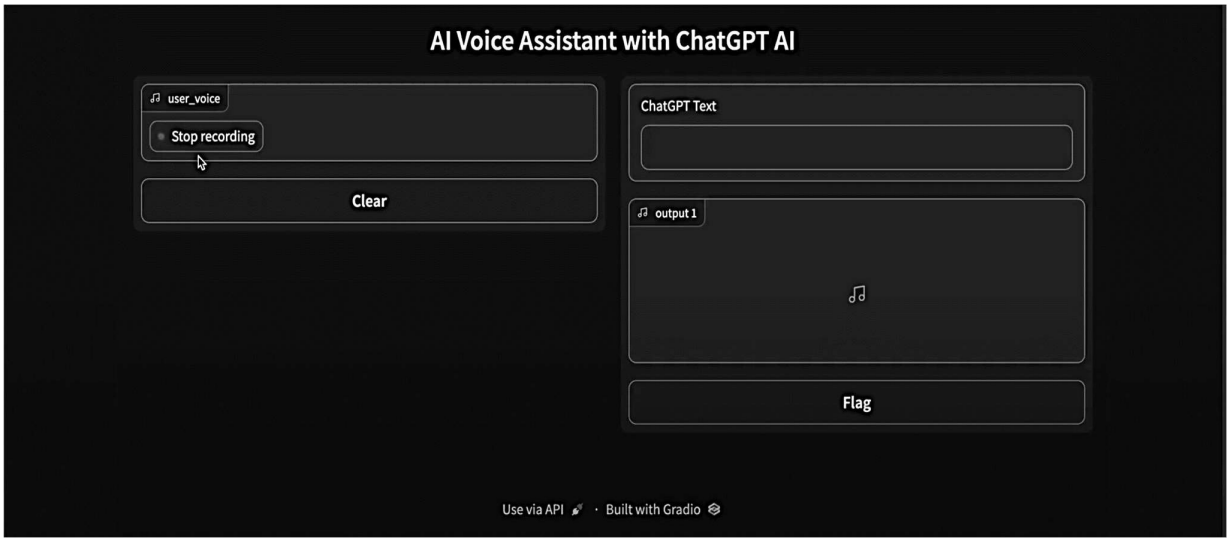
WORKING OF OUR MODEL

The device's microphone receives the audio input from a voice assistant and transmits it to the cloud for processing. The NLP algorithm in the cloud first converts the audio signal into a digital signal before analyzing it to comprehend the user's intent and extract relevant information.

The NLP calculation utilizes AI models, like profound brain organizations, to recognize designs in the sound information and characterize it into explicit purposes or activities, like setting an update, playing music, or getting climate data.

The voice assistant responds to the user through the speaker of the device by accessing external services or devices, such as music streaming services, smart home devices, or search engines, once the user's intent has been understood.
Machine learning models are also used by voice assistants to learn from user interactions and improve over time in terms of accuracy and performance. They are also able to integrate with other AI technologies, such as natural language generation and computer vision, to provide a user experience that is more individualized and seamless.

Our chatbot is an AI-based system that enables users to interact with the latest AI tools like ChatGPT through a voice-based interface. Users can activate the chatbot using the Whisper API from OpenAI, and interact with it using natural language.
The chatbot utilizes the Gradio API to provide a user-friendly interface, allowing users to easily input and receive information.
The chatbot's key features include text-to-speech conversion and voice-to-text conversion, enabling seamless communication between the user and the system.


AI Voice Assistant with ChatGPT AI
user_voice
Stop recording
Clear
ChatGPT Text
output 1
Flag
Use via API
Built with Gradio


user_voice
0:00
Clear
ChatGPT Text
SRM University is a private university located in Chennai, Tamil Nadu, India. Founded in 1985, it offers undergraduate, postgraduate and doctoral programs in Engineering, Management, Medicine and Health Sciences, Arts and Humanities, and Science and Technology. It has four campuses situated in Kattankulathur, Vadapalani, Ramapuram and NCR (National Capital Region) Ghaziabad. The university has been accredited by the University Grants Commission (UGC) and has also been awarded a 'Category 1' status by the Ministry of Human Resource Development, Government of India.
output 1
0:35
Flag

SOCIAL RELATED PROBLEM TO BE SOLVED

• Lack of **accessibility** to mental health services, especially in rural or remote areas

• Long **wait times** for therapy appointments, causing delays in receiving support

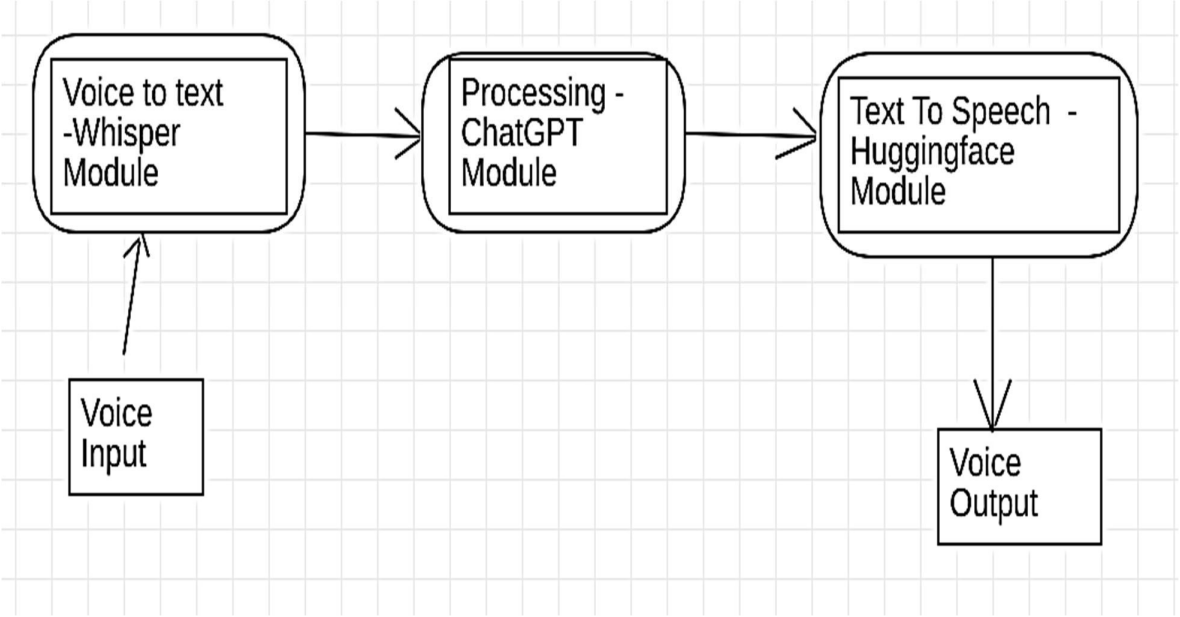
• **Stigma** surrounding mental health and seeking therapy, preventing individuals from accessing services

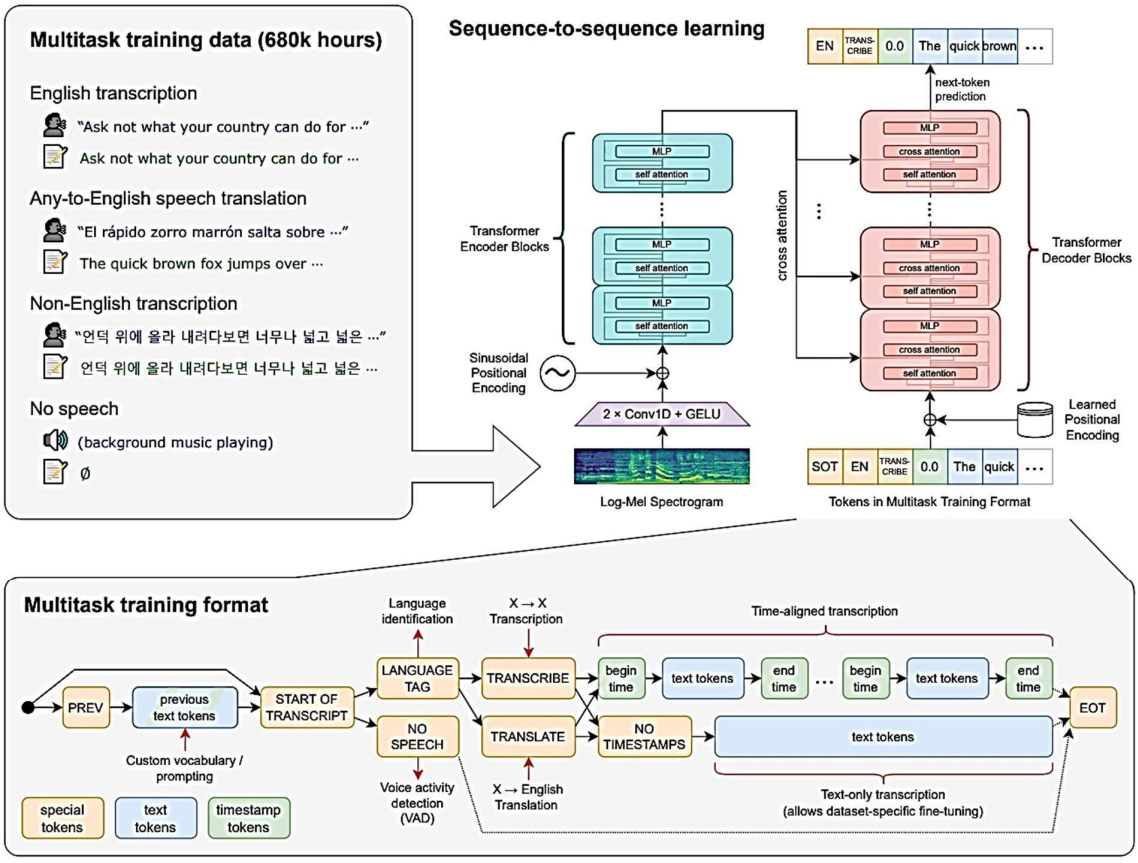
• **High costs** associated with traditional therapy services, making them inaccessible to many individuals

• Limited **availability** of therapy services, especially for those who work during typical office hours.

• Limited **knowledge** and **awareness** of mental health and therapy services among the general population.

ARCHITECTURE DIAGRAM



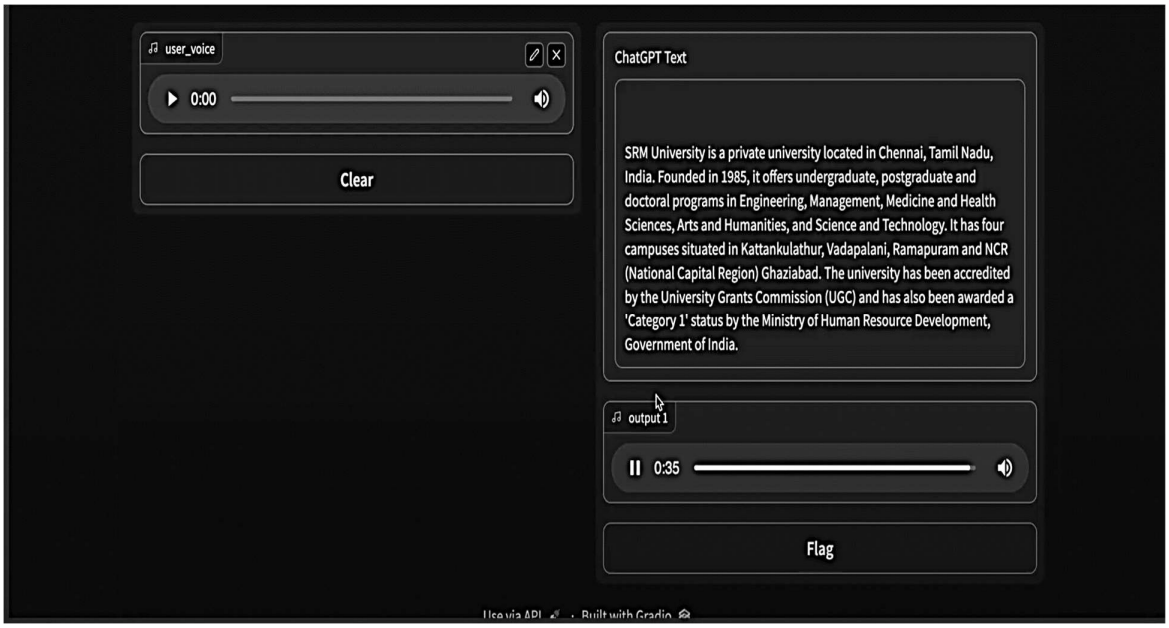


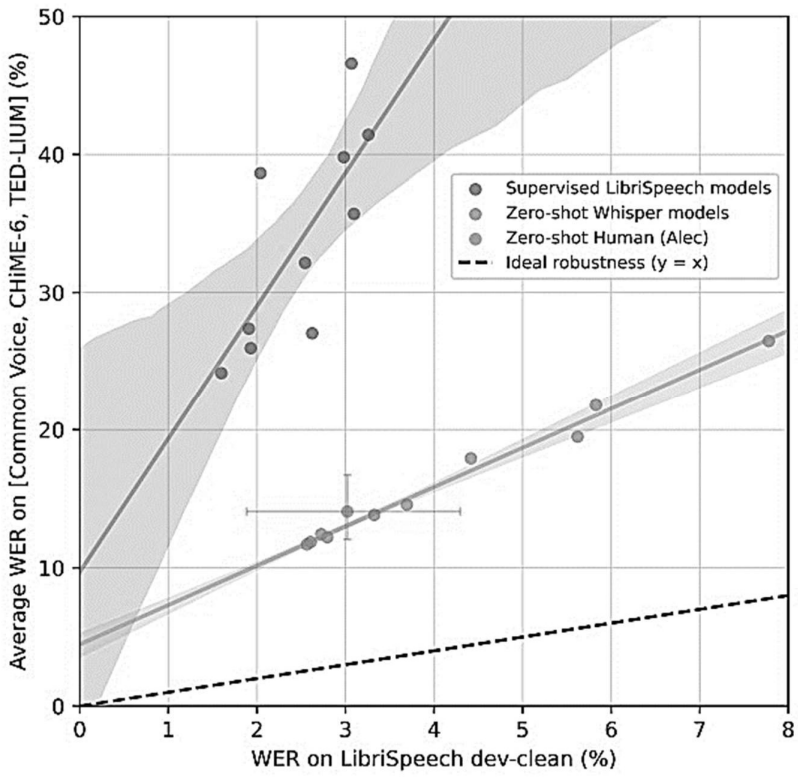
CODE

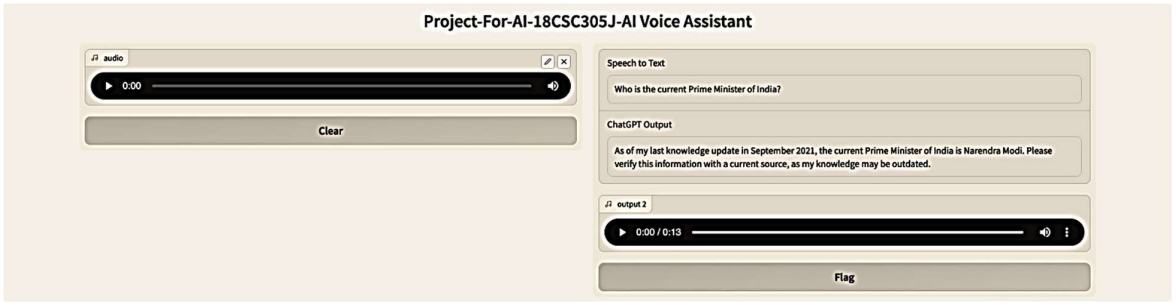
```
!pip install -q TTS
!pip install -q openai-whisper
!pip install -q gradio
!pip install -q openai
import whisper
import gradio as gr
import openai
from TTS.api import TTS
TTS.list_models()
model_name = TTS.list_models()[9]
tts = TTS(model_name)
model = whisper.load_model("medium")
openai.api_key = ''
def listen(user_voice):
    user_message = model.transcribe(user_voice)["text"]
    return user_message
def wake_bot(user_voice):
    wake_word = 'hello ava'
    is_wake = False
    while True:
      init_input = listen(user_voice)
      if(init_input.lower() == 'hello ava'):
        voice_chat(voice_chat)
'''
Input: user audio text
Output: Bot audio text
'''
def voice_chat(user_voice):
    messages = [
    {"role": "system", "content": "You are a kind helpful assistant to help
out the users"},
    ]
    u_message = listen(user_voice)
    #reply = user_message
    messages.append(
        {"role": "user", "content": u_message},
    )
    print(messages)
    chat = openai.ChatCompletion.create(
        model="gpt-3.5-turbo", messages=messages
```

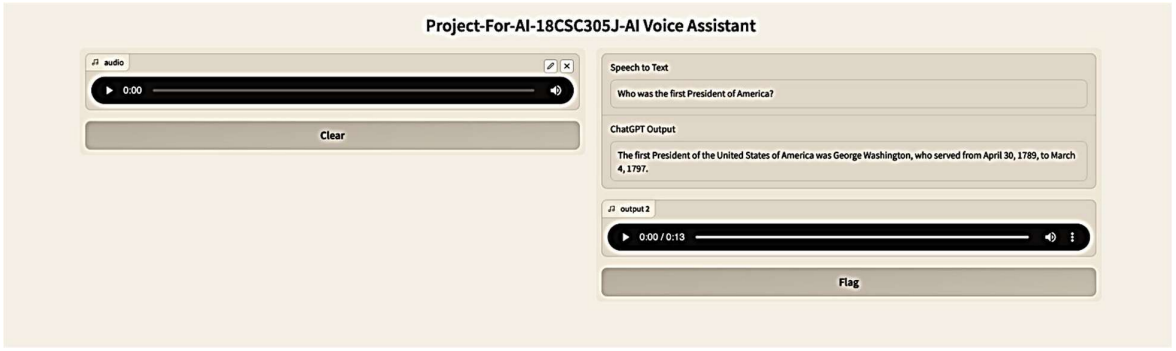
```
    )
    reply = chat.choices[0].message.content
    messages.append({"role": "assistant", "content": reply})
    tts.tts_to_file(text=reply, file_path="output.wav")
    return(reply, 'output.wav')
text_reply = gr.Textbox(label="ChatGPT Text")
voice_reply = gr.Audio('output.wav')
gr.Interface(
    title = 'Voice Based Chatbot',
    fn=voice_chat,
    inputs=[
        gr.inputs.Audio(source="microphone", type="filepath")
    ],
    outputs=[
        text_reply,  voice_reply
    ], live = True).launch(debug = True)
```

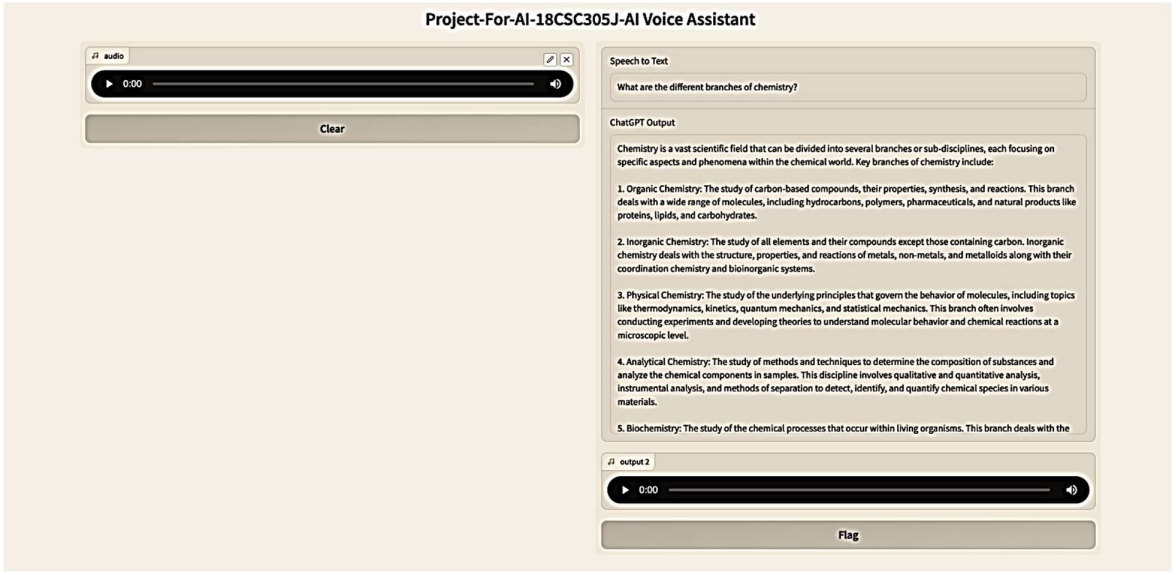
OUTPUT


user_voice
0:00
Clear
ChatGPT Text
SRM University is a private university located in Chennai, Tamil Nadu, India. Founded in 1985, it offers undergraduate, postgraduate and doctoral programs in Engineering, Management, Medicine and Health Sciences, Arts and Humanities, and Science and Technology. It has four campuses situated in Kattankulathur, Vadapalani, Ramapuram and NCR (National Capital Region) Ghaziabad. The university has been accredited by the University Grants Commission (UGC) and has also been awarded a 'Category 1' status by the Ministry of Human Resource Development, Government of India.
output 1
0:35
Flag


Supervised LibriSpeech models
Zero-shot Whisper models
Zero-shot Human (Alec)
Ideal robustness (y = x)
Average WER on [Common Voice, CHiME-6, TED-LIUM] (%)
WER on LibriSpeech dev-clean (%)
0
10
20
30
40
50
0
1
2
3
4
5
6
7
8


Project-For-AI-18CSC305J-AI Voice Assistant
audio
0:00
Clear
Speech to Text
Who is the current Prime Minister of India?
ChatGPT Output
As of my last knowledge update in September 2021, the current Prime Minister of India is Narendra Modi. Please verify this information with a current source, as my knowledge may be outdated.
output 2
0:00 / 0:13
Flag


Project-For-AI-18CSC305J-AI Voice Assistant
audio
0:00
Clear
Speech to Text
Who was the first President of America?
ChatGPT Output
The first President of the United States of America was George Washington, who served from April 30, 1789, to March 4, 1797.
output 2
0:00 / 0:13
Flag


Project-For-AI-18CSC305J-AI Voice Assistant
audio
0:00
Clear
Speech to Text
What are the different branches of chemistry?
ChatGPT Output
Chemistry is a vast scientific field that can be divided into several branches or sub-disciplines, each focusing on specific aspects and phenomena within the chemical world. Key branches of chemistry include:
1. Organic Chemistry: The study of carbon-based compounds, their properties, synthesis, and reactions. This branch deals with a wide range of molecules, including hydrocarbons, polymers, pharmaceuticals, and natural products like proteins, lipids, and carbohydrates.
2. Inorganic Chemistry: The study of all elements and their compounds except those containing carbon. Inorganic chemistry deals with the structure, properties, and reactions of metals, non-metals, and metalloids along with their coordination chemistry and bioinorganic systems.
3. Physical Chemistry: The study of the underlying principles that govern the behavior of molecules, including topics like thermodynamics, kinetics, quantum mechanics, and statistical mechanics. This branch often involves conducting experiments and developing theories to understand molecular behavior and chemical reactions at a microscopic level.
4. Analytical Chemistry: The study of methods and techniques to determine the composition of substances and analyze the chemical components in samples. This discipline involves qualitative and quantitative analysis, instrumental analysis, and methods of separation to detect, identify, and quantify chemical species in various materials.
5. Biochemistry: The study of the chemical processes that occur within living organisms. This branch deals with the
output 2
0:00
Flag


Project-For-AI-18CSC305J-AI Voice Assistant
audio
0:00
Clear
Speech to Text
Write a poem about fuzzy logic in 20 words.
ChatGPT Output
In realms of gray, certainty wanes,
Fuzzy logic blooms, where soft truth reigns,
Boundless thoughts connect, untamed,
Precision fades, yet wisdom gained.
output 2
0:13 / 0:13
Flag

USE CASES

How <u>Our</u> Chatbot Can Help People with Disabilities?

Chatbots can be staggeringly useful for individuals with incapacities. The following are some ways chatbots can help:

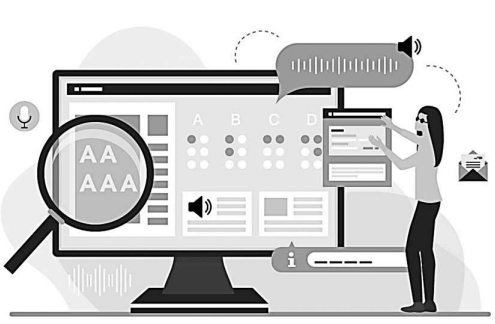
1. *Communication*: Chatbots can be used to interact with others by people with disabilities that affect their ability to speak or communicate effectively. For instance, an individual with a conference hindrance can utilize a chatbot to speak with a meeting individual who may not know communication via gestures.

2. *Assistance*: People with disabilities who may have difficulty independently completing certain tasks can benefit from chatbots' assistance. A chatbot, for instance, could assist a person with a visual impairment in navigating a website or carrying out an online transaction.

3. *Reminders*: Chatbots can be customized to send suggestions to individuals with inabilities, like taking prescription or going to arrangements.

4. *Support from within*: Chatbots can offer profound help to individuals with incapacities who might be battling with emotional wellness issues or different difficulties. They can listen and offer suggestions for resources or coping mechanisms.

5. *Accessibility*: Chatbots can also make it easier for people with disabilities to use them. A chatbot can, for instance, be programmed to provide alternative text for images, making the content more accessible to people who are blind or visually impaired.



MODULES IMPLEMENTED

We have implemented various modules in our chatbot to make it user-friendly and efficient. The modules are as follows:

Voice-to-text module: This module uses the Whisper API of OpenAI to convert the user's voice into text, which the chatbot can process.

Text-to-speech module: This module uses a text-to-speech Huggingface library to convert the chatbot's responses into speech.

ChatGPT module: This module uses the OpenAI ChatGPT API to generate responses to the user's queries or messages.

Gradio interface: This module provides a user-friendly interface for the chatbot, where users can input their queries or messages and get responses in both text and voice formats.

CONCLUSION

Our chatbot has the potential to **significantly improve accessibility** for people with disabilities, enabling them to access the latest AI tools through a more user-friendly and efficient interface.

Future directions for the project include expanding the **chatbot's capabilities** and improving its accessibility, potentially by incorporating more advanced voice recognition and natural language processing techniques.

Though current hardware and software does not ensure hundred percent assurance, but with further growth in technology we have a surety that it will be improved more.

We welcome feedback and engagement from the audience to help further improve the chatbot and its impact on accessibility.

REFERENCES

https://chat.openai.com/

https://www.ibm.com/topics/chatbots

https://en.wikipedia.org/wiki/Artificial_intelligence

https://ieeexplore.ieee.org/document/ai-for-diables