International Islamic University Chittagong

PROJECT REPORT

Course Name: Tools and Technologies for Internet Programming
Course Code: CSE-3532

Submitted By

Name: Nazrana Nahreen
Student ID: C231444
Section: 5BF
Semester: Spring 2025
Date of Submission: 10 May 2025
Project Team Members: Solo

Submitted To

Course Instructor: Sara Karim
Designation: Adjunct Lecturer, Department of CSE, IIUC

PROJECT LAYOUT

Phase 1-Setup the brain of the Doctor (Multimodal LLM)

- *Setup GROQ API key*
- *Convert image to required format*
- *Setup Multimodal LLM*

Phase 2-Setup voice of the patient

- *Setup Audio recorder (ffmpeg & portaudio)*
- *Setup Speech to text-STT-model for transcription*

Phase 3-Setup voice of the Doctor

- *Setup Text to Speech-TTS-model (gTTS & ElevenLabs)*
- *Use Model for Text output to Voice*

Phase 4-Setup UI for the VoiceBot

- *VoiceBot UI with Gradio*

TOOLS AND TECHNOLOGIES
- *Groq for AI Inference OpenAI Whisper (Best open source model for Transcription)*
- *Llama 3 Vision (Open source by Meta)*
- *gTTS & ElevenLabs (Speech to Text)*
- *Gradio for UI*
- *Python*
- *VS Code*