Graphic Era Hill University
DEHRADUN • BHIMTAL • HALDWANI

PROJECT AND TEAM INFORMATION

Project Title

DocBot: AI-Powered Medical Assistance Chatbot

Student/Team Information

Team Name: **Tech Resolutions** Team Mentor: **Mr. Amit Gupta**	**Tech Resolutions**
Tanmay Chauhan (Team Lead) Student ID: 23011566 Email: chauhantanmay253@gmail.com	
Abhay Kanojia Student ID: 230115137 Email: abhaykanojia51@gmail.com	
Ayush Chand Student ID: 230112536 Email: ayushchand862@gmail.com	

PROJECT PROGRESS DESCRIPTION

Project Abstract

DocBot is an intelligent, text-based medical chatbot designed to simulate doctor-patient interactions, offering preliminary guidance on common health concerns with a specialization in neurological conditions such as brain tumors. Leveraging natural language processing and TF-IDF vectorization, DocBot interprets user queries to deliver accurate, understandable medical advice. Its image-based diagnostic capability is now fully functional, integration deep-learning models trained on annotated brain MRI scans to enable preliminary detection of anomalies directly from user-provided images. This dual-modal system improves accessibility to empowering patients with timely, reliable information before consultation with a healthcare professional. Users can upload MRI scans of the brain and receive predictions about the likelihood of a tumor, improving early detection.

Updated Project Approach and Architecture

1. **Frontend Interface:** A responsive web client built with pure HTML, CSS, JavaScript enabling secure text chat with optional file upload.
2. **Backend:** A Flask microservice handling routing user queries to the NLP module or imaging pipeline.
3. **NLP Module:**
    - **Input Preprocessing:** Clean text by removing noise, lowercasing, and stopwords.
    - **TF-IDF Vectorization:** Transforms cleaned text into feature vectors using a persisted TF-IDF model.
    - **Response Generation:** Matches input vectors to doctor-annotated responses via nearest neighbor search.
4. **Imaging Pipeline:** Convolutional neural network trained on labeled MRI scans for tumor detection, deployed as a separate Flask endpoint. Users can upload brain MRI scans and receive instant predictions.
5. **Database:** Libraries used include scikit-learn, Flask, TensorFlow, and JQuery for minimal frontend enhancements.

Tasks Completed

Task Completed	Team Member
Collected and preprocessed doctor-patient dataset Cleaned and tokenized patient text inputs and removed stopwords. Trained and saved TF-IDF vectorizer; integrated model loading in backend. Supported frontend integration by resolving TF-IDF object and JSON serialization issues.	⇨ **Tanmay Chauhan**
Collected and preprocessed brain MRI dataset; applied resizing and normalization Annotated images with tumor labels; vectorized image features Built CNN classifier for tumor detection; completed validation and deployed model	⇨ **Abhay Kanojia**
Developed Flask backend for both text and image handling	⇨ **Abhay Kanojia & Ayush Chand**
Designed and styled chat interface Implemented message rendering and file uploader	⇨ **Ayush Chand**

Challenges/Roadblocks

1. **Limited Healthy MRI Examples:** We had many more tumor scans than healthy ones, so the model sometimes learned tumor patterns too strongly. To fix this, we created new healthy images by rotating and mirroring existing ones.
2. **Varied Patient Descriptions:** People use different words and slang to describe symptoms, so matching their input to the right response was tricky. We plan to add more medical synonyms and clean up terms before processing.
3. **Connection issues b/w Frontend and Backend:** At first, the chat app couldn't talk smoothly to our server because of browser security settings and data format problems. We standardized our replies to always use simple JSON and fixed server settings, so messages and image files pass through without errors.

Tasks Pending

Task Pending	Team Member (to complete the task)
All done	

Project Outcome/Deliverables

- **Text-based Chatbot:** Fully functional Flask API with TF-IDF NLP engine and interactive web UI.
- **Brain Tumor Classifier:** CNN model capable of binary classification (tumor vs healthy) with evaluation metrics.
- **MRI Upload and Prediction Feature:** Fully integrated. Users can upload brain scans and receive predictions.

Progress Overview

This project is now completed. Users can chat with DocBot and upload MRI brain scans to receive accurate predictions. Text-chat functionality is stable, and the CNN-based diagnostic model is fully integrated with the image upload interface. The system has been tested internally for accuracy and reliability. The remaining work involves deployment polishing and minor documentation improvements.

Codebase Information

- **Repository:** https://github.com/tanmaychauhan12/DocBot-AI-Powered-Medical-Assistance-Chatbot
- **Key Commits:** Add TF-IDF model training script, implemented Flask endpoints for text chat, initial CNN model and evaluation scripts.

Testing and Validation Status

Test Type	Status (Pass/Fail)	Notes
Unit Tests (NLP Module)	Pass	95% coverage on preprocessing functions
Integration Tests (API + UI)	Pass	Handles both text and images uploads
CNN Model Validation	Pass	90% accuracy on hold-out MRI test set
End-to-End Chat Simulation Tests	Pass	Verified multi-turn dialogue and image inference

Deliverables Progress

Deliverable	**Status**
Text Chatbot (Flask + UI)	Completed
Brain Tumor Detection Model	Completed
MRI Upload & Inference API	Completed
Documentation & User Guide	Completed