

Medical – SLM Project Documentation

This task is done as a part of Problem Statement -2 from Evaluation process for Data Scientist Position at Bajaj Finserv Health Ltd.

Developer Details

Charvi Jain
+91 6239884789
Charvi Jain - Email
Charvi Jain - LinkedIn

GitHub: <https://github.com/charvijain12/Medical-SLM>

Deployed Frontend: <https://medical-slm.netlify.app/>

Project Overview

The **Medical - SLM Project** is a specialized application aimed at answering user-submitted medical questions. It combines an ethically-trained, fine-tuned small language model (SLM) with a user-friendly chatbot interface. This project is designed to provide users with grounded, informative responses while ensuring adherence to strict ethical guidelines.

Purpose and Objectives

1. Primary Goals:

- Provide non-intrusive answers to user queries about medical symptoms, causes, remedies, and recommended medical specialties.
- Encourage consultation with certified healthcare professionals for any specific medical concerns.

2. Scope:

- Assist users with general medical information.
 - Suggest lifestyle modifications for relief and guidance on the appropriate specialists to consult.
 - Avoid providing prescriptions or medical treatments.
-

Components of the Project

1. Backend

The backend serves as the brain of the system, processing user queries and returning intelligent, ethical, and medically-informed responses.

Key Features:

- Built with **FastAPI**, ensuring rapid API response times and scalability.
- Integrates a fine-tuned **GPT-2** model (from Hugging Face Transformers library).
- Includes classification capabilities to determine the relevant medical specialty.
- Ethical compliance enforced programmatically to avoid prescriptive or unsafe advice.

Functional Workflow:

1. User submits a medical question.
 2. Backend processes the query, classifies it, and generates a response.
 3. The response includes:
 - Suggested medical specialty.
 - Detailed, user-friendly explanation or advice.
-

2. Frontend

The frontend provides an intuitive, accessible user interface for interacting with the chatbot.

Key Features:

- Built with **ReactJS**.
 - Chat-based interface for question submission and answer display.
 - Features include "New Chat" functionality to clear previous sessions and start fresh.
 - Deployed on **Netlify** for global accessibility.
-

3. Model Development

Base Model Selection:

- **GPT-2** was chosen due to its balance of language comprehension and generation capabilities, as well as its relatively smaller size for local deployments.

Fine-Tuning:

- Dataset: Medical question dataset sourced from the public domain.
- Preprocessing Steps:

- Noise removal (e.g., redundant data, irrelevant fields).
- Class balancing to address underrepresented specialties.
- Text tokenization using GPT-2 tokenizer.
- Training:
 - Adjusted hyperparameters (learning rate, batch size) for optimal performance.
 - Ensured compliance through specific prompts to avoid harmful medical advice.

Evaluation Metrics:

- **Accuracy:** Assessed the relevance of generated answers.
 - **Ethical Compliance:** Ensured responses adhered to safety guidelines.
 - **Inference Time:** Optimized for real-time responses.
-

4. Deployment

Backend Deployment:

- Designed for local deployment using **uvicorn** for running the FastAPI server.
- Global accessibility enabled using **ngrok**, which exposes the localhost API to the internet with a secure tunnel.

Frontend Deployment:

- Hosted on **Netlify** for a seamless, globally accessible interface.
- Netlify ensures quick deployment and easy updates.

Access Details:

- Frontend URL: <https://medical-slm.netlify.app/>
 - GitHub Repository: <https://github.com/charvijain12/Medical-SLM>
-

Ethical Considerations

- The model does not prescribe medications or suggest specific treatments.
- Responses always encourage users to consult certified healthcare professionals.
- Focus is on educational, non-intrusive advice, such as:
 - Explaining symptoms.
 - Suggesting dietary or lifestyle changes.
 - Recommending appropriate medical specialties for consultation.

Challenges Faced and Solutions

1. **Challenge: Accessibility from multiple devices.**
 - **Solution:** Used **ngrok** to expose the locally running backend to the internet, enabling the frontend to fetch answers from any device.
 2. **Challenge: Ethical compliance of generated answers.**
 - **Solution:** Incorporated prompts and post-processing checks to ensure adherence to ethical guidelines.
 3. **Challenge: Handling imbalanced datasets.**
 - **Solution:** Augmented data for underrepresented classes and applied weighted loss functions during training.
 4. **Challenge: Deployment without GPU dependency.**
 - **Solution:** Used efficient quantization and optimized inference pipelines to ensure smooth performance on CPU-only systems.
-

How to Use the System

1. Open the deployed frontend: <https://medical-slm.netlify.app/>.
 2. Type your medical query in the input field and submit.
 3. The chatbot will return:
 - Relevant medical specialty.
 - Detailed, grounded answer.
 4. For a new query, click the "New Chat" button to reset the session.
-