

# **HEALTH AI : Intelligent Healthcare Assistant** **Using IBM Granite**

## **PROJECT DOCUMENTATION**

### **1.Introduction**

Project title : **Health AI**

Team leader : **Gayathri.B**

Team member : **Abitha.K**

Team member : **Abitha.S**

Team member :**Ayshwariya.V**

### **2. Project Overview**

The "Health AI" project is a generative AI application that provides basic medical insights and information. Built on IBM's Granite model, the application serves as a healthcare assistant, offering functionalities like disease prediction and personalized treatment plan suggestions. The project's core purpose is to make health information more accessible and user-friendly, while a crucial disclaimer emphasizes that the information is for guidance only and not a substitute for professional medical advice

### **3.Features:**

#### **Disease Prediction:**

Users input a list of symptoms, and the model analyzes them to suggest possible medical conditions and general medication. This feature leverages the model's reasoning capabilities to connect disparate symptoms to likely diagnoses.

## **Personalized Treatment Plans:**

Given a medical condition, along with the user's age, gender, and medical history, the AI generates a customized treatment plan. This includes both home remedies and medication guidelines, showcasing the model's ability to generate tailored, context-aware responses

## **4.Core Functionalities**

The application has two main features:

### **Disease Prediction:**

Users can enter a list of symptoms, and the AI model will provide a list of possible medical conditions and general medication suggestions.

### **Treatment Plans:**

Based on a specified medical condition, age, gender, and medical history, the application generates a personalized treatment plan, including home remedies and general medication guidelines.

It's important to note that the application includes a clear disclaimer that the information provided is for informational purposes only and users should always consult a healthcare professional for a proper diagnosis and treatment.

## **5.Setup Instructions Prerequisites:**

- Python 3.x
- Gradio framework
- Transformers & Torch libraries
- Google Colab with T4 GPU access
- GitHub account

### **Installation Process:**

1. Open Google Colab and create a new notebook.
2. Change runtime type → select **T4 GPU**.

3. Install dependencies:
4. !pip install transformers torch gradio -q
5. Load and run the Citizen AI code.
6. Access the Gradio app link to interact with the model.
7. Download project files and upload them to GitHub for version control. **5. Folder**

**Structure** citizen\_ai

```
|— app.py          # Main Gradio application
|— model_config.py # IBM Granite model setup
|— dashboard.py    # Dashboard logic
|— requirements.txt # Required dependencies
└— README.md       # Project description
```

## 6. Technology Stack

The project leverages a number of powerful tools and libraries to function.

### Generative AI Model:

- The core of the application is the IBM Granite model, specifically the granite-3.2-2b-instruct version, which is a lightweight and fast model fine-tuned for reasoning tasks. The model processes user inputs to generate personalized and data-driven medical guidance.

### Frameworks:

- **Gradio:**  
The user interface (UI) is built with the Gradio framework, which makes it easy to create and share interactive web applications for machine

learning models. The UI consists of two tabs: "Disease Prediction" and "Treatment Plans".

- **Hugging Face transformers:**

This library is used to load and work with the IBM Granite model and its tokenizer.

- **Python:**

The entire application logic is written in Python, using libraries like torch for handling the model's tensors and GPU compatibility.

- **Deployment:**

- The application is designed to be deployed on Google Colab. Google Colab provides a free environment with access to a T4 GPU, which is essential for running the large AI model efficiently.

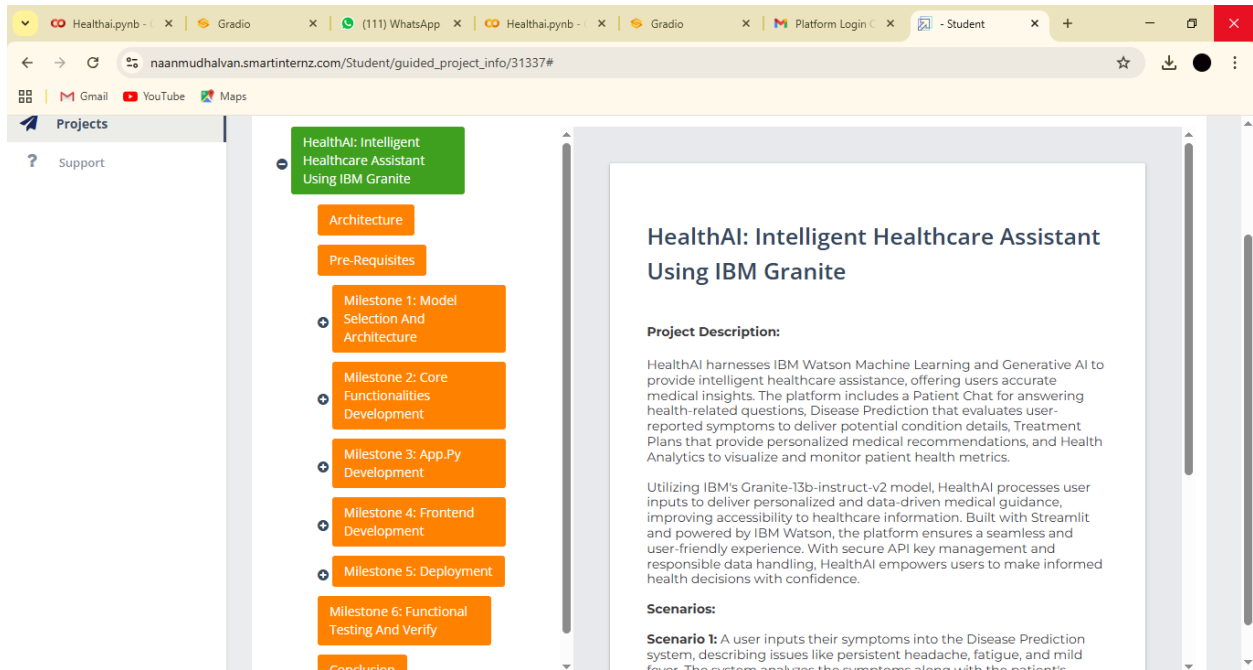
## **7. Project Workflow**

Building this project involved a clear, step-by-step process:

- i. Exploring the Project Portal
- ii. Choosing the AI Model.
- iii. Running the Application on Google Collab
- iv. Uploading the Project to GitHub

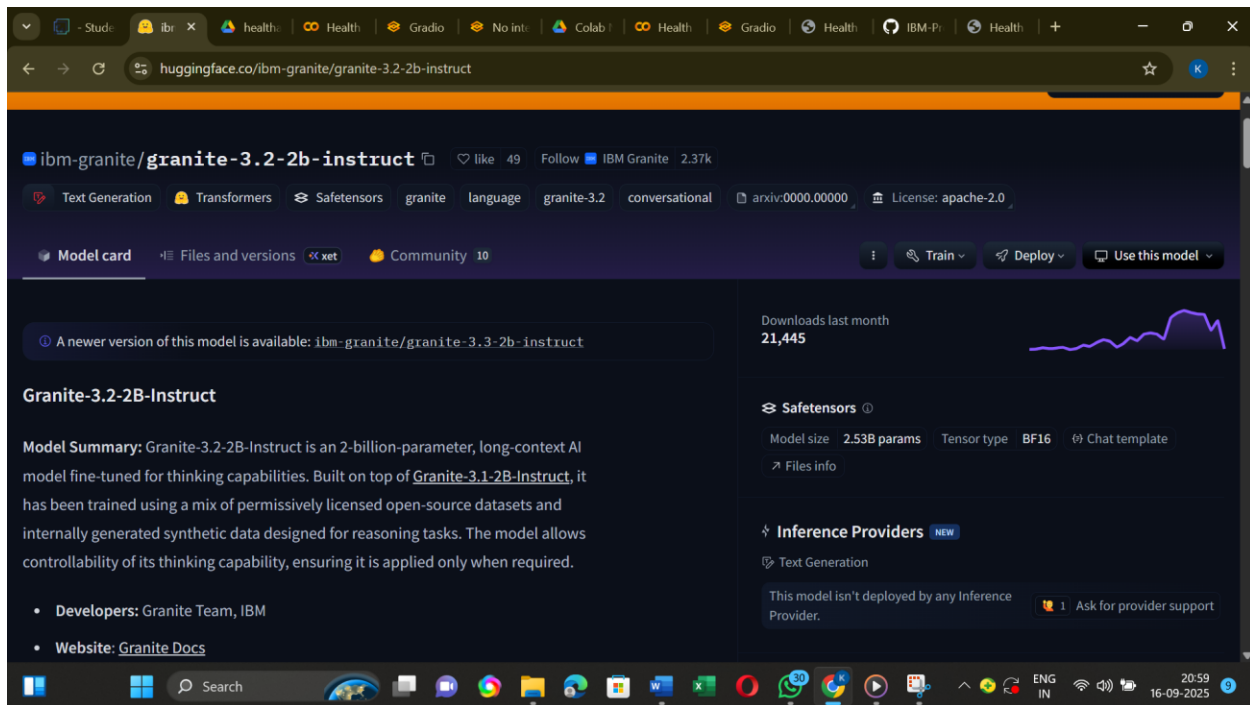
### **1. Exploring the Project Portal:**

The project began by accessing the "Naan Mudhalvan - SmartInternz" portal to understand the project requirement and details.



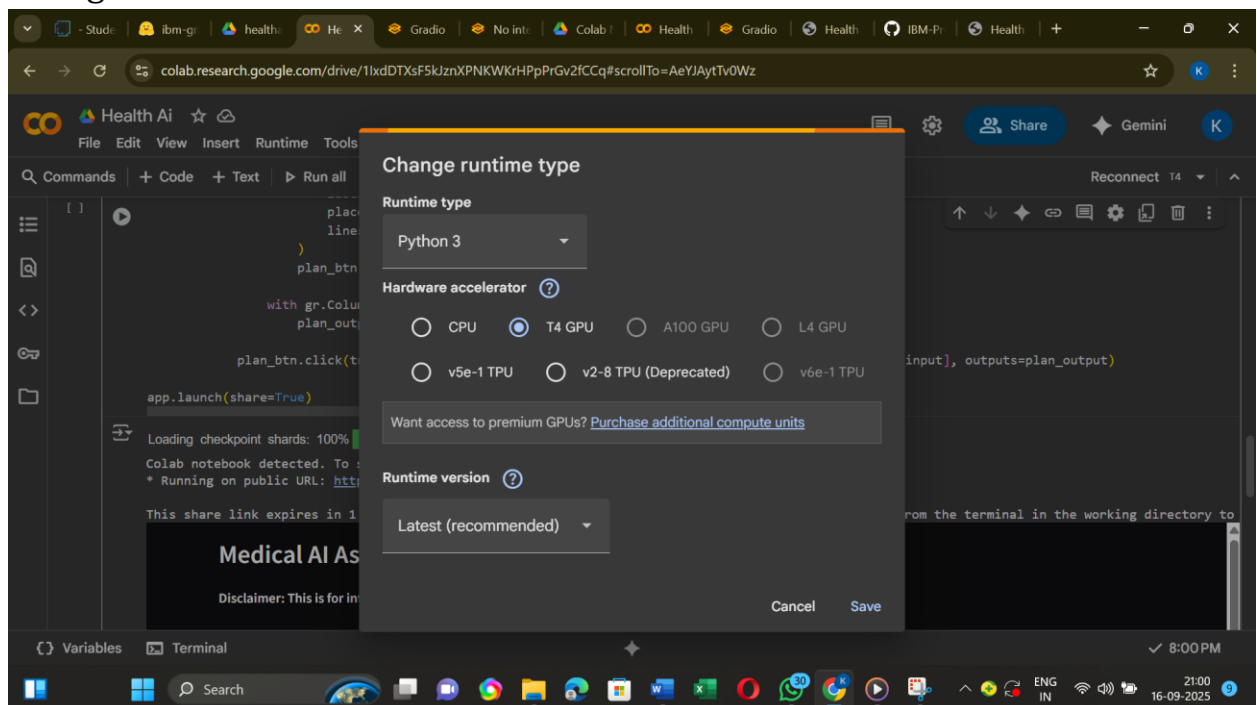
## 2.Choosing the AI Model:

An IBM Granite model was selected from Hugging Face, a platform that hosts and allows collaboration on public AI models and datasets. The ibm-granite/granite-3.2-2b-instruct model was chosen for its performance and efficiency.

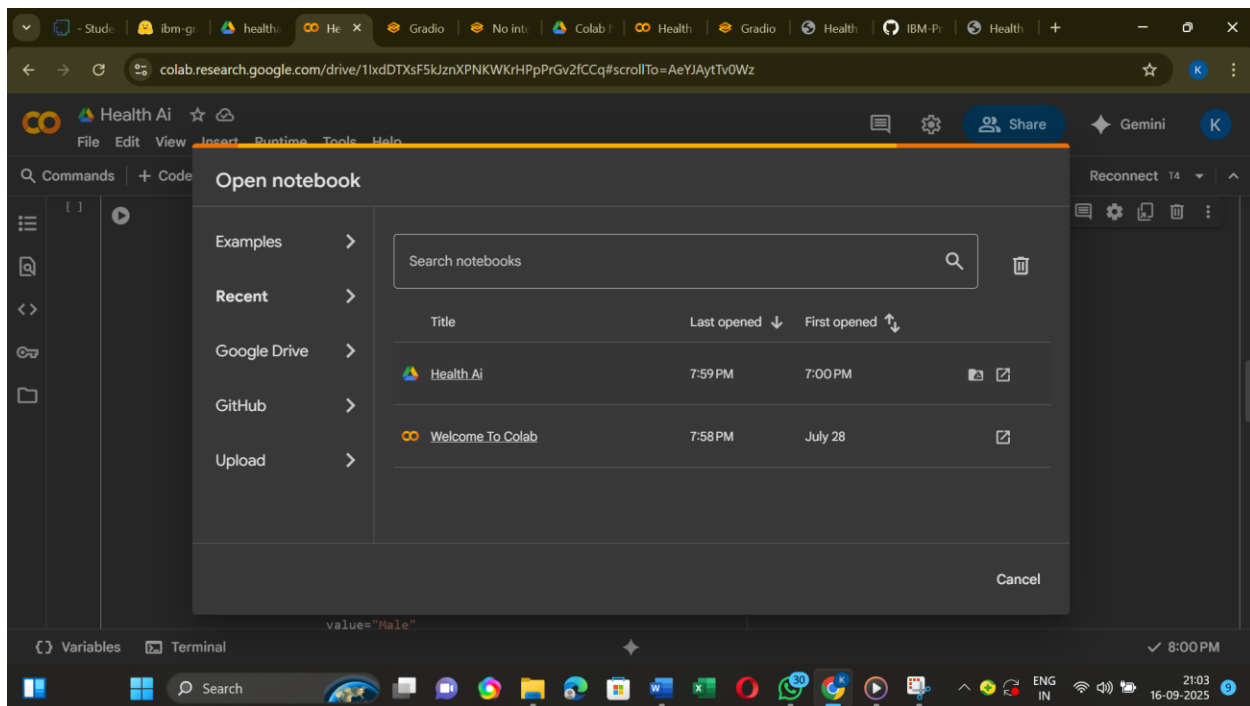


### 3. Running the Application on Google Colab:

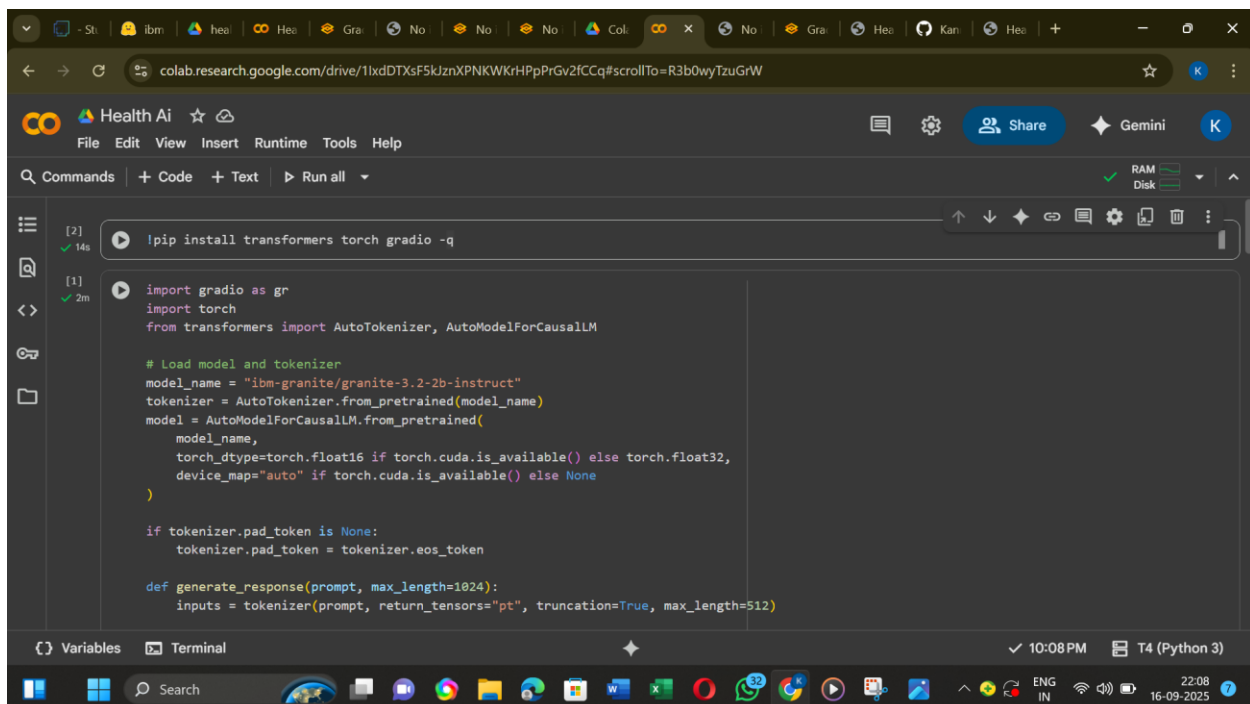
- A new notebook was created in Google Colab and the runtime was configured to use a T4 GPU. Then Save.



- Go to file, click Open notebook.



- Required libraries like transformers, torch, and gradio were installed.



- The Python code, which handles loading the model, defining the generation functions, and creating the Gradio interface, was added to the notebook and executed.

The screenshot shows a Google Colab notebook interface. The code in the cell is as follows:

```
[1]
import gradio as gr
import torch
from transformers import AutoTokenizer, AutoModelForCausalLM

# Load model and tokenizer
model_name = "ibm-granite/granite-3.2-2b-instruct"
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForCausalLM.from_pretrained(
    model_name,
    torch_dtype=torch.float16 if torch.cuda.is_available() else torch.float32,
    device_map="auto" if torch.cuda.is_available() else None
)

if tokenizer.pad_token is None:
    tokenizer.pad_token = tokenizer.eos_token

def generate_response(prompt, max_length=1024):
    inputs = tokenizer(prompt, return_tensors="pt", truncation=True, max_length=512)

    if torch.cuda.is_available():
        inputs = {k: v.to(model.device) for k, v in inputs.items()}
```

The notebook interface includes a top bar with 'Health Ai' branding, a menu (File, Edit, View, Insert, Runtime, Tools, Help), and a 'Share' button. The left sidebar shows file explorer icons. The bottom status bar indicates 'T4 (Python 3)' and the time '22:09 16-09-2025'.

The screenshot shows the same Google Colab notebook with the code continued:

```

    inputs = {k: v.to(model.device) for k, v in inputs.items()}

    with torch.no_grad():
        outputs = model.generate(
            **inputs,
            max_length=max_length,
            temperature=0.7,
            do_sample=True,
            pad_token_id=tokenizer.eos_token_id
        )
```

The notebook interface is consistent with the previous screenshot, showing the same top bar, sidebar, and bottom status bar. The time in the bottom right corner is now '09:15 18-09-2025'.



```
google colab - Search x Health ALIpynb - Colab x Gradio x +
https://colab.research.google.com/drive/1yvWjUOWB2THvysx8ypIDyBKIMHjdxdf#scrollTo=qPuSVNmDtKJP

Health ALIpynb ☆
File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all
RAM Disk
[1] 2m
response = tokenizer.decode(outputs[0], skip_special_tokens=True)
response = response.replace(prompt, "").strip()
return response

def disease_prediction(symptoms):
    prompt = f"Based on the following symptoms, provide possible medical conditions and general medication suggestions. Always emphasize the importance of consulting a doctor for"
    return generate_response(prompt, max_length=1200)

def treatment_plan(condition, age, gender, medical_history):
    prompt = f"Generate personalized treatment suggestions for the following patient information. Include home remedies and general medication guidelines.\n\nMedical Condition:"
    return generate_response(prompt, max_length=1200)

# Create Gradio interface
with gr.Blocks() as app:
    gr.Markdown("# Medical AI Assistant")
    gr.Markdown("Disclaimer: This is for informational purposes only. Always consult healthcare professionals for medical advice.")

    with gr.Tabs():
        with gr.TabItem("Disease Prediction"):
            with gr.Row():
                with gr.Column():
                    symptoms_input = gr.Textbox(
                        label="Enter Symptoms",
                        placeholder="e.g., fever, headache, cough, fatigue...",
                        lines=4
                    )
                    predict_btn = gr.Button("Analyze Symptoms")

            with gr.Column():
                prediction_output = gr.Textbox(label="Possible Conditions & Recommendations", lines=20)

        predict_btn.click(disease_prediction, inputs=symptoms_input, outputs=prediction_output)

    with gr.TabItem("Treatment Plans"):
        with gr.Row():
            with gr.Column():
                condition_input = gr.Textbox(
                    label="Medical Condition",
                    placeholder="e.g., diabetes, hypertension, migraine...",
                    lines=2
                )
                age_input = gr.Number(label="Age", value=30)
                gender_input = gr.Dropdown(
                    choices=["Male", "Female", "Other"],
                    label="Gender",
                    value="Male"
                )
                history_input = gr.Textbox(
                    label="Medical History",
                    placeholder="Previous conditions, allergies, medications or None",
                    lines=3
                )
                plan_btn = gr.Button("Generate Treatment Plan")

            with gr.Column():
                plan_output = gr.Textbox(label="Personalized Treatment Plan", lines=20)

        plan_btn.click(treatment_plan, inputs=[condition_input, age_input, gender_input, history_input], outputs=plan_output)

app.launch(share=True)
```

```
google colab - Search x Health ALIpynb - Colab x Gradio x +
https://colab.research.google.com/drive/1yvWjUOWB2THvysx8ypIDyBKIMHjdxdf#scrollTo=qPuSVNmDtKJP

Health ALIpynb ☆
File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all
RAM Disk
[1] 2m
prediction_output = gr.Textbox(label="Possible Conditions & Recommendations", lines=20)

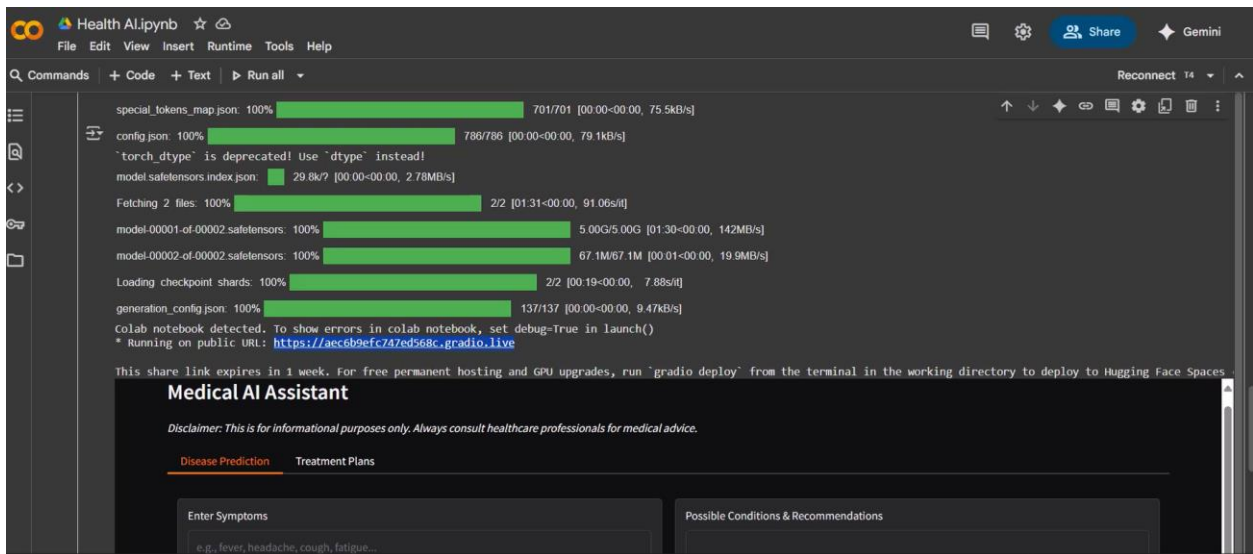
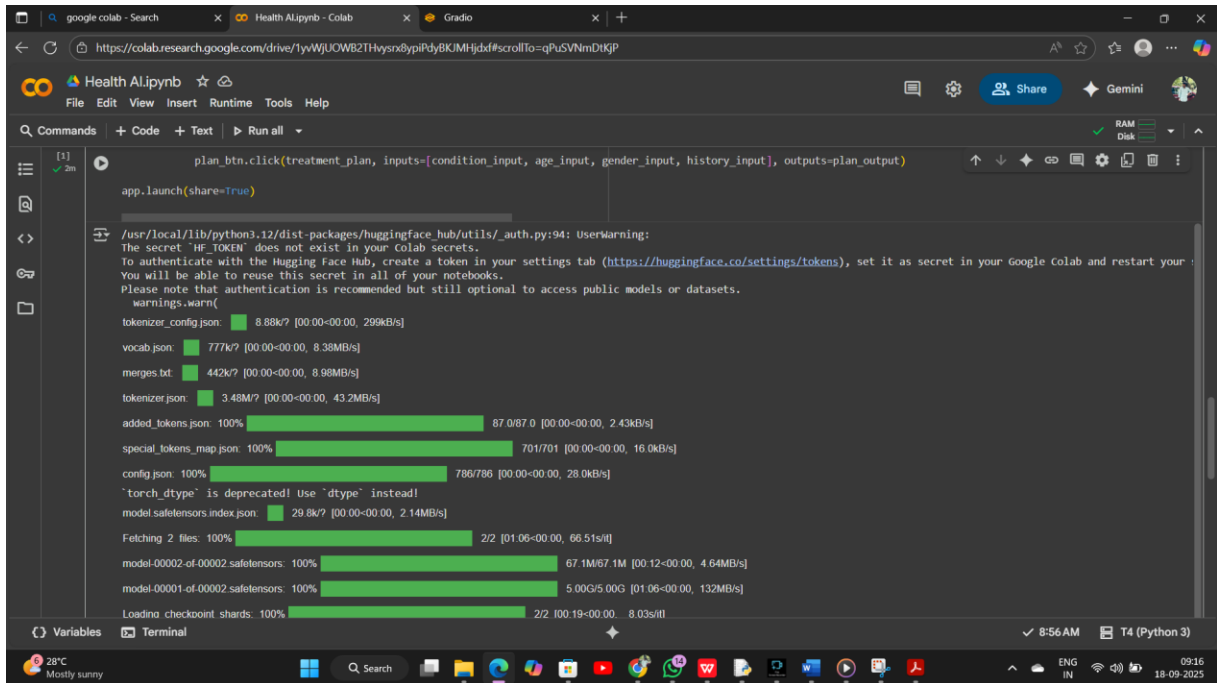
predict_btn.click(disease_prediction, inputs=symptoms_input, outputs=prediction_output)

with gr.TabItem("Treatment Plans"):
    with gr.Row():
        with gr.Column():
            condition_input = gr.Textbox(
                label="Medical Condition",
                placeholder="e.g., diabetes, hypertension, migraine...",
                lines=2
            )
            age_input = gr.Number(label="Age", value=30)
            gender_input = gr.Dropdown(
                choices=["Male", "Female", "Other"],
                label="Gender",
                value="Male"
            )
            history_input = gr.Textbox(
                label="Medical History",
                placeholder="Previous conditions, allergies, medications or None",
                lines=3
            )
            plan_btn = gr.Button("Generate Treatment Plan")

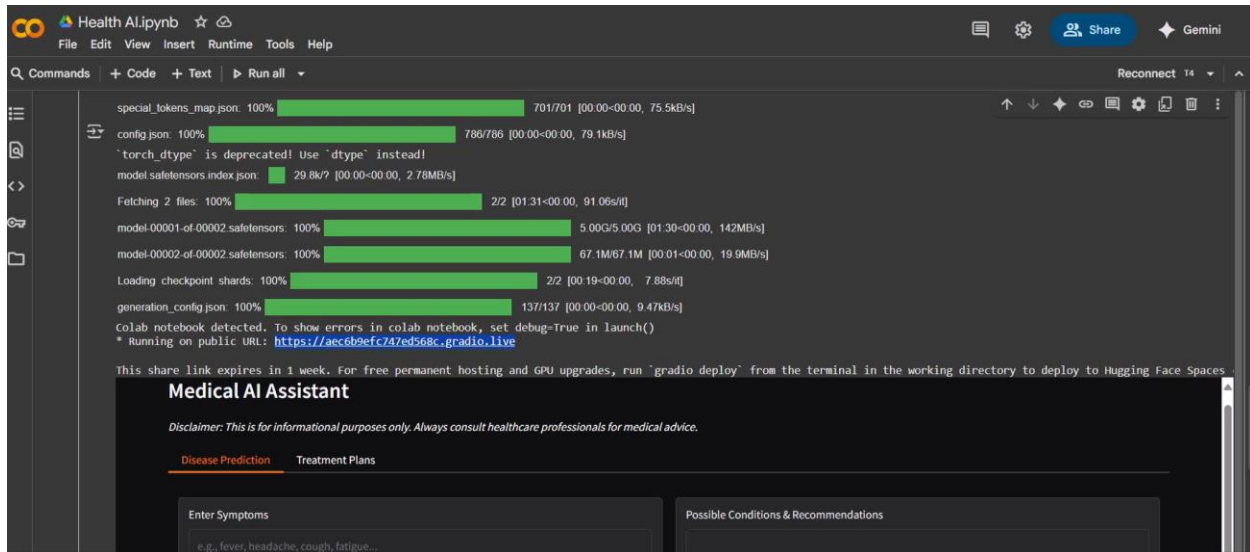
        with gr.Column():
            plan_output = gr.Textbox(label="Personalized Treatment Plan", lines=20)

        plan_btn.click(treatment_plan, inputs=[condition_input, age_input, gender_input, history_input], outputs=plan_output)

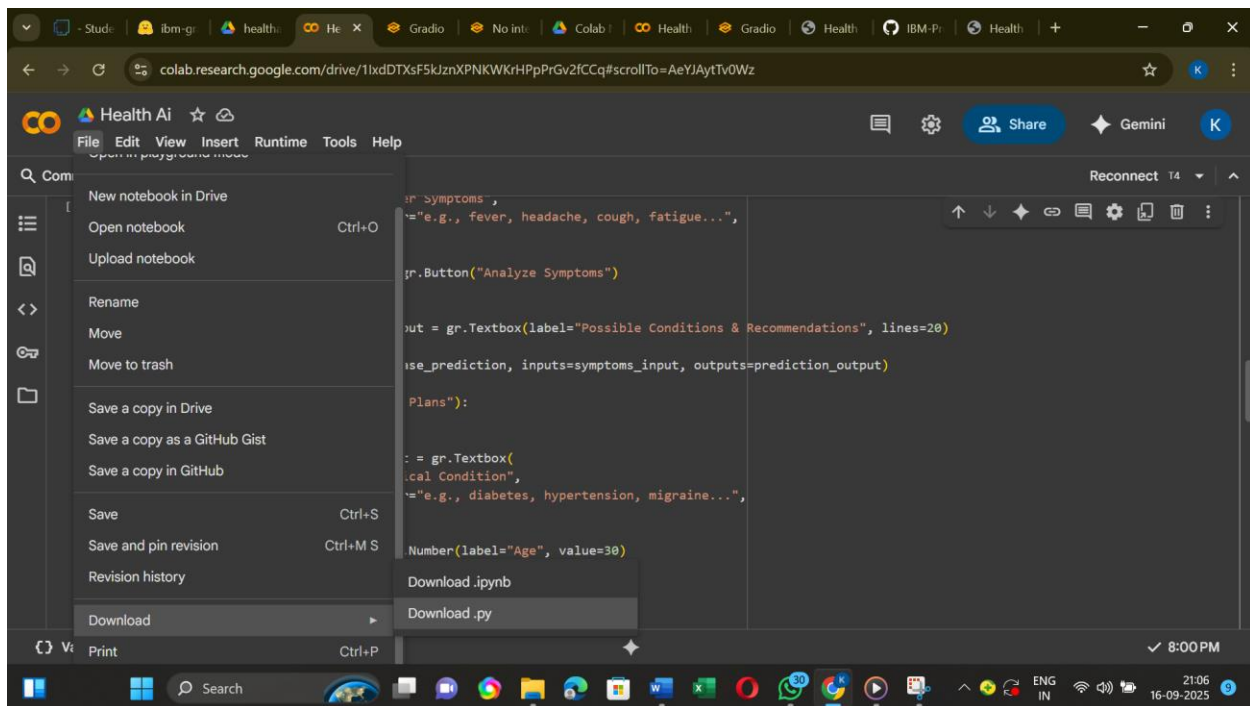
app.launch(share=True)
```



- The application launched, providing a public URL to access the "Health AI" interface.

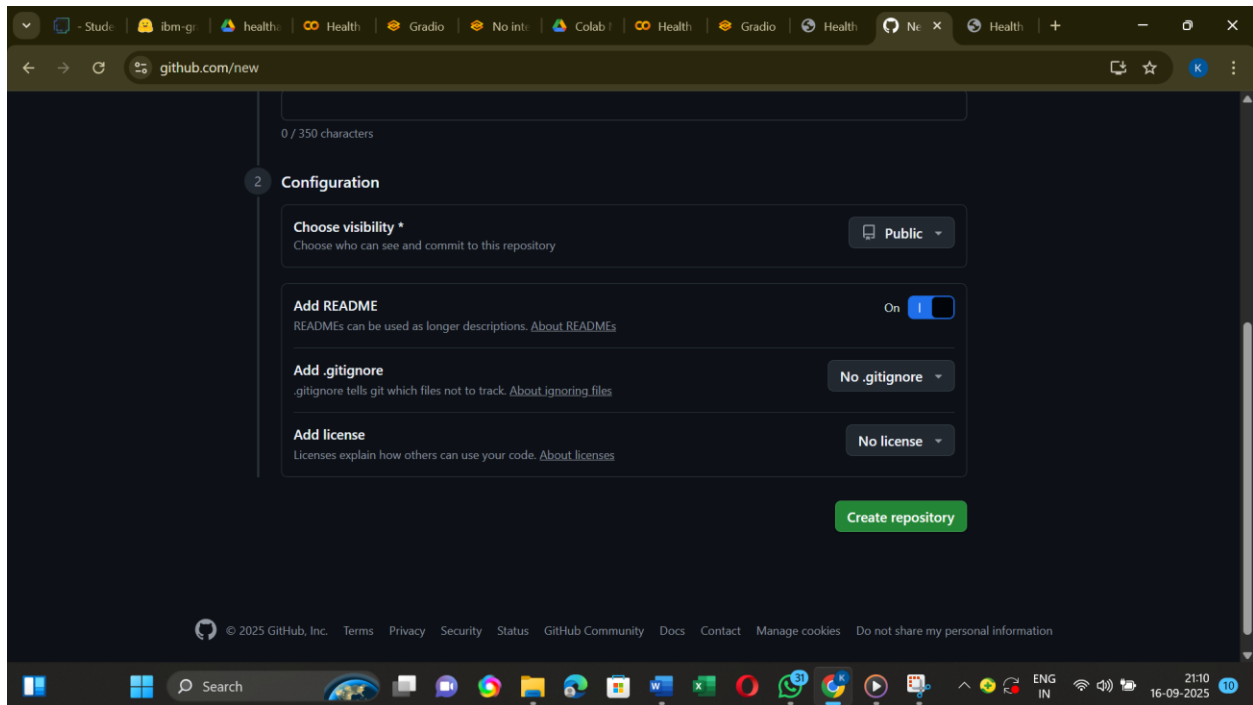


- Go to file and download the project as Download.py.

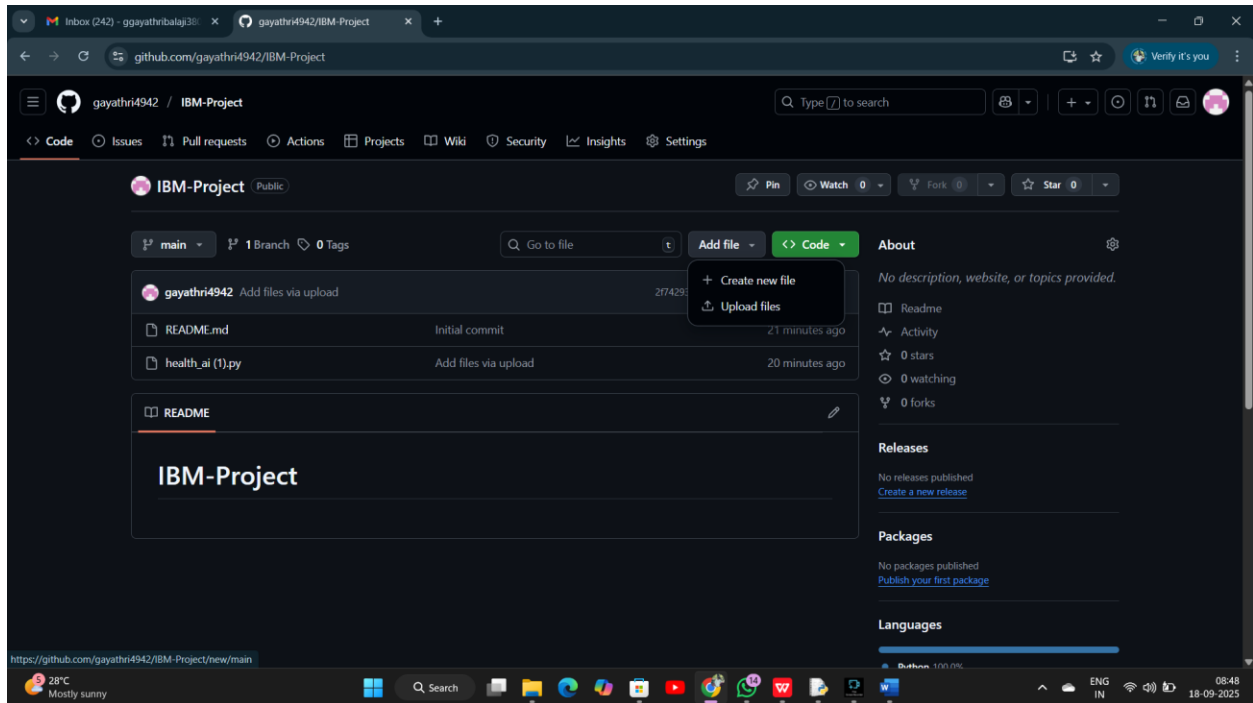


## 8. Uploading the Project to GitHub:

- The Python code from the Google Colab notebook was downloaded as a .py file.
- Use your own repository name.
- In configuration, README should be turned On.
- A new repository was created on GitHub to store the project files.



- The downloaded Python file was then uploaded to the GitHub repository.



## 9.Interface: A Gradio-based Medical AI Assistant

The user interface for the "Health AI: Intelligent Healthcare Assistant" is built using the Gradio framework. Gradio provides a clean, easy-to-navigate, and interactive web application that allows users to interact with the underlying AI model without needing any coding knowledge.

### Overall Layout:

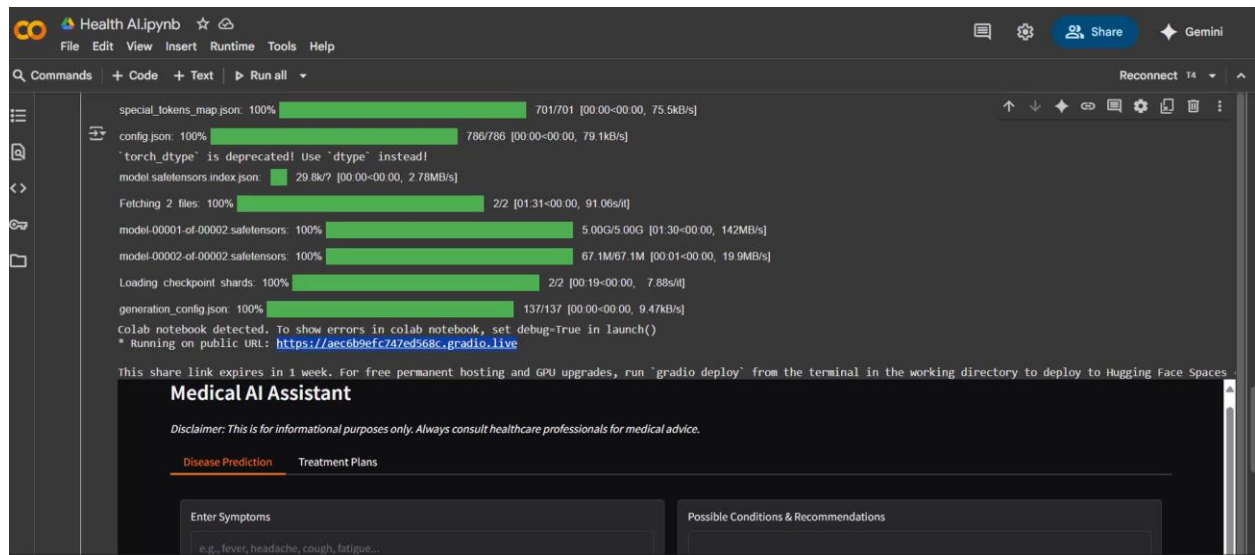
The application is structured using `gr.Blocks()`, which serves as a container for the entire UI. The main page features a prominent title and a crucial medical disclaimer to ensure responsible use.

- **Title:** A large, centered heading reads "# Medical AI Assistant".
- **Disclaimer:** A markdown text below the title clearly states: "Disclaimer: This is for informational purposes only. Always consult healthcare professionals for medical advice." This is a vital component that sets appropriate user expectations from the start.

The core functionality is organized into a tabbed layout using `gr.Tabs()`, which separates the two main features of the application.

## Tab 1: "Disease Prediction"

This tab is designed for users who want to understand possible medical conditions based on their symptoms. The layout is simple and intuitive, with an input area and a corresponding output area.

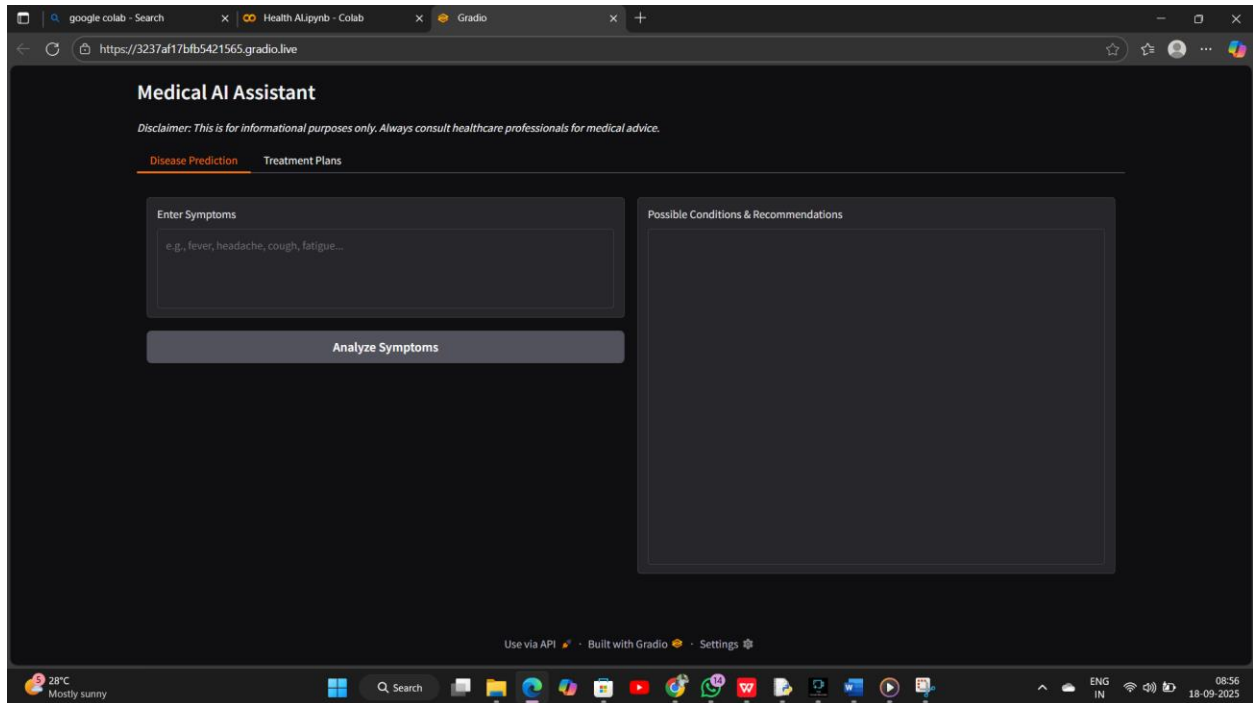


- Input (gr.Textbox):

Label: "Enter Symptoms"

Placeholder: "e.g., fever, headache, cough, fatigue..."

Functionality: This is a multi-line text box where the user can type in a list of their symptoms. The placeholder text provides clear examples to guide the user.



- Button (gr.Button):

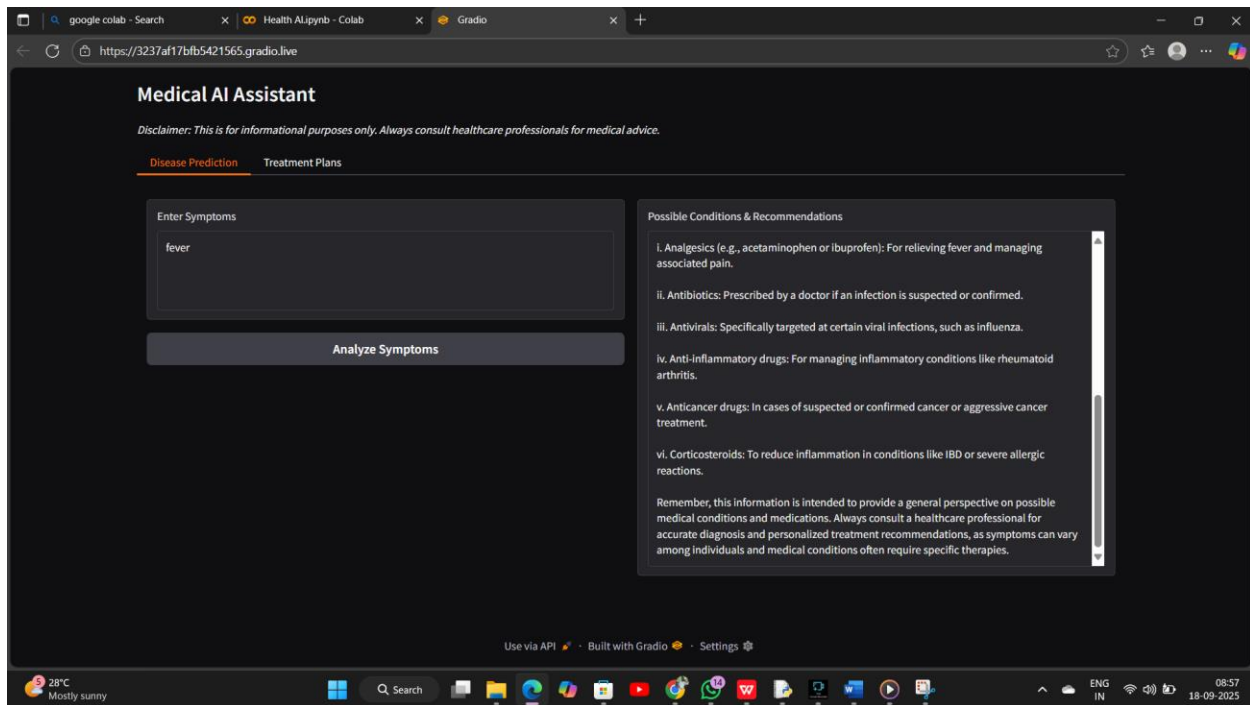
Label: "Analyze Symptoms"

Functionality: Clicking this button sends the text from the symptoms input box to the disease\_prediction function in the backend.

- Output (gr.Textbox):

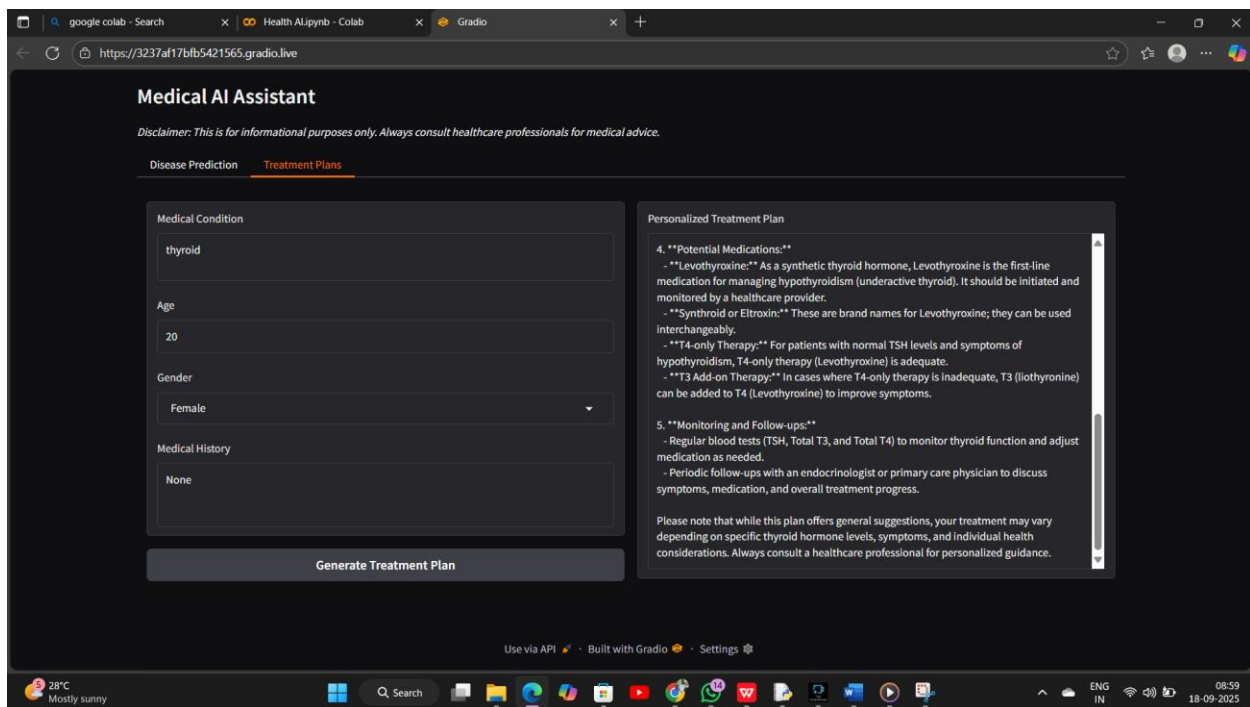
Label: "Possible Conditions & Recommendations"

Functionality: This large, read-only text box displays the response generated by the AI model. The model's output, which includes a list of potential conditions and general recommendations, appears here for the user to read.



## Tab 2: "Treatment Plans"

This tab is for users who have a diagnosed or known condition and want to explore personalized treatment suggestions. It collects more specific information to provide a tailored response.





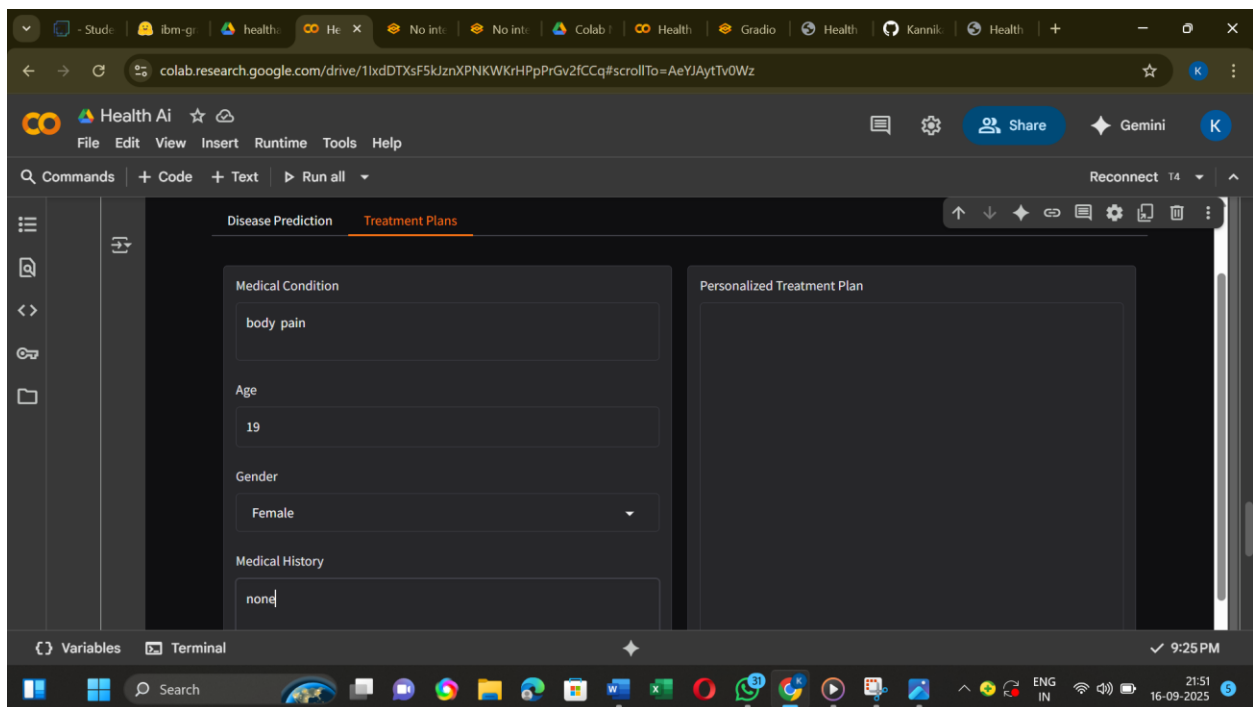
- Input Fields (gr.Textbox, gr.Number, gr.Dropdown):

*Medical Condition (gr.Textbox):*

Label: "Medical Condition"

Placeholder: "e.g., diabetes, hypertension, migraine..."

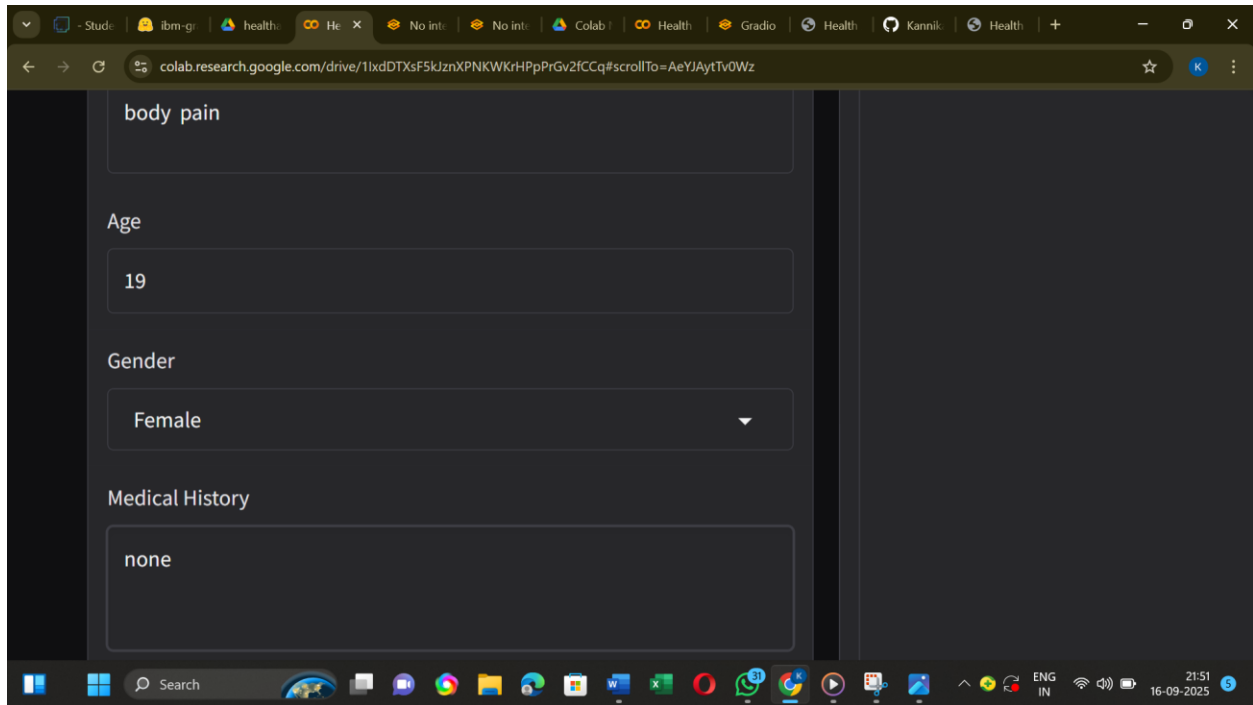
Functionality: A text box for the user to enter their specific medical condition.



*Age (gr.Number):*

Label: "Age"

Functionality: A numerical input field for the user's age, which is a key factor in many medical treatments.



*Gender (gr.Dropdown):*

Label: "Gender"

Choices: "Male", "Female", "Other"

Functionality: A dropdown menu to select gender.

The screenshot shows a web browser window with a dark theme. The address bar displays a Google Drive link. The form is titled 'body pain' and contains the following fields:

- Age:** A text input field containing the value '19'.
- Gender:** A dropdown menu with 'Female' selected.
- Medical History:** A multi-line text area containing the value 'none'.

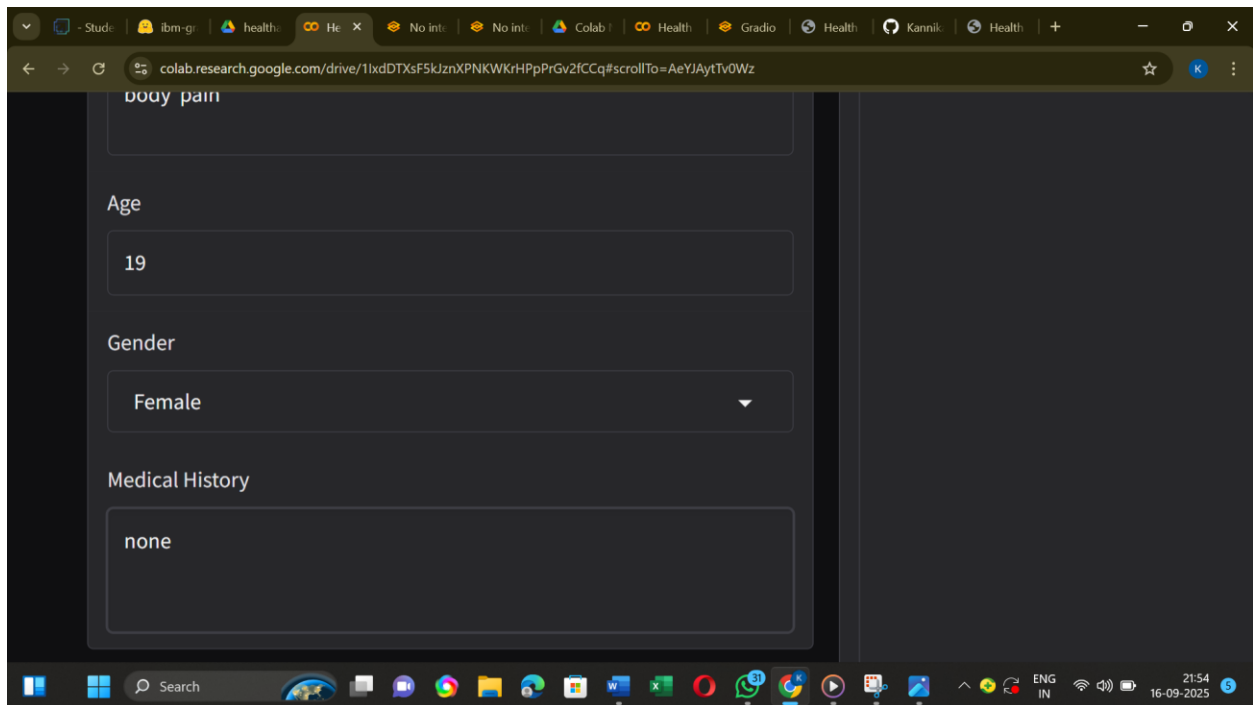
The Windows taskbar is visible at the bottom, showing the search bar and various application icons. The system clock indicates the time is 21:54 on 16-09-2025.

*Medical History (gr.Textbox):*

Label: "Medical History"

Placeholder: "Previous conditions, allergies, medications or None"

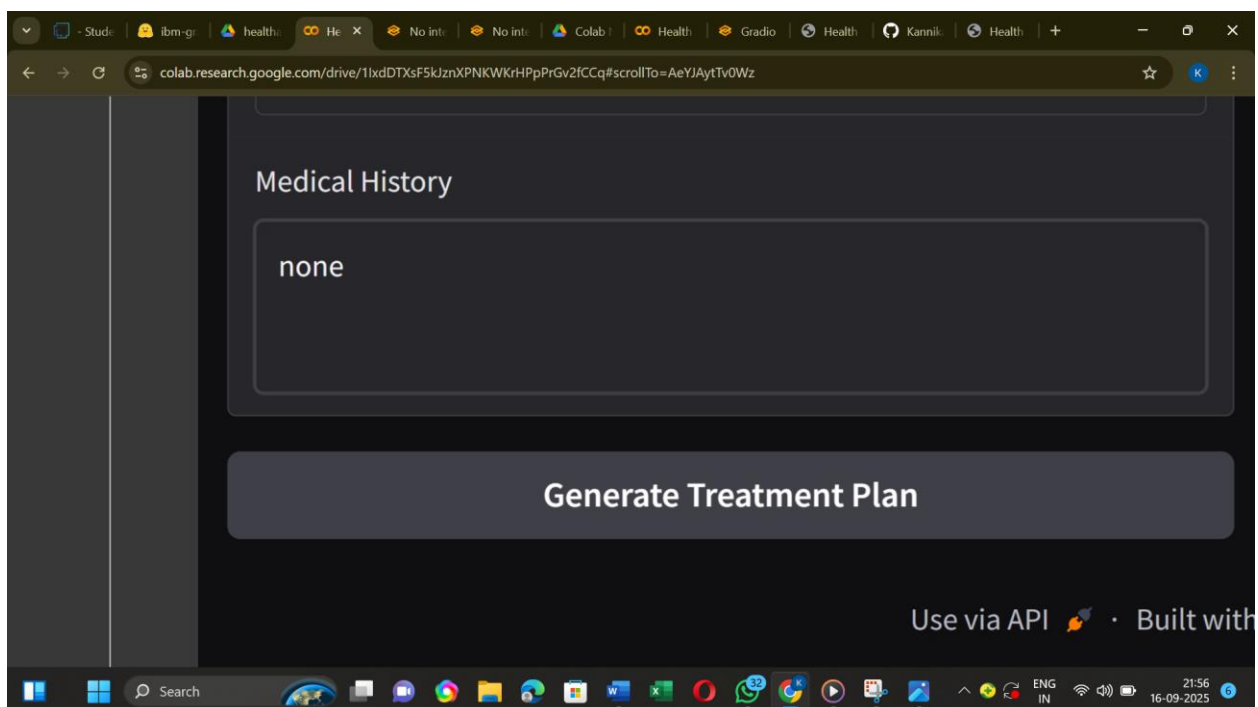
Functionality: A multi-line text box for the user to provide details about their health background.



*Button (gr.Button):*

Label: "Generate Treatment Plan"

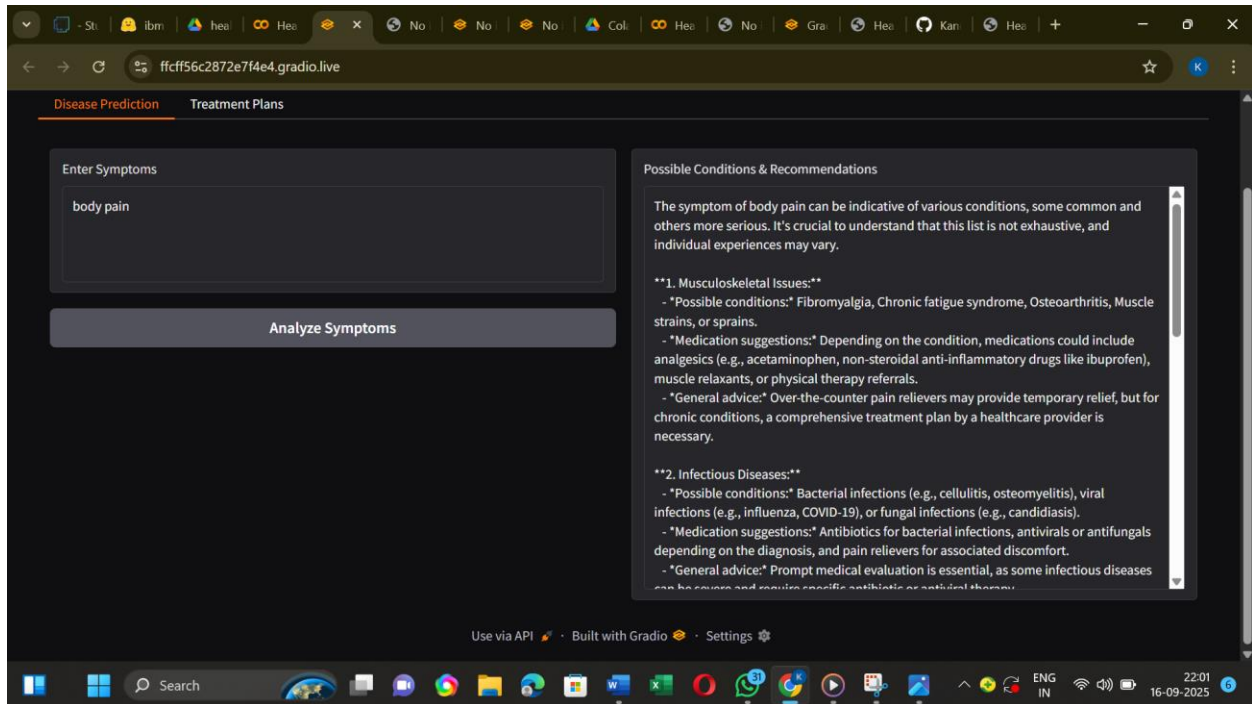
Functionality: This button triggers the `treatment_plan` function in the backend, passing all the entered information to the AI model.



*Output (gr.Textbox):*

Label: "Personalized Treatment Plan"

Functionality: A large, scrollable text box that displays the AI-generated treatment plan, including home remedies and medication guidelines.



## 10. User Flow:

1. The user lands on the application page.
2. They can choose between the "Disease Prediction" and "Treatment Plans" tabs.
3. In the chosen tab, they fill out the input fields with their specific information.
4. They click the "Analyze Symptoms" or "Generate Treatment Plan" button.
5. The Gradio interface sends this information to the Python backend, where the IBM Granite model processes the request.
6. The AI's response is returned and displayed in the designated output textbox.

7. The user can then modify their inputs and generate new responses as needed.

## 11. Known Issues (HealthAI)

- **Limited Dashboard Customization:** HealthAI's current dashboards provide only basic visualizations with restricted flexibility for customization based on user needs.
- **Dependency on Colab GPU Resources:** The application relies on Google Colab's free GPU, which can be inconsistent and may limit scalability and uptime.
- **No Role-Based Access System:** There is currently no authentication system to differentiate between citizens (patients) and officials (health workers/administrators), which restricts secure access to certain features.

## 12. Future Enhancements (HealthAI)

- **Role-Based Access Control (RBAC):** Develop a secure authentication and authorization system to allow differentiated access for citizens (patients) and officials (e.g., healthcare professionals, administrators).
- **Multilingual Interface Support:** Integrate support for multiple languages to make HealthAI more inclusive and accessible to a broader population.
- **Enhanced Analytics Dashboards:** Expand the analytics capabilities with advanced, interactive, and customizable dashboards to better support decision-making.
- **Cloud Platform Deployment:** Migrate HealthAI from Google Colab to more robust cloud platforms such as IBM Cloud or AWS to enhance performance, reliability, and scalability.

