

---

☐ **Generative AI Project using IBM Cloud – HEALTHAI**




☐ **Project Documentation Format**

---

## 1. Introduction

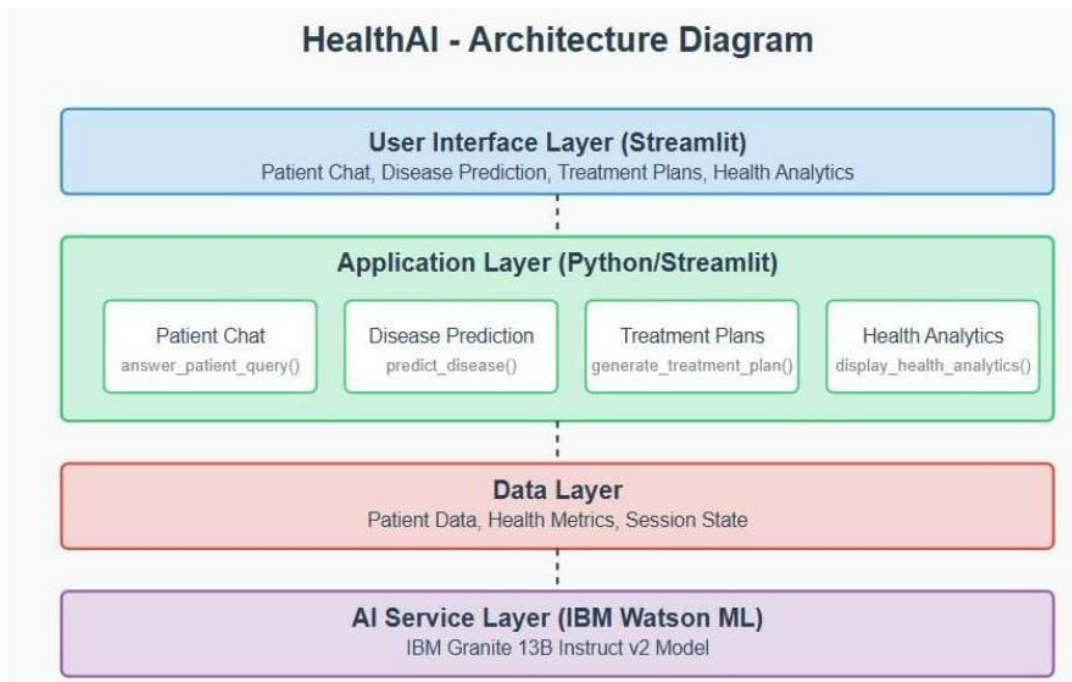
- **Project Title: HEALTHAI: Intelligent Healthcare Assistant using IBM Granite (Generative AI with IBM Cloud)**
  - **Team Members:**
    - **Donduboina Divya Siva Naga Malleswari(Team Leader – Development & Integration):**  
Led the complete development of the HEALTHAI application, including IBM Granite integration, Streamlit-based UI design, module creation, and model API handling.
    - **Althy Sri Vidya Lakshmi , Angirekula Bhavana (Model Interaction & Testing):**  
Contributed by assisting in prompt design, testing the AI model outputs across modules like Disease Prediction and Health Chat, and refining interactions with IBM Granite.
    - **Akula Sujan (UI Structuring & Feature Enhancement):**  
Supported in designing user flow, organizing the Streamlit interface across all modules, and suggesting improvements in user interaction and feature behavior.
- 

## 2. Project Overview

- **Purpose:**  
To build a Generative AI-based healthcare assistant using IBM Granite, capable of answering health queries, predicting diseases, suggesting treatments, and displaying analytics.
  - **Features:**
    -  AI Health Chat using IBM Granite
    - ☐ Disease Prediction from user symptoms
    -  Treatment Plan Suggestions
    -  Health Analytics Dashboard
    - ☐ Centralized shared model for performance optimization
- 

## 3. Architecture

- **Frontend:**  
Built using **Streamlit** for a clean and responsive web interface. Each feature is modularized for easy navigation via sidebar.
- **Backend & Model:**
  - No traditional backend. All logic handled in Streamlit using Python.
  - Uses **IBM Granite 3.3B Instruct model** from Hugging Face: `ibm-granite/granite-3.3-2b-instruct`
  - Supports both API and **local model loading** (`granite/` folder).
- **Shared Model Loader:**  
The `shared_model.py` file centrally loads and shares the AI model across modules to prevent memory crashes and redundancy.



## 4. Setup Instructions

### Prerequisites

- Python 3.10+
- pip
- Hugging Face account and token
- Installed model files if using local (`granite/` folder)

### Installation



git clone

[https://github.com/ammu914/HealthAI-](https://github.com/ammu914/HealthAI-IntelligentHealthcare-Assistant-Using-IBM-Granite)

IntelligentHealthcare-Assistant-Using-IBM-Granite

cd Health-ai

pip install -r requirements.txt

### Environment Variables

Create a .env file in the root folder:

HUGGINGFACEHUB\_API\_TOKEN=hf\_EPkOkQWaTrYYRwbVgrfzpiTWNrSADVyjnd

✓ .env file must be excluded in .gitignore.

---

## 5. Folder Structure

Health-ai/

- ├─ app.py           # Main entry point
- ├─ shared\_model.py   # Shared AI model instance
- ├─ patient\_chat.py   # AI Health Chat module
- ├─ disease\_prediction.py # Disease Prediction logic
- ├─ treatment\_plans.py   # Treatment Plan suggestions
- ├─ health\_analytics.py   # Analytics module
- ├─ requirements.txt   # Python dependencies
- ├─ .env           # API token (not pushed to GitHub)
- ├─ granite/       # [Optional] Local model folder
- └─ assets/       # Logos and screenshots

---

## 6. Running the Application

### For Hugging Face API:

streamlit run app.py

### For Local Model:

Ensure granite/ folder contains the downloaded model and tokenizer files.

In shared\_model.py, update:

model\_path = "./granite"

---



## 7. API Documentation

### Endpoint:

<https://api-inference.huggingface.co/models/ibm-granite/granite-3.3-2b-instruct>

### Method: POST

### Headers:

```
{  
  "Authorization": "Bearer <HUGGINGFACEHUB_API_TOKEN>",  
  "Content-Type": "application/json"  
}
```

### Example Request:

```
{  
  "inputs": "What are the symptoms of diabetes?"  
}
```

### Example Response:

```
{  
  "generated_text": "Common symptoms of diabetes include frequent urination..."  
}
```

---

## 8. Authentication

- Hugging Face token is securely stored in .env
- .env is excluded via .gitignore
- App is currently public and stateless (no user login)
- Streamlit or Firebase Auth can be added in future

---

## 9. User Interface


- Built entirely with **Streamlit**
- Sidebar for navigation
- Text/chat inputs for interaction
- Visual graphs and health tips in Analytics

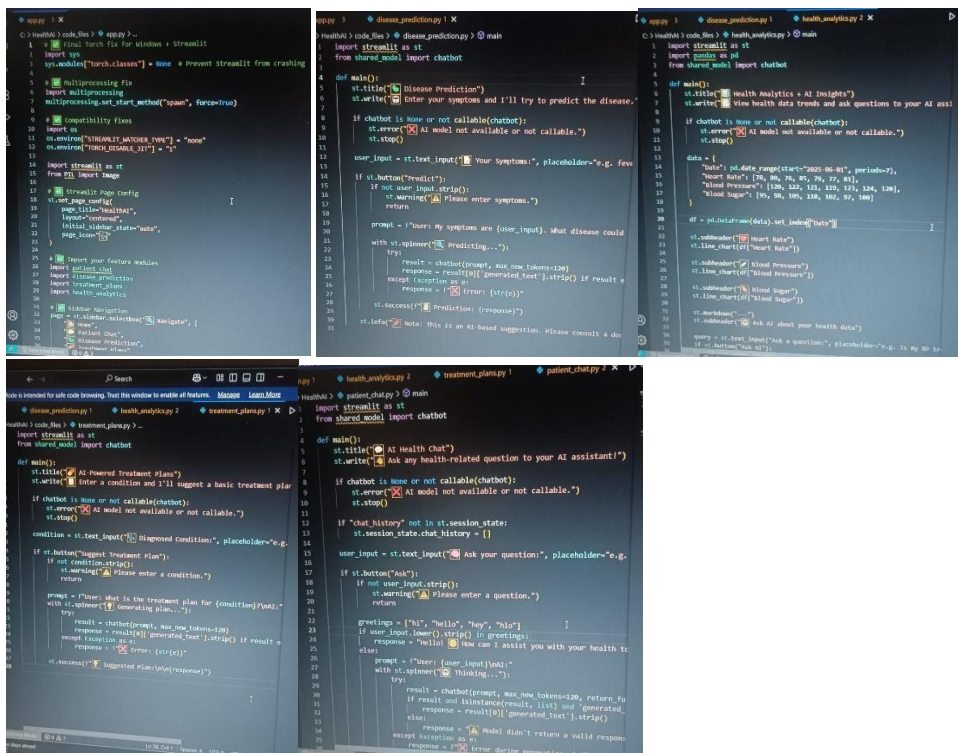
- Centralized theme and branding

## 10. Testing

- ☒ Manual testing across all modules
- ☒ Model tested with varied prompts and edge cases
- ☒ Handled errors for invalid inputs and model timeouts

## 11. Screenshots or Demo

-  [Demo Video on YouTube](#)
- **INPUTS ( CODES ) :**



```

# disease_prediction.py
def main():
    st.title("Disease Prediction")
    st.write("Enter your symptoms and I'll try to predict the disease.")
    if chatbot is None or not callable(chatbot):
        st.error("AI model not available or not callable.")
        st.stop()
    user_input = st.text_input("Your symptoms, placeholder: e.g. fever")
    if st.button("Predict"):
        if not user_input.strip():
            st.warning("Please enter symptoms.")
            return
        prompt = f"User: My symptoms are {user_input}. What disease could it be?"
        with st.spinner("Predicting..."):
            try:
                result = chatbot(prompt, max_new_tokens=100)
                response = result[0] if isinstance(result, list) else result
            except Exception as e:
                response = f"Error: {str(e)}"
            st.success(f"Prediction: {response}")
    st.info("Note: This is an AI-based suggestion. Please consult a doctor.")

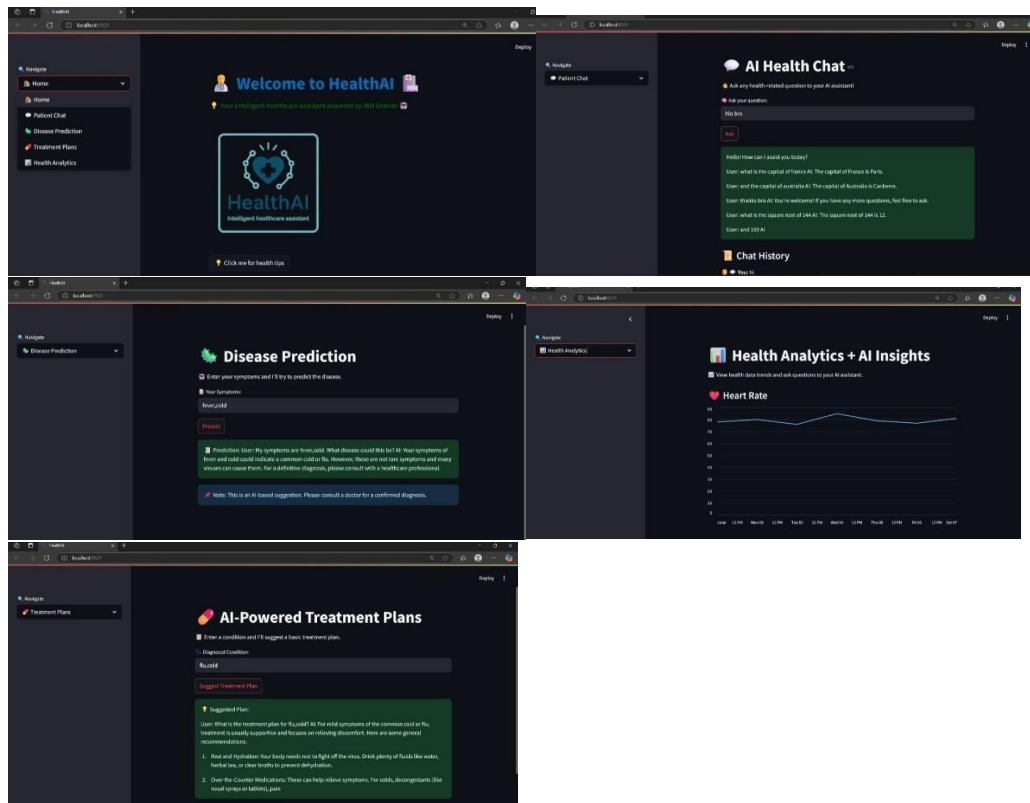
# health_analytics.py
def main():
    st.title("Health Analytics & AI Insights")
    st.write("Enter health data to track and ask questions to your AI assistant.")
    if chatbot is None or not callable(chatbot):
        st.error("AI model not available or not callable.")
        st.stop()
    data = {
        "Date": st.date_input("Date (YYYY-MM-DD)", period=7),
        "Heart Rate": [70, 80, 90, 100, 110, 120, 130],
        "Blood Pressure": [120, 130, 140, 150, 160, 170, 180],
        "Blood Sugar": [90, 95, 100, 105, 110, 115, 120]
    }
    df = pd.DataFrame(data, index=range(1, len(data)+1))
    st.subheader("Heart Rate")
    st.line_chart(df["Heart Rate"])
    st.subheader("Blood Pressure")
    st.line_chart(df["Blood Pressure"])
    st.subheader("Blood Sugar")
    st.line_chart(df["Blood Sugar"])
    st.markdown("Ask AI about your health data")
    query = st.text_input("Ask a question: e.g. Is my BP too high?")
    if st.button("Ask AI"):
        pass

# treatment_plan.py
def main():
    st.title("AI Powered Treatment Plan")
    st.write("Enter a condition and I'll suggest a basic treatment plan.")
    if chatbot is None or not callable(chatbot):
        st.error("AI model not available or not callable.")
        st.stop()
    condition = st.text_input("Diagnosed Condition, placeholder: e.g. flu")
    if st.button("Suggest Treatment Plan"):
        if not condition.strip():
            st.warning("Please enter a condition.")
            return
        prompt = f"Patient: What is the treatment plan for {condition}? Please generate a plan."
        with st.spinner("Generating plan..."):
            try:
                result = chatbot(prompt, max_new_tokens=100)
                response = result[0] if isinstance(result, list) else result
            except Exception as e:
                response = f"Error: {str(e)}"
            st.success(f"Suggested Plan: {response}")

# patient_chat.py
def main():
    st.title("AI Health Chat")
    st.write("Ask any health-related question to your AI assistant!")
    if chatbot is None or not callable(chatbot):
        st.error("AI model not available or not callable.")
        st.stop()
    if "chat_history" not in st.session_state:
        st.session_state.chat_history = []
    user_input = st.text_input("Ask your question, placeholder: e.g. How can I assist you with your health?")
    if st.button("Ask"):
        if not user_input.strip():
            st.warning("Please enter a question.")
            return
        greetings = ["hi", "hello", "hey", "hlo"]
        if user_input.lower().strip() in greetings:
            response = "Hello! How can I assist you with your health?"
        else:
            prompt = f"User: {user_input} \n AI: "
            with st.spinner("Thinking..."):
                result = chatbot(prompt, max_new_tokens=100, return_full_result=True)
                response = result[0] if isinstance(result, list) else result
            st.info("Model didn't return a valid response. Reason: " + str(result[1]))
    st.session_state.chat_history.append([user_input, response])
    for msg in st.session_state.chat_history:
        st.chat_message(msg[0]).write(msg[1])
    st.session_state.chat_history.clear()

```

- **OUTPUT :**



## 12. Known Issues

- ☐ Generic model outputs due to lack of medical domain fine-tuning
- ☐ Internet dependency when using Hugging Face API
- ☐ No data persistence (currently stateless app)

## 13. Future Enhancements

- ☒ Add user authentication and patient record storage
- ☒ Deploy on IBM Cloud / Hugging Face Spaces
- ☒ Multilingual prompt support
- ☒ Mobile version of the app
- ☒ Integrate with real-time health APIs or EHRs