

CITIZEN AI WITH IBM

Project Documentation

– Introduction

- Project Title: Citizen AI with IBM
- Team Leader: BHARAKATHUL ELMiya S
- Team Member: PAARKAVI M
- Team Member: ALRIFA H
- Team Member: DHILSHA B
- Team Member: YOPIKASRI S

– Project Overview

• Purpose: The purpose of Citizen AI is to provide quick and reliable information about government services and civic issues. It also collects citizen feedback and presents it through simple dashboards for officials, enabling smarter decision-making and better citizen engagement.

• Features:

- Conversational Interface: Citizens can interact naturally and get instant answers.
- Sentiment Analysis: Tracks public opinion and feedback.
- Citizen Feedback Loop: Helps governments listen to people's voices and adapt policies.
- Dashboards: Provides visual insights for officials to make data-driven decisions.
- Gradio Interface: Simple web-based application for demonstration.

3. Architecture

Frontend (Gradio): A user-friendly interface for citizens to interact with the AI system.

Backend (Google Colab + Python): Hosts the application, runs the AI model, and processes inputs.

LLM Integration (IBM Granite Models): The core AI engine that generates responses and summaries.

Version Control (GitHub): Ensures project files are stored, tracked, and shared easily.

4. Setup Instructions

Prerequisites:

- Python Programming Knowledge
- Gradio Framework
- IBM Granite Model Access (via Hugging Face)
- Google Colab with T4 GPU
- GitHub Account Steps:
 1. Access the Naan Mudhalvan Smart internz portal.
 2. Choose an IBM Granite model from Hugging Face.
 3. Run the application in Google Colab with required libraries.
 4. Upload final project files to GitHub.
- 5. Folder Structure
 - app/- Backend logic and integration.
 - ui/ – Gradio app interface files.
 - citizen AI.py– Main application file.
 - model loader.py – Handles IBM Granite model integration.
 - dashboard.py – Visualization of citizen feedback.
- 6. Running the Application
 1. Open Google Colab and load the project notebook.
 2. Install dependencies and configure runtime with GPU.
 3. Run the notebook cells to start the Gradio app.
 4. Access the provided link to interact with Citizen AI.
- 7. API Documentation

Citizen AI provides endpoints for:

 - Asking questions about government services.
 - Uploading feedback for sentiment analysis.
 - Viewing summarized policies.
 - Accessing dashboards and reports.
- 8. Authentication

For the demo, Citizen AI runs in an open setup. In real deployments, authentication methods like API keys, OAuth2, and role-based access would be used.

9. User interface

The Gradio interface is clean and simple with:

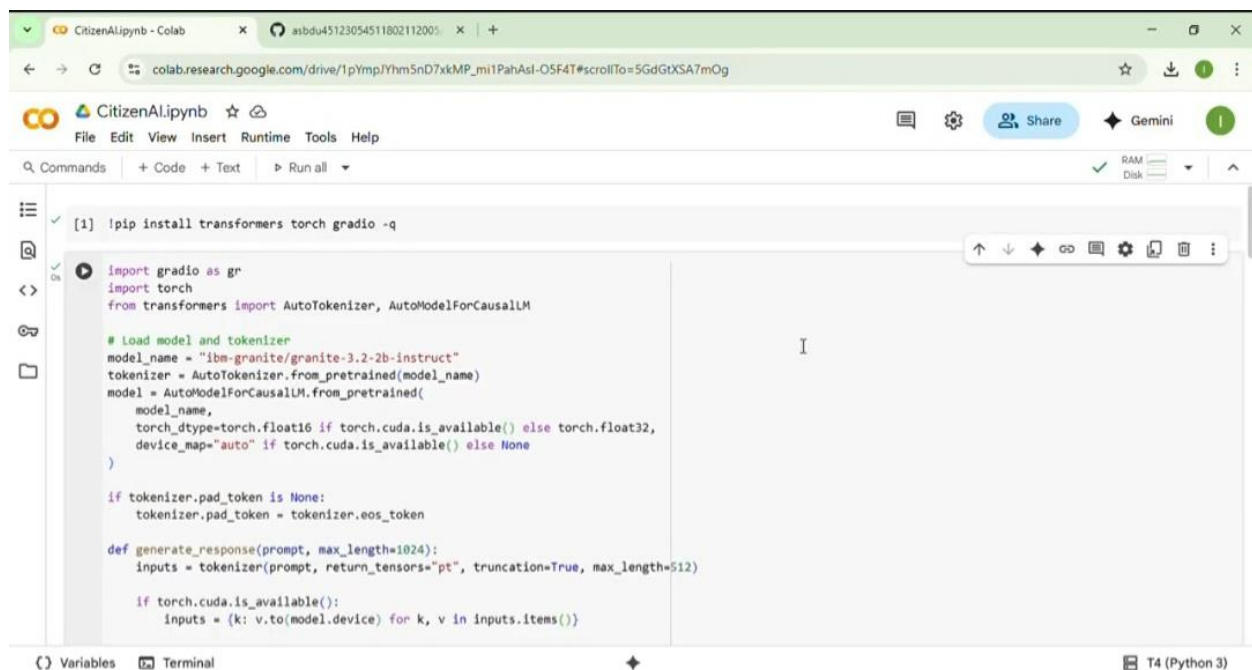
- A chat box for queries.
- Dashboard views for officials.
- Options to upload feedback and documents.

10. Testing

Testing included:

- Unit testing for AI responses.
- Manual testing of the Gradio interface.
- Edge case handling with unexpected inputs.

11. Screenshots



The screenshot shows a Google Colab notebook environment. The top bar includes the Colab logo, the notebook name 'CitizenAI.ipynb', and a share button. The menu bar contains 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. Below the menu is a toolbar with 'Commands', '+ Code', '+ Text', and 'Run all'. The main code editor displays the following Python code:

```
[1] !pip install transformers torch gradio -q

import gradio as gr
import torch
from transformers import AutoTokenizer, AutoModelForCausalLM

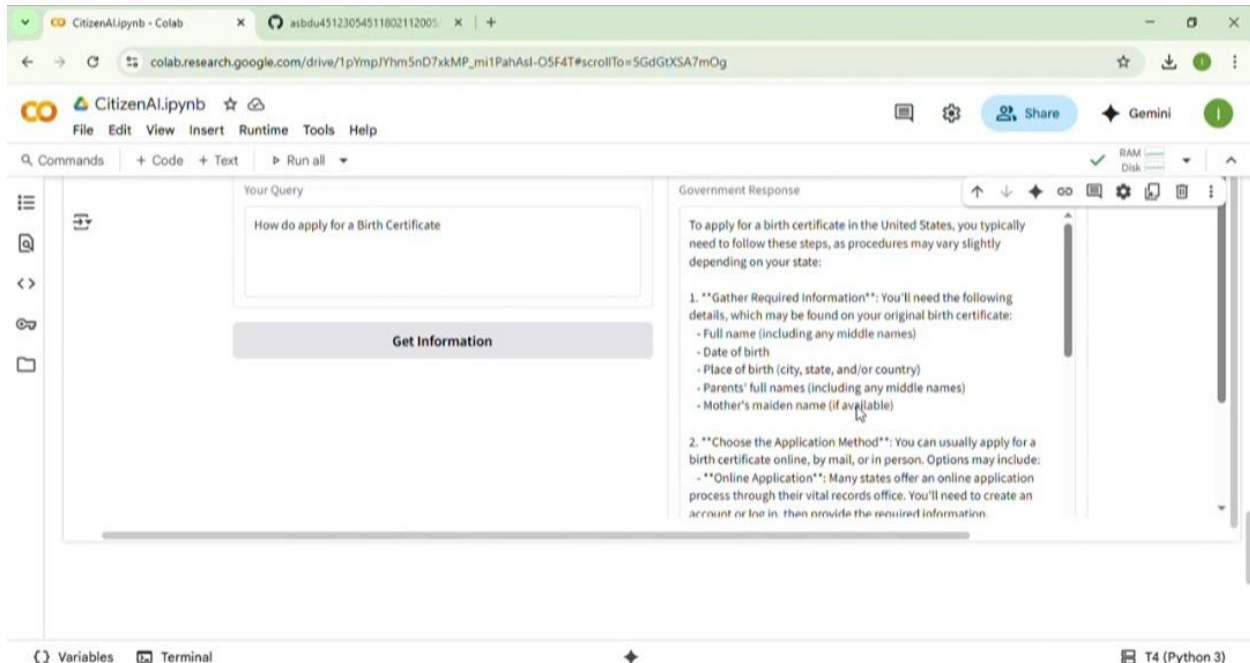
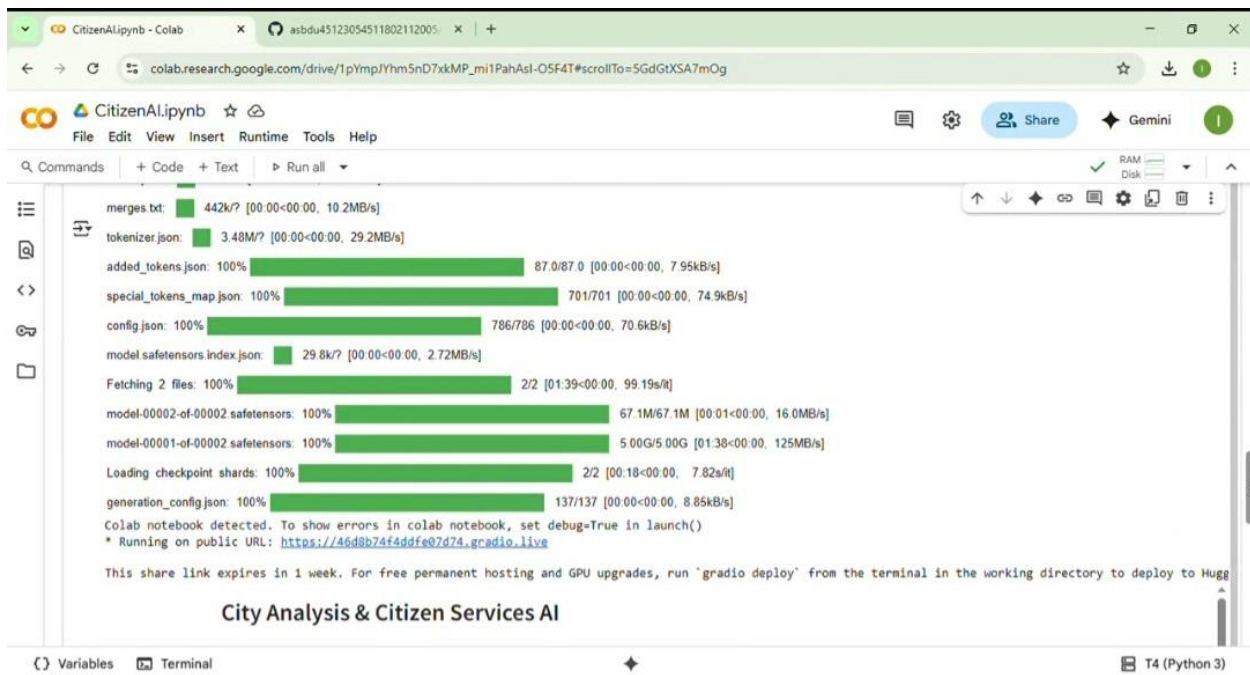
# Load model and tokenizer
model_name = "ibm-granite/granite-3.2-2b-instruct"
tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForCausalLM.from_pretrained(
    model_name,
    torch_dtype=torch.float16 if torch.cuda.is_available() else torch.float32,
    device_map="auto" if torch.cuda.is_available() 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 bottom of the notebook shows tabs for 'Variables' and 'Terminal', and a status bar indicating 'T4 (Python 3)'.



12. Known Issues

- Limited scope due to demo environment.
- Requires internet for Colab runtime and Hugging Face model.

13. Future Enhancements

- Integrate advanced analytics for deeper insights.
- Expand support for multiple languages.
- Deploy on cloud platforms for real-world scalability.