Project Initiation Phase

Project Overview

The City Analysis & Citizen Services AI project is designed to provide two key services:

1. City Analysis – Offers insights into crime index, traffic accidents, and safety assessments for any given city.

2. Citizen Services – Provides government-related information such as policies, public services, and civic queries.

The project leverages the IBM Granite LLM with Hugging Face Transformers and is deployed on Google Colab with T4 GPU runtime to enable faster and more efficient response generation. A simple and interactive Gradio UI allows end users to interact with the system seamlessly.

Objectives

To deliver a unified AI-driven platform for safety analysis and government information services.

To improve citizen engagement by answering queries in natural language.

To ensure transparency, accessibility, and safety awareness through technology.

Scope

Covers analysis of global cities (crime index, accidents, safety).

Provides government-related responses (policies, public services, civic issues).

Accessible through web interface (Gradio).

Can be scaled further for mobile deployment, real-time APIs, and dashboards.

Project Planning Phase

Planning Activities

1. Requirement Analysis

Functional Requirements:

Input: City name or citizen query.

Output: AI-generated detailed analysis/response.

Gradio-based user interface with tabs for City Analysis and Citizen Services.

Non-Functional Requirements:

Fast response (GPU-supported).

Accurate and clear output.

Scalability for multiple users.

2. Technology Stack

Programming Language: Python

Frameworks: Hugging Face Transformers, Gradio

Platform: Google Colab with T4 GPU runtime

Libraries: Torch (PyTorch), AutoTokenizer, AutoModelForCausalLM

3. Milestones

Phase 1: Setup AI model and tokenizer.

Phase 2: Develop city analysis function.

Phase 3: Implement citizen services query handling.

Phase 4: Build and test Gradio interface.

Phase 5: Deploy prototype for users.

Project Execution Phase

Implementation Details

The IBM Granite Model is loaded with Hugging Face. Tokenizer ensures compatibility with user prompts.

A response generation function is implemented with temperature control, padding, and truncation.

Two core modules are developed:

City Analysis: Generates safety statistics and assessment.

Citizen Interaction: Responds to policy-related queries.

Gradio Interface provides two tabs with text inputs and outputs, allowing users to switch between analysis and services.

The project runs in Google Colab GPU runtime, ensuring faster inference compared to CPU.

Challenges Faced

Handling long inputs required token truncation.

Multi-user load caused slight latency issues.

Managing GPU memory efficiently was necessary for stable performance.

Execution Output

Users can input city names (e.g., Mumbai) to get safety analysis.

Users can input queries (e.g., “How to apply for a voter ID?”) to get structured government responses.

The interface is fully functional, interactive, and user-friendly.

Project Monitoring & Closure Phase

Monitoring and Testing

Unit Testing: Checked response generation for short and long queries.

Performance Testing: Verified GPU vs CPU response times.

Load Testing: Tested with multiple sequential queries.

User Acceptance Testing (UAT): Ensured interface is simple and understandable.

Results

Average response time on GPU: 3–5 seconds.

Stable performance under sequential requests.

Acceptable performance under concurrent usage.

Conclusion

The project successfully achieved its objectives of providing AI-based city analysis and citizen services assistance. It demonstrated the potential of integrating large language models with civic governance and public services.

Future Enhancements

Integration with real-time government APIs for dynamic updates.

Multilingual support for wider inclusivity.

Mobile app version for on-the-go access.

Data visualization dashboards (charts, graphs, heatmaps).

Improved scalability through cloud deployment (AWS/GCP/Azure).