Landmark Lens Project Overview

🌐 **Frontend (What Users See)**

- **Technology Used**: React + Vite + Tailwind CSS + Framer Motion

🔸 **What is What:**

- **React**: A JavaScript library used to build dynamic and responsive user interfaces.
- **Vite**: A fast frontend build tool that helps run and compile React apps quickly.
- **Tailwind CSS**: A utility-first CSS framework used to design beautiful and responsive layouts easily.
- **Framer Motion**: A React-based library used to add smooth animations and transitions to UI components.

🔸 **Purpose:**

- Lets users upload an image.
- Displays detected landmark details: name, location, description, and map.
- Adds styling and animations for a professional look and feel.

🔧 **Backend (The Brain Behind the Scenes)**

- **Technology Used**: Python (Flask or FastAPI)

🔸 **What is What:**

- **Python**: The main programming language used to write backend logic.
- **Flask**: A lightweight Python web framework used to create APIs for smaller applications.
- **FastAPI**: A modern, high-performance Python framework used to build APIs faster and with automatic documentation.

🔸 **Purpose:**

- Receives the uploaded image from the frontend.
- Uses an AI model to generate a caption describing the image.
- Matches the caption with a known landmark in the dataset.
- Sends back landmark metadata like name, GPS coordinates, pincode, and description.

📁 **Dataset (Stored Landmark Information)**

- **Format**: JSON file (e.g., landmarks.json)

🔸 **What is What:**

- **JSON (JavaScript Object Notation)**: A lightweight data format used for storing and sharing structured data.
- **landmarks.json**: A file that stores information about landmarks such as:
    - Name (e.g., Taj Mahal)
    - Latitude & Longitude (GPS coordinates)
    - City, country, and pincode
    - A short description (e.g., "A white marble mausoleum in Agra")

🔸 **Purpose:**

- Helps the backend compare the AI-generated caption to known landmark data.
- Acts as a mini database for matching landmark names and returning relevant details.

🧠 **AI/ML Component (The Smart Part)**

- **Model Used**: BLIP (Bootstrapping Language-Image Pretraining) from Hugging Face

🔸 **What is What:**

- **AI (Artificial Intelligence)**: Technology that mimics human intelligence to make decisions.
- **ML (Machine Learning)**: A part of AI that learns from data to make predictions or recognize patterns.
- **BLIP Model**: A pre-trained deep learning model that can generate captions for images.
- **Hugging Face**: A platform that hosts open-source AI/ML models including BLIP.

🔸 **Purpose:**

- Automatically generates a short description of the uploaded image (e.g., "a large white marble structure").
- Enables the backend to intelligently understand image content and identify the matching landmark.

This complete system allows users to simply upload an image and receive detailed landmark information powered by modern web technologies and artificial intelligence.