

AI Legal Advisor — Project Overview

Project: AI Legal Advisor (Indian Law — IPC & CrPC, extendable to IT Act)

1. Problem Statement

Many citizens find it difficult to locate and understand which Indian law sections apply to a given offence. Legal texts such as the Indian Penal Code (IPC) and the Code of Criminal Procedure (CrPC) are comprehensive but hard to search and interpret for non-experts. The project builds an AI-powered RAG (Retrieval-Augmented Generation) Legal Advisor that accepts a plain-English description of a crime, retrieves the most relevant statutory sections, and produces clear, structured explanations (including punishments, procedure references, and practical next steps).

2. Objectives

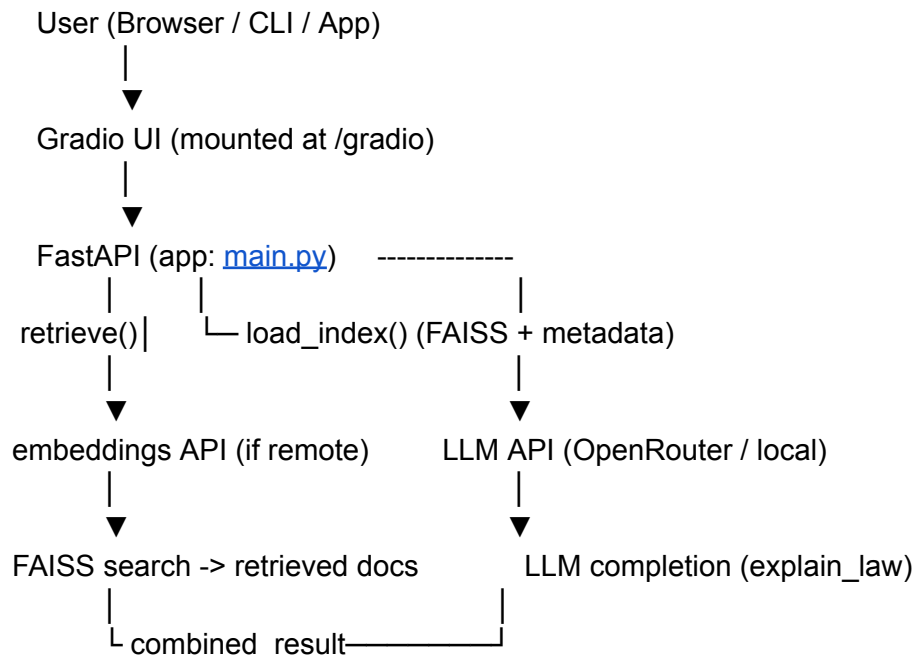
- Provide accessible, accurate, and structured explanations for user crime queries in India.
- Combine semantic search (vector embeddings + FAISS) with an LLM for natural-language output (RAG pattern).
- Support queries across multiple acts: IPC, CrPC, and optionally the IT Act.
- Offer a user-friendly UI (Gradio) and a REST API (FastAPI) for integration.
- Optimize for latency, safety, and clarity — include bailable/compoundable metadata and confidence scores.

3. Database

Data Sources

- **IPC:** Structured JSON/CSV containing section id, title/offense, description, punishment, metadata.
- **CrPC:** Structured JSON/CSV with section, section_name, description, procedural notes.

4. System Architecture



5. Tech Stack

- **Backend API:** FastAPI (Python)
- **UI:** Gradio (Blocks, Chatbot)
- **Embeddings / LLM:** OpenRouter SDK (or local models like sentence-transformers + local LLM)
- **Vector DB:** FAISS (faiss-cpu)
- **NLP:** spaCy `en_core_web_sm` (entity extraction / normalization)
- **Data processing:** pandas

6. File Structure (recommended)

```
legal-advisor/
├── main.py          # FastAPI + Gradio mount
├── embeddings.py    # Build/load FAISS index
├── rag_engine.py    # retrieve() and explain_law()
├── convertcrpc.py  # conversion script (CSV -> JSON)
├── data/
│   ├── ipc_sections.json
│   ├── crpc_sections.json
│   └── it_act_sections.json
├── ipc_index.faiss
└── ipc_data.json    # saved metadata
```

```
|— requirements.txt
|— .env           # OPENROUTER_API_KEY etc.
|— README.md
```

7. Future Enhancements

- Add **case law** and judgments (with citations) to enrich explanations.
- Add **multilingual support** (Hindi first) and localized UI.
- Add **fine-tuned LLM** on Indian legal QA pairs for authoritative tone.
- Add **explainability features**: highlight the exact excerpt in statute used for the answer.