

DocuMind Enterprise – Detailed 4-Week Execution Plan

Project Type: RAG-Based SOP Assistant

Team Size: 3 Interns

Budget: ₹0 (100% Free & Local Stack)

Team Structure:

- Team Lead: Arpit
- Intern A (Backend): Janhavi
- Intern B (Data / QA): Mathangi

WEEK 0 – PREPARATION & ALIGNMENT

- GitHub basics, Python setup, VS Code
- Install Ollama and test local LLM
- Understand RAG basics

WEEK 1 – DOCUMENT INGESTION

Objective: Convert PDFs into vector embeddings.

- PDF parsing
- Chunking strategy
- Embedding generation
- Store in ChromaDB

Responsibilities:

Arpit: Architecture, chunking strategy

Janhavi: Ingestion scripts, setup

Mathangi: Validate extraction & chunks

WEEK 2 – RETRIEVAL & RAG LOGIC

Objective: Question answering from documents only.

- Retriever setup
- Prompt guardrails
- Hallucination refusal

Responsibilities:

Arpit: Prompt & RAG chain

Janhavi: Backend logic

Mathangi: QA & testing

WEEK 3 – FASTAPI & STREAMING

Objective: Expose system as API.

- /ingest and /chat endpoints
- Streaming responses
- Citations

Responsibilities:

Arpit: API contract & review

Janhavi: FastAPI implementation

Mathangi: API testing

WEEK 4 – POLISH & DEMO

Objective: Internship-ready delivery.

- Dockerization
- README
- Demo preparation

Final Outcome:

- Working RAG chatbot
- Zero-cost AI stack
- Clean GitHub repo