HARSH JAIN

Bengaluru / Lucknow | +91-9091125822 | harshf9@proton.me
harshf9.space | github.com/harsh-f9 | linkedin.com/in/harshf9

AI engineer and M.Sc. Computer Applications (AI/ML) student at IIIT Lucknow. Skilled in APIs, SDKs, and RAG, with project experience in forecasting, retrieval, knowledge graphs, and engineering drawing systems.

TECHNICAL SKILLS & CURRENT FOCUS

Languages & scripting: Python, SQL, JavaScript, Bash

Backend & integration: FastAPI, Pydantic, REST APIs, SDKs, Celery

AI, retrieval & data: RAG, PyTorch, XGBoost, LSTM, LangGraph, spaCy, GLiNER, BM25, vector search, OpenCV

Databases: PostgreSQL, SQLite, SQLCipher, Neo4j, Qdrant, Redis

Infrastructure & tools: Git, Docker, Kubernetes, Linux, AWS, Kafka, Datadog, ripgrep, ast-grep

AI development: SDKs and APIs, OpenAI Responses and Decisions, Claude Code

Development workflow: Use Claude Code to work efficiently and manage context to reduce token usage.

Currently learning: Decision Audio Models, CRMs, and harness engineering.

EXPERIENCE

Data Analyst Intern | Climate Resilience Observatory, Govt. of U.P. | Sep 2025 - Jun 2026

Lucknow, UP | FloodCast live | github.com/harsh-f9/floodcast

- Architected and deployed FloodCast, a flood early-warning system covering 378 gauge stations across 69 districts, with 7-day, multi-tier risk forecasts.
- Built automated ETL pipelines for five climate and geospatial data sources and a hybrid LSTM + XGBoost forecasting workflow.
- Integrated an LLM copilot for natural-language queries over reservoir telemetry, historical anomalies, and forecast risk levels.

Account Manager | Sunrise Traders | Pune | Nov 2022 - Apr 2024

- Managed financial accounting, ledger reconciliation, and reporting using Excel, Tally, and BUSY.

SELECTED PROJECTS

Airgap Graph | Offline Entity Graph & Engineering Drawing Intelligence | 2026 - Present

- Built an offline-first platform that extracts entities and co-occurrence relationships, with an interactive force-directed graph, neighborhood exploration, and source-level character-offset provenance.
- Developed a blueprint vision pipeline using aspect-preserving scaling, overlapping image tiles, adaptive zoom, and sub-patch localization for architectural and engineering diagrams.
- Added visual proof crops, coordinate mapping, and confidence checks for grounded diagram answers; secured local persistence with encrypted storage and optional SQLCipher.

Retrieval Funnel Toolkit | Codebase Indexing & Architecture Analysis | 2026

github.com/harsh-f9/retrieval_funnel_toolkit

- Built a six-stage code analysis pipeline using ast-grep and ripgrep to produce structured dependency graphs from Git repositories.
- Used Pydantic schema validation and PageRank-based ranking to prioritize architectural context; reported benchmark reduced LLM input from 8.7M to 2.1M tokens (75.86%).

EDUCATION & RECOGNITION

M.Sc. Computer Applications (AI/ML), IIIT Lucknow | Aug 2025 - Present

B.Sc. (General), North Bengal St. Xavier's College, Siliguri

Delhi Public School, Siliguri

Meta x Scaler Reinforcement Learning Hackathon: Top 100 finalist, Bengaluru.

IIT JAM Mathematics 2025: Qualified, All India Rank 786.