Gauri Nigam

nigamgauri24@gmail.com Bangalore, India
github.com/nigamgauri gaurinigam.vercel.app linkedin.com/in/gauri-nigam

Skills

Languages Python (async), TypeScript, C++, SQL, Solidity
AI LangChain, LangGraph, RAG, agentic and multi-agent orchestration, LLM evals and observability, prompt engineering, LLM APIs (Anthropic, Gemini), vector databases, Hugging Face, PyTorch, MLOps
Web3 Foundry (Forge, Anvil, Cast), EVM execution model, gas mechanics, Chainlink oracles, fuzz and invariant testing, Slither static analysis, reentrancy / access control / front-running
Backend & Cloud FastAPI, Next.js, React, PostgreSQL, MongoDB, Redis, Docker, AWS (EC2, S3, Lambda), CI/CD, Git

Experience

Software Development Engineer Intern Aug 2026 - Present
Krowd *Bangalore, India*

- Took a UK-only transaction platform into a second region under legal data-residency constraints, designing per-request region binding with Python ContextVars and region-keyed Postgres connection routing so one FastAPI deployment serves both regions from physically separate databases.
- Traced a silent cross-region data bug to the Next.js server-side proxy hop, where six routes building their own requests dropped the region header and returned UK data for US queries; shipped a shared forwarding helper with regression tests.
- Hardened authentication after an internal audit, adding token verification on unguarded routes, JWT issuer and token-type validation, and session revocation on logout and password change, with attack tests proving the API cannot be bypassed outside the UI.

AI Engineering Intern Dec 2025 - Mar 2026
TechCirkle *Remote*

- Cut average RAG query latency by 35% on an internal benchmark by tuning adaptive chunking and retrieval strategies across a LangChain and Hugging Face embeddings pipeline.
- Reduced LLM token usage by 40% through prompt compression and KV-cache reuse, and built an eval harness tracking factual recall and context relevance as documents changed underneath the index.

Achievements

- **Open Source:** Haystack #11828 (21k+ stars), fixed an embedding precision bug breaking document equality; traceAI #184, fixed a span-overwrite bug silently corrupting multi-turn trace data, merged with the founder.
- **Hackathons:** 2nd of 62,000+ teams at the TVS Credit EPIC Hackathon, India's largest student hackathon; finalist at the ChurnZero Hackathon, IIT Kharagpur (1,300+ teams).
- **Competitive Programming:** Codeforces Expert (1769, top 2% globally); CodeChef 3-star (1749).

Projects

LLMEval
LLM-as-judge evaluation framework scoring RAG and LLM outputs on faithfulness, relevance, hallucination and coherence through structured JSON grading prompts. Validated against human ratings with Cohen's Kappa to confirm the scoring pipeline held up. Live.

UrbanMind
Multi-agent LangGraph workflow with intent-classification, recommendation and rendering nodes, including fallback and replanning logic, wired to an interactive floor-plan visualization. Live.

GreenLight Go
Real-time ambulance detection and adaptive traffic-signal control using YOLOv8 and OpenCV, at 97.3% mAP and 30 FPS. Quantized to INT8 for CPU-only edge deployment with no accuracy regression.

Education

B.Tech in Computer Science & Engineering, Institute of Engineering and Technology, Lucknow Expected 2027
CGPA: 8.8/10