

Sohaib

AI Application Engineer · Senior Software Engineer | LLM Apps · RAG · Agents | Python · FastAPI · Node.js · AWS

Gurgaon / New Delhi, India | 7701849698 | ssohaib4@gmail.com

linkedin.com/in/sohaib-07106a131 | github.com/pseudog0d | s0ha1b.netlify.app

SUMMARY

Software engineer with 4+ years shipping production backend and full-stack systems, now building LLM applications: retrieval-augmented generation, agentic workflows, tool calling, schema-constrained output and evaluation harnesses. Backend foundation in Python, FastAPI, Django, Node.js and React/Next.js over PostgreSQL, MySQL and MongoDB — REST APIs, auth, async event-driven pipelines, microservices, Docker and AWS. Ships AI systems that stay debuggable in production.

EXPERIENCE

Hughes Network Systems

Feb 2024 – Present

Senior Software Development Engineer · via Teleperformance

Gurgaon, India

- Own end-to-end delivery of enterprise web applications and AI-integrated APIs (React / Next.js with FastAPI / Node.js), from technical design through production rollout, improving page-load performance 23%.
- Designed REST and streaming APIs over PostgreSQL / MySQL / MSSQL with authentication, authorization and audit logging, improving backend throughput 13% and lowering latency under concurrent load.
- Built async, event-driven pipelines (Celery, Redis, WebSockets) for inference-style and embedding workloads, and containerized microservices (Docker, nginx) with CI/CD on AWS, cutting deployment setup 20%.

Handshake AI — Contract

2026 – Present

AI Training & Evaluation Engineer

Remote

- Author verifiable software-engineering evaluation tasks used to train and benchmark large language models, spanning security, machine learning, data processing / ETL and digital forensics domains.
- Build automated graders, deterministic test harnesses and static preflight checks in Python and pytest, enforcing reproducible CI-verified scoring for every task.

IndiaMART

Mar 2022 – Feb 2024

Software Development Engineer

Noida, India

- Built high-traffic React.js / Next.js surfaces backed by Node.js / Express and Python services with REST and third-party API integrations, achieving 1.5x faster TTI and improved SEO.
- Developed scalable backend services and relational data models (MySQL / PostgreSQL) with optimized queries powering search- and recommendation-style features, improving operational efficiency 18%.

PROJECTS

Autonomous AI Job-Application Agent

2025 – Present

Python · FastAPI · Gemini · Groq · Next.js · PostgreSQL

Personal

- Built an end-to-end autonomous agent that scrapes job sources, vets fit, tailors resumes and drafts outreach, orchestrating Google Gemini and Groq LLMs with automatic provider fallback for reliability.
- Engineered prompt pipelines with strict structured-JSON output, theme-coverage gates and anti-hallucination guardrails, plus retrieval and scoring over JD data and an LLM pre-screening vetter.
- Shipped a Next.js + FastAPI review dashboard with async run history and logging for human-in-the-loop oversight of agent decisions.

TECHNICAL SKILLS

AI / LLM: LLM Application Development, Retrieval-Augmented Generation (RAG), Prompt Engineering, AI Agents / Agentic Workflows, Tool & Function Calling, Structured Output (JSON Schema), Embeddings, Vector Search, LLM Evaluation, Google Gemini, Groq, OpenAI API

Languages: Python, JavaScript, TypeScript, SQL, Java

Backend: FastAPI, Node.js, Express.js, Django, Flask, REST APIs, WebSockets, AsyncIO, Celery, Microservices, Auth (JWT, OAuth2)

Data: PostgreSQL, MySQL, MSSQL, MongoDB, Redis, pgvector, Kafka, Airflow, ETL, Pandas, Query Optimization

Frontend & Cloud: React.js, Next.js (SSR), Redux, AWS (EC2, S3), Azure, Docker, nginx, Linux, CI/CD, GitHub Actions, Git

Practices: System Design, Distributed Systems, Event-driven Architecture, Agile, Pytest, Testing

EDUCATION & CERTIFICATIONS

B.Tech, Computer Science & Engineering — Amity University, Delhi, 2018–2022 · **Certifications:** Microsoft Azure Developer Associate (AZ-204), CCNA