Akshit Gupta

f20231136@hyderabad.bits-pilani.ac.in | +91 99585 71609 | linkedin.com/in/akshit-gupta | github.com/akshit9162

Education

BITS Pilani, Hyderabad Campus CGPA: 8.12
B.E., Electrical and Electronics Engineering; Minor in Computing and Intelligence *2027*

Delhi Public School, Sushant Lok, Gurgaon 95.2% | 96.6%
Class XII | Class X (CBSE) *2023 | 2021*

Technical Skills

Languages: Java, Python, JavaScript, C++, C#, C, SQL
Backend & Systems: Spring Boot, FastAPI, REST, microservices, Kafka, Redis, PostgreSQL, Oracle PL/SQL, Docker, AWS
Tooling & Testing: Git, GitHub Actions, GitLab CI/CD, CMake, Linux/macOS, JUnit, pytest, Playwright, ASan/UBSan
AI & Search: RAG, semantic/vector search, hybrid (dense + BM25) retrieval, reranking, LLM agents, tool calling, PyTorch
AI-Assisted Engineering: Claude Code for coding, debugging and test generation, with every change reviewed and tested
Coursework: Data Structures & Algorithms, Operating Systems, OOP, DBMS, Machine Learning, AI

Experience

Upsure.AI May 2026 – Jul 2026
AI/ML Engineering Intern

- Architected an event-driven **FNOL claims pipeline** (**Kafka**, Kogito rules on an **Apache AGE graph**, FastAPI state machine) routing cases to **ML models**, re-engaging users via SSE and queueing edge cases for time-bounded **human review**.
- Built the **telemetry SDK** and PostgreSQL **audit-logging** module for an internal observability platform, with **SHA-256 hash-chaining** for tamper-evident compliance logs and non-blocking instrumentation via daemon threads.
- Deployed a CPU-optimized **FastAPI** service (**ONNX Runtime** + RapidOCR) for damage detection and **multilingual OCR**: precision **0.56** → **0.83**, OCR latency ∼**14s** → ∼**4s**.
- Integrated Gitleaks, Trivy and SonarQube into **CI** across multiple **GitLab** repositories, cutting manual testing time by **80%**; added a **code-fix agent** that **retrieves** relevant source files and tests LLM patches in isolated git worktrees.
- Built an **autonomous LLM QA agent** (**tool calling**, **structured outputs**) that writes Playwright, Postman and k6 test suites and **self-heals**: failing tests go back to the LLM, which patches them or **classifies** them as real bugs.

Coffeee.io May 2025 – Jul 2025
Software Development Intern

- Built an **AI-powered resume parser** in **Python** with **spaCy** and **NLTK NER** to extract data from **PDF/DOCX** files.
- Added **regex** and **fuzzy matching**, batch processing, JSON/CSV export, **SQLite** storage and a **Streamlit** UI.

Projects

Distributed Auction & Order Platform | akshit9162.github.io/Distributed-Order-System

- Engineered **Java/Spring Boot microservices** (order, inventory, payment) with **SAGA orchestration** and compensating transactions; **Redis** distributed locks, **JPA optimistic locking** (**0 oversells** across 50 concurrent orders), **idempotency keys** and **Kafka** dead-letter topics; k6 **p95 137 ms** on AWS.
- Added an **auction service**: second-price **proxy bidding** with bid increments as a pure domain model; **optimistic locking** with jittered retry keeps **1,000 concurrent bids** identical to serial replay (**379 conflicts** retried).
- Hands winners to the order SAGA via a **transactional outbox** (at-least-once, **idempotent consumer**); **observer**-based live **JavaScript** page over **Server-Sent Events** with an order tracker; **41 JUnit tests** in CI.

tickflow: Low-Latency Order-Matching Engine (C#) | akshit9162.github.io/tickflow

- Built a **price-time-priority order book** behind an **ASP.NET Core** REST + **WebSocket** API on a lock-free **single-writer** engine thread: **0.3 µs p50** matching, **476K orders/s**.
- Crash-safe via the **command pattern** and an **event-sourced** CRC32 write-ahead log with **group commit** (**55×** per-order fsync throughput): **0 of 43,808** acked orders lost across 10 **kill -9** tests.
- Trades fan out to **observers** (WebSocket feed, SQL **repository**); live **JavaScript** trading screen with backoff reconnects; **28 tests** in CI.

Grounded RAG Platform | akshit9162.github.io/Grounded-RAG-Platform

- Built a **citation-grounded RAG** service (**FastAPI**, Claude) with **hybrid BM25 + dense retrieval** fused by reciprocal rank, **cross-encoder reranking** and semantic caching, so every answer cites its sources; **Dockerized** on AWS.
- Built an **evaluation harness** (**RAGAS**) scoring retrieval and generation separately (**1.0 context recall**, **0.947 answer relevancy**), plus a **guardrail** that flags answer sentences unsupported by the retrieved sources.

Achievements

- **Codeforces Expert** (max 1606) • LeetCode 1627, 400+ problems solved • **NTSE** Scholar, State Rank 155, Haryana.