

NIRANJAN MURUGARASU

Email: niranjanmurugarasu@gmail.com

LinkedIn: niranjan-murugarasu-b8b706289

LeetCode: niranjanmurugarasu

Mobile: +91-9994066134

GitHub: Niranjan-Murugarasu024

HuggingFace: BlueMen

Portfolio: portfoliovercel-ebon-nu.vercel.app

Summary

Final-year B.Tech Artificial Intelligence and Data Science student building reliable,evaluated GenAI systems – from prompt design and multi-metric output-quality frameworks to full-stack deployment (FastAPI, Docker). Experience spans LLM-driven product design (adaptive prompt-based interview simulation, automated output-safety validation) and deep model customization (QLoRA/LoRA fine-tuning, published on Hugging Face). Strong in Python, PyTorch, and translating ambiguous problems into measurable AI-quality metrics; 500+ LeetCode problems solved.

Education

B.Tech in Artificial Intelligence and Data Science

Dr. N.G.P Institute of Technology, Coimbatore

2023–2027

CGPA: 8.65

Higher Secondary

Elgi Matric Hr. Sec. School, Coimbatore

2023

86%

Technical Skills

- **GenAI / LLMs:** Prompt Engineering, LLM Output Evaluation & Validation, RAG, Vector Databases, LoRA/QLoRA Fine-Tuning, Transformers
- **ML / Data Science:** PyTorch, Scikit-learn, Supervised/Unsupervised Learning, NLP, Feature Engineering
- **Cloud & MLOps:** GCP, Docker, Kubernetes, CI/CD, Model Deployment
- **Programming & Databases:** Python, Java, SQL
- **Frameworks & Tools:** FastAPI, Streamlit, Git/GitHub
- **Core CS:** Data Structures & Algorithms, OOPs, DBMS, Operating Systems, System Design
- **Mathematics:** Probability, Statistics, Linear Algebra, Optimization

Internship

Cloud Intern, Accent Techno Soft

June 2025

Gained practical industry exposure in cloud engineering and architecture. Developed hands-on proficiency with Google Cloud Platform (GCP). Focused on cloud resource management and cloud security policies.

Graphic Designer Intern, Visaiithalam Solutions

October 2023 – February 2024

Delivered branding and marketing creatives across client projects under deadline constraints

Projects

1) MedRx – Fine-Tuned Clinical Drug & Treatment Recommendation LLM

2026

- Built a post-inference output-validation layer for a fine-tuned clinical LLM, running automated drug-allergen cross-reactivity and dosage range checks against a structured (Pydantic-parsed) schema to catch unsafe model outputs before they reach an end user.
- Designed a 7-gate data-quality filtering pipeline (PHI scrubbing,MD5 deduplication, drug-class-templated synthetic augmentation) that reduced 172K-record raw corpus to 13.7K high-quality training samples (92% rejection rate) – directly defining what "high-quality" meant for the fine-tuning task.
- Fine-tuned Microsoft's Phi-4-mini-Instruct (3.8B) via 4-bit QLoRA (NF4) with Unsloth, training an 8.9M-parameter LoRA adapter (0.23% of base) on the resulting 13,728-sample clinical-reasoning dataset.
- Root-caused and resolved a Triton JIT compiler segfault specific to Ada Lovelace GPUs via native stack-trace analysis; published the resulting model publicly on Hugging Face.

2) SwiftDesk AI – Intelligent Ticket Triage, SLA Escalation & Auto-Routing System 2026

- Engineered an automated multi-tier support routing backend in Java 25 and Spring Boot 3, implementing a keyword-based NLP triage engine for instant category classification, severity resolution, and customer priority dishonesty (trap) detection.
- Designed a real-time load-balancing assignment engine across L1, L2, and L3 engineering tiers based on active workload capacity ratios (`active_tickets / max_capacity`), featuring automated priority upgrade triggers upon SLA escalation.
- Architected a persistent relational data model in PostgreSQL 17 via Spring Data JPA and HikariCP, capturing full lifecycle audit logs across `tickets`, `ticket_history`, and transactional `email_logs` with non-destructive database seeding.
- Built a dynamic React 19 and Material-UI customer/support dashboard connected via an Axios API service with automatic mock-fallback resilience for offline testing and seamless RESTful state synchronization.

3) PrepAI — AI-powered Interview Simulation Platform 2025

- Designed and iterated prompt templates for Llama 3.3 70B (via Groq) to generate context-aware interview questions and adaptive follow-ups conditioned on job descriptions and prior candidate responses.
- Built a 5-metric evaluation framework (technical accuracy, communication, business impact, confidence, structured thinking) to score AI-generated candidate responses, translating qualitative LLM output into a structured 0–10 rubric with STAR-method feedback.
- Shipped the system on a full-stack architecture (Next.js, Node.js/Express) with Groq-hosted Llama 3.3 70B inference, generating detailed, actionable performance reports per candidate.

4) Scrutinize – Event-Driven GitHub Audit & Durable Queue Engine 2026

- Engineered an event-driven GitHub App in TypeScript/Node.js (Probot) to audit PR review thoroughness; decoupled incoming webhooks into a durable job queue, achieving under 50ms fast-ack responses to eliminate GitHub webhook timeouts.
- Designed an idempotent queue architecture using SQLite `UNIQUE` header constraints to silently drop duplicate deliveries, paired with atomic row-level claims (`UPDATE ... WHERE status IN ('pending','failed')`) to prevent race conditions during concurrent worker sweeps.
- Developed a deterministic rule engine calculating diff-based reading speed floors (0.4s per line, 30s minimum) and detecting AI co-author signatures; utilized Octokit pagination to bypass default 30-item API caps across PR commits, reviews, and inline comments.
- Implemented a periodic self-healing sweep daemon for automatic crash recovery (resumes jobs stuck processing over 5 minutes or failed under 5 attempts) and migrated storage to Turso Cloud SQLite (`@libsql/client`) to prevent data loss on ephemeral PaaS containers.

5) Medicl: Enterprise-Grade HIPAA-Compliant SaaS Platform 2025

- Automated missing value handling, outlier detection, and schema validation.
- Integrated anomaly detection using Autoencoders and Isolation Forests.
- Built CLI and Streamlit-based visualization interface.

Co-Curricular Activities

- Participated in the Slack Agent Builder Challenge, developing **StalePR**, an AI-powered GitHub pull request management assistant.
- Launched and author the monthly LinkedIn newsletter **The AI Edge**, sharing insights on artificial intelligence, emerging technologies, and industry trends.
- Presented a research paper on an AI-powered Healthcare Data Handling Platform at **Coimbatore Institute of Technology (CIT)**, Coimbatore.
- Presented a research paper on **Solar Technology** at **KPR Institute of Engineering and Technology**, Coimbatore.
- Participated in the **Illuminate Entrepreneurship Bootcamp** organized by the E-Cell, IIT Bombay, focusing on innovation, startup ideation, and entrepreneurship.

Certifications

- AI Fluency Framework – **Anthropic** 2025
- Generative AI Fundamentals – **Databricks** 2026
- Elements of AI – **University of Helsinki** 2026
- GenAI Powered Data Analytics Job Simulation – **Tata** 2026