

Krishna Gupta

Bhopal, Madhya Pradesh, India | +91-9993153109 | hg497kg@gmail.com
[linkedin.com/in/krishnaofficialgupta](https://www.linkedin.com/in/krishnaofficialgupta) | github.com/KGupta171025 | kgupta.is-a.dev

PROFESSIONAL SUMMARY

Software Engineer & AI Systems Developer | B.Tech Data Science (Class of 2027) ranking in the **97th National Percentile (611/800, 100% in Coding, 87% in Computer Fundamentals) in the Aon CoCubes National Employability Assessment** with certified High Job Suitability for Software Engineer roles. Progressive industry experience architecting production **FastAPI, Django, and Flask** microservices, distributed AI systems (LLMs, RAG, Agentic AI, Prompt Engineering), automated testing frameworks (Pytest, Postman, Locust), and CI/CD-enabled cloud infrastructure on AWS and Docker. Proven expertise delivering high-throughput RESTful APIs, PostgreSQL-backed services, and LangChain/LangGraph multi-agent systems with 99.9% uptime. Passionate about software architecture, quality engineering, test automation, and building scalable, robust distributed solutions.

EDUCATION & STANDARDIZED ASSESSMENT

Oriental Institute of Science and Technology

Bhopal, Madhya Pradesh, India

Bachelor of Technology (B.Tech) in Data Science (Class of 2027)

Aug 2023 – May 2027

• **Aon CoCubes National Employability Assessment (Jun 2026):** Overall Score **611/800** (“Excellent”) | **97th National Percentile** • 98th Regional Percentile | **100% Score in Coding (5/5) • 87% in Computer Fundamentals (5/5)** | Certified High Job Suitability for **Software Developer & Software Engineer** roles.

TECHNICAL SKILLS

Programming & Core CS: Python (Advanced, mandatory – OOP, DSA, Data Structures & Algorithms), SQL (PostgreSQL, MySQL, Snowflake), Java (Core), JSON/XML, Object-Oriented System Design (HLD/LLD), Agile/Scrum, Jira

Backend & Web Frameworks: **FastAPI, Django, Flask**, RESTful API Architecture, Microservices, Node.js, React.js, Next.js, API Gateways (Nginx), Event-Driven Architecture, Webhooks, Message Queues, Distributed Caching (Redis), ETL/ELT Pipelines

Generative AI & Agentic AI: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG / Hybrid RAG), Prompt Engineering, Agentic AI, Autonomous Multi-Agent Systems, LangChain, LangGraph, Azure OpenAI, AWS Bedrock, Hugging Face Transformers, FAISS / pgvector, Vision-Language Models (VLMs), Model Alignment (RLHF/DPO)

Test Automation & Quality Engineering: Pytest (automated test frameworks & fixtures), API Testing, Functional Testing, Integration Testing, End-to-End (E2E) Testing, Performance Testing (Locust), Regression Testing, Test Plans & Defect Analysis, Postman Collections & Newman

Cloud, DevOps & MLOps: AWS (EC2, S3, IAM, CloudWatch), Microsoft Azure, Docker Containerization, CI/CD Pipelines (GitHub Actions), MLflow (Tracing, Evaluation, Prompt Templates, AI Gateways, Agent Server), Linux Server Administration, Cloudflare Workers

Machine Learning & Data Engineering: PyTorch, TensorFlow, Scikit-learn, Time-Series Forecasting (LSTM), Deep Neural Networks (CNNs), Computer Vision (OpenCV, EasyOCR), Dimensional Modeling (Star/Snowflake Schemas), Semantic Search, Pandas, NumPy, Knowledge Graphs

WORK EXPERIENCE

NexoraMind Tech

Remote

AI Engineer Intern – Product Engineering & AI Automation (6-Month Program)

Sep 2026 – Present

- Architecting and developing production-grade **FastAPI, Django, and Flask** microservices and intelligent **agentic AI workflows** using LangChain/LangGraph, LLMs, and RAG pipelines; writing **Pytest automated test suites** (API testing, integration, functional) ensuring >95% test coverage and system reliability.
- Building scalable **Hybrid RAG** semantic retrieval engines and **REST API** integration layers; validating request/response contracts with **Postman**, enforcing JSON schema validation, and executing end-to-end regression test cycles in CI/CD pipelines.
- Designing robust **AI evaluation benchmarks**, performance test harnesses (Locust), and safety guardrails to audit inter-agent communication, token consumption, and ensure deterministic execution compliance in enterprise-scale deployments.

Ethara.AI

Remote

LLM & AI Post Training Intern (Paid Internship)

Feb 2026 – May 2026

- Benchmarked and evaluated 50,000+ LLM outputs within **LLMOps workflows**, applying structured **Prompt Engineering** and automated validation pipelines to improve model alignment, precision, and safety guardrails by 15% across diverse input distributions.
- Designed and automated scalable **Python** data preprocessing, feature extraction, and **functional test** validation pipelines, reducing manual QA overhead by 20+ hours/week; documented test plans and collaborated with AI research teams on **defect root-cause analysis**.
- Formulated structured capability rubrics for **multi-agent AI** systems and reinforcement learning (RLHF) environments; co-authored technical documentation on dataset curation best practices in Agile sprint workflows.

Kanchan Pvt Ltd – Web Development Wing (Sapphire)

Remote

Full Stack Development Intern (Paid Internship)

Oct 2025 – Feb 2026

- Built and deployed full-stack **REST API** services for review analytics platform “RevU Social” (revu.social) using Node.js & Python; validated all endpoints via **Postman** functional and integration test collections, documented API contracts, and resolved 30+ defects in Agile sprints.
- Optimized **PostgreSQL** and **MySQL** relational database schemas, query execution plans, and indexing strategies, reducing query latency by 25%; wrote **complex SQL** queries for cross-system data validation and business intelligence.
- Implemented **JWT** authentication, API rate limiting, and deployed **Docker**-containerized microservices on Linux cloud servers with **CI/CD** integration, maintaining 99.9% uptime across production environments.

KEY TECHNICAL PROJECTS

Stock Market Future – AI-Powered Financial Forecasting & Analytics Platform 2024 – Present
Technologies: Python, FastAPI, PyTorch (LSTM + Transformers), Scikit-learn, MLflow, REST APIs, Pytest, React 19, Firebase, Docker
Live Platform: smf-ai.web.app | GitHub: github.com/KGupta171025/Stock-Market-Future

- Engineered a full-stack financial platform processing 10,000+ daily market ticks; implemented dual LSTM and Transformer models for time-series equity forecasting with automated inference pipelines, reducing prediction latency to <200ms.
- Architected decoupled **FastAPI microservices** with comprehensive **Pytest unit and integration test coverage**; validated REST API endpoints with Postman collections, enforced input data integrity, and automated regression test suites.
- Integrated **MLflow tracking and model registry** to monitor model drift, evaluation metrics, and hyperparameter runs; automated containerized deployments via Docker and GitHub Actions CI/CD to Firebase Hosting with zero downtime.

KALKI 1.5 – Autonomous Multi-Agent AI OS & MLOps Platform 2024 – Present
Technologies: Python, FastAPI, LangGraph, LangChain, LLMs, VLMs, Azure OpenAI, AWS Bedrock, FAISS, Pytest, MLflow, Postman, CI/CD, Docker
Live Demo: kalki.hg497kg.workers.dev | GitHub: github.com/KGupta171025/KALKI-1.5

- Engineered an autonomous multi-agent AI operating system integrating LLMs and VLMs via LangGraph and LangChain with Agentic AI workflows; developed automated Prompt Engineering evaluation pipelines and RAG agent benchmarks.
- Architected a low-latency Hybrid RAG pipeline with dense FAISS vector embeddings and in-memory caching; validated retrieval precision, recall, and end-to-end API correctness using Postman collections and Pytest fixtures, cutting latency 40%.
- Integrated MLflow tracing, prompt evaluation templates, and guardrail logging to audit multi-agent token usage, safety, and determinism; containerized services in Docker with CI/CD deployment pipelines for repeatable releases.

Score Vision AI (FOG Tech) – Computer Vision & Document OCR Extraction Platform 2025 – Present
Technologies: Python, OpenCV, EasyOCR, Sentence-Transformers, FAISS, FastAPI, React, Docker, REST APIs, Pytest
Live Demo: scorevision-ai.hg497kg.workers.dev | GitHub: github.com/KGupta171025/FOG_Technologies

- Developed an intelligent OCR and computer vision pipeline achieving 92% text extraction accuracy; designed a decoupled FastAPI microservice architecture with Docker to handle high-throughput image processing at scale.
- Implemented automated functional and API tests (Pytest, Postman) to validate OCR output quality, FAISS vector similarity search correctness, and sub-200ms semantic matching SLA compliance across regression cycles.

Razorpay Recovery – Smart Payment Recovery & FinTech Transaction Pipeline 2025
Technologies: Python, SQL, PostgreSQL, REST APIs, Webhooks, Event-Driven Architecture, Pytest, Data Pipelines
Live Demo: razor-recovery.hg497kg.workers.dev | GitHub: github.com/KGupta171025/Razor-Recovery

- Constructed an automated event-driven transaction recovery pipeline capturing failed payment webhooks in real time with idempotent processing; wrote integration and functional Pytest suites to validate pipeline reliability, recovering 18% of abandoned checkouts.
- Designed analytical star-schema databases in PostgreSQL, wrote complex SQL queries to aggregate failure patterns, automate notification triggers, and validate business rule logic through structured data testing workflows.

HACKATHONS & EVENTS

HackerRank Orchestrate – October Edition 2026 Oct 2026 • HackerRank
Selected as a verified Participant in **HackerRank Orchestrate – October Edition 2026**, a competitive coding and AI orchestration challenge focused on algorithm design, API development, and intelligent workflow automation.

Nuroen Agent Forge – In-Person AI Agent Hackathon Oct 2026 • Bengaluru
Participated in an in-person AI agent hackathon organized by **Nuroen.ai** at OJone, Bengaluru – built governed, automated agentic workflows using LLMs and multi-agent orchestration frameworks.

Full Stack Web Development Challenge 2026 – Code-A-Nova × AptiMaster Aug 2026
Completed the **Full Stack Web Development Challenge 2026** by Code-A-Nova and AptiMaster AI – assessed on full-stack architecture, REST API design, and modern web development best practices.

AICTE OIB-SIP Data Science Internship – Star Performer Award May – Jun 2026 • Oasis Infobyte
Recognized as **Star Performer** among all cohort participants in the AICTE-sanctioned Oasis Infobyte SIP (OIB/M1/IP301) – awarded for exceptional dedication and outstanding contributions throughout the 1-month Data Science program.

Nestlé Nesternship Program – E-learning Completion 2026 • Nestlé
Completed the **Nestlé E-learning 2026 | Resilience** module as part of the Nesternship – a globally recognized corporate learning initiative by Nestlé Needs YOUTH for emerging professionals.

Bluestock Fintech – Software Development Engineer (SDE) Intern Feb – Mar 2026 • Remote
Selected as **SDE Intern** at Bluestock Fintech (ID: BFSD234758) – contributed to software development in a dynamic fintech environment, gaining hands-on experience with real-world financial software workflows.

CERTIFICATIONS & ASSESSMENTS

- Aon CoCubes Assessment:** 97th %ile (611/800, 100% Coding, Certified SE)
 - Generative AI with AWS** (AWS, Jun 2026)
 - Neural Networks & CNNs** (LinkedIn Learning, Jun 2026)
 - Data Science & Analytics** (HP LIFE, Jun 2026)
- AWS Certified Developer Associate** (Infosys / AWS, Jun 2026)
 - Machine Learning with Python** (IBM SkillsBuild, Jun 2026)
 - Gemini for Google Workspace** (Google Cloud / Simplilearn, Dec 2025)
 - Deloitte Australia Data Analytics Simulation** (Forage, Sep 2025)

RELEVANT COURSEWORK

Data Structures & Algorithms • Object-Oriented Programming • Database Management Systems (RDBMS) • Software Engineering & Testing • Computer Networks & Distributed Systems • Operating Systems • Machine Learning & Deep Learning • Big Data Engineering • Cloud Computing (AWS/Azure)