

# Kunal Sharma

Software Engineer — Machine Learning Specialist

[kunalsharma13579kunal@gmail.com](mailto:kunalsharma13579kunal@gmail.com) | +91-639-339-2926 | [linkedin.com/in/kunalsharma](https://linkedin.com/in/kunalsharma) | [github.com/MrKunalSharma](https://github.com/MrKunalSharma)  
Eldeco Garden Estate Raipurva, Kanpur, Uttar Pradesh

## PROFESSIONAL SUMMARY

Computer Science student at IIIT Kota with 8.8/10 CGPA and proven expertise in developing production-grade machine learning systems delivering 85-99.95% accuracy. Ranked in top 5% on LeetCode with 300+ algorithmic solutions. Generated measurable business impact including 85% processing time reduction and sub-100ms response latency, seeking Software Engineer position to drive revenue growth through scalable ML solutions.

## CORE SKILLS

Machine Learning — Data Structures — Algorithms — Neural Networks — Python — C++ — JavaScript — SQL — API Integration — Database Systems — Performance Optimization

## EDUCATION

**Bachelor of Technology in Computer Science** — IIIT Kota — August 2022 – Present

- Maintained 8.8/10.0 CGPA, ranking in top 5% of 200+ student cohort
- Specialized in: Data Structures, Algorithms, Machine Learning, Operating Systems, Database Systems
- Executed 15+ technical projects demonstrating practical application of theoretical concepts

## WORK EXPERIENCE

**HacktheChain 2.0 Hackathon** — February 2024 — [\[Link\]](#)

- Prototyped emergency response system recognized in top 20 from 100+ competing teams
- Integrated Geolocation API with Firebase, reducing emergency response time by 15% (from 8 to 6.8 minutes)
- Optimized React Native performance, decreasing application load time by 30% (from 3.2s to 2.24s)
- Led cross-functional team of 4 engineers, delivering fully functional mobile solution within 48-hour deadline

## PROJECTS

**Enterprise RAG Knowledge Assistant** — PostgreSQL, FastAPI, ChromaDB, Ollama — September 2025 — [\[GitHub\]](#)

- Developed production-ready RAG system achieving 500ms query latency with 276x performance improvement through intelligent caching, enabling real-time document Q&A for 100+ concurrent users
- Built end-to-end document processing pipeline supporting PDF/DOCX/TXT formats with semantic vector search (ChromaDB), PostgreSQL backend, and JWT authentication for enterprise-grade security
- Integrated local LLM (Ollama) with context-aware retrieval system, Prometheus monitoring, and Streamlit UI, reducing knowledge retrieval time by 90% compared to manual document search

**Enterprise LLM Fine-Tuning Framework** — Transformers, PEFT, FastAPI, Kubernetes — September 2025 — [\[Demo\]](#) — [\[GitHub\]](#)

- Developed parameter-efficient fine-tuning system using LoRA reducing memory footprint by 75% and training time by 80% for 2.7B parameter models
- Implemented 7-metric evaluation suite including factuality and toxicity detection, ensuring model safety with 98% harmful content filtering accuracy
- Architected microservices platform with authentication, rate limiting, and auto-scaling supporting multi-model A/B testing for production deployments

**Real-time Fraud Detection MLOps System** — FastAPI, Docker, Prometheus, Scikit-learn — August 2025 — [\[Demo\]](#) — [\[GitHub\]](#)

- Developed ML pipeline analyzing 284,807 transactions with 99.95% fraud detection accuracy, preventing potential losses of \$2.5M annually
- Established A/B testing framework with 80/20 traffic split and real-time drift monitoring across 30 features, improving model reliability by 25%
- Containerized microservices architecture processing 100+ concurrent requests with 98ms average latency and 94.83% success rate

## TECHNICAL COMPETITIONS

- **LeetCode:** Attained 1849 rating (top 5%) by solving 300+ algorithmic challenges (August 2022–Present) [\[Profile\]](#)

- **CodeChef:** Secured 3-star ranking (1619 rating) through completion of 115+ competitive programming problems (January 2024–Present) [\[Profile\]](#)
- **Code 360:** Earned 1687 rating (top 5% nationally) by mastering 125+ data structure challenges (September 2023–Present) [\[Profile\]](#)
- **GeeksforGeeks:** Ranked 318 nationally with 3-star recognition for solving 100+ algorithmic problems (October 2023–Present) [\[Profile\]](#)

## SKILLS

- **Programming:** Python (3+ years), C++ (3+ years), C (3+ years), JavaScript (2+ years), SQL (2+ years)
- **Machine Learning:** Computer Vision, YOLOv8, TensorFlow, Keras, Scikit-learn, Neural Networks, NLP, Classification, Regression, MLOps
- **Data Analysis:** NumPy, Pandas, Matplotlib, Statistical Analysis, Data Visualization, Feature Engineering, Vector Databases (ChromaDB)
- **Development Tools:** Docker, FastAPI, REST APIs, Git, VS Code, Jupyter, Streamlit, LangChain, Prometheus, Google Cloud
- **Computer Science:** Data Structures, Algorithms, System Design, Microservices Architecture, Real-time Systems, Cloud Deployment

## CERTIFICATIONS

- Earned Knight status on LeetCode with 5 achievement badges in Data Structures and Algorithms (January 2025)
- Qualified for Google Arcade Gold League with 14 verified technical skill badges (November 2024) [\[Profile\]](#)
- Received Specialist and Achiever badges on Code360 for algorithm proficiency (December 2023)