

# Shubham Kumar Gupta

+91-9376091240 | kshubham04907@gmail.com | shubham.cv | linkedin.com/in/shubhamgupta04907 | github.com/kshubham090

## Education

### Amity University

Bachelor of Technology in Computer Science, AI & ML Specialization

Noida, India

Aug 2023 – Jul 2027

*Relevant Coursework:* Data Structures & Algorithms, Artificial Neural Networks, CNNs, RNNs, LSTMs, Computer Vision, Database Management Systems, Operating Systems, Object-Oriented Programming

## Experience

### WINNIIO / LifeAtlas

AI Engineering Intern

Sweden (Remote)

May 2026 – Present

- Built the initial voice-interview prototype on Retell AI / Twilio: voice in, Claude profile extraction, ranked matches via pgvector + Voyage rerank.
- In parallel, rebuilt the full voice pipeline from scratch, cutting cost ~20X vs Retell AI while keeping the same flow.
- Contributed to the ZeroClaw auth proxy; worked across edge-native AI, Digital Twin, and 3D Gaussian Splatting (3DGS) for spatial intelligence on the LifeAtlas platform.

### Stakrid Logistics Private Limited

Founder, Lead Engineer

Noida, India

Jan 2025 – Jan 2026

- Built a last-mile logistics platform: 40+ REST API endpoints (Java, Spring Boot) for order tracking and delivery workflows. Cut manual processing 80% and API latency from 800ms to under 200ms via query optimization and connection pooling.
- Deployed on GCP with CI/CD, load balancing, and monitoring; integrated Supabase, payment gateways, and SMS APIs with retry logic.

### IET Student Chapter, Amity University

Technical Team Member

Noida, India

Mar 2024 – Jun 2025

- Built and ran technical workshops on software development and data structures for 100+ students.

## Projects

### Agent Reliability Stack

Python, Claude API, llm, sentence-transformers, NumPy, Click, pytest, GitHub Actions

Jun 2026 – Present

- Eval harness: CI-based behavioral regression testing for agent outputs.
- LLM gateway / proxy: cost attribution, semantic caching, and model routing across providers.
- Guardrails middleware: pre-action validation, retry-with-repair, and a kill switch.

### Chakra47: Autonomous OS for Physical AI

Python, LangGraph, YOLOv8, Open Policy Agent, Qwen 70B, Claude API

Oct 2025 – Present

- Four-layer autonomous OS: LangGraph multi-agent swarm (open-sourced as Chakra47-AgenticSwarm), world-model long-term memory and state, neuro-symbolic governance, hybrid reasoning (Qwen 70B offline + Claude online).
- Governance layer gates agent actions before they execute to cut hallucination-driven decisions: YOLOv8 edge inference, Open Policy Agent rule validation, SHA-256 audit trail.
- Working proof of concept live on GitHub.

### Military Deployment Decision System — Python, CNN, Claude API, Rules Engine

Feb 2026

- Built a tactical AI analysis system with CNN-based threat detection, Claude Sonnet reasoning, and a hard-coded Rules of Engagement validation layer to prevent unsafe deployment recommendations.
- Presented at AI Impact Summit 2026 (Government of India); received validation from defense tech stakeholders.

### Real-Time Posture Analysis System — Python, OpenCV, MediaPipe, CNNs

Jan 2026

- Multi-person pose estimation pipeline running at 25+ FPS on CPU via model quantization and frame-skip strategies; rule-based anomaly detection over pose keypoints for ergonomic risk flagging.

## Technical Skills

**Languages:** Python, Java, SQL, Bash

**DSA:** Arrays, Linked Lists, Trees, Graphs, Dynamic Programming, Recursion, Sorting, Sliding Window, Two Pointers

**AI/ML:** Deep Learning, Computer Vision, CNNs, RNNs, LSTMs, Neuro-Symbolic AI, TensorFlow, PyTorch, YOLOv8, OpenCV, MediaPipe

**Agentic & Backend:** LangGraph, LangChain, Claude API, RAG, pgvector, Voyage rerank, Open Policy Agent, Spring Boot, REST API Design, JDBC, Spring Data JPA

**Cloud & DevOps:** Google Cloud Platform, Docker, CI/CD, Supabase, Firebase

**Concepts:** System Architecture, OOP, Design Patterns, Database Design, Distributed Systems

## Certifications

NCA-AIIO: AI Infrastructure & Operations Associate (NVIDIA, In Progress) · Machine Learning Specialization (Stanford/Coursera) · Java Spring Framework 6 with Spring Boot 3 (Udemy) · Build and Secure Networks in Google Cloud (Google Cloud Skills Boost)