Umang Sharma

umangxrsharma@gmail.com — +91 9172294130 — github.com/0xzcr —
linkedin.com/in/umang-sharma-35923433b/ — x.com/umxngz —

Education

VIT-AP University Expected May 2028
Bachelor of Technology in Computer Science and Engineering **CGPA: 7.95/10**
Relevant Coursework: Data Structures and Algorithms, Operating Systems, Database Management Systems, Computer Networks, Machine Learning, Deep Learning, Computer Organization and Architecture

Technical Skills

Languages: Python, Java, C++, JavaScript, SQL
Backend & Infrastructure: FastAPI, Docker, Redis, PostgreSQL, REST APIs, WebSockets, Linux, Git
AI Systems: PyTorch, Transformers, ONNX Runtime, Qdrant, LangGraph, LangChain, MLflow
Developer Tools: GitHub Actions, Docker Compose, CLI Tooling, Logging & Observability

Experience

Open Source Contributor — MobileFaceNet Nov 2025 – December 2025

- Added ONNX Runtime support for MobileFaceNet to simplify deployment across multiple platforms.
- Reduced inference latency through graph optimization and post-training quantization.
- Collaborated through GitHub pull requests, issue discussions, and code reviews to improve deployment workflows.

Independent Product Developer(Startup) April 2026 – Present

- Building a collaborative CLI platform that enables multiple developers and AI agents to share synchronized context across distributed workspaces, eliminating context fragmentation during collaborative software development.
- Designed a Redis-backed session management and synchronization layer to coordinate concurrent updates, maintain shared agent state, and resolve conflicting context modifications across multiple clients.
- Developed repository-aware CLI workflows for codebase analysis, architecture comparison, and automated design reviews by integrating version control metadata with multi-agent reasoning pipelines.

Projects

Zuna [GitHub] JavaScript, Web APIs, PDF.js, Vercel

- Built a local-first browser reading companion that turns extractable PDFs and TXT files into listening sessions; extracts documents in bounded 8-page batches and begins playback before full-file processing.
- Shipped a zero-build static application with local resume state, seeking, speed controls, and on-device document data; no login, database, upload API, or server-side book processing.

Zoosh [GitHub] TypeScript, Next.js, Supabase, Prava, Cashfree

- Developed a group-expense platform that converts shared balances into one net settlement payment per debtor across groups, outings, expenses, and balance workflows.
- Integrated authenticated ledger flows with Supabase, passkey-protected payment authorization via Prava, Cashfree payment and payout rails, and Linq/email settlement notifications across 4 external payment/notification services.

Real-time Content Moderation [GitHub] Python, DistilBERT, PyTorch, ONNX Runtime, FastAPI, MLflow

- Built an end-to-end toxic-content classifier over 159,571 comments, reaching 96.08% accuracy, 0.9753 ROC-AUC, and 0.9125 macro F1 with 83.53% toxic recall.
- Optimized DistilBERT through ONNX graph rewrites and dynamic INT8 quantization, reducing model size 74.9%, reaching 24.9 ms p99 latency at batch 1, and delivering up to 2.92x p50 speedup while limiting prediction drift to 0.0001.

Leadership

Organizing Committee — Google Developer Groups (GDG) On Campus

- Coordinated technical events, workshops, and hackathons for the campus developer community.
- Collaborated with cross-functional teams to manage logistics, speaker coordination, and event execution.

Achievements

- Open source contributor to AI infrastructure projects.
- Built independent AI infrastructure and developer tooling projects focused on production systems.
- Building Google Developer Groups on campus online committee on Discord, acting as a Developer Relations lead.