Aarush Dubey

r.kdubey2004@gmail.com | +91 82181 37799 | linkedin.com/in/aarush-dubey-068698320 | github.com/aarushdubey | Amaravati, AP, India

EDUCATION

SRM University, AP 2024 – 2028

Bachelor of Technology in Computer Science and Engineering **CGPA: 8.9 / 10.0**

- **Core Coursework:** Data Structures & Algorithms, Object-Oriented Programming (Java, C++), Database Management Systems (SQL), Operating Systems, Computer Networks, Software Engineering (SDLC, Agile), Machine Learning, Discrete Mathematics.

INDUSTRY CERTIFICATIONS

- **MongoDB Associate Developer** VERIFIED CREDLY – Database schema architecture, aggregation pipelines, indexing strategies, and high-throughput Node.js CRUD optimization.
- **Advanced Java Concepts** VERIFIED COURSERA • EDUREKA – Modern Java engineering: Functional Programming, Streams API, Lambda Expressions, Concurrency, and Multithreading.
- **Generative AI Fundamentals** VERIFIED – Large Language Model orchestration, Prompt Engineering, Retrieval-Augmented Generation (RAG) architectures, and vector embeddings.

OPEN SOURCE CONTRIBUTIONS & RESEARCH

Hugging Face Transformers (PR #49957) – Multimodal Vision-Language Lead Review Oct 2026

- Diagnosed and patched a PIL dimension order inversion in `get_image_size` directly affecting bounding box normalization in GOT-OCR2 vision-language models; formally triaged by core maintainer to domain vision leads.
- Implemented comprehensive regression test suites guaranteeing coordinate precision and aspect-ratio invariance across non-square images.

CNCF Meshery (PR #22591 & #22588) – Cloud Native Manager Backend & CLI Oct 2026

- Engineered robust database error propagation and JSON decoding in Go server backend (`GetMesheryK8sContexts`), eliminating silent synchronization failures during Kubernetes context discovery.
- Mitigated CLI socket exhaustion by standardizing HTTP client timeout policies across `mesheryctl` (100% CI pass across 40 checks).

AISTATS 2027 Research: Fair Conformal Classification for Financial Transactions Oct 2026

- Formulated **FairTransCP**, addressing the *coverage–equity paradox* in financial ML by jointly calibrating demographic coverage equity and prediction set-size disparity; mathematically proved Theorem 1 (theoretical worst-group coverage lower bound).
- Implemented regularized threshold optimization atop classical tree ensembles (XGBoost, LightGBM), cutting set-size disparity by **34%** vs. Mondrian CP while ensuring 100% regulatory auditability (FCRA, ECOA, EU AI Act compliant).

PROFESSIONAL EXPERIENCE

Full-Stack Software Engineer – Quint Beauty Nov 2025 – May 2026

Production E-Commerce Platform (quintbeauty.com) *Remote*

- Architected full-stack e-commerce platform serving **2,000+ monthly active users**; developed responsive storefront (**React.js, TypeScript**) and modular REST API backend (**Node.js, Express.js, MongoDB**) with 20+ reusable components.
- Integrated **Razorpay payment gateway** with automated webhook reconciliation processing 200+ monthly transactions; engineered role-based Admin Portal managing 150+ SKUs, customer records, and dynamic order states.
- Optimized MongoDB queries and implemented component code-splitting, raising Google Lighthouse score from 62 to **95+** and slashing load times to under 1.8 seconds.

Research Intern – Quantum-Resistant Cryptography Jun 2026 – Aug 2026

SRM University-AP (Post-Quantum Cloud Storage Auditability) *Remote*

- Engineered a quantum-resistant Remote Data Integrity Checking (AB-RDIC) protocol in **C and Python** using **ML-DSA-44 (CRYSTALS-Dilithium)** via `liboqs` , protecting distributed cloud storage from quantum forgery.
- Benchmarked tag generation, proof construction, and verification latency over ring $\mathbb{Z}_q$, validating soundness and collision resistance.

KEY ENGINEERING PROJECTS

TransactAI – Explainable Financial Classification Pipeline (GitHub) GHCI 2025 Hackathon

- Engineered automated ML pipeline achieving macro F1-score of **1.00** using XGBoost ensembles and engineered transaction metadata; integrated SHAP values for transparent, auditable inference – selected as **Top-15 Finalist nationally** out of 1,200+ teams.

Guardian PDF – High-Performance Document Intelligence Engine (Live Demo) 2026

- Architected 3-tier RAG application (**React + FastAPI + ChromaDB**) handling 5,000+ queries for 300+ users; developed custom **C++ text extraction module** delivering **7.5x throughput improvement** over standard parsers; deployed via Docker CI/CD with 85% test coverage.

TECHNICAL SKILLS

Languages:	**Java** (Streams, Concurrency, OOP), **Python**, **C/C++**, **TypeScript**, **JavaScript (ES6+)**, **Go**, **SQL**
AI & Machine Learning:	LLM Orchestration, Prompt Engineering, RAG Architectures, Conformal Prediction, PyTorch, Hugging Face, Scikit-Learn, XGBoost, SHAP
Web & Backend:	Node.js, Express.js, FastAPI, Flask, React.js, Next.js, RESTful API Architecture, JWT Authentication, Microservices
Databases & DevOps:	MongoDB, PostgreSQL (Supabase), MySQL, SQLite, ChromaDB, Docker, AWS (EC2, S3), Git/GitHub, CI/CD Pipelines, PyTest
CS Fundamentals:	Data Structures & Algorithms, Object-Oriented Design, Operating Systems, Computer Networks, Agile/Scrum Life Cycle