Aarush Dubey

Vijayawada, AP, India | +91 82181 37799 | r.kdubey2004@gmail.com | linkedin.com/in/aarushdubey | github.com/aarushdubey | www.aarushdubey.com

EDUCATION

SRM University-AP — Bachelor of Technology in Computer Science and Engineering — 2024 – 2028 | Vijayawada, AP, India

Cumulative GPA: 8.9 / 10.0 | **Core Coursework:** Data Structures & Algorithms, Object-Oriented Programming (Java/C++), Operating Systems, Database Management Systems (SQL), Computer Networks, Machine Learning, Cloud Systems, Software Engineering.

OPEN SOURCE CONTRIBUTIONS & MACHINE LEARNING RESEARCH

Hugging Face Transformers — Core AI & Vision Infrastructure PR #50475 MERGED PR #49957 APPROVED — Oct – Nov 2026

- Diagnosed PyTorch JIT compiler defects in `torch.compile(fullgraph=True)` across **55 core model architectures** (Whisper, BART, OPT, Wav2Vec2) by gating LayerDrop random tensor draws; authored regression test suites **merged into upstream main by maintainers**.
- Patched coordinate inversion and bounding box normalization defects in `image_utils.py` for **GOT-OCR2 vision-language models** (PR #49957 Approved), ensuring precise spatial alignment across non-square images.

CNCF — Meshery — Cloud Native Computing Foundation Management Plane PR #22926 MERGED PR #22973 MERGED — Oct – Nov 2026

- Resolved silent failure bugs and nil dereference panics in component registry publishing with Google Sheets integration; authored Go unit test suites with **100% CI pass** (PR #22926 Merged).
- Patched empty-slice boundary crashes in CLI suggestions and expanded Levenshtein distance test suites; standardized HTTP client timeout and connection pooling across `mesheryctl` (PR #22973 Merged).

Fair Conformal Classification for Financial ML (AISTATS 2027 Submission) — First Author — Oct 2026

- Formulated **FairTransCP**, resolving the coverage-equity paradox in financial ML by jointly calibrating demographic coverage and prediction set-size disparity; mathematically proved **Theorem 1** (worst-group coverage lower bound).
- Implemented regularized threshold optimization over XGBoost/LightGBM ensembles, cutting set-size disparity by **34%** vs. Mondrian CP while ensuring 100% compliance with financial regulatory audits (FCRA, ECOA, EU AI Act).

ENGINEERING & PRODUCTION EXPERIENCE

Full-Stack Developer Intern — Quint Beauty (Commercial E-Commerce Platform) — Nov 2025 – May 2026 | Remote

- Architected and shipped a production e-commerce platform serving **2,000+ Monthly Active Users**; built responsive UI (React, TypeScript) and modular REST API services (Node.js, Express, MongoDB) across 20+ components.
- Engineered a cryptographic **HMAC SHA-256 webhook reconciliation pipeline** with idempotency keys, eliminating payment drops; indexed schemas and code-split assets, raising Google Lighthouse score from 62 to **95+** (load times <1.8s).

Research Intern — Quantum-Resistant Cryptography — SRM University-AP — Jun 2026 – Aug 2026 | Remote

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

FLAGSHIP ENGINEERING PROJECTS

LabClarity AI — Research Microscopy Image Enhancer — Live Demo | GitHub — 2026

- Engineered a life sciences image restoration system combining Poisson-Anscombe VST denoising, 2x/4x Lanczos-4 optical super-resolution, and Groq LLaMA 3.3-70B multimodal analysis for non-diagnostic research specimens.
- Developed Flask REST API and responsive laboratory workbench with real-time Tenengrad sharpness telemetry and draggable split-slider; deployed cloud service on Render with sub-second inference latency.

Guardian PDF — High-Performance Semantic Document Intelligence — Live Demo | GitHub — 2026

- Architected a 3-tier RAG application (FastAPI, ChromaDB, React, Docker) serving 5,000+ semantic search queries; built compiled C++ memory-mapped parser achieving a **7.5x throughput improvement** (1.6s vs 12.0s) over standard parsers.

TransactAI — Explainable Financial ML Classification Pipeline — GitHub | Top-15 National Finalist (GHCI 2025) — 2025

- Engineered automated ML pipeline achieving macro F1-score of **1.00** using tuned XGBoost ensembles; integrated SHAP feature attribution for transparent model explainability — selected as **Top-15 Finalist nationally out of 1,200+ teams**.

TECHNICAL SKILLS, LEADERSHIP & VERIFIED CREDENTIALS

Languages: Go, Python, Java (Streams, Concurrency, OOP), C/C++, TypeScript, JavaScript (ES6+), SQL, HTML/CSS, Bash

Systems, Cloud & DevOps: Docker, Kubernetes, CI/CD (GitHub Actions), Render Cloud, AWS (EC2/S3), Linux, REST APIs, Microservices, PyTest

AI / ML & Computer Vision: PyTorch, Hugging Face, OpenCV, LLM Orchestration, RAG Pipelines, Conformal Prediction, Scikit-Learn, XGBoost, SHAP

Databases: MongoDB, PostgreSQL, SQLite, ChromaDB (Vector DB), Redis

Leadership & Honors: Vice President, Smart Tech Club (200+ members; organized 10+ hackathons/workshops) | GHCI 2025 National Finalist | GDG Code4Her Finalist

Certifications: MongoDB Certified Associate Developer (Credly) | Advanced Java (Coursera) | Generative AI Fundamentals