Ali Sufiyan Khan

B.Tech Computer Science & Engineering, VIT Vellore · Class of 2027

alizaydsab@gmail.com · Vellore, India
LinkedIn · github.com/sting-raider · Portfolio

PROFILE

Computer Science undergraduate building backend systems, ML pipelines and low-level tools, with every claim backed by a measurement: throughput, latency, test coverage or evaluation pass rate.

PROJECTS

Amazon ML Challenge 2026 · *Business entity resolution* GitHub

Python, CatBoost, Modal, C++

- Built an end-to-end pipeline matching **1.73M** business records across three sources, using only the supplied data.
- Fused word and character 3/4-gram retrieval with a **57-feature** CatBoost matcher; final score **0.967** macro F0.5.
- Scored **265M** candidate pairs on 8 parallel cloud workers; ported feature extraction to C++ with byte-identical output.

AegisFlow · *Hybrid network intrusion detection* GitHub

Python, Scikit-learn, FastAPI, Redis, PostgreSQL, Docker

- Combined a calibrated classifier, Isolation Forest, autoencoder anomaly scoring and Suricata signals into one detector.
- Correlates detections into incidents behind a FastAPI backend with Redis and PostgreSQL, containerized with Docker.
- Raised batched throughput **23×**, from 153.5 to **3,496.8 flows/s**; verified by **238 tests** at **84% backend coverage**.

RoleAtlas · *Distributed job discovery & ranking* GitHub

Next.js, TypeScript, NATS JetStream, PostgreSQL, Docker

- Built streaming ingestion and normalization over NATS JetStream for 32 job sources, with provenance on every record.
- Added ranking and persisted search sessions; retained **3,128** records and verified **250** ranked jobs end to end.
- Cut first-result latency from **~24 s to 57 ms** with indexed responses, plus live coverage reporting across sources.

Project Dante · *Safety-gated agentic commerce* GitHub

Python, LLMs, Rules Engine, Razorpay Test Mode

- Built an LLM shopping agent that compiles natural-language requests into typed, machine-checkable constraints.
- Gated every payment and refund behind deterministic policy checks, executed through Razorpay's sandbox.
- Passes **68 intent**, **50 prompt-injection** and **28 financial-safety** evaluation cases.

Plaid · *Automated ROM-to-native recompiler* GitHub

Rust, C, Static Analysis, x86-64 / ARM64

- Discovers Nintendo 64 guest code and translates it into native x86-64 / ARM64 executables, with no interpreter or JIT.
- Differential-tested against a reference emulator, with explicit fault paths and documented per-game coverage.

Hybrid Multi-Cloud Kubernetes

K3s, AWS, GCP, Docker, Tailscale, Cloudflare Tunnels, Locust

- Designed a K3s cluster across local, AWS and GCP nodes, linked by Tailscale and exposed via Cloudflare Tunnels.
- Load-tested cloud-hosted services at **1,000+ concurrent users** with **0% observed request failures**.

Sivana · *Audio fingerprinting & song recognition* GitHub

Rust, FFT, Spectrograms, SQLite

- Rust CLI that extracts spectrogram peaks, forms time-frequency hashes and identifies songs from short clips.
- Indexed SQLite fingerprint store with song enrollment, offset-based match scoring and audio resampling.

SKILLS

Languages	Python, TypeScript, JavaScript, C, Rust, Java, SQL
Backend & Data	FastAPI, Next.js, React, PostgreSQL, SQLite, Redis, NATS JetStream
ML	PyTorch, Scikit-learn, CatBoost, FFT / audio DSP
Cloud & Tools	Docker, Kubernetes (K3s), AWS, GCP, Modal, Linux, Git, CI, Locust

EDUCATION

Vellore Institute of Technology · *B.Tech, Computer Science and Engineering*

Aug 2023 – May 2027

Coursework: Data Structures & Algorithms, Design & Analysis of Algorithms, Operating Systems, DBMS, Artificial Intelligence

CERTIFICATION

OCI 2025 Certified Generative AI Professional · *Oracle University* Verify

Aug 2025 – Aug 2027