SAUMAY KILLA

516-968-1066 — saumaykilla@gmail.com — LinkedIn — Portfolio — GitHub

PROFESSIONAL SUMMARY

Full-stack and **Founding Engineer** with experience building scalable **Generative AI platforms**, production systems, and data-driven ML solutions. Skilled in Python, React, and AWS, with hands-on experience **training ML models**, optimizing pipelines, and automating workflows. Proven ability to deliver high-performance applications while **reducing infrastructure costs** and improving reliability.

TECHNICAL SKILLS

Languages: JavaScript, TypeScript, Python, SQL, C/C++
Frontend: React, Next.js, Redux, Zustand, Tailwind CSS, WCAG 2.2
Backend: Node.js, Express, FastAPI, Flask, REST APIs, Microservices
Databases: PostgreSQL, MySQL, MongoDB, Supabase
Cloud / DevOps: AWS (Lambda, EC2, S3, API Gateway, SQS), Docker, CI/CD
AI / LLMs: LangChain, RAG, Prompt Engineering, Gemini

PROFESSIONAL EXPERIENCE

HumAInority   New York, NY
Full-Stack Developer (Founding Engineer)   Sep 2024 – Present

- Built a scalable **Generative AI platform** serving **100+ users**, owning frontend, backend, and cloud architecture.
- Architected **serverless AWS Lambda** microservices using S3, API Gateway, and SQS, cutting **infrastructure costs by 40%**.
- Developed Python microservices handling **50K+ monthly API requests** with authentication and robust error handling.
- Achieved **sub-200ms frontend load times** using modular React and TypeScript architecture.
- Reduced release cycles by **60%** using GitHub Actions CI/CD, Docker, and automated testing.
- Collaborated with stakeholders to translate business goals into deployable **MVP features**.

Wiz Freight   Chennai, India
Frontend Developer   Apr 2022 – Jul 2023

- Developed a reusable, accessible component library, improving feature development **speed by 30%**.
- Rebuilt marketing website with React and Tailwind, improving load speed by **40%** and conversions by **15%**.
- Optimized Redux state management, reducing redundant **network calls by 25%**.
- Worked in a cross-functional Agile environment, increasing **sprint velocity by 20%**.

HighRadius   Remote
Software Developer   Aug 2021 – March 2022

- **Trained ML models** to predict invoice clearance dates, improving payment forecasting for finance teams.
- Automated client-specific Excel file generation, reducing manual data processing **time by 40%**.
- Identified key data features influencing customer payment behavior, **improving model accuracy**.
- Collaborated with data engineers to optimize data pipelines, enhancing efficiency and reliability.

NOTABLE PROJECTS

AI Segment Analysis Engine (Generative AI, RAG)   *Python, LangChain, Gemini*

- Built a prompt-driven analysis engine, accelerating reporting **speed by 5x** with automated summaries.

Mock Interview Voice Agent   *Python, LiveKit, LLMs*

- Engineered a real-time voice-based interview agent, reducing interview **prep effort by 60%**.

AI-Powered UI Generation System   *Python, Next.js, AWS, Gemini*

- Designed an **event-driven AI system** generating production-ready UI code; optimized concurrency via AWS SQS.

EDUCATION

New York University, SPS   New York, NY
Master's in Management Information Systems   May 2025

SRM University   Chennai, India
Bachelor of Technology in Computer Science   May 2022