SAUMAY KILLA

Brooklyn, NY | 516-968-1066 | saumaykilla@gmail.com | linkedin.com/in/saumaykilla

PROFESSIONAL SUMMARY

Full stack engineer with 3 years building production systems from 0-to-1. Architected and shipped AI platform serving 2,000+ users as founding engineer. Reduced infrastructure costs 40% through serverless architecture. Expert in React, Python, AWS, and rapid MVP development. Thrive in early-stage environments where speed, ownership, and technical judgment directly impact business outcomes.

TECHNICAL SKILLS

Languages: JavaScript (ES6+), TypeScript, Python, SQL, C, C++
Frontend: React.js, Next.js, Redux, Zustand, Tailwind CSS, Responsive Design, Web Performance Optimization
Backend: Node.js, Express.js, FastAPI, Flask, REST APIs, WebSocket, Microservices Architecture
Database: PostgreSQL, MySQL, MongoDB, Supabase, Database Design
Cloud & DevOps: AWS (Lambda, EC2, S3, API Gateway, SQS, Amplify), Docker, CI/CD, GitHub Actions
AI & Automation: LiveKit, OpenAI API, GenAI Integration, Web Scraping (Playwright), ML Pipelines
Tools & Practices: Git, GitHub, Agile/Scrum, Jira, Figma, Postman, System Design, API Development

PROFESSIONAL EXPERIENCE

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

- Architected and launched scalable AI platform from scratch serving 2,000+ users, owning complete system design including frontend architecture, backend services, cloud infrastructure, and deployment workflows
- Reduced infrastructure costs by 40% by building serverless microservices on AWS Lambda with optimized S3, API Gateway, and SQS workflows, eliminating expensive managed services while improving scalability
- Achieved sub-200ms frontend load times through modular React system with TypeScript, implementing code splitting, lazy loading, and performance monitoring best practices
- Built secure Python microservices handling 50K+ monthly API requests with FastAPI, integrating third-party tools while implementing authentication, rate limiting, and comprehensive error handling
- Reduced release cycles by 60% by establishing CI/CD pipeline with GitHub Actions, automated testing, and zero-downtime deployments using Docker containerization
- Established engineering foundations including coding standards, documentation practices, database schemas, and technical architecture patterns adopted across the team

WIZ FREIGHT   Chennai, India
Frontend Developer   *Apr 2022 – Jul 2023*

- Rebuilt marketing website with React and Tailwind CSS, increasing load speed by 40% and contributing to 15% improvement in lead conversions through optimized rendering and client-side performance
- Developed reusable component library used across product teams, cutting feature development time by 30% and establishing consistent design system and UI patterns
- Optimized Redux state management architecture, reducing redundant network calls by 25% and improving application time-to-interactive metrics
- Increased sprint velocity by 20% through effective cross-functional collaboration with design and backend teams in Agile environment, consistently delivering features ahead of schedule

HighRadius   REMOTE
Software Engineer (ML Focus)   *Aug 2021 – Mar 2022*

- Built and tuned machine learning models for invoice prediction, improving forecasting accuracy by 16% for enterprise finance clients through feature engineering and model optimization
- Automated Excel report generation with Python scripts, reducing manual reporting work by 40% and eliminating human error in client deliverables
- Optimized ML feature pipelines in collaboration with data engineering team, cutting model training time by 25% through efficient data processing and feature selection
- Supported end-to-end ML workflows including dataset cleaning, feature engineering, model evaluation, and performance monitoring using Python, Pandas, and Scikit-Learn

Notable Projects & Initiatives

AI-Powered Sales Order PDF Extractor: Built FastAPI tool leveraging AI and text extraction to automate data entry from PDF sales orders, cutting manual effort by 90% and improving data accuracy by 25%

Live Ticker Price Tracking Dashboard: Developed real-time trading dashboard using Playwright and Next.js for live stock data scraping and visualization with seamless multi-ticker performance updates

AI Segment Analysis Engine: Created prompt-driven insight generator using LangChain and Gemini APIs, delivering 5x faster business reporting through automated analysis of product, customer, and strategy data

Education

New York University – School of Professional Studies — New York, NY
Master's Degree in Management Information Systems — *Sep 2023 – May 2025*

SRM University (Sri Ramaswamy Memorial University) — Chennai, India
Bachelor of Technology in Computer Science — *Jul 2018 – Jul 2022*

Certifications

Full Stack Web Development with MERN Stack & GenAI 2025 | Next.js 15 & React - The Complete Guide | The Complete Web Developer Course 3.0