VAIBHAV SRIVASTAVA

vaibhav.srivastava405@gmail.com | LinkedIn® Profile | GitHUB | +91 872 628 3209

Professional Summary

Software Developer with 3+ years of experience building scalable backend systems using FastAPI, Flask, and AWS. Proven expertise in microservices, distributed systems, and database optimization, with a track record of improving performance by 30–60%. Passionate about solving large-scale engineering challenges, from high-concurrency APIs to AI/ML integrations.

Skills

Technical Skills: Python, FastAPI, Flask, Django, REST/gRPC APIs, SQLAlchemy, PostgreSQL, MySQL, Redis (Caching, Pub/Sub), AWS Cloud Services, Docker, GitHub Actions CI/CD, Serverless, Microservices, Event-Driven Systems, Distributed Systems, Load Balancing, Pytest, RBAC, OpenAI GPT, Pgvector (vector DB), LLM Agent

Work Experience

Facts Online | Software Developer   May 2023 - Present

- Built high-traffic APIs using FastAPI and Flask, reducing latency by 40% through Redis caching and query optimization.
- Designed a **microservice architecture** handling 10K+ RPM (requests per minute) for a financial platform, improving scalability.
- Collaborated with frontend teams to integrate ReactJS dashboards, enabling real-time analytics for 50K+ users.
- Led **production deployments** with zero-downtime using Docker + AWS, maintaining 99.9% uptime SLA.
- Developed 8+ modular **REST APIs** (FastAPI/Flask) with OOP principles, reducing code redundancy by 35%.

Research & Analytics Solutions | Python Developer   Nov 2022 - Apr 2023

- Built scalable web scraping pipelines using **Selenium WebDriver** to extract data from dynamic, JavaScript-heavy sites with login and pagination.
- Processed and transformed large datasets using **Pandas**, **NumPy**, and **regex**, storing results in **PostgreSQL/MongoDB** for real-time analytics and reporting.

Projects

FintrackPro – Loan Management Platform   Feb 2025 - Present

- Engineered a scalable fintech backend using Flask, React.js, and SQLAlchemy; deployed on AWS (EC2, RDS, Lambda).
- Integrated **Razorpay payment system** to streamline secure transaction processing for loan management.
- Reduced deployment time 50% using Dockerized **CI/CD pipelines**.
- Optimized PostgreSQL queries, cutting report generation time 30% for 100K+ loan records.

My AI Twin – AI-Powered Professional Representative & Networking Agent   Jul 2024 - Sep 2025

- Developed a full-stack AI agent using Python and FastAPI, reducing response latency by 40% with async processing.
- Integrated DeepSeek LLM and a RAG pipeline to analyze LinkedIn profiles to improve response by 60%.
- Automated lead capture via Pushover API, ensuring 100% follow-up and cutting response time to under 5 Mins..
- Designed modular JSON-based tool calling, reducing feature dev time by 30% and enabling smooth deployment on Hugging Face Spaces.

Certifications

- AWS Cloud technical Essentials (Coursera)
- Full Stack Data Science 2.0 (iNeuron.ai)
- AI For All : From Basic To GenAI (NVIDIA)

Education

GLA University, Mathura   **July 2017 - May 2021**
Bachelor of Technology in Civil Engineering Mathura, UP, India

- Relevant coursework: Data Structures, Algorithms (self-taught)

Keywords: ReactJS (Familiarity), JavaScript, HTML5, AWS (Lambda, EC2, RDS, S3), OpenAI GPT, Vector databases, LLM fine-tuning, JWT, OAuth2, Celery, Websocket, asynchronous

Key Projects & Contributions

Zonal Excellence System – Employee KPI Analytics Platform

Built an API-driven platform to display zone-wise and employee-wise performance metrics with ranking filters.

- **Tech Stack:** Python, FastAPI, Linux, Docker, REST APIs, PostgreSQL, LDAP Authentication, ORM
- **Achievement:** Enabled business teams to access real-time KPI insights, reducing reporting delays by 80%.

Unified Health Check – Real-Time Server Monitoring (Automation)

Developed a modular backend to monitor Vodafone Idea's nationwide servers in real-time with API-driven health reports.

- **Tech Stack:** Python, FastAPI, PostgreSQL, Linux, Docker, REST APIs, React.js
- **Achievement:** Reduced server downtime detection time by 60%, improving operational efficiency across 22 telecom circles.

Network Monitoring System – Real-Time Device & Port Status

Developing an asynchronous system to monitor device and port status via SNMP and ping, with real-time updates using PostgreSQL triggers and listeners.

- **Tech Stack:** Python, Asyncio, SNMP, PostgreSQL, DSA, System Design, FastAPI Websocket
- **Achievement:** Designed scalable architecture enabling instant device status updates for a dynamic monitoring dashboard.

AI Chat Bot with RAG – Context-Aware AI Assistant

Developed a GenAI-powered assistant by fine-tuning LLaMA-2 with Retrieval-Augmented Generation (RAG) to provide real-time, context-aware responses using vector databases.

- **Tech Stack:** Python, Langchain, LLaMA-2, RAG, Redis (Vector DB), PostgreSQL, Docker, embeddings
- **Achievement:** Improved response accuracy by 60% compared to rule-based systems through Redis vector caching and dynamic model pipeline optimization; reduced response latency by 25% with production deployment.

Ed-Tech-AI – Intelligent Quiz Generation & Learning Analytics Platform

Developed an AI-powered platform to auto-generate personalized quizzes and track learner performance through intelligent analytics.

- **Tech Stack:** Python, GPT-3.5, FastAPI, Tesseract OCR, NLP, PostgreSQL, Redis, REST APIs
- **Achievement:** Generated 1,000+ adaptive quizzes with GPT-3.5, boosting learner engagement by 85%. Achieved 70% noise reduction in text extraction via custom OCR+NLP pipeline and delivered 95% accuracy in real-time performance tracking.

Automated Web Scraping & Data Extraction

Created Selenium-based scrapers handling authentication, pagination, and dynamic JavaScript content.

- **Tech Stack:** Python, Selenium, PostgreSQL
- **Achievement:** Automated data collection from 20+ sources, reducing manual data entry efforts by 90%.

IPTV Live Streaming Platform – Razorpay Integration

Developing a live streaming platform for cable providers to deliver channel packages similar to Zee/Star networks.

- **Tech Stack:** Python, Django, Razorpay API, PostgreSQL, Asynchronous
- **Achievement:** Implemented secure online payment gateway via Razorpay, enabling seamless subscription purchases and real-time payment status updates.