

Varun Kumar Andigari

varunkumarandigar@gmail.com | +1 (513) 258 9411 | [LinkedIn](#) | [GitHub](#)

Summary

Software Engineer with a proven track record of designing, developing, scaling, and deploying secure, high-availability software solutions. Strong expertise in Python, microservices architecture, RESTful APIs, cloud-native development, and full SDLC execution across cross-functional Agile environments. Experienced in building scalable backend systems and integrating AI, Generative AI, and Large Language Models (LLMs) into software applications to enhance automation, intelligent workflows, and user experiences. Adept at transforming technical requirements into clean, maintainable code while driving performance optimization, reliability, and API integrations.

Technical Skills

Backend & Programming: Python, JavaScript, TypeScript, FastAPI, Flask, Django, AsyncIO, Multithreading, Multiprocessing, SQLAlchemy, Pydantic, Pytest, Scripting & Automation, OOP, Data Structures & Algorithms, Poetry, Pip

AI, Generative AI & LLM: Large Language Models, OpenAI API, Prompt Engineering, Retrieval-Augmented Generation (RAG), Vector Search, Semantic Search, LangChain, AI Agents, Function Calling, Model Context Protocol (MCP), Context Management

Machine Learning & Data: Supervised & Unsupervised Learning, Scikit-learn, Pandas, NumPy, Data Preprocessing, Feature Engineering, Model Evaluation, Data Analysis, Data Visualization, Jupyter Notebook

Microservices & API Development: RESTful Web Services, REST APIs, API Gateway, OpenAPI/Swagger, JSON, XML, Service Discovery, Distributed Systems, Event-Driven Architecture, OAuth 2.0, JWT, API Authentication & Authorization

Cloud & DevOps: AWS (EC2, S3, RDS, Lambda, CloudWatch), Docker, Kubernetes, Jenkins, CI/CD, Terraform, Containerization, Infrastructure as Code (IaC), Cloud Deployment, Monitoring & Logging

Data Infrastructure, Messaging & Testing: PostgreSQL, MySQL, MongoDB, Redis, SQL, Data Modeling, Query Optimization, Apache Kafka, RabbitMQ, Celery, Async Processing, Pytest, API Testing, Postman, TDD

Development Tools & Engineering: Git, Jira, Confluence, Linux, Agile/Scrum, System Design, Claude Code (Agentic Workflows), Debugging & Troubleshooting

Professional Experience

Software Engineer, Capgemini

01/2026 – Present | Remote, USA

- Designed microservices using Python, FastAPI, Flask, and Pydantic, reducing API response times by 30%; implemented circuit-breaker patterns for downstream service calls to improve application resilience and prevent cascading failures.
- Implemented RESTful APIs and gRPC services using Python, optimizing communication between distributed systems through FastAPI, asyncio, and asynchronous programming to improve latency, throughput, and overall application performance.
- Optimized database performance using PostgreSQL, SQLAlchemy, and Redis, implementing caching strategies and CQRS (Command Query Responsibility Segregation) to efficiently scale data access patterns and support high-volume workloads.
- Integrated Generative AI and LLM APIs into Python applications using LangChain, RAG, and prompt engineering to build a context-aware document Q&A system that retrieved indexed content and generated accurate, grounded responses.
- Developed an MCP (Model Context Protocol) server using Node.js to expose internal APIs and business data as reusable tools for LLM-powered applications, implementing tool schemas, request validation, and secure service integration.
- Automated CI/CD pipelines (Jenkins, ArgoCD) and Kubernetes deployments for Python/AI workloads, slashing deployment cycles by 60% with zero downtime while securing APIs via OAuth 2.0 and RBAC.

AI Software Engineer Intern, AgileRadioCom LLC

05/2025 – 12/2025 | Remote, USA

- Deployed PyTorch ML models (EKF, online LSTM) into production for real-time state estimation, forecasting, and autonomous tuning, integrating GenAI/LLMs to interpret model outputs.
- Engineered a Python/Flask AI backend with Redis caching and WebSocket dashboards (Chart.js), streaming multi-API data to process 2.2M+ records and render real-time predictions, accuracy metrics, and LLM insights.
- Resolved a 146 GB log-growth production outage, freeing 157 GB of disk space through root-cause analysis while simultaneously refactoring a 1,200-line monolithic codebase into a modular Python architecture for scalable AI/ML workflows.

Software Engineer, AT&T

12/2022 – 07/2024 | Hyderabad, India

- Engineered Python and Node.js microservices (FastAPI, Flask, REST APIs) to automate telecom network provisioning and activation workflows across OSS/BSS systems, boosting data processing efficiency by 30% and reducing manual operational overhead.
- Integrated automated test suites for functional, non-functional, and regression testing using Python, PyTest, Selenium, and Maven, increasing test coverage by 50% and reducing regression test cycles by 30%.
- Leveraged GenAI and LLMs to automate network incident analysis, correlate troubleshooting data, and deliver actionable recommendations, significantly streamlining operations and network service assurance.
- Automated Python test execution and deployment via Jenkins CI/CD pipelines with automated release gates, reducing manual effort by 40% and ensuring Dev/QA environment consistency.
- Optimized PostgreSQL queries, Python services, and API processing workflows supporting network inventory, service-order, and provisioning data, reducing average API response time by 25% and improving performance for high-volume telecom transactions.

Education

Master of Engineering, University of Cincinnati

08/2024 – 05/2026 | Cincinnati, OH

Computer Science - GPA: 3.8/4.0