

Sai Chandu Machavarapu

MLOps Engineer

+1 9403444378 • saichandumachavarapu7@gmail.com • [LinkedIn](#) • [GitHub](#) • [Portfolio](#)

SUMMARY

MLOps Engineer with expertise building scalable, cloud-native ML and infrastructure platforms across Kubernetes, KVM Linux, and public cloud environments, supporting highly available applications and distributed model-serving systems. Experienced designing CI/CD pipelines for automated model deployment, integrating MLflow model registries and RAG-based retrieval with FAISS vector databases, and troubleshooting complex infrastructure and application performance issues. Applied Python, Golang, and BASH to build operational tooling, infrastructure-as-code workflows, and AI-driven automation, while securing AI-enabled workloads. Delivered measurable improvements including 99.9% availability, 40% faster incident response, 2x higher deployment frequency, and model deployment time cut from hours to under 10 minutes.

WORK EXPERIENCE

Koch Industries

Site Reliability Engineer

NYC, NY

Mar 2026 - Present

- Architected highly scalable solutions across KVM Linux, Kubernetes, and public cloud providers, configuring subnetting, routing, NAT, and IPv4/IPv6 connectivity for analytics platforms supporting 10+ microservices and maintaining 99.9% availability.
- Managed and deployed distributed systems across OpenStack and Kubernetes environments, while monitoring operating-system and application performance, troubleshooting complex infrastructure issues, and correcting reliability defects that reduced incident response time by approximately 40%.
- Secured platform environments through nftables, firewalls, SSH, PGP, PKI, and Multi-factor authentication, strengthening web security and protecting production services against unauthorized access and configuration exposure.
- Wrote and maintained custom infrastructure tools using Python, Golang, and BASH, automating operational diagnostics and system-management workflows that improved deployment efficiency across 10+ services.

Cerner Corporation

Cloud Platform Engineer

North Kansas City, MO

Mar 2025 - Feb 2026

- Engineered containerized healthcare data-processing platforms using Docker and Kubernetes, supporting 4 production environments while managing reliable application deployments and improving consistency across distributed clinical workloads.
- Configured and troubleshooted core network services including DHCP and ARP, validated subnet connectivity and routing paths, and corrected complex network issues across production environments to reduce service disruption risks.
- Automated software delivery and system-management workflows with Ansible, Jenkins, and Gradle, improving release velocity by approximately 20% while maintaining reliable deployment processes across diverse platform environments.
- Participated with security and engineering teams to strengthen web security, SSH, PKI, PGP, and Multi-factor authentication controls, monitored infrastructure risks, and motivated cross-functional remediation efforts that reduced vulnerability findings by approximately 30%.

Citibank

Associate DevOps Engineer

Bangalore, India

Jun 2021 - Jul 2024

- Managed CI/CD workflows using Jenkins and GitLab CI, automated application build and release processes, and supported secure healthcare platform deployments that improved release velocity by approximately 20%.
- Provisioned and configured infrastructure through Ansible and Bash, participated in infrastructure automation initiatives, and supported Azure and AWS environments across 4 production environments to improve configuration consistency.
- Built and maintained containerized healthcare workloads using Docker and Kubernetes, troubleshooted application and infrastructure issues, and improved deployment reliability for clinical data-processing services by approximately 30%.

PROJECTS

Self-Service Internal Developer Platform

- Built a self-service developer platform using Backstage with a service catalog and scaffolding templates, integrating Terraform-based golden-path CI/CD workflows to cut new-service bootstrap time from a full day to under 20 minutes.

MLOps Pipeline for Model Deployment & RAG-Based Retrieval

- Built an automated CI/CD pipeline integrating an MLflow model registry with Kubernetes-based serving infrastructure and a FAISS vector database for RAG-based retrieval, cutting model deployment time from a manual multi-hour process to under 10 minutes and achieving sub-200ms query response time.

SKILLS

Programming & Scripting: Python, Bash, Java, Scala

Cloud Platforms: AWS, Microsoft Azure, Google Cloud Platform (GCP), Virtual Networks, Cloud Infrastructure Design

Containerization & Orchestration: Docker, Kubernetes, GKE, EKS, ECS, OpenShift, AWS Fargate, Container Registries, Microservices Architecture

CI/CD & Infrastructure Automation: Jenkins, GitHub Actions, ArgoCD, Gradle, Terraform, Backstage, Golden-Path CI/CD Templates, Infrastructure as Code

Monitoring, Security & AI/MLOps: Prometheus, Grafana, CloudWatch, Splunk, IAM, Cloud Key Vault Services (AWS/Azure), Secrets Management, MLflow, FAISS, Vector Databases, RAG (Retrieval-Augmented Generation), Model-Serving Infrastructure, NoSQL Databases

EDUCATION

University of Texas at Tyler

Master of Science, Computer and Information Science

Aug 2024 - May 2026

Tyler, TX