

SAI CHANDU MACHAVARAPU

Junior DevOps / Cloud Engineer | AI Platform & MLOps (Entry-Level)

📍 Tyler, TX, USA | 📞 +1 940-344-4378 | ✉️ saichandu@mail.com | 🔗 [LinkedIn](#) | 🐙 [GitHub](#)

PROFESSIONAL SUMMARY

Junior DevOps and Cloud Engineer with a strong software engineering background and hands-on experience in **AWS, Docker, Kubernetes, Terraform, CI/CD, and Linux automation**. Proven ability to build **production-style cloud infrastructure**, automate deployments, and operate containerized workloads. Additionally experienced in deploying **LLM-based RAG applications**, bridging DevOps and AI platforms with a focus on **scalability, reliability, and cost-aware production systems**. Actively seeking **Junior DevOps, Cloud Engineer, Platform Engineer, or Junior MLOps / AI Platform** roles in the United States.

EDUCATION

The University of Texas at Tyler, Tyler, TX

Master of Science in Computer and Information Science — May 2026

CGPA: 3.52

KKR & KSR Institute of Technology & Sciences, Guntur, India

Bachelor of Technology in Computer Science (AI Specialization) — May 2024

CGPA: 3.2

TECHNICAL SKILLS

Cloud & DevOps: AWS (EC2, EKS, S3, IAM, VPC), Docker, Kubernetes, Terraform, GitHub Actions, CI/CD

AI Platform / MLOps (Applied): LLMs, RAG Pipelines, LangChain, Hugging Face, Vector Databases (FAISS), Model Deployment

Programming & Automation: Python, Bash, SQL, Linux

Monitoring & Operations: CloudWatch, Logging, Deployment Debugging, Performance Optimization

Tools & Practices: Git, REST APIs, YAML, Infrastructure as Code, Secure Secrets Management

WORK EXPERIENCE

Deep Learning Intern — Codegnan IT Solutions Pvt Ltd (APSCHE Collaboration)

Nov 2023 – Apr 2024

- Applied engineering principles and AI techniques to build end-to-end ML pipelines, improving model accuracy by 18%.
- Developed interactive FastAPI + Streamlit dashboards for predictive models, reducing inference latency by 40%.
- Implemented object-oriented software solutions, ensuring maintainable, high-quality code with proper testing and validation.
- Processed structured and unstructured datasets (web scraping, PDFs), improving ML-ready dataset quality.
- Collaborated with peers and stakeholders, solving complex AI/ML problems and delivering projects on time.

AWS AI/ML Virtual Internship (AICTE & EduSkills)

Mar 2022 – May 2022

- Built ML pipelines using **scikit-learn, PyTorch, and AWS SageMaker**
- Deployed scalable inference endpoints with **CI/CD automation**, logging, and monitoring
- Applied **cloud best practices** for reliability, reproducibility, and model versioning
- Improved model performance and pipeline stability through systematic testing

PROJECTS

Production CI/CD & Cloud Deployment Project

Tech: AWS, Docker, GitHub Actions, Linux, Python

- Built a full **CI/CD pipeline** to build, test, and deploy containerized Python services
- Automated deployments on AWS with secure IAM roles and rollback strategies
- Reduced manual deployment effort and improved system reliability

Infrastructure as Code with Terraform (AWS)

Tech: Terraform, AWS (VPC, EC2, IAM, Load Balancer)

- Designed complete AWS infrastructure using **Terraform modules and variables**
- Provisioned VPCs, subnets, security groups, EC2 instances, and IAM policies
- Enabled repeatable, scalable cloud environments using IaC best practices

Kubernetes Deployment & Scaling

Tech: Kubernetes, Docker, AWS

- Deployed containerized workloads to Kubernetes clusters
- Configured services, health checks, autoscaling, and rolling updates
- Debugged pod failures and optimized resource usage

LLM-Based RAG AI Platform

Tech: Python, LangChain, Hugging Face, FAISS, Docker, AWS

- Built a **Retrieval-Augmented Generation (RAG)** system for document-based Q&A
- Integrated vector search with LLM inference for domain-specific queries
- Containerized and deployed the application with CI/CD automation
- Focused on **production concerns**: latency, cost, security, and scalability

CERTIFICATIONS

- Microsoft Certified: Azure AI Fundamentals (AI-900)
- Microsoft Certified: Azure Administrator Associate
- Oracle Vector AI Search – Certified Specialist
- Scientific Computing with Python – freeCodeCamp
- Oracle Database on AWS – Architect Professional