Vaishnavi Pangare

+91 9607486905   vaishnavipangare1304@gmail.com   LinkedIn   Medium   GitHub   Portfolio   Pune

SUMMARY

Versatile engineer with expertise in DevOps, cloud infrastructure, data engineering, data science, and AI/ML, backed by hands-on projects, startup experience, and certifications. Proven in modernizing systems to scalable microservices, building CI/CD pipelines, ETL workflows, ML models, and analytics dashboards across AWS, Azure, and GCP. Eager to deliver production-grade solutions in dynamic environments.

EDUCATION

AISSMS Institute of Information Technology,Pune   **11/2022 – 06/2026**

Bachelors of Technology in Artificial Intelligence and Data Science

TECHNICAL SKILLS

- **Cloud & Infrastructure:** AWS (ECS, Fargate, EC2, S3, VPC, IAM, Route 53, DynamoDB, CloudWatch), GCP(Big Query), Azure(Azure Static Web Apps, Azure Postgres, Azure Cloud Functions, Azure Container Apps, Azure App Service)
- **DevOps /& MLOps:** Docker, Kubernetes, Terraform, Jenkins, GitHub Actions, MLflow, Elasticsearch, Kibana, GIt, GitHub
- **Databases /& Data Tools :** MySQL, PostgreSQL, Neo4j, Qdrant,PowerBI, Tableau, Dagster, Data Wrangling, Statistical Analysis, EDA , Looker, Mage-AI
- **Operating Systems :** RHEL, Ubuntu, Linux Administration
- **Programming and Scripting:** Python, Shell Scripting, YAML, Groovy, BASH

EXPERIENCE

Hosteze – [Founding Member and DevOps Cloud Engineer]   **05/2023 – 11/2024**

- Architected and deployed AWS cloud infrastructure for student housing platform using S3, DynamoDB, Amplify, ECS, CloudWatch, and IAM to support roommate and room search functionality at scale
- Containerized applications using Docker and engineered streamlined CI/CD pipelines to streamline deployment processes and reduce operational overhead

Afto Dynamo – [DevOps Engineer]   **07/2025 – Present**

- Architected and executed infrastructure modernization from monolithic VM deployment to microservices architecture on Azure, decomposing **8+ tightly coupled services** (Elasticsearch, Chatwoot, Flask ETL apps, image scraper, client frontends/backends) into isolated **Azure Container Apps** with autoscaling, **Static Web Apps**, and **managed Elasticsearch**, slashing monthly cloud spend from **$3,000 to $859 (71% reduction)** while minimizing system failures through service isolation
- Engineered production **CI/CD pipelines with GitHub Actions** orchestrating multi-environment deployments (Staging, Production) with Container Apps' blue-green deployment strategy, reducing deployment time by **60%** and achieving **95% uptime SLA**
- Implemented comprehensive observability using **Azure Monitor and Application Insights** across all deployed services with custom metrics and distributed tracing; conducted **50+ Locust load tests** validating performance under **10K+ concurrent requests**
- Developed **Python automation scripts** for database schema migration between legacy and redesigned models with validation mechanisms; built **ETL pipeline** processing restaurant menu data with **Gemini 2.5 Pro AI-powered** description/subcategory generation, extracting structured inventory data into Elasticsearch and PostgreSQL for **10K+ product records**
- Migrated open-source Chatwoot service from VM to containerized microservices architecture, troubleshot and resolved **SSO authentication issues**, fixed complex CORS/API gateway configurations for multi-tenant admin systems, and designed **VNet-based security architecture** with private endpoints for production workloads

PROJECTS

Orange Disease Detection System - MLOps Project   **Repository**

- Built CNN-based disease classification system processing **1600+ orange leaf images** using VGG19 transfer learning with custom dense layers, achieving **94% accuracy** across 4 disease classes (canker, black spot, greening, healthy)
- Orchestrated production-ready MLOps pipeline with **MLflow experiment tracking**, containerized Flask API deployment, and automated CI/CD using GitHub Actions, **reducing model deployment time by 75%**
- Deployed scalable inference service on **AWS ECS** with Docker containerization, enabling real-time disease detection through web interface with **sub-second response times**

GraphLM - Knowledge Graph Generation Platform   **Repository**

- Engineered end-to-end knowledge graph platform processing research PDFs into structured, confidence-scored graphs using **Neo4j** and **Qdrant vector database**, reducing AI hallucination rates by **60%** and enabling verifiable fact extraction for research workflows

- Evaluated and benchmarked multiple LLM architectures including **Ollama's LLaMA 3.2** for local deployment and **OpenAI models** for production, architecting scalable microservices with **Next.js frontend** and containerized backend implementing chunked embedding pipelines to extract entity relationships with **95%+ accuracy**
- Deployed production system on **Azure App Service** with automated CI/CD pipelines, delivering interactive graph visualization and chat interface enabling researchers to extract insights **3x faster** than traditional document review methods

Three-Tier News Application - Kubernetes Deployment   Repository

- Constructed scalable **3-tier React.js news application** with structured data management using **MySQL database**, enabling efficient content delivery and user engagement through modern web architecture
- Deployed containerized application on **Amazon EKS cluster** using **multistage Dockerfile optimization** and **Helm charts**, achieving **zero downtime deployments** with automated rollouts and rollbacks
- Enhanced deployment efficiency by implementing **Kubernetes-native orchestration strategies**, reducing manual configurations by **80%** and enabling dynamic scalability to adapt to workload changes

NYC Taxi Data Engineering Pipeline - End-to-End ETL Project   Repository

- Constructed end-to-end data engineering pipeline processing **NYC Yellow Taxi trip records** using **Mage AI ETL orchestration**, transforming raw datasets into star schema data model with **7 dimension tables and 1 fact table**, enabling efficient analytical queries across **100K+ trip records**
- Formulated dimensional modeling strategy with custom Python transformations extracting temporal features (hour, day, weekday), geospatial coordinates, and categorical mappings (rate codes, payment types), achieving normalized schema **reducing data redundancy by 40%**
- Established scalable cloud infrastructure on **GCP** with Storage Buckets for raw data, Compute Engine VM hosting Mage AI pipelines, and **BigQuery data warehouse** for analytics processing; synthesized production SQL analytics layer and crafted interactive **Looker dashboards** visualizing trip patterns, revenue metrics, and geographic distribution

Amazon Sales Data Analysis and Dashboard   Repository

- Processed and cleaned Amazon sales dataset using Python (Pandas) to ensure data accuracy and performed exploratory data analysis to uncover key trends in sales performance, customer segmentation, and seasonal patterns
- Developed interactive dashboards in both Power BI and Tableau featuring 15+ visualizations (bar charts, pie charts, funnel analysis, radial gauges, geographic maps) with dynamic slicers enabling stakeholders to explore sales trends across regions, categories, and time periods
- Delivered actionable insights identifying top-performing products, regional sales disparities, and customer behavior patterns that informed strategic recommendations for inventory optimization

CERTIFICATIONS

AWS Certified AI Practitioner	**09/2025 – 09/2028**
Google Cloud Certified Associate Cloud Engineer	**11/2024 – 11/2027**
GitHub Actions	**07/2024 – 07/2027**

EXTRACURRICULAR ACTIVITIES

Scholarship Holder @Lila Poonawalla Foundation	**11/2022 – Present**
Technical Writer @Medium	**09/2022 – Present**
Cloud Lead at Google Developers Group, AISSMS IOIT	**09/2024 – 09/2025**
Volunteer at AWS User Group Pune	**02/2025 – Present**