HEMISH VEERABOINA

(951) 316-0972 ◇ MA, USA
hemish.veer@gmail.com ◇ LinkedIn ◇ GitHub

SUMMARY

DevOps / SRE Engineer with 4+ years of experience building large-scale data pipelines, Kubernetes-based deployments, CI/CD automation, observability platforms, cloud infrastructure, and MLOps workflows across AWS, Azure, and on-prem environments. Hands-on experience with Docker, Kubernetes, Helm, Terraform, Jenkins, GitLab CI, Prometheus, Grafana, MLflow, BentoML, and distributed Python systems, with strong focus on reliability, automation, ML platform engineering, and GPU-based LLM inference infrastructure.

SKILLS

Languages	Python, SQL, PySpark, Go
Data Engineering	Dask, Databricks, Hadoop, HDFS
Orchestration	Prefect, Airflow, Mage
Kubernetes	K8s, K3s, Minikube, GKE, AKS, Helm, CRDs, Controllers, Operators
Backend	FastAPI, Django, Pydantic, SQLAlchemy 2.0, REST APIs
Cloud & DevOps	AWS, Azure, Docker, Terraform, Ansible, Git, GitOps, ArgoCD, GitLab CI, Jenkins
Observability	Prometheus, Grafana, PromQL, CloudWatch, ELK
Databases & Storage	PostgreSQL, MongoDB, MySQL, SQLite, S3, ADLS Gen2
ML / AI Systems	MLflow, BentoML, vLLM, Triton Inference Server, Ray Serve

EXPERIENCE

Data Engineer / DevOps Engineer — Oct 2024 – Present
National Internet Observatory, Northeastern University — *MA, USA*

- Built distributed data pipelines in Python using Marimo notebooks, Prefect, and Dask to migrate 5–10 million records per minute from MongoDB to PostgreSQL, using Pydantic for schema validation and SQLAlchemy 2.0 for ORM-based ingestion.
- Containerized applications with Docker and deployed them to Kubernetes using Helm, Git, and GitLab CI, standardizing CI/CD workflows for distributed pipelines running on a daily cadence across Kubernetes cluster environments.
- Implemented Kubernetes-based monitoring and observability for data pipelines and dashboard services using Prometheus, Grafana, and PromQL, building metrics-driven dashboards to track pipeline health, throughput, failures, resource utilization, and application traffic across daily workloads processing 5B+ records.
- Built a research-facing visualization platform on a DMZ-hosted VM using Django for authentication and invite-based user provisioning, with a FastAPI/Pydantic API layer and Polars-powered live queries over PostgreSQL, enabling secure self-service analytics for 100+ research users.

Python Data Engineer — Aug 2024 – Oct 2024
Adobe — *CA, USA*

- Architected Project AJAX, an end-to-end event-driven pipeline in Python to extract content from Apache AEM via Solr, parsing 5,000+ pages in under 10 minutes for downstream processing and ML-ready data generation.
- Extended the pipeline with automated data preparation workflows for incremental ingestion, historical archiving, multi-format exports, and on-demand filtering, delivering training-ready datasets to AI assistant teams across Acrobat, Photoshop, Lightroom, Firefly, and VEGA in under 5 minutes.

Data Engineer Intern — Oct 2023 – Dec 2023
Cloud Data Works — *TX, USA*

- Engineered Azure Data Factory pipelines with Databricks notebooks and PySpark to ingest and transform Supabase data into ADLS Gen2, enabling 10-second refresh cycles for near real-time student performance analytics.

- Modeled analytics datasets in Azure Synapse Analytics using SQL and CETAS over ADLS Gen2 data processed with Databricks and PySpark, engineering derived features from Supabase source tables to reduce custom report preparation time by 65% for coach and student reporting.

Solution Delivery Associate / Site Reliability Engineer Jan 2020 – Aug 2022
Deloitte Touche Tohmatsu Limited *Hyderabad, India*

- Built a serverless AWS security pipeline with Lambda, Step Functions, S3, Glue, and CloudWatch to ingest and transform daily Security Hub findings, organizing alerts from GuardDuty, Inspector, and Macie into cataloged, queryable datasets used by QuickSight for infrastructure risk visibility and operational reporting.
- Developed custom Terraform templates and Ansible playbooks to automate security remediation across New York Life Insurance cloud infrastructure, contributing to a 22% improvement in security posture through scalable infrastructure changes, incident response, and root-cause troubleshooting aligned with client SLA/SLO targets.
- Engineered Jenkins CI/CD pipelines for RHEL-based environments hosted on EC2 to support blue-green deployment workflows, reducing deployment downtime by 40% and enabling safer releases, faster rollback, and improved reliability across First American's enterprise infrastructure.

PROJECTS & RESEARCH

Benchmarking GPU-Based LLM Inference on Kubernetes with vLLM, Triton, and Ray Serve. Evaluated throughput, p50/p95 latency, time-to-first-token, concurrency, autoscaling behavior, and GPU utilization using Docker, Helm, Prometheus, Grafana, and DCGM telemetry.

End-to-End MLOps Pipeline for Automated Review Workflows. Building an end-to-end MLOps pipeline for automated review applications using Python, MLflow, BentoML, Prometheus, and Grafana, covering experiment tracking, model registration, serving, monitoring, and model version upgrades.

PUBLICATIONS

Estimation in Deregulated Environments with Spark and Big Data for Power Tracing Apr 2023

Facial Emotional Recognition Using Deep Convolutional Neural Networks Sep 2021

Learning Based Approach for Hindi Sentiment Analysis Using Naive Bayes Classifier Aug 2020

CERTIFICATIONS

AWS Certified Solutions Architect – Associate

Terraform Associate

EDUCATION

MS in CS, San Jose State University Aug 2022 – May 2024
San Jose, CA

BE in CSE, M.V.S.R Engineering College Jul 2016 – Aug 2020
Telangana, India