Sai Krishna Sriram

Boulder, CO | saikrishna.sriram@colorado.edu | +1 720 233 6444 | LinkedIn | GitHub | Portfolio

Sr. Data Engineer | AWS • Python • Snowflake • Cloud & Big Data Pipelines

PROFESSIONAL SUMMARY

Data Engineer with 5+ years of experience designing and deploying scalable, cloud-native data solutions on AWS. Proven expertise in Python, SQL, Terraform, and modern data platforms (Databricks, Snowflake, Airflow) to build high-performance ETL/ELT workflows, real-time pipelines, and analytics-ready datasets. Adept at optimizing cost, latency, and reliability—consistently delivering 20–40% efficiency gains in cross-functional, high-impact projects.

EDUCATION

UNIVERSITY OF COLORADO BOULDER – M.S. in Data Science | GPA: 3.9/4.0   **May 2025**
Relevant Courses: Data Mining, Big Data Arch, Machine Learning, Computer Vision, Neural Nets & Deep Learning

TECHNICAL SKILLS

Programming & Data Processing: Python (pandas, NumPy, Flask), R, SQL, Scala, Java, C#, C++, Bash/Shell
Cloud Platforms & Services: AWS (S3, ECS, Lambda, Glue, API Gateway, SQS, CloudWatch, EKS, Kinesis), GCP (BigQuery), Azure (Data Factory, DevOps, AI Services), Terraform, MLflow, CloudFormation
Data Engineering & Orchestration: Airflow, Dagster, Spark, PySpark, dbt, Databricks
Data Warehousing & Databases: Snowflake, Redshift, PostgreSQL, MySQL, HiveQL
DevOps & Containerization: Docker, Kubernetes, Jenkins, GitHub Actions
Monitoring & Logging: CloudWatch, Prometheus, Grafana, Datadog
Additional: Elasticsearch, REST APIs, Feature Store Provisioning, MuleSoft, LLM-Fine Tuning, RAG, Graph-RAG
Domain Expertise: Scientific Data Processing, Real-time Log Analytics, ETL Automation

CERTIFICATIONS

- **AWS Certified Data Engineer – Associate *(Credly)* | Generative AI with LLMs, DeepLearning.AI *(Coursera)***

RELEVANT EXPERIENCE

AI/ML ENGINEER CO-OP   **Boulder, US**
CLD-9   Jan 2025 – Present

- Owned **end-to-end LLM pipeline** on **AWS Bedrock** & **FastAPI**—managing data ingestion, feature engineering, model lifecycle, and versioning via **MLflow**. Collaborated with **cross-functional engineering** and **analytics teams** to integrate ML-driven services into production data workflows.
- Improved predictive accuracy by 35% and cut inference latency by 25%; built observability pipelines with **Kinesis, Firehose**, and **CloudWatch** to automate anomaly detection and accelerate root-cause resolution by 40%.

AI ENGINEER INTERN   **Boulder, US**
ASANTe   Jan 2025 – May 2025

- Designed and built real-time **ETL pipelines** in **Python/PostgreSQL** on **AWS RDS** for a rewards management system—processing streaming redemption events, reducing pipeline latency by 40%, and boosting financial reporting accuracy by 18%.
- Developed **Flask & gRPC** microservices alongside **Tableau** dashboards to visualize daily rewards **KPIs** (redemptions, balances, error rates), enabling product and finance teams to make data-driven optimizations.
- Containerized microservices using **Docker** and deployed them on **AWS EKS Kubernetes clusters** with rolling updates, ensuring scalable, zero-downtime operation for ETL and KPI-tracking workflows.

ETL DEVELOPER   **Hyderabad, India**
ABSOLUTELABS.   Sep 2022 – Jul 2023

- Architected **production-grade ETL pipelines** using **PySpark**, **MapReduce**, **Hive**, and **Airflow** on **Databricks**, processing 5M+ records/day and increasing throughput by 20%.
- Integrated multi-AZ **MySQL (Aurora)** and **SOAP/REST APIs**; implemented **Jenkins CI/CD** and data-profiling workflows, improving data accuracy by 20% and reducing sync failures by 30%. Implemented AWS CloudWatch and Datadog alerts for real-time monitoring of data jobs and API integrations.

DATA ENGINEER - Engineer I   **Hyderabad, India**
QUALCOMM   Jul 2020 – Aug 2022

- Designed and maintained **high-throughput ETL pipelines** in Python & SQL for ingesting multi-terabyte hardware test-log data into QlikView dashboards; built automated precision/recall/F1 metric tracking, improving test coverage accuracy by **25%**.

- Led the implementation of a **Kafka**–**Elasticsearch** streaming analytics platform for real-time anomaly detection, reducing incident triage time by **30%** and improving operational uptime.
- Deployed **AWS S3 archival and ECS task scheduling** using Terraform, integrated with SQS-based alerts, streamlining data retention compliance and cutting manual intervention by **40%**.
- Containerized streaming analytics and **ETL workloads** in **Docker** and orchestrated deployments via **Kubernetes**, enabling efficient scaling, fault tolerance, and streamlined CI/CD delivery.

DATA ENGINEER – Graduate Intern   **Hyderabad, India**
QUALCOMM   Jan 2020 – Jun 2020
- Developed **Python-based automation scripts** to preprocess and normalize hardware performance data, reducing engineering validation time by **20%**.
- Created proof-of-concept **Scala streaming ingestion** pipelines for low-latency data visualization, laying the foundation for production-grade real-time dashboards.

MACHINE LEARNING INTERN   **Hyderabad, India**
CAREER LAUNCHER   Aug 2019 – Oct 2019
- Engineered **ML models** (regression, clustering, classification) in Python (pandas, scikit-learn) for real-time stock price prediction, boosting forecast accuracy by **18%**.
- Automated **REST API**–**to**–**PostgreSQL** ETL pipelines, cutting preprocessing by **40%**, and built **Tableau dashboards** to surface correlations and market trends, improving investment decision-making by **12%**.

CLOUD & IoT DATA ENGINEERING INTERN   **Hyderabad, India**
TATA CONSULTANCY SERVICES   Jan 2019 – Jun 2019
- Built an **IoT telemetry ingestion pipeline** using Raspberry Pi 3B+, DHT22 & MQ-135 sensors, **Flask REST APIs, AWS EC2/S3 storage**, and **PostgreSQL analytics**.
- Created **Grafana real-time dashboards** to monitor environmental trends and system health, enabling proactive anomaly detection for field deployments.

IoT DATA ENGINEERING INTERN   **Hyderabad, India**
ELECTRONICS CORPORATION OF INDIA LIMITED (ECIL)   Jan 2019 – Mar 2019
- Developed an **Arduino Uno**–**based weather monitoring system** with DHT11 & BMP180 sensors, streaming to Firebase Realtime DB for cloud-based storage.
- Automated **data cleansing and aggregation** in Python (pandas, matplotlib) and built a Flask + Chart.js dashboard for real-time visualization used by engineering teams.

PROJECTS
Scalable Data Lake Architecture
- Designed a **Hadoop-backed data lake** on **AWS S3, Glue, Athena,** and **HDFS**—authored HiveQL tables and partition-pruning policies to improve query speed by 30% and cut storage costs by 20%. Automated dynamic schema discovery and ingestion of structured and semi-structured sources using **PySpark** and AWS Glue jobs, enabling real-time analytics for downstream consumers.

GenAI Chatbot with LangChain & OpenAI
- Developed a **production-grade chatbot** using **OpenAI GPT-4** and **Lang**Chain, implementing retrieval-augmented generation (RAG) and intent classification; deployed in **Docker containers** on **Kubernetes** with **Jenkins CI/CD**. Instrumented **MLflow** for experiment tracking and set up **Prometheus/Grafana monitoring** for latency, throughput, and drift, reducing unplanned downtime by 40%.

RESEARCH EXPERIENCE

Performance Analysis of Optimization Algorithms GA, PSO, and ABC based on DWT-SVD Watermarking – **IEEE**
- Conducted comparative performance analysis of three evolutionary optimization algorithms — **Genetic Algorithm (GA), Particle Swarm Optimization (PSO),** and **Artificial Bee Colony (ABC)** — applied to a **DWT-SVD**-based digital image watermarking technique.

AWARDS & ACHIEVEMENTS
- **President, Advanced Academic Center (AAC) -** Led academic and technical initiatives, organized workshops and peer-learning programs during undergrad.
- **Second Runner-Up, *DaSSA Hackathon*, University of Colorado Boulder (2024) -** Developed a real-time data-driven solution leveraging machine learning for predictive analytics.
- **Finalist, *AAKAR IIT Bombay Hackathon* (2019) -** Competed in an ethical hacking challenge, developing security assessment and penetration testing strategies for real-world systems.
- **Black Belt, *Karate Budokan* -** Recognized for discipline, leadership, & commitment to excellence in martial art.