

SAI SHRIYA SURLA

☎ 334-849-0900 | ✉ shriya.surla@gmail.com | 🌐 [Sai Shriya Surla](#) | 📄 [Sai Shriya Surla](#) | 🏠 [AWS AI Practioner](#) | 🌐 [website](#)

SUMMARY

Data-driven engineer with a Masters in Computer Science, **AWS Certified AI Practitioner** with experience building and **optimizing ETL pipelines, analyzing large-scale datasets, and improving system workflows**. Proven ability to **translate complex technical issues into actionable insights, with hands-on experience in SQL, Python, and cloud-based systems**. Skilled in cross-functional collaboration, process improvement, and delivering reliable, scalable data solutions.

TECHNICAL SKILLS & CERTIFICATIONS

Generative AI & LLMs: LangChain, LangGraph, RAG Pipelines, Context Engineering, Prompt Optimization, OpenAI API.

Data Analytics: SQL (advanced), Snowflake, dbt, Dimensional Modeling, Data Warehousing, Data Quality & Validation, MLflow.

Frameworks & Libraries: Scikit-learn, XGBoost, PyTorch, Keras, TensorFlow, Pandas, NumPy, Express.js, React Native, jQuery

Data Warehousing & Modeling: Snowflake, dbt, Dimensional Modeling (Star/Snowflake Schemas), Enterprise Data Warehousing

Data Engineering: ETL/ELT Pipelines, Data Quality Checks, Validation Frameworks, Reporting Datasets, Medallion Architecture

Certifications: Python for Data Science (**IBM**), Data Scientist Master's Certificate (**Simplilearn**), AI Practioner (**AWS**)

Cloud, Infrastructure and languages: AWS (S3), Windows environments, Git, REST APIs, CI/CD fundamentals, python(advanced)

WORK EXPERIENCE

TempoDyn – AI Engineer | Sunnyvale, California

June 2025 – Present

- **Design & Implementation:** Developed and implemented core components of **Generative AI solutions**, specifically **RAG pipelines** using **LangChain** and **LangGraph** to process 1,000+ dynamic templates.

- **Context Engineering:** Engineered sophisticated **context engineering features** and optimized **prompt templates**, improving structured extraction reliability to 95%+ for business-critical workflows.

- **Collaboration & Integration:** Collaborated with cross-functional teams to integrate LLM outputs into a **microservice architecture**, ensuring seamless data flow between AI agents and downstream audit applications.

Sri Sathya Sai Institute - Data Engineer Intern | Karnataka, India (Remote)

May 2025 – November 2025

- **Engineered and maintained** analytics-focused ETL/ELT pipelines using **SQL, Python, Spark, dbt, and Snowflake**, processing **500K+ records across multiple healthcare datasets**

- **Modeled and architected** 20+ fact and dimension tables using star schema design to support scalable business reporting and downstream analytics. **Implemented automated** data quality frameworks including schema validation, reconciliation checks, and anomaly detection, improving reporting accuracy to **99%+ consistency**.

- **Optimized** complex SQL transformations and warehouse queries, reducing data processing time by **30%** for large-scale analytical workloads.

Al Tomouh IT LLC. - Testing Intern | Muscat, Oman

July 2023 – October 2023

- **Engineered and executed** 58+ structured test cases for a large-scale e-municipality platform supporting **150+ online citizen services**, ensuring functional accuracy pre-launch

- **Validated and analyzed** system requirements in collaboration with developers and business analysts, identifying and documenting critical defects prior to production release

- **Improved release readiness** by producing detailed defect reports, root-cause documentation, and test summaries for stakeholder review

Apparel Group UAE - Software Developer Intern | Dubai, United Arab Emirates

June 2022 – August 2022

- Led a team of 5 interns in developing core features for the **"Studio" page** of a large-scale e-commerce platform.

- Analyzed behavioral data from **100,000+ users** to inform UI/UX decisions and improve engagement.

PROJECTS

Market Shock Digital Twin | Python, FastAPI, Kafka, TimescaleDB/ClickHouse, Redis, React | [[GitHub](#)]

- Built** an event-driven market replay and portfolio simulation platform with **Kafka-backed time-warp playback (1x/10x/100x)** and REST APIs for scenario evaluation. **Implemented** idempotent ingestion, deduplication, and replay correctness checks; added Redis caching and query optimization for low-latency analytics reads.

- Deployed** via Docker and instrumented with Prometheus; load-tested replay throughput and API p95 latency under concurrent users.

Insurance Claims Analytics & Provider Performance Dashboard | Snowflake, SQL, dbt, Python, Data Modeling | [[GitHub](#)]

- **Developed** a claims analytics pipeline for insurance data, transforming raw beneficiary, inpatient, outpatient, and provider datasets into analytics-ready fact and dimension tables. **Modeled** healthcare data using a star-schema-like design, enabling efficient querying of provider utilization, reimbursement trends, and beneficiary coverage patterns. Wrote **complex SQL aggregations** and joins to compute KPIs such as average reimbursement per claim, claim counts per provider, and unique beneficiary distributions.

- Validated data quality through row-count reconciliation and schema checks across ingestion layers, ensuring end-to-end data integrity. Produced analysis outputs suitable for BI dashboards and reporting, supporting insurance business decision-making around provider efficiency and cost drivers.

Formula 1 Analytics Lakehouse | Python, PySpark, SQL, Delta Lake, Unity Catalog | [[GitHub](#)]

- Built a **scalable pipeline** for race analytics, implementing **concurrent API pulls** with **rate limiting** and **schema enforcement** to improve **data freshness** and cut end-to-end ingestion **latency** by **45%** using **Python, Databricks, and Delta Lake**.

- Engineered **Silver** and **Gold layers** to flatten nested race results into analytics-ready tables, applying **deduplication** to improve query performance by **~35%** and enable driver and constructor insights using **PySpark, Spark SQL, Delta Lake, and Unity Catalog**.

EDUCATION

Indiana University Bloomington | *Master of Science in Computer Science*

December 2025

Heriot Watt University Dubai, Scotland | *Bachelor of Science in Computer Science Hons. (AI)*

May 2023