

Shreyash Sahare

Boulder, CO | work.shreyash1001@gmail.com | 720-255-4636 | LinkedIn | GitHub

SKILLS

Programming & Scripting: Python (Pandas, NumPy), SQL, Bash, Git
Data Engineering & Tools: Apache Airflow, Spark (PySpark, Spark SQL), Parquet, ETL Pipelines, CI/CD, Docker
Cloud Platforms: AWS (Glue, Lambda, S3, Athena), Azure (Databricks, Delta Lake), GCP (BigQuery)
Databases: MySQL, PostgreSQL (OLTP), Google BigQuery (OLAP), MongoDB (NoSQL)
Data Governance: Role-Based Access Control, IAM (AWS), Azure Active Directory

EXPERIENCE

Void Robotics Marathon, FL (Remote)
Robotics Engineering Intern 08/2025 - Present

- Containerized perception modules with **Docker**, applying **CI/CD principles** to build **reproducible** data processing pipelines
- Processed **50+ images** with **computer vision techniques** in a structured workflow, ensuring accurate outputs for automation
- Implemented **real-time data ingestion and processing** with **YOLO + SLAM**, strengthening experience with performance tuning

Mercor San Francisco, CA (Remote)
Data Science Expert (Independent Contractor) 08/2025 - 09/2025

- Delivered **20+ data projects** across forecasting and predictive modeling, reinforcing **Python data transformation expertise**
- Built structured datasets and evaluation workflows to benchmark reasoning models, **improving validation coverage by 27%**
- Applied **data wrangling** and **pipeline design principles** to standardize workflows across multiple projects

Parlay Finance Boulder, CO (US)
Data Scientist 01/2025 - 05/2025

- Optimized **ETL pipelines** for loan applications, integrating **class rebalancing** and **anomaly detection**
- Improved data quality and **reduced false approvals by 22%** through validation and monitoring checks
- Collaborated with **3 teams** to deliver **technical docs** and **SHAP explainability** to support stakeholder reviews

National Institute of Technology, Tiruchirappalli Tiruchirappalli (India)
Undergraduate Research Intern 06/2022 - 08/2022

- Devised **ResCoWNet CNN** to denoise OCT scans, boosting **SSIM by 12%** over baseline models for **clearer retinal imaging**
- Employed dual-tree complex wavelet transform (**DT-CWT**), reducing speckle noise by **30%** and enhancing **OCT scan clarity**

TECHNICAL PROJECTS

Scaleflow - Supply Chain ETL Analytics | [GitHub] 02/2025 - 05/2025

- Architected RAG pipelines (**LangChain, GPT-4o**) to semantically search **3M+** trade records for **supply chain insights**
- Orchestrated **3 ETL pipelines (Airflow + Docker)** to ingest **1.5M+ records** for **supply chain analytics**
- Reduced ETL runtime by **58%** via multithreaded batching, enabling downstream access for **hourly reporting and ML tasks**

YouTube Ad Campaign Optimization | [GitHub] 12/2024 - 02/2025

- Deployed a serverless ETL pipeline (**AWS Glue + Lambda**) to structure **419K+** YouTube records for **ad targeting**
- Optimized S3-Athena queries, reducing latency by **40%** to enable **real-time content analytics** over **70K+ records**
- Streamlined **S3 parquet triggers**, enabling **2x faster ad reporting** for campaign optimization

Wine Quality Prediction with MLOps | [GitHub] 10/2023 - 11/2023

- Designed a **modular MLOps pipeline** with **user-triggered predictions** for beverage manufacturers
- Logged **30+** experiments via **MLflow** UI for **ElasticNet** tuning, reducing RMSE by **18%** and training time by **27%**
- Deployed via **Docker** on **AWS EC2** with GitHub actions, delivering predictions in **under 3s** per input

EDUCATION

University of Colorado Boulder 08/2023 - 05/2025
Master of Science in Data Science GPA: 3.73

- **Key Coursework:** Neural Networks and Deep Learning, Machine Learning, Big Data Analytics, Statistical Methods

National Institute of Technology, Tiruchirappalli 07/2019 - 05/2023
Bachelor of Technology in Electronics and Communication Engineering GPA: 3.22

- **Key Coursework:** Linear Algebra and Calculus, Statistical Theory of Communication, Deep Learning

CERTIFICATION

- **SnowPro Core Certification**, Snowflake - Aug 2025