

# Sakshi Ravindra Asati

Boulder, CO | 303-356-2393 | sakshiasati51@gmail.com

linkedin.com/in/sakshiasati | github.com/sakshiasati17 | https://sakshiasati17.github.io/SakshiAsati-Portfolio/

## EDUCATION

**University of Colorado Boulder**, Boulder, CO

Aug 2024 – May 2026 (Expected)

Master of Science in Data Science

- Teaching Assistant – Data Mining

**G H Raison College of Engineering**, Nagpur, India

Aug 2018 – May 2022

Bachelor of Engineering in Information Technology

## WORK EXPERIENCE

**Innovation & Entrepreneurship Initiative – CU Boulder**, Boulder, CO

Jan 2026 – Present

*Entrepreneurship Fellow*

- Built an automated Python ETL pipeline to collect and enrich 1,000+ CU Boulder alumni founder records from 5+ public sources using web scraping and API integrations.
- Designed normalized SQLite schema and implemented cleaning, enrichment, and deduplication workflows to produce high-quality, CRM-ready datasets.
- Generated Salesforce-ready exports and productionized weekly runs with logging, rate limiting, and error handling, replacing manual research with a fully automated process.

**Cognizant Technology Solutions**, Pune, India

May 2022 – May 2024

*Programmer Analyst*

- Maintained and stabilized production ETL pipelines ingesting high-volume, multi-format datasets into Hadoop, ensuring secure GDPR-compliant data handling.
- Refactored Apache Airflow workflows with retries and SLA monitoring, reducing pipeline failures and cutting manual intervention by 50%.
- Debugged distributed Spark job failures by tracing logs and exceptions, collaborating cross-functionally to restore stability for business-critical pipelines.
- Built Python data-quality checks and SQL validations to detect upstream issues early, reducing incident response time by 15% and improving data availability.

## PROJECTS

**Secure Medical AI — FastAPI • React • PostgreSQL • Redis • GPT-4 • Docker • GCP**

- Built a full-stack clinical documentation platform using FastAPI and React to summarize notes and assess patient risk with GPT-4, reducing documentation time and helping clinicians identify high-risk patients faster.
- Developed the backend with authentication (JWT/RBAC), audit logging, PostgreSQL, Redis caching, and background workers and deployed Docker services on GCP Cloud Run to support daily clinical workflows

**VisionStock — Retail Inventory Detection — YOLOv8 • FastAPI • PostgreSQL • Docker • GCP**

- Fine-tuned YOLOv8 on retail shelf images to detect and classify 34 product categories for automated inventory tracking.
- Built a FastAPI inference service and Streamlit dashboard to process images and return predictions in under 2 seconds, enabling near real-time stock checks during store operations.

**SupportSync — Real-Time CRM Ticketing System — MERN • Docker • GCP**

- Developed a full-stack CRM with JWT authentication and role-based access to create, assign, and track tickets for 100+ concurrent users through REST APIs.
- Integrated MongoDB Change Streams and an FAQ chatbot to push instant updates and reduce repetitive support questions; deployed Docker containers on GCP.

## TECHNICAL SKILLS

**Languages:** Python, SQL, JavaScript, R

**Backend/APIs:** FastAPI, Node.js, REST, JWT Auth

**Databases:** PostgreSQL, MySQL, MongoDB, SQLite, Redis

**ML/AI:** Scikit-learn, PyTorch, TensorFlow, XGBoost, YOLOv8

**GenAI/NLP:** GPT-4, LangChain, RAG, FAISS

**Data Eng/Infra:** Airflow, PySpark, Hadoop, ETL Pipelines

**Cloud/Tools:** Linux/Unix, Docker, AWS, GCP, CI/CD, Git, Eclipse, Streamlit, Tableau, Power BI

## LEADERSHIP

- **Treasurer, Data Science Student Association CU Boulder (DaSSA)**