

# Husayn El Sharif

## Senior Data Scientist • Machine Learning Engineer

📍 Marietta, GA    ✉ helsharif@gmail.com    ☎ (561) 247-1430    🌐 hesportfolio.com    📄 husayn-el-sharif

### Professional Summary

---

*Data Scientist and ML Engineer with 10+ years of experience developing data-driven solutions across environmental, engineering, and business domains. Ph.D. in Civil Engineering from Georgia Tech with expertise in machine learning, Generative AI, RAG, deep learning, geospatial analytics, and cloud platforms. Google Cloud Professional ML Engineer and Databricks Certified Generative AI Engineer Associate.*

### Technical Skills

---

**Programming:** Python, SQL, R, MATLAB

**Machine Learning:** Scikit-Learn, TensorFlow, PyTorch, XGBoost, Deep Learning

**Data Engineering & MLOps:** NumPy, Pandas, PySpark, MLflow, Airflow, Streamlit, Docker, Git

**AI Engineering:** Prompt Engineering, Fine-Tuning, Agentic RAG, Claude Code, Codex

**Cloud & Data Platforms:** Google Cloud, AWS, Databricks, Snowflake, Hugging Face

**Visualization:** Matplotlib, Seaborn, Plotly, Tableau, Power BI

**Geospatial Analysis:** ArcGIS, QGIS, GeoPandas, Google Earth Engine

### Certifications

---

- Google Cloud Professional Machine Learning Engineer (2026)
- Databricks Certified Generative AI Engineer Associate (2026)

### Selected Machine Learning Projects

---

#### Remote Sensing & ML for Water Quality Assessment at Lake Lanier

Built a Sentinel-2 and Google Earth Engine ML pipeline to estimate chlorophyll-a concentrations and detect harmful algal blooms, enabling near-real-time monitoring for Gwinnett County water operations. [View Project](#)

#### Agentic RAG Workflow for Building & Fire Code Review

Built a LangChain agentic RAG app with hybrid search for Cobb County building and fire code review, raising faithfulness from 0.26 to 0.97 with 0.77 context precision and 0.72 recall. [View Project](#)

#### Teacher Support Studio: Student Proficiency Prediction & AI Teacher Assistant

Built an educational ML system on 259K+ student interactions with leakage-safe XGBoost (0.70 ROC-AUC, 0.85 average precision), SHAP, and a grounded FastAPI/LangGraph teacher assistant. [View Project](#)

#### Smart Meter Risk Segmentation with Unsupervised ML

Built a k-means pipeline on high-frequency AMI data to segment smart meters by load and power-quality behavior, flagging high-risk assets for proactive utility maintenance. [View Project](#)

#### AI-Powered Retinal Disease Detection with Deep Learning & Computer Vision

Built a deep learning computer vision pipeline for four-class retinal disease detection, achieving 89% accuracy and 0.88 macro F1 to support automated screening and clinical decisions. [View Project](#)

# Professional Experience

---

Senior Data Scientist (Part-Time)

Sept 2025 – present

**Demand Delta**

Marietta, GA

- Developed a property lead-scoring pipeline using weather, permit, valuation, and geospatial data; outperformed a heuristic baseline with logistic regression and XGBoost, identifying a projected 12% revenue lift (~\$450K per 600 leads) and 3.3x higher PR-AUC. Added Platt scaling and SHAP for explainable, per-lead insights.
- Built an explainable RAG platform to match project requirements with qualified engineering firms, indexing 7K+ public profiles into ~70K retrieval documents using PostgreSQL/pgvector, embeddings, hybrid search, reciprocal rank fusion, and cross-encoder reranking.
- Built an AI-powered lead-enrichment pipeline with Python, Google Gemini, and the Google Places API to surface high-value industrial prospects for an engineering consultancy, replacing manual prospecting with a scored outreach list and accelerating lead qualification.

Senior Data Scientist (Research Engineer)

Aug 2012 – present

**Georgia Institute of Technology, School of Civil & Environmental Engineering**

Atlanta, GA

- Built Dockerized environments for geospatial ML workflows, supporting reproducible deployment of satellite-based environmental monitoring pipelines.
- Developed ML models integrating soil, crop, and climate data to predict crop yield and irrigation demand, enabling data-driven water resource planning and quantifying long range risks to agricultural production in the Apalachicola-Chattahoochee-Flint River Basin.
- Oversaw \$400K+ in annual federal and state research funding as Assistant Director of the Georgia Water Resources Institute.
- Mentored graduate researchers in data science and ML methods, supporting publications and conference presentations.

Engineering Consultant (Part-Time)

Aug 2022 – Aug 2025

**4Earth, Inc.**

Kennesaw, GA

- Developed physics- and data-driven system models and AI-Ops pipelines for real-time AWS IoT sensor analytics, improving engineering efficiency and reducing delivery time by 20%.

Data Scientist (Graduate Research Assistant)

May 2010 – Aug 2012

**Florida Atlantic University, Department of Civil, Environmental, & Geomatics Engineering**

Boca Raton, FL

- Built and validated ML models to reconstruct missing environmental data, improving accuracy and spatial consistency for hydrologic and climate analyses.

# Education

---

|       |                   |                                 |                |           |          |
|-------|-------------------|---------------------------------|----------------|-----------|----------|
| Ph.D. | Civil Engineering | Georgia Institute of Technology | Atlanta, GA    | GPA: 3.97 | May 2019 |
| M.Sc. | Civil Engineering | Georgia Institute of Technology | Atlanta, GA    | GPA: 3.97 | May 2016 |
| B.Sc. | Civil Engineering | Florida Atlantic University     | Boca Raton, FL | GPA: 3.81 | Aug 2011 |