

Husayn El Sharif

Senior Data Scientist • Machine Learning Engineer

Marietta, GA • helsharif@gmail.com • +1(561) 247-1430 • <https://helsharif.github.io>

Professional Summary

Senior Data Scientist and Machine Learning Engineer with 10+ years of experience building and deploying production-grade ML systems, analytics pipelines, and AI-powered solutions. Proven ability to translate complex data into actionable insights across geospatial, environmental, and applied ML domains.

Technical Skills

Programming	Python, SQL, R, MATLAB
Machine Learning	Scikit-Learn, TensorFlow, PyTorch, XGBoost, Deep Learning, Transfer Learning, Computer Vision
Data & MLOps	Numpy, Pandas, PySpark, MLflow, Docker, Airflow, dbt, Streamlit, Git
GenAI & LLMs	NLP, Prompt Engineering, Fine-Tuning, RAG, LangChain, CrewAI
Data Visualization	Matplotlib, Seaborn, Plotly, Tableau
Cloud Platforms	AWS, Google Cloud, Databricks, Snowflake, Hugging Face
Geospatial Analysis	ArcGIS, QGIS, GeoPandas, Google Earth Engine

Selected Projects

Remote Sensing & ML for Water Quality Assessment at Lake Lanier

- Developed a Dockerized ML pipeline for satellite-based water quality monitoring, enabling near-real-time detection of harmful algal blooms to support proactive water quality management for the Gwinnett County Dept. of Water Resources. [View Project](#)

Customer Churn Modeling & Retention Optimization with SparkML, MLflow, & Neural Networks

- Built scalable churn prediction models (Logistic Regression, Neural Networks) with SparkML, TensorFlow, and MLflow on telecom data to flag high-risk customers and drive targeted retention strategies. [View Project](#)

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

- Developed a deep learning computer vision pipeline for retinal disease classification, achieving 89% accuracy and 0.88 macro-F1 score across four diagnostic classes, supporting automated screening and clinical decision support. [View Project](#)

Multi-Agent GenAI System for Automated Literature Review

- Developed a multi-agent GenAI pipeline (Python, LangChain, LLMs) to automate hydrology literature analysis, reducing manual literature review time by ~70–80%. [View Project](#)

Professional Experience

Senior Data Scientist (Research Engineer)

Aug 2012 – Present

Georgia Institute of Technology, School of Civil & Environmental Engineering

Atlanta, GA

- Designed and deployed ML models for environmental and time-series prediction, enabling real-time water quality monitoring, flood risk assessment, and climate impact analysis through automated dashboards used by government partners.
- Applied ML to fuse NASA satellite imagery and climate time-series data, enabling more accurate basin-scale forecasting of crop yield, drought, irrigation, and streamflow.
- Oversaw \$400K+ in annual federal and state research projects as Assistant Director, Georgia Water Resources Institute.
- Mentored student researchers in data science methods, guiding successful publications, conference presentations, and career advancement.

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 time-to-market 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
Master of Science, Civil Engineering	Georgia Institute of Technology	Atlanta, GA	GPA: 3.97	May 2016
Bachelor of Science, Civil Engineering	Florida Atlantic University	Boca Raton, FL	GPA: 3.81	Aug 2011