HAMIDREZA ZORAGHEIN

hrz1365@gmail.com ⋄ linkedin ⋄ Website ⋄ +1(720) 878-2151

Senior Data Scientist

Data Scientist with 10+ years of experience transforming complex spatial datasets in energy, climate, and population domains into actionable predictive insights for agencies, utilities, and global organizations. Proven leader of end-to-end projects, from raw data engineering to production-ready analytics pipelines. Experienced in time-series forecasting, deep learning, MLOps and GeoAI, with a growing focus on integrating LLMs into workflows to drive new levels of automation, interpretation, and decision support.

SKILLS

Programming Languages:	Python, R, SQL
Deep Learning & LLM:	RAG, transformers, Prompt Engineering, LSTM, tensorflow, pytorch, FAISS, langchain, deepeval
Predictive Modeling:	XGBoost, prophet, pytorch-forecasting
Statistical Analysis:	Bayesian Statistics, Uncertainty Quantification, Anomaly Detection, Monte Carlo Simulation, Regression, Spatial Statistics
Machine Learning:	Decision trees, SVM, Time-series Forecasting, Regression, Clustering, Feature Engineering, Gradient Boosting, scikit-learn
MLOps & Automation:	MLflow (experiment tracking, model registry), Docker, Git/GitHub, CI/CD Pipelines (Github Actions), Jupyter Notebook (prototyping)
Cloud & Data Engineering Tools:	PySpark, Azure Databricks, AzureML, Azure Data Factory, Azure Synapse
Databases & Query Engines:	PostgreSQL, DuckDB
Visualization:	Tableau, shiny, datawrapper, streamlit, matplotlib, ggplot
Soft Skills:	Project Management, Team Collaboration, Client Engagement, Critical Thinking, Problem Solving, Communication, Adaptability

PROJECTS

EVRAG: Built a full **RAG** pipeline for EV charging insights, integrating **LangChain**, **FAISS**, and **Transformers**, evaluated via **DeepEval**, and operationalized with **Docker**, **GitHub Actions**, and **Streamlit**.

NYC Town and Gown, Climate Vulnerability, Impacts, and Adaptation Analysis: Built a spatial deep learning model **(LSTM, TensorFlow)** to **predict** NYC population vulnerability by age and race using Python and R; paper accepted for publication in PNAS.

EXPERIENCE

Senior Data Scientist/ML Engineer   Jan 2024 - Present
Rocky Mountain Institute   *Boulder, CO (remote)*

- Northeast Freight Corridor Charging Plan - in collaboration with the utility company National Grid and funded by U.S. Department of Energy ($**1.2M** project supporting multi-billion-dollar regional EV charging infrastructure investment plans)

  - Designed, led, and deployed **predictive simulation** models, with **uncertainty quantification** enabled, using **Geotab's vehicle telematics data** across 140 highway plazas in 9 Northeast states, enabling National Grid and several utility companies in the Northeast to plan multi-megawatt medium- and heavy-duty charging infrastructure in support of a national freight electrification strategy.

  - Developed a **modular** system to prepare, process, and visualize data for scalable, reusable analytics.

  - Co-led the development of a client-facing electrification roadmap, synthesizing technical analysis into **actionable** policy and investment guidance for utilities, regulators, and state agencies.

  - Co-led workshops to **communicate** the insights of the project to various **stakeholders**, ranging from utilities to state representatives in the Northeast.

- [GridUp](#) (funded by FedEx)
  - Led a **scalable**, **containerized** application for aggregating load demand forecasts based on **Replica's travel behavior data** to **2095** utility service territories and **3856** census places using **Dask**, **DuckDB**, and **Docker**, with **CI/CD** pipelines automated via **GitHub** Actions—enabling fast, distributed, and reproducible analytics.
  - Engineered a **granular spatio-temporal** temperature product by applying geospatial analysis with **xarray** to **unstructured**, **high-resolution**, **multi-dimensional** climate datasets, enabling quantitative assessment of seasonal effects on EV battery performance.

Data Scientist/ML Engineer Oct 2019 - Jan 2024
Population Council *New York, NY (remote)*

- Co-led the **collaborative** project of utilizing **AI-enabled** models for generating **weekly forecasts** of COVID-19 incidence rates at U.S. county level, with **uncertainty quantification**, over different forecast horizons. Built an end-to-end machine learning workflow (data collection, feature engineering, model optimization and validation) using **XGBoost** and **LSTM**. Ranked among the **top 5 accurate** forecasting models in the [US COVID19 Forecast Hub](#), delivering **50**-case-per-county improvements over the COVIDhub-ensemble (scikit-learn, XGBoost, tensorflow) used by CDC decision-makers.
- Led the launch of the Community Demographic Model, an interactive data hub hosting a wide array (15+) of data products and visualizations of socioeconomic mega trends (**matplotlib**, **shiny**, **ggplot**).
- Led the geospatial assessment of India's subnational vulnerability to climate hazards by standardizing/combining a diverse range of big **structured** (650K+ tabular census records) and **unstructured** (high-resolution remote sensing imagery) datasets (R: sf, ggplot, raster, Python: Arcpy).

Research Scientist Sept 2018 - Oct 2019
Frederick S. Pardee Center for International Futures *Denver, CO*

- Led a **predictive** model to forecast 10+ demographic characteristics (population size, gender, age structure, migration, etc) of U.S. states over 2010-2100 using **R** (**tidyverse**, **data.table**).
- Led a parallel geospatial model to generate **predictions** of the U.S. spatial distribution of population with **numerical optimization** (**10M** points) using Python (numpy, scipy, fiona, rasterio, multiprocessing).

Postdoctotal Fellow Aug 2017 - Aug 2018
National Center for Atmospheric Research *Boulder, CO*

- Led a **pattern recognition** model to predict subnational urbanization levels for the U.S., China, and India using R (tidyverse, devtools).
- Created interactive **visualization dashboards** for NCAR's demographic models using R's **shiny**.

EDUCATION

Ph.D. in Geographic Information Science, University of Colorado Boulder 2013 - 2017

Master of Geographic Information Systems, K.N. Toosi University of Technology Iran 2008 - 2011

Bachelor of Civil Engineering, K.N. Toosi University of Technology Iran 2004 - 2008

PROFESSIONAL CERTIFICATES

- Microsoft Azure Data Scientist: Cloud-Powered Skills (Microsoft and Coursera)
- Azure Data Fundamentals (Microsoft)
- Deep Learning Specialization (DeepLearning.AI)
- TensorFlow Developer (DeepLearning.AI)