Harsh Khatarkar

Betul, Madhya Pradesh | +919479392707 | harsh.khatarkar.1999@gmail.com |github.com/A-lice-ha

Skills

Programming/Scripting: R (geo-analysis, time-series forecasting, data visualization, data-pipelines), python (basic)

Machine Learning: Geo- and data- based classification, forecasting, clustering (Random Forest, SVM, XGBoost, GAM, SARIMA)

GIS & Remote-sensing: QGIS, Google Earth Engine, Google Earth Pro

Infrastructure: High-performance computing using CLI and RStudio on 256GB RAM workstation

Cloud & Version Control: Git, Network-Attached Storage, Nextcloud

Experience

Project Assistant

Indian Institute of Science, Bengaluru — Jun 2024-Mar 2025

- Built automated R pipeline for geospatial assessment of semi-arid ecosystems by harmonizing multiple satellite data, reducing runtime by ~ 67% (15 to 5 mins)
- Analyzed spatial/environmental trends in Google Earth Engine for site selection
- Localized global ecological thresholds and climate parameters for Indian context using contextual data modeling
- Validated 50+ candidate sites using Google Earth Pro; ~50% adopted in project
- Applied basic spatial forecasting using stochastic cellular automata models
- Maintained documentation of methods and communicated outputs to research leads

Projects

Solar Energy Forecasting using NASA POWER API (Time-series & ML modeling) — R

- Cleaned and Harmonized 10 years (2013-2023) solar irradiance data for time-series forecasting
- Trained and compared models (Linear Regression, XGBoost, GAM, SARIMA, Random Forest) for next-year prediction
- Modularized forecasting workflow in R, including feature engineering and lagged variables
- Evaluated forecasts (RMSE, MAE) and visualized predictions for solar energy planning through data-driven insights

Geo-AI Classification for Environmental Pattern Detection — R & Google Earth Engine

- Used Google Earth Engine to extract spatial datasets for model input and preprocessing
- Built and evaluated supervised classification models (Random Forest, k-NN, SVM, XGBoost) for vegetation and floral class detection
- Performed preprocessing, exploratory data analysis, model tuning, and confusion matrix evaluation
- Developed foundational skills in spatial ML for pattern recognition and environmental risk mapping

Education

Indian Institute of Science Education and Research, Bhopal, MP — Aug 2017-May 2022

Bachelor in Science and Master in Science (Dual Degree) - *Biological Sciences (Relevant Coursework: Biostatistics)*

Learning

- Machine Learning foundations and deployment (forecasting, classification, clustering)
- Retrieval-Augmented Generation (RAG) models and LLM-based workflows
- Neural-Networks using TensorFlow and Keras