

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 512GB RAM workstation

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

Experience

Project Assistant

Jun 2024-Mar 2025

Indian Institute of Science, Bengaluru

R, Google Earth Engine, Google Earth Pro, QGIS

- Designed an automated R pipeline for geospatial assessment of semi-arid ecosystems, reducing runtime by ~67% (15 min → 5 min)
 - Extracted and analyzed environmental trends using Google Earth Engine for ecological site selection and spatial decision-making
 - Localized global ecological tipping points for Indian landscapes using contextual geospatial modeling
 - Validated over 50 candidate grassland sites using Google Earth Pro; ~50% accepted for field deployment
 - Applied stochastic cellular automata to simulate vegetation spread and degradation across landscapes
 - Maintained reproducible documentation and communicated results through visual reports and internal briefs
-

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

Stochastic Cellular Automata for Vegetation Dynamics

R

- Developed a stochastic cellular automaton on a 50x50 grid to simulate vegetation expansion and loss
- Implemented probabilistic growth and death transitions based on Von Neumann neighborhood
- Ran multi-step simulations and visualized spatial dynamics of vegetation under varying conditions

- Strengthened understanding of local interaction-based ecological modeling and spatiotemporal analysis
-

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