



Fundamentals of Machine Learning for Earth Science

April 20, 27, & May 4, 2023

10:00-11:30 (English) or 14:00-15:30 (Spanish) EDT (UTC-4)

Artificial intelligence and machine learning have grown in popularity in recent decades as a result of advances in high-performance computing and open-source software. At the core, machine learning provides a statistical inference based on the inputs provided by the user, in which algorithms learn relationships between input data and output results. The complexity of these algorithms allows for the discovery of patterns and trends invisible to the human analyst, making it important to create analysis-appropriate input for these models to ensure that they answer the questions we are asking. This training will provide attendees an overview of machine learning in regards to Earth Science, and how to apply these algorithms and techniques to remote sensing data in a meaningful way. Attendees will also be provided with end-to-end case study examples for generating a simple random forest model for land cover classification from optical remote sensing. We will also present additional case studies to apply the presented workflows using additional NASA data.

Part 1: Overview of Machine Learning

Trainers: Jordan A. Caraballo-Vega, Caleb Spradlin, Jian Li, Jules Kouatchou

- Overview of Machine Learning
- Importance of Machine Learning targeted towards Earth Science
- Usability of Machine Learning
- Software to Support Machine Learning
- Machine Learning Applications
- Hands on Jupyter Notebook Exercise: Load and Visualize Data
- Post-session assignment
- Q&A Session

Part 2: Training Data and Land Cover Classification Example

Trainers: Jordan A. Caraballo-Vega, Caleb Spradlin, Jian Li, Jules Kouatchou

- Download the training data
- Exploratory data analysis
- Extracting training data from tabular dataset
- Extracting training data from raster dataset
- Training and inference of tabular and raster dataset
- Metrics and model evaluation
- Hands on Jupyter Notebook Exercise: MODIS Water Classification Case Study
- Post-session assignment
- Q&A Session



ARSET empowers the global community through remote sensing training.



Part 3: Model Tuning, Parameter Optimization, and Additional Machine Learning Algorithms

Trainers: Jordan A. Caraballo-Vega, Caleb Spradlin, Jian Li, Jules Kouatchou

- Overview of model tuning
- Overview of parameter optimization
- Exercise to optimize existing model
- Overview of model explainability and interpretability
- Overview of additional machine learning algorithms
- Hands on Jupyter Notebook Exercise: Improvements to MODIS Water Classification Model
- Post-session assignment
- Q&A Session



ARSET empowers the global community through remote sensing training.