NCAR
OPERATED BY UCAR

Education, Engagement
& Early-Career Development

# Introduction to Machine Learning for Earth System Scientists

**Gunter Leguy and Mariana Cains (EdEC)**

CGD Staff: Cecile Hannay, Kirsten Mayer, Katie Dagon
CISL Staff: Daniel Howard, David John Gagne, Negin Sobhani, Jessica Amaya, Charlie Becker

**edec.ucar.edu**


*This material is based upon work supported by the NSF National Center for Atmospheric Research, a major facility sponsored by the U.S. National Science Foundation and managed by the University Corporation for Atmospheric Research. Any opinions, findings and conclusions or recommendations expressed in this material do not necessarily reflect the views of NSF.*

# Machine learning bootcamp goals and expectations

**Goals**

- Set a **fundamental Machine Learning (ML) knowledge** basis using examples from the Earth system sciences for scientists and researchers with little to no knowledge on the topic of ML
- Get an idea **how ML might be applied in your work**

**You will…**

- Learn ML terminology and theoretical foundation
- Learn about ML technique and models
- Learn to evaluate and critique ML

**You won't…**

- Become a ML expert
- Know the mathematics behind ML concepts
- Know how to code ML algorithms

# Introduction to Machine Learning for Earth System Scientists

## *Curriculum couples lectures with hands-on exercises.*

- **Lecture 1: Supervised Machine Learning Basics**
  - Machine Learning Basics and Artificial Neural Networks
- **Lecture 2: Earth System Science Application**
  - Predict phase of ENSO with ANN
  - Convolutional Neural Networks
  - Predict phase of ENSO with CNN
- **Lecture 3: Explainable Artificial Intelligence (XAI)**
  - XAI Overview
  - XAI Evaluation Metrics and Methods
  - Apply XAI to ENSO ANN and CNN
- **Lecture 4: Pre-trained ML Models**
  - Image Detection using a pre-trained model (1 hour)
  - Running pre-trained AI4WP/Climate models in inference mode
- **Lecture 5: Decision Trees and Tree Ensembles**
  - Introduction to Decision Trees
  - Decision Tree Ensembles
  - Exercise: Decision Trees for ENSO
- **Lecture 6: CAMulator + Ask an AI expert**
  - CAMulator
  - Is machine learning appropriate for your science question?
  - AI/ML efforts across NSF NCAR

# Additional Resources and Opportunities

*Note: Inclusion of content does not signify endorsement of said content by the EdEC nor NSF NCAR.*

## NSF Unidata (requires free account)

NSF Unidata is a community data facility supporting a diverse community of education and research institutions with the common goal of sharing geoscience data and the tools to access and visualize that data.

- Machine Learning Foundations in the Earth Systems Sciences | NSF Unidata
- Machine Learning Applications in the Earth Systems Sciences | NSF Unidata
- Machine Learning Analysis in the Earth Systems Sciences | NSF Unidata
- Supervised Machine Learning Readiness | NSF Unidata

## National Energy Research Scientific Computing Center

NERSC regularly hosts trainings, hackathons, conferences, and more. Each event has its own registration and attendance requirements. Adding an event to your calendar does not constitute registration.

- https://www.nersc.gov/news-and-events/calendar-of-events

## ESS Applications of AI/ML

- Seeing the Atmosphere through Machine Learning | Education, Engagement & Early-Career Development
- LEAP Machine Learning Journal Club playlist | YouTube

## Machine Learning for Environmental Engineering

Explore and test state-of-the-art machine learning methods applied to environmental sciences and engineering challenges.

- Machine Learning for Environmental Engineering | Columbia Plus