
SCIWS15 - Machine Learning for Small, Uncertain, and Sparse Data Sets



Saturday, 10 December 2022



07:00 - 15:00



McCormick Place - S102cd

Workshop Agenda

Demonstrate the utilization and importance of unsupervised machine learning to analyze geoscience data

The world's most valuable resource is no longer oil. It is data. Mining geoscience data is a key new research and application area in our field. Recently, a series of novel ML tools have been developed to analyze diverse datasets related to Earth sciences, including climatic, geologic, geophysics, geothermal, carbon storage, oil/gas, and wildfire applications. Importantly, geologic datasets are often sparse or have missing values. Also, they are often small; therefore, data-hungry ML (e.g., deep learning) methods cannot provide accurate predictions. This workshop will talk about appropriate ML tools for handling sparse, uncertain, and small datasets. Specially, we will demonstrate various unsupervised and physics-informed unsupervised ML methods. Moreover, the workshop will also discuss the pros and cons of alternative machine learning methods for different real-world problems. The workshop will include seminar talks and hands-on real-time demonstrations of existing case studies. The workshop will be suitable and valuable for anyone regardless of their machine learning experience by providing materials at introduction, intermediate and expert levels.

Event For

Atmospheric Sciences, Earth and Planetary Surface Processes, Earth and Space Science Informatics, Hydrology, Nonlinear Geophysics

Type

Scientific Workshop

Submitter

[Bulbul Ahmmed](#)

Los Alamos National Laboratory

Presenters

[Bulbul Ahmmed](#)

Los Alamos National Laboratory

[Velimir monty V Vesselinov](#)

Los Alamos National Laboratory

Learning Objectives

1. We will demonstrate tutorials for k -means clustering, principal component analysis, non-negative matrix factorization (NMF), NMF with customized k -means clustering, and physics-informed NMF. Also, we will go over the brief mathematics of each method. Furthermore, we will discuss the advantages and disadvantages of each method for geologic data.

Ask a question or comment on this session (not intended for technical support questions).

Have a question or comment? Enter it here.

Section: Scientific Workshops

Sections: Nonlinear Geophysics

Sections: Hydrology

Sections: Earth and Space Science Informatics

Sections: Earth and Planetary Surface Processes

Sections: Atmospheric Sciences

Type: Scientific Workshop