

Machine Learning in Geoscience: Riding a Wave of Progress

2nd Annual Machine Learning in Solid Earth Geoscience Conference; Santa Fe, New Mexico, 18–22 March 2019



Machine learning offers a new way to use the petabytes of available geoscience data to tackle complex, unsolved problems. Last March, conference attendees discussed the latest findings and future applications of machine learning in geosciences. Credit: iStock.com/metamorworks

By Daniel T. Trugman, Gregory C. Beroza, and Paul A. Johnson © 3 May 2019

The geosciences are data rich, with petabytes of readily and publicly available data. This availability, combined with the complexity of unsolved problems in the field, has motivated vigorous interest in the

application of machine learning (ML) techniques. ML offers a new “lens” for viewing data and scientific hypotheses that differs from the perspective of traditional domain expertise. Initial uses of ML have tended to be limited in scope and isolated in application, but recent efforts to promote benchmark geoscientific data sets and competitions promise to propel broader, deeper, and increasingly coordinated and collaborative efforts.

This year, we were astonished by a flood of interest, with attendees representing institutions from across the globe and a range of career stages.

We organized the second Machine Learning in Solid Earth Geoscience Conference

(<https://www.lanl.gov/calendar/2019/March/0318-geoscience-conference.php>) to promote discussion of the latest findings and to brainstorm for the future. The response for the first conference, held in 2018, was enthusiastic but modest. This year, we were astonished by a flood of interest, with attendees representing institutions from across the globe and a range of career stages—graduate students mingled with midcareer and senior scientists. Conference organizers sought to facilitate a common understanding of ML by arranging a set of tutorials to begin the conference.

Contributed presentations tended to fall along research themes particularly amenable to ML. Seismic event detection was prominently featured, but presentations also highlighted research in seismic imaging: a subject that is more typically the bastion of classical inverse theory

(<https://www.geos.ed.ac.uk/homes/acurtis/InverseTheory.html>). Other contributions dealt with the analyses of laboratory shear experiments, which provided a wealth of data to explore the physics of earthquake nucleation at the smallest of scales.

One session leveraged a mixture of geodetic and seismic data to explore the dynamics of volcanoes and geysers. Landslides, ground motion prediction, and human-induced seismicity

(<https://eos.org/articles/catching-oklahomas-tiny-tremors-in-the-act>) were among other hazards receiving attention. Still other presentations demonstrated that ML representations of subsurface fractures can provide new insights for both energy production and the evolution of tectonic faults. Researchers also used ML to capture geologic and geomechanical properties.

The plenary presentation focused on the theoretical underpinnings of deep learning (<https://eos.org/editors-vox/deep-learning-a-next-generation-big-data-approach-for-hydrology>), a successful, but notoriously opaque, technique. By connecting deep neural networks (<https://eos.org/opinions/three-steps-to-successful-collaboration-with-data-scientists>) to the problem of spline fitting ([https://en.wikipedia.org/wiki/Spline_\(mathematics\)](https://en.wikipedia.org/wiki/Spline_(mathematics))), this presentation provided a conceptual framework for improved understanding of this approach.

Attendees agreed to expand future conferences to include other fields, with the hope of bringing fresh perspectives and inspiration to long-standing geoscience problems.

The conference concluded with a discussion of future directions for ML in the geosciences. Given the rapid expansion of research in this field, attendees expressed concern that machine learning was becoming overhyped, and they worried that a precipitous fall in enthusiasm would follow, as happened during the ML “winter (https://www.theregister.co.uk/2018/02/08/second_ai_winter/)” of the late 20th century. A renewed emphasis on technical rigor, with common benchmarks, and an emphasis on repeatability, as is common in computer sciences, could help avoid this.

Attendees agreed to expand future conferences to include other fields, with the hope of bringing fresh perspectives and inspiration to long-standing geoscience problems. ML is a rapidly evolving set of approaches that are incompletely understood. However, its demonstrated power suggests that we are still riding the wave and that ML’s application in the geosciences will continue to grow rapidly for the foreseeable future.

—Daniel T. Trugman (dtrugman@lanl.gov (<mailto:dtrugman@lanl.gov>)), Earth and Environmental Sciences Division: Geophysics, Los Alamos National Laboratory, N.M.; Gregory C. Beroza, Department of Geophysics, Stanford University, Calif.; and Paul A. Johnson, Earth and Environmental Sciences Division: Geophysics, Los Alamos National Laboratory, N.M.

Citation: Trugman, D. T., G. C. Beroza, and P. A. Johnson (2019), Machine learning in geoscience: Riding a wave of progress, *Eos*, 100, <https://doi.org/10.1029/2019EO122671>. Published on 03 May 2019.

Text © 2019. The authors. [CC BY-NC-ND 3.0](https://creativecommons.org/licenses/by-nc-nd/3.0/)

Except where otherwise noted, images are subject to copyright. Any reuse without express permission from the copyright owner is prohibited.