MACHINE LEARNING

Training machines in Earthly ways

Geoscientists are training computers to learn from a wide range of geologic data and, in the process, the machines are teaching geoscientists about the workings of Earth.

Chris Marone

Machine learning (ML) — a rapidly growing set of computational techniques and adaptable decision-making algorithms — enables the identification of patterns that are difficult or impossible for humans to detect, at speeds that are hundreds to thousands of times faster than traditional computational techniques. The worldwide explosion of digital data, coupled with expanding computational power, has fuelled intense efforts to build fast algorithms that can learn from data to solve complex problems. ML is having a significant impact in areas ranging from automatic music identification to global financial forecasting and image processing, and is being adopted widely in industrial applications and scientific research. And now machines are being trained in Earthly ways. A conference[1] entitled 'Machine Learning in Solid Earth Geosciences' was held in February in Santa Fe, New Mexico, USA. The 125 participants discussed the applications of ML to topics as diverse as where to drill in an oil field, carbon sequestration and human-induced seismicity.

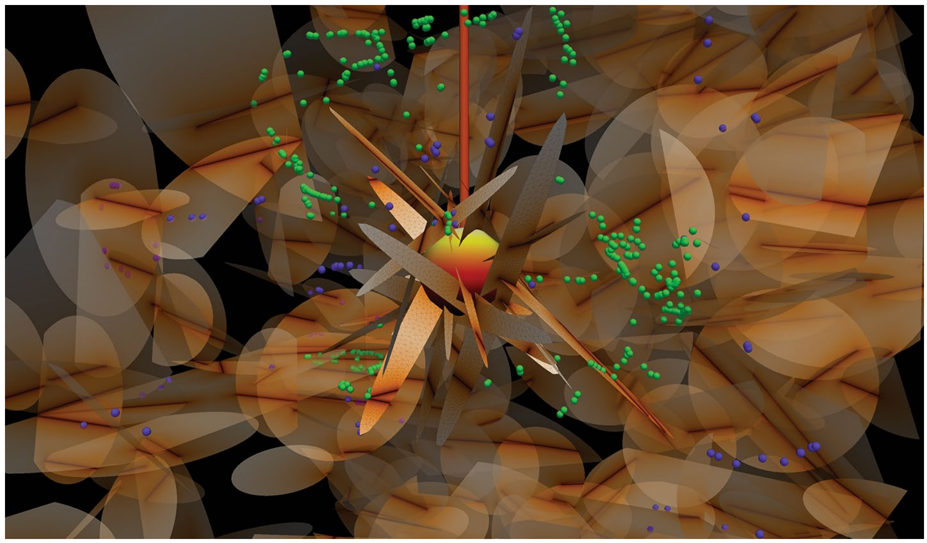

Fig. 1 | Machines trained in the geosciences. Xenon gas seepage (green dots) and fracture planes (grey and brown ellipses) coming from a blast cavity (central ball) identified using ML-based emulators trained to predict conditions in a simulated nuclear test[4]. Discussions during the 'Machine Learning in Solid Earth Geoscience' conference[1] highlighted numerous applications of ML to large geoscience datasets, including models developed to predict fluid and gas leakage from underground sites. Credit: Image courtesy of J. Hyman, Los Alamos National Laboratory

ML algorithms are astonishingly good at handling big datasets with unknown noise and uncertainty. This makes them particularly impactful in the geosciences, where large arrays, continuous monitoring and high-resolution simulations produce huge, complex datasets. Particular effort has been aimed at using automated techniques to develop more comprehensive earthquake catalogues. By borrowing from techniques developed for music identification, we can rapidly identify and fingerprint earthquake signals that might previously have gone unnoticed, and automatically search archived and real-time seismic data for related events (Gregory Beroza, Stanford University, USA). This approach could transform the way that seismologists build earthquake catalogues, as well as how scientists investigate earthquake physics and fault mechanics.

An exciting development in the field of earthquake physics is the application of ML to earthquake forecasting, which has been tested in the laboratory and is being applied to tectonic faults. Continuous seismic signals emanating from laboratory faults look like noise to the human eye. However, by using ML to map statistical attributes of the elastic radiation emitted during shearing, laboratory earthquakes can be predicted with remarkable fidelity[2,3]. The technique predicts the full spectrum of lab stick-slip modes, including slow-slip events and fast, dynamic rupture. It is also being tested on tectonic faults, with surprising results: preliminary work shows that seismic data can be used to retrospectively predict both the amount of fault displacement and the timing of episodic slip events on the Cascadia subduction interface that separates the North American and Juan de Fuca plates (Claudia Hulbert, Bertrand Rouet-Leduc and Paul Johnson, Los Alamos National Laboratory, USA).

Building a crowd-sourced earthquake early warning system is another promising ML application. A smart phone app can now monitor movements using the accelerometer we all carry — our mobile or cellular phones (Qingkai Kong and Richard Allen, University of California, Berkeley, USA). The app uses ML to distinguish seismic signals from our daily activities and it learns from our individual characteristics and lifestyles to distinguish real seismic activity from our normal movements. When a mobile phone detects seismic shaking, it alerts a cloud-based server where the spatiotemporal distribution of alerts and the acceleration record are used to locate the event and broadcast a warning, or ignore it as a false alarm.

ML models are also being developed to predict flow, optimize hydraulic fracturing, and detect seepage from carbon sequestration and nuclear non-proliferation sites (Gowri Srinivasan and Jeffrey Hyman, Los Alamos National Laboratory, USA). Preliminary results show that ML can achieve similar levels of accuracy to more traditional, complex numerical solutions[4] (Fig. 1). Yet, ML can do this in just a few seconds, compared to a few days for traditional methods, which may lead to real-time decision making at drilling and production sites. Similarly, ML is being used to reduce error in groundwater and contaminant flow models by using improved geologic models and more realistic descriptions of the underlying physics of flow and channel dynamics (Albert Valocchi, University of Illinois, USA). The improved accuracy of flow channelling and groundwater usage will aid sustainable water use, by avoiding excessive pumping and groundwater depletion that can occur with poorly calibrated models. And ML is also leading to the discovery of previously unknown processes: it seems that the topology of a fracture network plays a far more important role in determining primary flow pathways than physical attributes such as permeability, contrary to conventional wisdom[5] (Hari Viswanathan and colleagues, Los Alamos National Laboratory, USA).

High-resolution images of Earth's core–mantle boundary region are being developed by using multi-scale models and ML to analyse seismic waves. A mathematical technique based on micro-local analysis can be combined with ML to incorporate vast sets of often low-quality seismograms to produce ultra-high-resolution images of Earth's mantle[6] (Maarten de Hoop, Rice University, USA). Rather than using computationally expensive simulations, the technique uses self-adapting algorithms known as neural networks to reveal a new level of layered complexity near the Earth's core–mantle boundary that may reflect a previously unknown phase transition from a mineral called perovskite to post-perovskite.

Vast areas of Earth's surface can also be rapidly mapped by applying ML to remote sensing data. The combination of airborne geophysics and Landsat imagery to classify local geology and lithology[7] highlights the value of spatially diverse ML training data in order to generate geologically plausible predictions[8] (Anya Reading, University of Tasmania, Australia). The technique is finding particular application in industry where the rapid and automated assessment of large, remote sensing datasets can aid in locating lithologic boundaries in areas that are inaccessible or impractical to study with traditional mapping techniques.

The conference[1] on 'Machine Learning in Solid Earth Geoscience' put into sharp focus the modern pedagogical view that teaching and learning are fundamentally intertwined, and that to be good at one requires a focus on both. ML techniques have had immediate impact in many areas where learning is enough, such as the rapid identification of our favourite music. But ML is also having an impact as a tool for teaching, emerging as an important research direction that is likely to produce a new era in scientific discovery. ❐

Chris Marone

Department of Geosciences, Pennsylvania State University, University Park, PA, USA.
e-mail: marone@psu.edu

Published online: 23 April 2018
https://doi.org/10.1038/s41561-018-0117-5

References

1. *Machine Learning in Solid Earth Geosciences* (Center for Nonlinear Studies, Los Alamos National Laboratory, 2018); https://go.nature.com/2GFLYO0
2. Rouet-Leduc, B. et al. *Geophys. Res. Lett.* **44**, 9276–9282 (2017).
3. Rouet-Leduc, B. et al. *Geophys. Res. Lett.* **45**, 1321–1329 (2018).
4. Valera, M. et al. *Computat. Geosci.* http://doi.org/cm3q (2018).
5. Karra, S., O'Malley, D., Hyman, J. D., Viswanathan, H. S. & Srinivasan, G. *Phys. Rev. E* **97**, 033304 (2018).
6. Yu, C., Day, E. A., de Hoop, M. V., Campillo, M. & van der Hilst, R. D. *J. Geophys. Res. Solid Earth* **122**, 10364–10378 (2017).
7. Kuhn, S., Cracknell, M. & Reading, A. *ASEG Extended Abstracts* http://doi.org/cm3r (2018).
8. Cracknell, M. J. & Reading, A. M. *Computat. Geosci.* **63**, 22–33 (2014).