Preface to the Focus Section on Machine Learning in Seismology

by Karianne J. Bergen, Ting Chen, and Zefeng Li

Machine learning (ML) is a collection of algorithms and statistical models that enable computers to extract relevant patterns and information from large data sets. Unlike physical modeling approaches, in which scientists develop theories based on physical laws to compare with real-world observations, ML approaches learn directly from data without explicitly reasoning about the underlying physical mechanisms.

ML algorithms are often categorized into supervised and unsupervised learning (see fig. 2 in Kong *et al.*, 2018). Supervised learning algorithms build a model from existing labeled input data with the goal of predicting the labels of new data. A key challenge in supervised ML is to ensure the model learns general patterns in the data, rather than memorizing the known examples. Thus in practice, the available data are split into two subsets, called training and test sets, which are used to learn the model and measure the model performance, respectively (see fig. 3 in Kong *et al.*, 2018). In contrast, unsupervised learning is used to discover the hidden patterns or structures in a given data set without using labeled data. Major tasks in unsupervised learning include cluster analysis, which separates objects into groups so that the objects within the same group share similar features, and dimensionality reduction, which learns a low-dimensional feature representation for the data.

In the last decade, there have been rapid advances in a class of ML methods called deep neural networks, or deep learning (Jordan and Mitchell, 2015; LeCun *et al.*, 2015). Given enough data, these deep networks can learn very complex, nonlinear relationships between input data and prediction targets, leading to accurate predictions on a number of challenging tasks. Deep neural network architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are widely applied in image recognition, natural language processing, and robotic control tasks.

Although seismologists used ML algorithms for decades to analyze seismic signals (Dysart and Pulli, 1990; Dai and MacBeth, 1995; Ohrnberger, 2001), over the last few years, there has been rapid growth in research activity on ML applications in seismology. This is driven by several factors: the increasing size of seismic data sets, improvements in computational power, new algorithms and architectures (e.g., deep neural networks), and the availability of easy-to-use open-source ML frameworks (Abadi *et al.*, 2016).

The renewed interest in ML is evidenced by a number of new data-science-focused conference sessions and activities in the seismology community. The annual meetings of the Seismological Society of America and the American Geophysical Union included well-attended sessions on ML in seismology since 2016 and 2017, respectively. In 2018, Los Alamos National Laboratory (LANL) hosted a dedicated conference on ML in Solid Earth Geoscience, and the 2018 Conference on Neural Information Processing Systems, a top artificial intelligence conference, included a workshop session on geophysical applications. Data science challenges such as the Wenchuan aftershock detection competition (Fang *et al.*, 2017) and the LANL earthquake prediction competition (Kaggle Research Prediction Competition) attracted the interest of hundreds of teams, including both seismologists and data scientists.

Earthquake detection and phase picking have been popular targets for automation with ML (Yoon *et al.*, 2015; Perol *et al.*, 2018; Ross, Meier, and Hauksson, 2018; Zhu and Beroza, 2018), with the task often framed as a classification problem: labeling signals as earthquake versus noise, or as P-, S-phase versus noise. Seismologists also made efforts to use ML in earthquake early warning systems due to its relatively low online operation cost and high generalization capability (Kong *et al.*, 2016; Li *et al.*, 2018). In seismic hazard assessment, ML has been used to develop ground-motion prediction equations (GMPEs) that are nonlinear and include unconventional factors (Alavi and Gandomi, 2011; Derras *et al.*, 2012; Trugman and Shearer, 2018). Kong *et al.* (2018) provide a more detailed review of recent applications of ML in earthquake detection and phase picking, earthquake early warning, ground-motion prediction, seismic tomography, and earthquake geodesy.

This focus section of *Seismological Research Letters* contains 16 original articles covering a range of ML applications in seismology. Five articles use CNN-based approaches for various seismic data processing steps in the earthquake monitoring pipeline. CNNs are a class of supervised deep neural networks that have been shown to be successful for a range of prediction tasks (see the focus section article by Dokht *et al.*, 2019, for a good introduction to CNNs). Their applications in this focus section include event detection, phase picking and identification, phase association, hypocenter location, and event characterization.

Dokht *et al.* (2019) present CNNs for event detection and phase detection using training data from the western Canada earthquake catalog. The first CNN discriminates classifies spectrograms as noise or earthquake, and a second CNN identifies P and S phases from the synchosqueezing wavelet transform of the detected earthquake signals. The phase identification network can also estimate timing of phase arrivals. The authors report that the use of time–frequency features results in fewer false detections and better *P*-phase identification than the time-domain representation of signals, which has been used in previous CNN-based detection, for example, Perol *et al.* (2018). Woollam *et al.* (2019) apply a simple CNN architecture to classify seismic phases for a local seismic network in northern Chile. Compared to other studies with extensive training data (Ross, Meier, and Hauksson, 2018; Zhu and Beroza, 2018), this work shows that, with a small training data set and appropriate extra processing, CNNs may still be useful in producing more accurate and consistent phase picks.

McBrearty *et al.* (2019) link seismic arrivals on a pair of stations using a CNN, in comparison to the work by Ross, Yue, *et al.* (2019) who used an RNN to aggregate the arrivals on multiple stations simultaneously. The authors applied this technique to a large data set in Chile and obtain a success rate of 80%, which supplements current standard travel-time-based association methods. Kriegerowski *et al.* (2019) use a CNN to learn a mapping from multistation normalized waveforms to hypocenter location. The trained CNN locates events in the 2008 western Bohemia earthquake swarm to within hundreds of meters of double-difference catalog locations. *P*-phase picks for a reference station are required in the current implementation. The authors demonstrate that the initial layers of the CNN learn features that can be used for event detection. The authors made their code available to the seismology community (see Data and Resources in Kriegerowski *et al.*, 2019).

Lomax *et al.* (2019) extend the work of Perol *et al.* (2018) and present ConvNetQuake_INGV, a CNN capable of detecting earthquakes and inferring their magnitude, distance, depth, and azimuth with 50-s-long waveforms on individual stations. The method achieves very strong performance for event detection, and moderate performance for characterization. The possibility of simultaneously detecting and estimating a number of event parameters is an important development toward real-time monitoring, but the authors caution that additional research is necessary to ensure that ML-based monitoring algorithms are robust and generalize well to regions where training data are limited. The authors also made their code available to the seismology community (see Data and Resources in Lomax *et al.*, 2019).

Two contributions focus on detecting or discriminating specific types of seismic events. Nakano *et al.* (2019) train a CNN to discriminate between tremor, local earthquakes, and noise in data collected by ocean-bottom seismometers along Nankai trough. As in Dokht *et al.* (2019), the input data in this work are represented in the time–frequency domain, and the authors report a small performance gain from applying asymmetric pooling along the time axis only. Bregman and Rabin (2019) use an ML-based approach to detect repeating seismic events in an aftershock sequence. They use diffusion maps to learn a low-dimensional embedding for the seismic signals and identify aftershocks based on their proximity to the mainshock in the embedding space. The diffusion map approach is also proposed as a tool for validating events identified with waveform correlation.

In the applications above, ML is being used to develop new data-driven methods for seismic signal processing in traditional seismic networks. Three articles in the focus section use ML to analyze alternative data sources for earthquake detection and early warning applications. Kong *et al.* (2019) introduce several ML algorithms implemented in the MyShake system. With smartphone recorded ground-motion data, they tried artificial neural network (ANN), CNN, and density-based clustering algorithm DBSCAN to classify an incoming signal as earthquake or other sources. Once an earthquake is confirmed, a random forest is used to estimate the earthquake magnitude. Wang *et al.* (2019) detect earthquakes directly from scanned images of old analog seismograms using an ML approach. In this work, the authors bypass the step of digitizing the data, and demonstrate the feasibility of detecting on scanned analog seismograms directly. Zhou *et al.* (2019) present a method to detect laboratory earthquakes in rock experiment data. The proposed method extracts features from a time series and constructs two dictionaries of feature vectors representing labquakes and noise. Classification is performed by selecting the dictionary with the best fit to the data.

Handling noisy data is an ongoing challenge across seismic signal processing applications. Three articles address this challenge; these papers use ML for denoising seismic data or present preprocessing methods that improve the performance of ML algorithms on noisy data. Zhang *et al.* (2019) present an unsupervised method for denoising seismic without labeled training data; this contrasts with other recent efforts that used a supervised learning approach for denoising (Zhu *et al.*, 2018). This approach learns a dictionary of atomic waveforms from data; iteratively fitting these dictionary waveforms to the data allows the noise to be discarded. This approach is applied to synthetic and field data and shows potential for improving signal visibility. Cortés *et al.* (2019) use empirical mode decomposition to standardize noisy seismic data to improve performance of a hidden Markov model (HMM) classifier in volcanoseismic recognition systems. The use of this denoising approach allows the authors to apply an HMM trained on data from one station on Deception Island to detect and classify events in seismic data collected at another station 900 m away. Yuan *et al.* (2019) distinguish the state of Chimayó geyser, discriminating eruptions and precursory seismic signals, using a random forest classifier applied to a collection of time series features. The authors demonstrate that the use of an autoregressive model for denoising improves the classification performance.

The use of supervised ML for ground-motion modeling is the subject of two focus section articles. Khosravikia *et al.* (2019) predict peak ground acceleration from three input parameters: magnitude, distance, and V_{S30} (time-averaged shear-wave velocity to 30 m depth), using a simple ANN model. The results show high-prediction accuracy, demonstrating ML has the potential to complement current methods using traditional GMPEs. Khoshnevis and Taborda (2019) use an ANN as a surrogate for physics-based earthquake ground-motion simulations. A pool-based query-by-committee active learning method is employed to train the ANN. Using a regional-scale simulation data set with variable anelastic models and corresponding peak ground velocity, it has been shown that the proposed active learning approach is better than passive learning for reducing computation by requiring less training data. Finally, one contribution to the focus section relates to surface-wave tomography. Ortega *et al.* (2019) apply Markov chain Monte Carlo (MCMC) to select dispersion curves, a task which is typically performed manually by an expert. The authors first use existing velocity structure models to obtain prior dispersion curve for various focal mechanisms. Then, they use MCMC to tune the dispersion curves to improve the fit to the data.

The seismology community is starting to explore the use of state-of-the-art ML methods to advance seismology and earthquake science. The contributions in this focus section demonstrate the potential of ML as tools for scientific discovery across a broad range of seismological applications.

ACKNOWLEDGMENTS

The authors would like to thank all the authors for their contributions to this focus section. The authors are grateful to reviewers for their constructive and timely feedback. The authors thank *SRL* Editor-in-Chief Zhigang Peng for inviting us to be guest editors of this focus section and *SRL* Managing Editor Mary George for managing the focus section.

REFERENCES

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.* (2016). Tensorflow: A system for large-scale machine learning, *12th USENIX Symposium on Operating Systems Design and Implementation*, Vol. 16, 265–283.

Alavi, A. H., and A. H. Gandomi (2011). Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing, *Comput. Struct.* **89,** nos. 23/24, 2176–2194, doi: 10.1016/j.compstruc.2011.08.019.

Bregman, Y., and N. Rabin (2019). Aftershock identification using diffusion maps, *Seismol. Res. Lett.* doi: 10.1785/0220180291.

Cortés, G., R. Carniel, M. Á. Mendoza, and P. Lesage (2019). Standardization of noisy volcanoseismic waveforms as a key step toward station-independent, robust automatic recognition, *Seismol. Res. Lett.* doi: 10.1785/0220180334.

Dai, H., and C. MacBeth (1995). Automatic picking of seismic arrivals in local earthquake data using an artificial neural network, *Geophys. J. Int.* **120,** no. 3, 758–774.

Derras, B., P. Y. Bard, F. Cotton, and A. Bekkouche (2012). Adapting the neural network approach to PGA prediction: An example based on the KiK-net data, *Bull. Seismol. Soc. Am.* **102,** no. 4, 1446–1461.

Dokht, R. M. H., H. Kao, R. Visser, and B. Smith (2019). Seismic event and phase detection using time-frequency representation and convolutional neural networks, *Seismol. Res. Lett.* doi: 10.1785/0220180308.

Dysart, P. S., and J. J. Pulli (1990). Regional seismic event classification at the NORESS array: Seismological measurements and the use of trained neural networks, *Bull. Seismol. Soc. Am.* **80,** no. 6B, 1910–1933.

Fang, L., Z. Wu, and K. Song (2017). SeismOlympics, *Seismol. Res. Lett.* **88,** no. 6, 1429–1430, doi: 10.1785/0220170134.

Jordan, M. I., and T. M. Mitchell (2015). Machine learning: Trends, perspectives, and prospects, *Science* **349,** no. 6245, 255–260, doi: 10.1126/science.aaa8415.

Khoshnevis, N., and R. Taborda (2019). Application of pool-based active learning in physics-based earthquake ground-motion simulation, *Seismol. Res. Lett.* doi: 10.1785/0220180296.

Khosravikia, F., P. Clayton, and Z. Nagy (2019). Artificial neural network-based framework for developing ground-motion models for natural and induced earthquakes in Oklahoma, Kansas, and Texas, *Seismol. Res. Lett.* doi: 10.1785/0220180218.

Kong, Q., R. M. Allen, L. Schreier, and Y. W. Kwon (2016). MyShake: A smartphone seismic network for earthquake early warning and beyond, *Sci. Adv.* **2,** no. 2, e1501055, doi: 10.1126/sciadv.1501055.

Kong, Q., A. Inbal, R. M. Allen, Q. Lv, and A. Puder (2019). Machine learning aspects of the MyShake global smartphone seismic network, *Seismol. Res. Lett.* doi: 10.1785/0220180309.

Kong, Q., D. T. Trugman, Z. E. Ross, M. J. Bianco, B. J. Meade, and P. Gerstoft (2018). Machine learning in seismology: Turning data into insights, *Seismol. Res. Lett.* **90,** no. 1, 3–14, doi: 10.1785/0220180259.

Kriegerowski, M., G. M. Petersen, H. Vasyura-Bathke, and M. Ohrnberger (2019). A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms, *Seismol. Res. Lett.* doi: 10.1785/0220180320.

Kaggle Research Prediction Competition, LANL Earthquake Prediction, available at https://www.kaggle.com/c/LANL-Earthquake-Prediction/data (last accessed January 2019).

LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep learning, *Nature* **521,** no. 7553, 436, doi: 10.1038/nature14539.

Li, Z., M.-A. Meier, E. Hauksson, Z. Zhan, and J. Andrews (2018). Machine learning seismic wave discrimination: Application to earthquake early warning, *Geophys. Res. Lett.* **45,** 4773–4779, doi: 10.1029/2018GL077870.

Lomax, A., A. Michelini, and D. Jozinović (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network, *Seismol. Res. Lett.* doi: 10.1785/0220180311.

McBrearty, I. W., A. A. Delorey, and P. A. Johnson (2019). Pairwise association of seismic arrivals with convolutional neural networks, *Seismol. Res. Lett.* doi: 10.1785/0220180326.

Nakano, M., D. Sugiyama, T. Hori, T. Kuwatani, and S. Tsuboi (2019). Discrimination of seismic signals from earthquakes and tectonic tremor by applying a convolutional neural network to running spectral images, *Seismol. Res. Lett.* doi: 10.1785/0220180279.

Ohrnberger, M. (2001). Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia, *Doctoral Thesis*, Institut für Geowissenschaften, Universität Potsdam.

Ortega, R., D. Carciumaru, E. Huesca, and E. Gutierrez (2019). Automatic selection of dispersion curves based on a weighted probability scheme, *Seismol. Res. Lett.* doi: 10.1785/0220180282.

Perol, T., M. Gharbi, and M. Denolle (2018). Convolutional neural network for earthquake detection and location, *Sci. Adv.* **4,** no. 2, doi: 10.1126/sciadv.1700578.

Ross, Z. E., M.-A. Meier, and E. Hauksson (2018). *P* wave arrival picking and first-motion polarity determination with deep learning, *J. Geophys. Res.* **123,** 5120–5129, doi: 10.1029/2017JB015251.

Ross, Z. E., Y. Yue, M.-A. Meier, E. Hauksson, and T. H. Heaton (2019) PhaseLink: A deep learning approach to seismic phase association, *J. Geophys. Res.* **124,** doi: 10.1029/2018JB016674.

Trugman, D. T., and P. M. Shearer (2018). Strong correlation between stress drop and peak ground acceleration for recent M 1-4 earthquakes in the San Francisco Bay area, *Bull. Seismol. Soc. Am.* **108,** no. 2, 929–945, doi: 10.1785/0120170245.

Wang, K., W. L. Ellsworth, G. C. Beroza, G. Williams, M. Zhang, D. Schroeder, and J. Rubinstein (2019). Seismology with dark data: image-based processing of analog records using machine learning for the Rangely earthquake control experiment, *Seismol. Res. Lett.* doi: 10.1785/0220180298.

Woollam, J., A. Rietbrock, A. Bueno, and S. D. Angelis (2019). Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network, *Seismol. Res. Lett.* doi: 10.1785/0220180312.

Yuan, B., Y. J. Tan, M. K. Mudunuru, O. E. Marcillo, A. Delorey, P. M. Roberts, J. D. Webstre, C. Gammans, S. Karra, G. Guthrie, and P. A. Johnson (2019). Using machine learning to discern eruption in noisy environments: A case study using CO_2-driven cold-water geyser in Chimayó, New Mexico, *Seismol. Res. Lett.* doi: 10.1785/0220180306.

Zhang, C., M. van der Baan, and T. Chen (2019). Unsupervised dictionary learning for signal-to-noise ratio enhancement of array data, *Seismol. Res. Lett.* doi: 10.1785/0220180302.

Zhou, Z., Y. Lin, Z. Zhang, Y. Wu, and P. Johnson (2019). Earthquake detection in 1-D time series data with feature selection and dictionary learning, *Seismol. Res. Lett.* doi: 10.1785/022018315.

Zhu, W., and G. C. Beroza (2018). PhaseNet: A deep-neural-network-based seismic arrival time picking method, *Geophys. J. Int.* **216,** no. 1, 261–273, doi: 10.1093/gji/ggy423.

Zhu, W., M. S. Mousavi, and G. C. Beroza (2018). Seismic signal denoising and decomposition using deep neural networks, available at https://arxiv.org/abs/1811.02695 (last accessed January 2019).

Yoon, C. E., O. O'Reilly, K. J. Bergen, and G. C. Beroza (2015). Earthquake detection through computationally efficient similarity search, *Sci. Adv.* **1,** no. 11, e1501057, doi: 10.1126/sciadv.1501057.

Karianne J. Bergen
Department of Earth and Planetary Sciences
Harvard University
Cambridge, Massachusetts 02138 U.S.A.
karianne_bergen@fas.harvard.edu

Ting Chen
Los Alamos National Laboratory
Los Alamos, New Mexico 87545 U.S.A.

Zefeng Li
Seismological Laboratory
California Institute of Technology
Pasadena, California 91125 U.S.A.

Published Online 13 February 2019