Introduction to this special section: Machine learning applicat

Simon Shaw[1], Arvind Sharma[2], Richard Baraniuk[3], and Baishali Roy[1]

https://doi.org/10.1190/tle38070510.1.

Since the last time *The Leading Edge* devoted a special section to "Data analytics and machine learning" in March 2017, the topic has continued to generate excitement in the field of exploration geophysics and beyond. Today, machine learning is finding new applications within a multitude of industries on an almost daily basis as companies invest heavily in this technology.

The prominence of machine learning stems principally from its ability to solve complex problems using only data. A machine can be trained with just input data and a choice, or description, of an output that we want to predict. The popularity of machine learning is also explained by the abundance of open-source libraries, a world that can seem awash with data, and ready access to online storage and computing power. There are also ubiquitous free online tutorials and training to help a novice data scientist get started.

Geoscientists are well positioned to become early adopters of the latest tools in data science. Most have enough background in linear algebra and statistics to understand the underlying mathematical concepts, and only rudimentary programming skills are required to begin making predictions with an existing model or to build a custom neural network. However, as networks become increasingly complicated with more and more layers, an intuitive understanding of what the network is doing, or how the model will behave when presented with new input data, can go beyond the understanding of even experts in the field of deep learning.

Just because one *can* generate a predictive model from an input data set, doesn't necessarily mean one *should*. It is widely understood that the successful application of machine learning requires large and diverse data sets with which to train the model. The terms "large" and "diverse" are generally poorly defined for anything other than relatively simple, well-structured data sets. Even with excellent training data, geoscientists are understandably reluctant to completely abandon physical principles and other laws of nature that have served them well for decades. However, machine learning need not replace traditional physics-based methods but instead can augment them to improve efficiency or effectiveness.

In this special section's first article, Russell presents a numerical study comparing deconvolution of seismic data with a machine learning solution, lifting the lid on how a neural network solves this classic geophysical problem. He concludes that when the underlying physics is well understood then conventional inverse methods generally lead to a better solution, but when the physical model does not adequately describe the real world, or the inverse problem is nonlinear, then machine learning could succeed where traditional methods fail.

Jin et al. utilize supervised machine learning to automatically detect fracture-hit events recorded by distributed acoustic sensing (DAS) in a monitor well from an unconventional reservoir. Rather than just use the raw or, in this case, low-frequency DAS data, the authors employ their domain knowledge to compute features that highlight the fracture hits and that score well on the test data. They show that, even for a relatively small data set, a model can be successfully trained to detect the events, a process that is time-consuming and highly subjective when done manually.

Zheng et al. present two examples of supervised deep learning and in both cases train their convolutional neural network (CNN) using only synthetic seismic data. Training with synthetics has several advantages including avoiding the time and expense of acquiring appropriate, labeled field data for training the network. In the first case study, the authors extend the CNN FaultNet architecture to automatically pick faults on seismic volumes with impressive results on field data. In the second case study, they attempt prestack seismic inversion of V_P, V_S, and density from a field data set using a model that is trained on 1D elastic prestack synthetics. They find that their results are sensitive to the parameterization of the wavelet used in training, but that the CNN can predict even the low-frequency information that is missing in field data.

Peters et al. compare and contrast the application of deep learning with traditional forward and inverse problems in geophysics. The two approaches share some numerical methods, objective functions, and regularization terms. The authors find that similar deep learning network architectures can be applied to different tasks in seismic interpretation, namely automatic horizon picking and interpolation of lithology between wells using seismic data.

Lastly, Li et al. present a new framework for seismic noise attenuation that combines essentially any method of separating signal from noise with an unsupervised dictionary learning and sparse inversion that predicts signal and noise simultaneously. The authors show how the scheme improves the noise attenuation results on two land seismic data sets contaminated with strong noise.

Those familiar with the Gartner Hype Cycle (https://www.gartner.com/en/research/methodologies/gartner-hype-cycle) may have an opinion on where machine learning plots on the Gartner curve. Depending on whom you ask and what the application is, we are somewhere between the "Peak of inflated expectations" and the "Plateau of productivity." As the articles in this special section show, a number of applications in geophysics are demonstrating real value by increasing efficiencies and improving the quality of predictions.

The level of activity in this field shows no sign of abating with another strong showing expected at the upcoming SEG Annual Meeting in San Antonio this September and a new SEG Advanced Modeling Corporation project taking shape that will focus on establishing data standards and examples to test machine learning applications. TLE

---

[1]ConocoPhillips, Houston, Texas, USA. E-mail: simon.shaw@conocophillips.com; baishali.roy@conocophillips.com.
[2]TGS, Houston, Texas, USA. E-mail: arvind.sharma@tgs.com.
[3]Rice University, Houston, Texas, USA. E-mail: richb@rice.edu.