Machines that make sense of the sky

This past April, astrophysicist Kevin Schawinski posted fuzzy pictures of four galaxies on Twitter, along with a request: Could fellow astronomers help him classify them? Colleagues chimed in to say the images looked like ellipticals and spirals—familiar species of galaxies.

Some astronomers, suspecting trickery from the computation-minded Schawinski, asked outright: Were these real galaxies? Or were they simulations, with the relevant physics modeled on a computer? In truth, they were neither, he says. At ETH Zurich in Switzerland, Schawinski, computer scientist Ce Zhang, and other collaborators had cooked the galaxies up inside a neural network that doesn’t know anything about physics. It just seems to understand, on a deep level, how galaxies should look.

With his Twitter post, Schawinski just wanted to see how convincing the network’s creations were. But his larger goal was to create something like the technology in movies that magically sharpens fuzzy surveillance images: a network that could make a blurry galaxy image look like it was taken by a better telescope than it actually was. That could let astronomers squeeze out finer details from reams of observations. “Hundreds of millions or maybe billions of dollars have been spent on sky surveys,” Schawinski says. “With this technology, we can immediately extract somewhat more information.”

The forgery Schawinski posted on Twitter was the work of a generative adversarial network, a kind of machine-learning model that pits two dueling neural networks against each other. One is a generator that concocts images, the other a discriminator that tries to spot any flaws that would give away the manipulation, forcing the generator to get better. Schawinski’s team took thousands of real images of galaxies, and then artificially degraded them. Then the researchers taught the generator to spruce up the images again so they could slip past the discriminator. Eventually the network could outperform other techniques for smoothing out noisy pictures of galaxies.

Schawinski’s approach is a particularly avant-garde example of machine learning in astronomy, says astrophysicist Brian Nord of Fermi National Accelerator Laboratory in Batavia, Illinois, but it’s far from the only one. At the January meeting of the American Astronomical Society, Nord presented a machine-learning strategy to hunt down strong gravitational lenses: rare arcs of light in the sky that form when the images of distant galaxies travel through warped spacetime on the way to Earth. These lenses can be used to gauge distances across the universe and find unseen concentrations of mass.

Strong gravitational lenses are visually distinctive but difficult to describe with simple mathematical rules—hard for traditional computers to pick out, but easy for people. Nord and others realized that a neural network, trained on thousands of lenses, can gain similar intuition. In the following months, “there have been almost a dozen papers, actually, on searching for strong lenses using some kind of machine learning. It’s been a flurry,” Nord says.

And it’s just part of a growing realization across astronomy that artificial intelligence strategies offer a powerful way to find and classify interesting objects in petabytes of data. To Schawinski, “That’s one way I think in which real discovery is going to be made in this age of ‘Oh my God, we have too much data.'”

Joshua Sokol is a journalist in Boston.

Caruana’s GAMs are not as good as AIs at handling certain types of messy data, such as images or sounds, on which some neural nets thrive. But for any data that would fit in the rows and columns of a spreadsheet, such as hospital records, the model can work well. For example, Caruana returned to his original pneumonia records. Reanalyzing them with one of his GAMs, he could see why the AI would have learned the wrong lesson from the admission data. Hospitals routinely put asthmatics with pneumonia in intensive care, improving their outcomes. Seeing only their rapid improvement, the AI would have recommended the patients be sent home. (It would have made the same optimistic error for pneumonia patients who also had chest pain and heart disease.)

Caruana has started touting the GAM approach to California hospitals, including Children’s Hospital Los Angeles, where about a dozen doctors reviewed his model’s results. They spent much of that meeting discussing what it told them about pneumonia admissions, immediately understanding its decisions. “You don’t know much about health care,” one doctor said, “but your model really does.”

**SOMETIMES, YOU HAVE TO EMBRACE** the darkness. That’s the theory of researchers pursuing a third route toward interpretability. Instead of probing neural nets, or avoiding them, they say, the way to explain deep learning is simply to do more deep learning.

Like many AI coders, Mark Riedl, director of the Entertainment Intelligence Lab at the Georgia Institute of Technology in Atlanta, turns to 1980s video games to test his creations. One of his favorites is Frogger, in which the player navigates the eponymous amphibian through lanes of car traffic to an awaiting pond. Training a neural network to play expert Frogger is easy enough, but explaining what the AI is doing is even harder than usual.

Instead of probing that network, Riedl asked human subjects to play the game and to describe their tactics aloud in real time. Riedl recorded those comments alongside the frog’s context in the game’s code: “Oh, there’s a car coming for me; I need to jump forward.” Armed with those two languages—the players’ and the code—Riedl trained a second neural net to translate between the two, from code to English. He then wired that translation network into his original game-playing network, producing an overall AI that would say, as it waited in a lane, “I’m waiting for a hole to open up before I move.” The AI could even sound frustrated when pinned on the side of the screen, cursing and complaining, “Jeez, this is hard.”
AI in Action: Machines that make sense of the sky
Joshua Sokol

Science 357 (6346), 26.
DOI: 10.1126/science.357.6346.26