

Machines Learning Astronomy

The new era of artificial intelligence & Big Data is changing how we do astronomy.

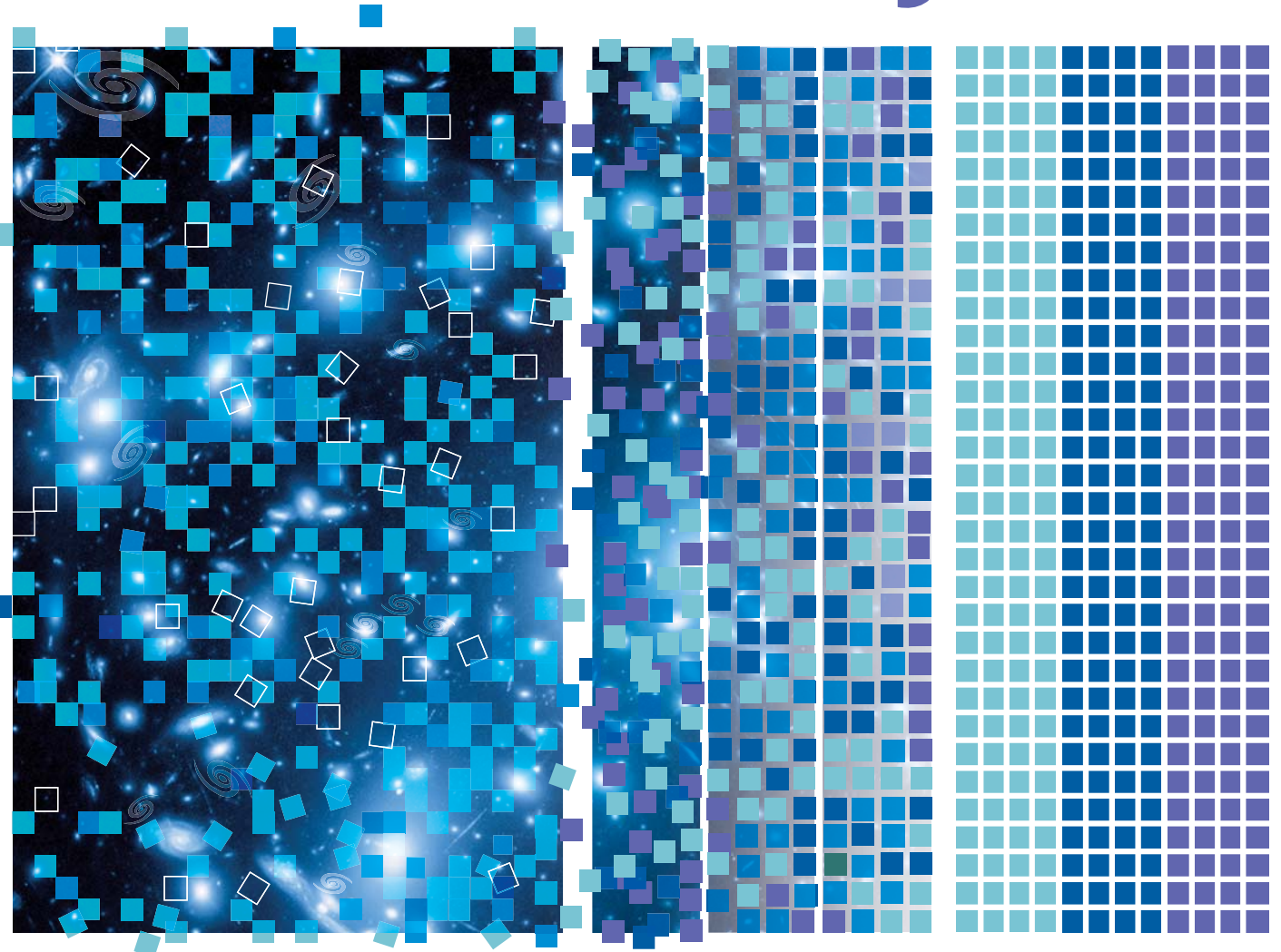


IMAGE: NASA / ESA / JENNIFER LOTZ / HFF TEAM (STSO); PAT COPPOLA / S&T

You probably use artificial intelligence a dozen times a day without realizing it. If you've recently scrolled through your Facebook feed, browsed Netflix videos, used iPhone's Siri, checked a relatively spam-free inbox, or just Googled a term, you've made use of complex computer algorithms that can learn from experience.

"Artificial intelligence," though correct, may be too fanciful a term for these processes — no computer today is capable of general intelligence or autonomy. Still, what the algorithms can do is possibly paradigm-shifting. Rather than humans programming a computer explicitly, the algorithms use data to construct their *own* mathematical models — sometimes ones too complex for human understanding. This implementation of artificial intelligence is known as *machine learning*, and it can filter through vast amounts of data.

Given the data tsunami facing astronomy (*S&T*: Sept. 2016, p. 14), it should come as no surprise that machine learning has popped up in just about every part of the celestial realm over the past few years (see box on page 28). From exoplanets and variable stars to cosmology, machine learning is going to play an ever-larger role in the coming decade of astronomical research.

Big Data, Big Opportunity

Machine learning isn't new — pioneers of the field hail back to the 1950s. But it was long discounted as impractical, requiring too much computational power to implement. Lack of data was also a problem: Like humans, machines build their internal models based on copious observations.

It was the advent of Big Data, as well as significant advances in computing, that finally enabled machine learning to take off. In one famous example, Andrew Ng (Stanford University), who led the Google Brain project, drew from 10 million YouTube videos to teach an algorithm to recognize — what else? — cats. Fortunately, the playing field isn't limited to internet whims. In astronomy, Big Data abounds.

"We are very firmly in the era of what we call 'survey astronomy,'" says Lucianne Walkowicz (Adler Planetarium).

The Sloan Digital Sky Survey (SDSS), which has imaged a third of the sky, qualified as Big Data when it began in 2000, but new and upcoming projects are dwarfing it. The standard-bearer is the Large Synoptic Survey Telescope (LSST), scheduled to begin science operations in 2022. It will generate the equivalent of the SDSS every night, monitoring 37 billion stars and galaxies in space and time to make a decade-long movie of the southern sky.

There are others as well: The Dark Energy Survey started charting hundreds of millions of galaxies in 2013; the Gaia satellite began mapping 1 billion stars in the Milky Way in 2014; and the Zwicky Transient Facility, due to see first light in late 2017, will be scanning 3,750 square degrees every hour. That's not to mention archival data from past surveys. The bytes have become numerous enough that, even with dozens of graduate students or thousands of citizen scientists, human eyes simply can't scan all the available data.



▲ **TO SEE CATS** Computer scientists at Google taught a powerful neural network to recognize, among other things, cat faces. The algorithm "learned" from 10 million 200×200-pixel images drawn from YouTube videos. When asked to render a cat, the algorithm produced this convincing image.

"People don't scale," notes Joshua Bloom (University of California, Berkeley), founder of the startup Wise.io. "And especially experts don't scale."

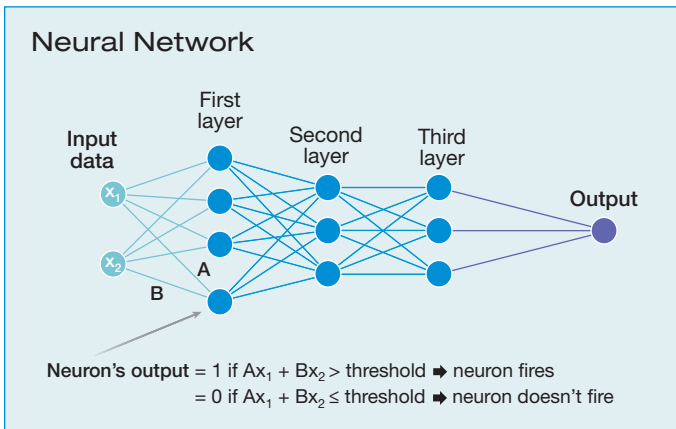
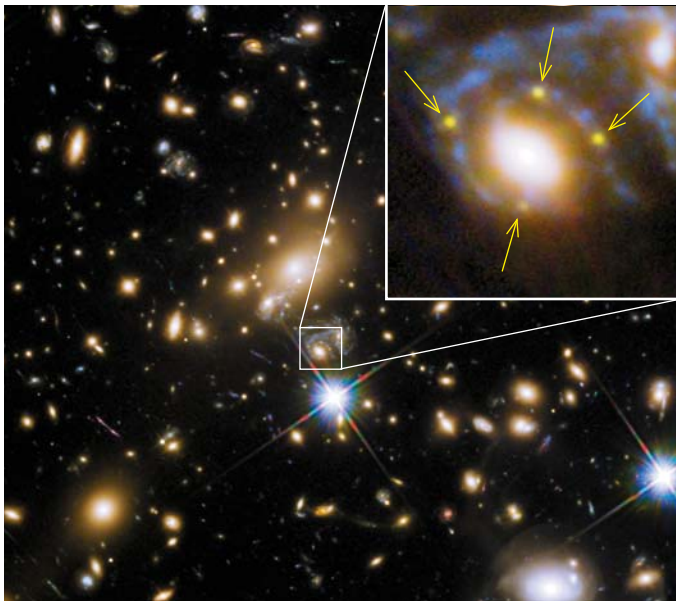
Brian Nord (Fermilab) encountered this problem of scale when he joined the Dark Energy Survey (DES) team, where 20 scientists spent months poring over 250 square degrees' worth of images. They were looking for the warped shapes that mark *strong gravitational lenses*, where clusters or massive galaxies bend the light from background objects.

About a thousand such lenses are already known, but the full, 5,000-square-degree DES offers the potential to triple that number, and LSST might reveal an order of magnitude more. A few of these lenses could magnify exceedingly rare — and cosmologically invaluable — distant supernovae. For this reason, the researchers made sure to look at *everything* in the preliminary data set, not applying any filters that might have eased the search, in order to avoid missing legitimate sources.

"It was a painful number of pixels to visually scan," Nord says. "I saw this, and I thought, 'Oh my God. This is so painful, there has got to be a better way.'"

MORE THAN JUST BIG

To qualify as "big" in the world of computer science, data need more than *volume*; they also need *variety* and *velocity*. Big Data is any huge amount of data that comes in a variety of formats (such as images, spectra, and times-series data) and must be dealt with in a timely manner.



Nord, who says he was inspired in part by Tesla's self-driving cars, set about eliminating humans from the process. He built and trained DeepLensing, a machine-learning algorithm technically known as a *convolutional neural network*, to recognize distorted galaxy images.

As the name implies, neural networks are loosely based on the myriad connections between neurons in the brain. Each neuron in the algorithm is a simple mathematical formula. Given a certain input, it performs a calculation to decide whether, in biological terms, to “fire.” During training, feedback loops within the code change a neuron’s firing capabilities based on sample data. Once training is finished, those feedback loops are turned off.

▲ **LUCKY HORSESHOE** Creating strong gravitational lenses, such as the one depicted here, requires a bit of good fortune. First, two celestial objects must line up in just the right way for one to gravitationally magnify the light coming from the other behind it. Then, actually finding a lens, especially one that appears small on the sky, takes a bit of serendipity — and a lot of work if humans are the ones looking.

◀ ▲ **FOUR-LEAF CLOVER** Astronomers observing a massive galaxy cluster as part of Hubble’s Frontier Fields project caught a supernova gravitationally lensed into four separate images (*inset*). A fifth image appeared several months later. Such celestial happenstances are incredibly rare and cosmologically valuable. The Dark Energy Survey could spot a few lensed supernovae over its five-year duration.

◀ **SKY NET** This schematic diagram portrays a simple neural network that starts with data, such as an image, and ends with an outcome, such as a classification. In between lie layers of neurons. The data enter every neuron in the first layer, and each neuron performs a simple calculation, weighting the data points (A and B are the weights for x_1 and x_2 , respectively) to determine whether it should fire. Each neuron’s decision then feeds into the next layer of neurons. The last layer of neurons then converges to produce an answer, such as “lens” or “not lens.”

HORSESHOE GALAXY: ESA / HUBBLE & NASA; LENSED SUPERNOVA: NASA / ESA / STSCI / UCLA; NEURAL NET: LEAH TISCIONE / 587

DeepLensing contains three sets of neural layers, where the output of each neuron in one layer acts as input to the neurons in the next. The layers act as filters, picking out features in the input image. For example, the first layer may separate dark regions from light ones; the next may highlight edges; the one after that may recognize shapes. Once trained, the last layer gives the final decision: lens or not lens.

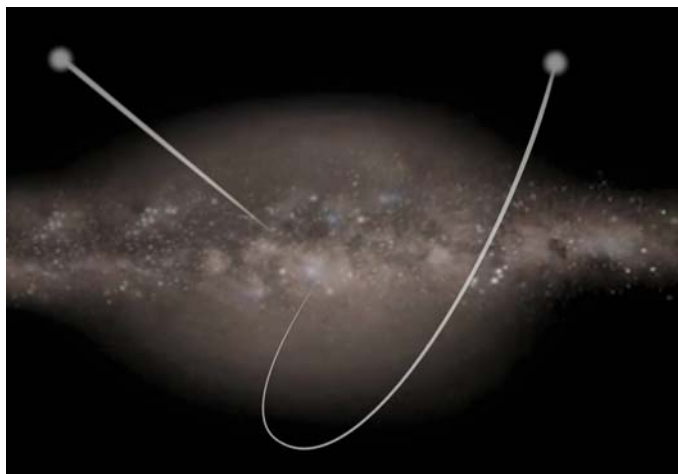
These networks are incredibly simple compared to the roughly 100 billion neurons in the human brain. In fact, it takes only a couple dozen lines of code to construct a basic neural network. Yet, mathematically, the result is linear algebra on a massive, sometimes incomprehensible scale.

Though DeepLensing is still a work in progress, it can already accomplish what other methods could not: quickly filtering thousands of input images while maintaining 90% accuracy in what it identifies as a lens in simulations. And DeepLensing isn't alone — similar, independently developed neural nets have since found and analyzed strong gravitational lenses millions of times faster than humans can.

Needles in the Haystack

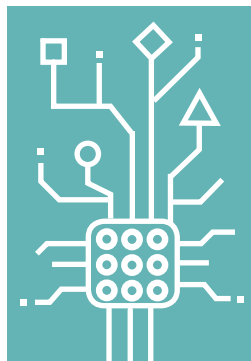
Machine learning is well suited to “needle-in-a-haystack”-type searches, so it was the tool that Elena Rossi (Leiden Observatory, The Netherlands) turned to when she wanted to study exceedingly rare *hypervelocity stars*.

These stars speed away from the Milky Way's center, probably ejected via a gravitational slingshot interaction with our galaxy's supermassive black hole. Some of them are on a course to escape the galaxy entirely. So far astronomers have found only about 20 hypervelocity stars, but amidst the 1 billion stars that Gaia is currently monitoring, Rossi expects to find at least a hundred more. Her research has shown that that would be enough to start using their trajectories to probe the shape of the dark matter cloud surrounding the Milky Way (*S&T*: Apr. 2017, p. 22).



▲ **RUNAWAY STARS** The conditions that allow a star to escape our galaxy, as pictured here in this artist's illustration, are exceedingly rare. Such *hypervelocity stars* are consequently difficult to find unless astronomers employ innovative methods.

TEST & VALIDATE



Machine-learning algorithms train on data, usually referred to as a *training set*. Once an algorithm has finished training, its accuracy is probed via a separate *validation set*. Both sets may consist of real data or, if large amounts of real data aren't available yet, simulated sources. Once the algorithm performs well on both the training and validation sets, it's ready for action.

But to map out a hypervelocity star's orbit within the dark matter halo, she first needs to know that star's motion through space. Gaia doesn't provide full, three-dimensional velocity information for every star it monitors — it only measures the motion toward or away from us, called *radial velocity*, of the very brightest stars. So not only did Rossi need to identify rare stars outnumbered a million to 1 by normal stars, she needed to do so with incomplete data.

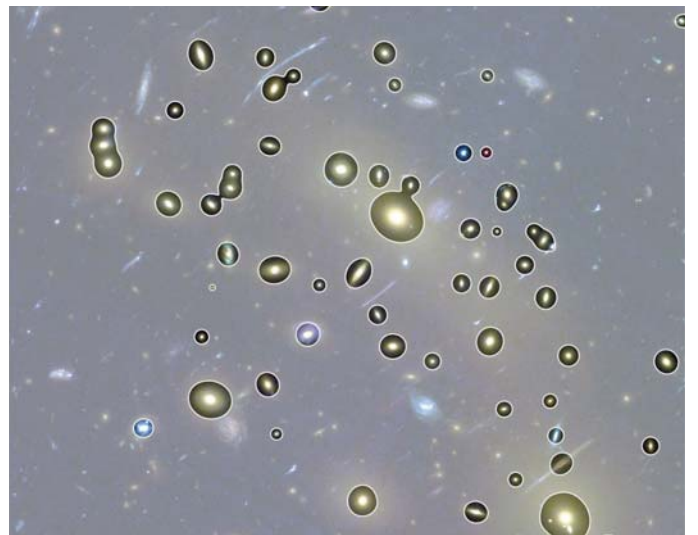
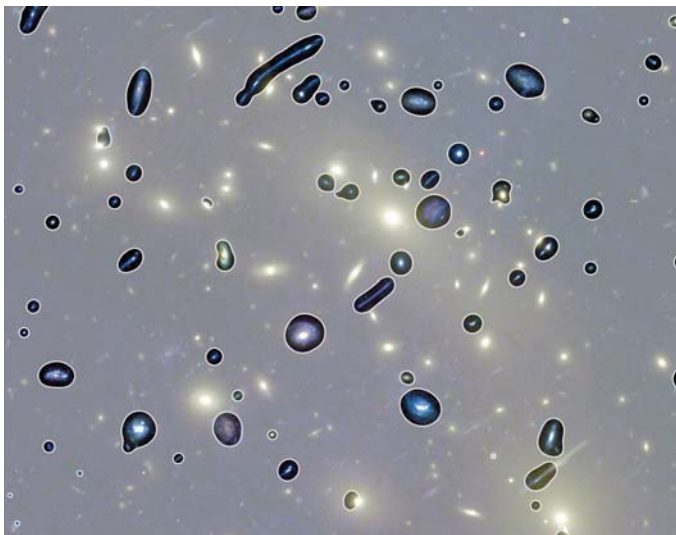
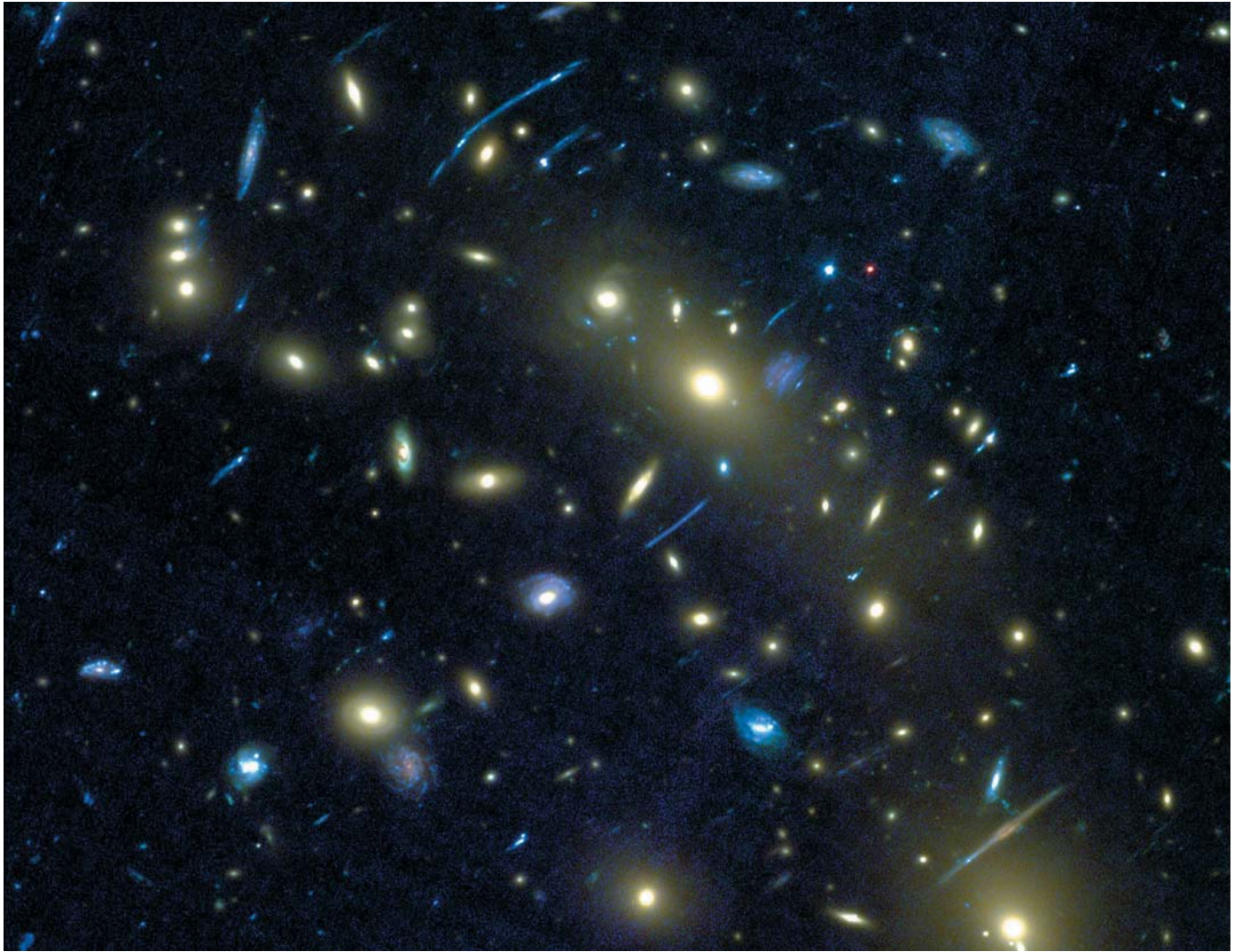
Machines are becoming capable of “systematizing serendipity.”

After some initial investigation, Rossi settled on a neural network of two layers with 119 and 95 neurons, respectively. The many neurons ensure that each layer is sufficiently complex to extract features from the data. Meanwhile, additional layers further filter and abstract the data, creating an increasingly flexible neural network. However, if the algorithm becomes too complex or too flexible, it begins to overfit: It tries so hard to learn from the training set that it becomes rigid, unable to generalize the lesson to other data.

Finding the right balance is often a matter of trial and error, and Rossi and her graduate student Tommaso Marchetti put the algorithm through hoops to learn how many layers and neurons would work best. On a fundamental level, though, it isn't clear why some architectures do better than other ones.

“We're still trying to understand our tool,” Rossi explains. “When I saw that our algorithm was giving us what we were looking for, at least in the right direction, I was very happy — but also surprised!”

The algorithm trawled through the first release of Gaia data, all 1 billion stars, to find 80 candidate hypervelocity stars — a reassuringly low number given the objects' rarity. Of these, 30 stars already had known radial velocities, and Rossi targeted 22 additional stars for follow-up observations. Ultimately, the team found six hypervelocity stars, a nice catch for the first go-round.



▲ **COMPUTING THE GALAXY ZOO** While citizen-science projects such as Zooniverse currently enable the classification of galaxies and other objects en masse, future Big Data surveys such as LSST will provide too much data for such methods to handle. Based on a Hubble Space Telescope image of the galaxy cluster MACS0416.1–2403 (*top*), Alex Hocking (University of Hertfordshire, UK) and colleagues taught a multi-part machine-learning algorithm to automatically recognize star-forming galaxies, including lensed ones, (*bottom left*) and ellipticals (*bottom right*).

The algorithm turned up another surprise: five runaway stars *not* coming from the Milky Way's center, each traveling between 400 and 780 kilometers per second (900,000 to 1.7 million mph). These stars may once have been part of binary systems in the Milky Way's disk that were ejected when their stellar partners went supernova. But such explosions don't typically eject stars at speeds this high. "Our algorithm picked up a very special case of this mechanism," Rossi says.

Gaia's next data release, which will help validate Rossi's finds, will come in April 2018.

Astronomers have had success honing machine learning to build samples of known, rare objects. But self-taught algorithms can do more than that — they can also discover entirely new types of celestial gems.

Detecting the unexpected comes second nature to humans, who excel at pattern recognition and can therefore easily pick out rare and unusual objects. Citizen science has reams of examples: green pea galaxies, Hanny's Voorwerp, and Tabby's Star (*S&T*: June 2017, p. 16), to name a few.

Now machines are becoming capable of, as Walkowicz puts it, "systematizing serendipity." Walkowicz is working with graduate student Daniel Giles (Illinois Institute of Technology) to train an algorithm that separates Kepler-observed stars into groups and ranks them by "weirdness." Using Tabby's Star as a test subject, Walkowicz and Giles are creating the tool to pick out Tabby's Star analogs in Kepler data and, eventually, in other surveys such as LSST.

Making Connections

Some are taking these programs even further — rather than finding needles in a haystack for future study, astronomers can apply machine learning to inspect all of the hay. Self-taught algorithms can make unforeseen connections between

Rather than finding needles in a haystack, astronomers are using machine learning to study all of the hay.

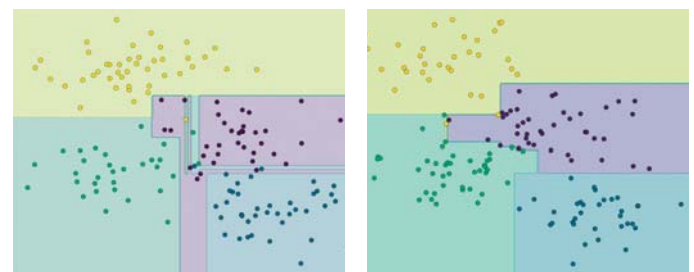
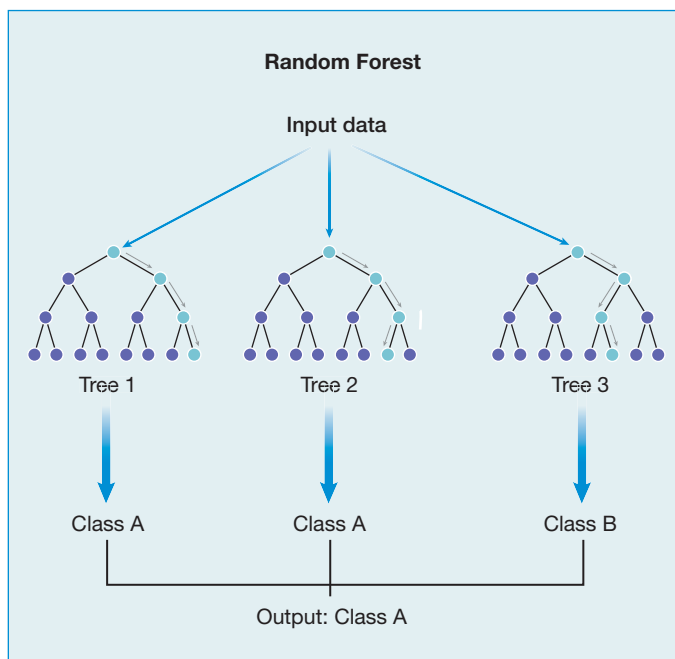
features in the data, enabling computers to classify and characterize objects en masse.

That ability may help solve one of the biggest problems facing the LSST. When the telescope comes online early next decade, it'll produce 15 terabytes' worth of brightness measurements every night, but it'll be missing something crucial: spectra. Spectral lines from the heavy elements that lace a star's gas reveal its physical properties, such as its surface temperature and gravity. However, follow-up spectroscopy will only be feasible for 0.1% of LSST-observed stars.

Nevertheless, astronomers can learn a lot about a star by its color, as well as by its light curve, which tracks the change in brightness over time. In 2015 Adam Miller, then a graduate student at University of California, Berkeley, and Joshua Bloom, his advisor, realized that machine learning could connect brightness measurements of variable stars to the physical properties normally gleaned from their spectra.

They conducted a proof of concept using a collection of decision trees, collectively known as a *random forest*. Each tree asks a series of questions to separate the variable stars into groups. The questions aren't programmed in; the trees decide the questions themselves based on the data they train on.

In this test case, the training set consisted of 9,000 variable stars observed in the Stripe 82 survey, a 315-square-degree field repeatedly imaged by the SDSS project. Follow-up spectroscopy came from the 6.5-meter Multiple Mirror Telescope in Arizona.



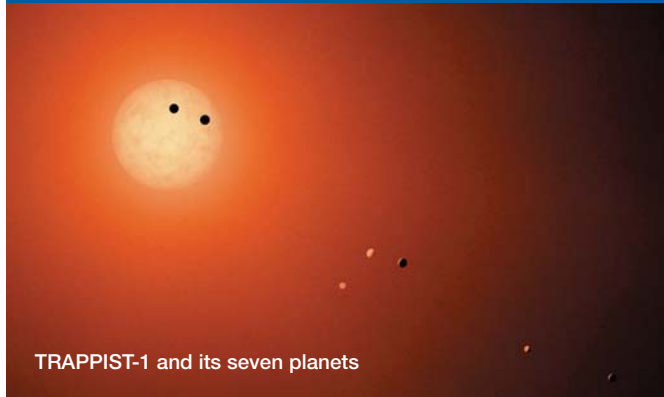
▲ **FROM TREES TO FORESTS** Random forest algorithms, such as the one shown in this simple schematic, are a collection of decision trees. Each tree is shaped slightly differently from its neighbors, asking different questions of the data and separating data points in different ways. The outputs from all the trees are averaged before providing an answer. As in neural networks, humans don't program the decision points — the trees determine what questions to ask from the data itself.

▲ **LOST IN THE WOODS** A decision tree asks questions to separate data — the more questions it asks, the more it divides the data (*left*). But a single series of decisions may miss the forest for the trees, carving up the training data so much that the algorithm is no longer useful for classifying new data sets. By averaging an ensemble of decision trees (*right*), a random forest algorithm reaches more robust conclusions.

Based solely on brightness measurements, the decision trees predicted the surface temperature, surface gravity, and abundance of heavy metals for 54,000 variable stars, achieving the same precision as low-resolution spectroscopy.

The result, which Bloom calls “sort of a weird head-scratcher,” is that machine learning could transform the LSST from an instrument that measures how variable stars change over time into a sort of spectrograph, measuring the stars’ spectral — and physical — properties.

AI-Aided Discoveries



TRAPPIST-1 and its seven planets

- ▶ Machine-learning algorithms separate true supernovae from bogus detections in the live data stream from the All-Sky Automated Survey for Supernovae (ASAS-SN). They are responsible, for example, for the detection of the most luminous supernova to date, ASASSN-15lh, which shines at 570 billion times the Sun’s luminosity (*S&T*: Nov. 2015, p. 12).
- ▶ Machine learning helped confirm the seventh planet orbiting the cool dwarf star TRAPPIST-1. A single transit had hinted at the world’s existence, but 70 days of additional Kepler observations and machine-learning analysis were crucial to verify the planet’s signal (*S&T*: Sept. 2017, p. 11).
- ▶ A machine-learning algorithm plucked 60 non-transiting hot Jupiter candidates from Kepler data by their reflected starlight. The find was unusual, since most Kepler exoplanets reveal themselves when they pass in front of their star. These candidates now await follow-up observations to confirm their hot Jupiter status.
- ▶ Astronomers have built the V-FASTR classifier for the Very Large Baseline Array to distinguish new radio-wave discoveries, such as fast radio bursts, from known pulsars and human-created radio interference with more than 98% accuracy.

“It’s like sitting in a room and hearing someone on the other side of the room singing,” Bloom says, “and you can tell how old they are, what their gender is, and what color their hair is from how they sing.”

The Black Box Problem

Despite its incredible potential, machine learning is just starting to take off in astronomy, and some of the delay is caused by sheer hesitation. “The generic problem with machine learning is that you always get an answer,” Bloom cautions. “And that’s really dangerous.”

Because machine learning can make connections and recognize patterns better than humans do, using these algorithms carries a significant risk: The answer an astronomer receives may be an answer — maybe even a wrong one — that the astronomer can’t understand.

Ashley Villar (Harvard University) ran into that complication when she was building what she calls a “home-brewed” neural network to better understand Type Ia supernovae. These are the flashes from white dwarfs that have reached their mass limit. All explode in a similar way, so they’re best known for their use as standard candles in cosmology. But the bursts are not identical, and understanding their differences can improve measurements of the expanding universe.

Villar studies how different amounts of heavy elements, or *metallicities*, might alter the detonations. The conditions in which those heavy elements exist are so extreme that they can’t be reproduced in the lab. So Villar is building a tiny neural network — two layers with six neurons total — to relate the heavy elements present in the star’s host galaxy (which gives an idea of the metallicity of the star itself) to the explosion’s spectrum.

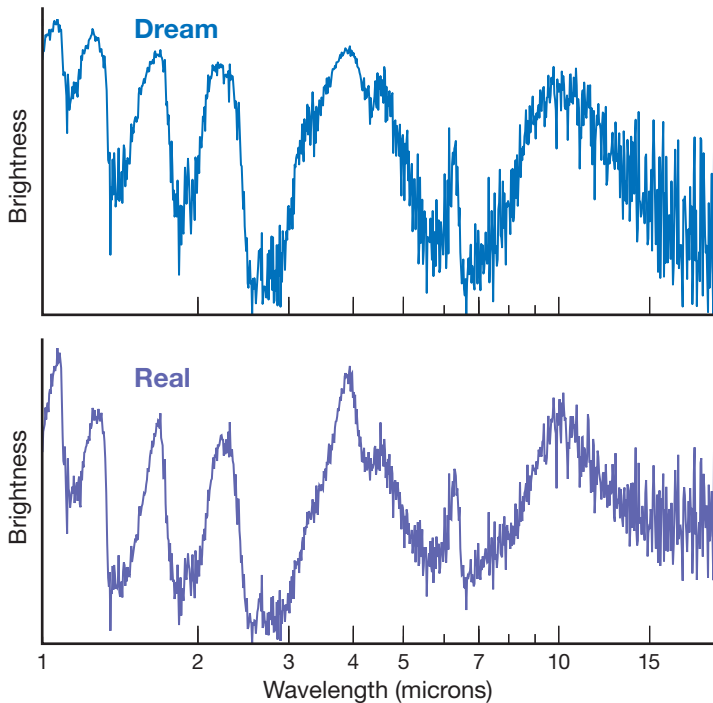
Villar trained the algorithm and it began producing output: When she fed in a Type Ia spectrum, it reported the progenitor’s metallicity. But how was it making the decision — and was it always the right one?

Answering that question is one of the biggest challenges facing machine learning today. “It’s being talked about,” Villar says, “but it’s hard to quantify. And in astronomy, it hasn’t been explored.”

One approach, which Villar has taken, is to black out parts of the spectrum and see how the calculation of the star’s metallicity reacts — basically, how wrong does the algorithm get when it’s missing bits of data? The more wrong it becomes, Villar figures, the more important that section of the spectrum was in determining the answer.

There’s another approach to explaining how an algorithm does what it does: letting it dream. It’s actually less fantastical than it sounds. As Ingo Waldmann (University College London) puts it, “Dreaming is just working backward.”

Waldmann had taught an algorithm called Robotic Exoplanet Recognition (ROBERT) to recognize molecules in an exoplanet’s transmission spectrum, measured from the light that passes through the sliver of atmosphere visible whenever a planet passes in front of its star. Here, it’s not the data that



▲ **DREAMING OF WATER** To see if ROBERT, the Robotic Exoplanet Recognition algorithm, had learned to spot the signal that water in an exoplanet atmosphere would imprint on its spectrum, Ingo Waldmann let ROBERT “dream.” When fed the label “water”, the algorithm came up with a depiction of a water spectrum that mimicked a real spectrum.

are complex, it’s the models: A computer might peruse half a terabyte’s worth of theoretical atmospheric models as it tries to match the patterns in a simple exoplanet spectrum, often taking days to reach a decision. In the coming age of dedicated exoplanet missions — including NASA’s Transiting Exoplanet Survey Satellite (TESS) and the European Space Agency’s Atmospheric Remote-sensing Exoplanet Large-survey (ARIEL) mission — that’s not good enough.

So Waldmann built a fast, three-layer neural network to recognize the imprints molecules leave in their exoplanet’s spectrum. Instead of flipping through hundreds of temperature profiles, molecular spectral lines, and cloud or haze possibilities, ROBERT simply learned the pattern that water takes in an exoplanet spectrum.

To test how the algorithm was learning that connection, Waldmann turned the algorithm around. Rather than feeding ROBERT a spectrum, he simply told it “water,” then let it produce its own idea of what an exoplanet spectrum with water would look like.

“When I first built ROBERT, it was too complex,” Waldmann says. “When I made it ‘dream,’ it had a really noisy spectrum. And then I realized . . . it was basically bored.” There were so many neurons that many of them weren’t

activating — they were just sitting there, producing noise. When Waldmann reduced the number of layers and neurons, the algorithm’s dreams crystallized, bringing forth a realistic portrayal of water’s spectral lines. The dreams showed that ROBERT now “understood” molecular patterns.

Even so, there’s a larger question: Does ROBERT also understand the underlying physics associated with those patterns? For example, in the process of learning the spectral lines created by water, did ROBERT also learn the temperature profiles of potentially water-carrying atmospheres?

“I think it should. There’s no reason why it shouldn’t,” Waldmann speculates. But the point is that he isn’t sure. “This is the problem with neural networks — you don’t know what they know.”

The trickiness of building and understanding a machine-learning algorithm and the great potential worth of its output are reflected in the response this kind of research receives. When Villar presented her supernova research at a meeting of the American Astronomical Society, she recalls, “Some people were really excited about this. They think it’s the end-all-be-all, it’s going to solve everything. And there are definitely people who completely reject it, they think it’s terrible.”

Even Bloom, perhaps one of astronomy’s biggest machine-learning proponents, says, “It is a whole pit of pitfalls.” In fact, he adds, “I give talks to lots of different groups, and the first and last thing I say is, ‘Don’t use machine learning unless you have to.’” Even so, many astronomers predict that machine learning will take on an important role in the field, perhaps becoming as essential as the telescope.

We might build algorithms that give us an answer, but it’s one that we aren’t capable of understanding.

In the coming decade, machine learning will no doubt replace or supersede some traditional analysis techniques. But it’s possible it could go a step further. What if the natural world is described by laws that are so complex that only machine-learning algorithms can describe the observations that future surveys obtain? We might build algorithms that give us an answer, but it’s one that we aren’t capable of understanding. “That’s kind of a crazy thought,” Bloom muses. But it’s the kind of thought that comes up when working with a set of tools that’s only beginning to be explored.

“From my perspective,” Bloom says, “it’s a little bit like being a kid in a candy store — before all the kids wake up.”

■ **MONICA YOUNG**, *Sky & Telescope*’s News Editor, looks forward to peering inside the black box of artificial intelligence.

• Try your own hand at machine learning — it’s not as hard as it might seem! To start experimenting with interactive tools for various skill levels, visit <https://is.gd/machinelearningastro>.