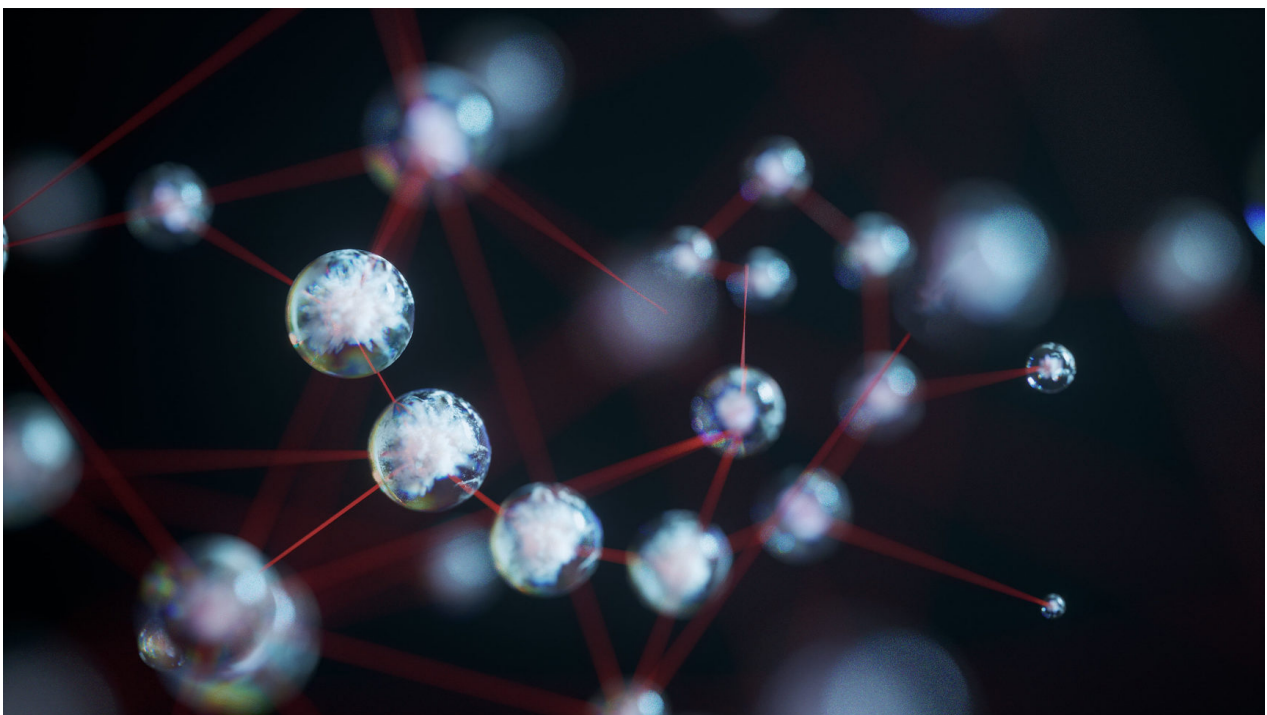




Job One for Quantum Computers: Boost Artificial Intelligence

The fusion of quantum computing and machine learning has become a booming research area. Can it possibly live up to its high expectations?

By George Musser



[Josef Bsharah](#) for Quanta Magazine

In the early '90s, [Elizabeth Behrman](#), a physics professor at Wichita State University, began working to combine quantum physics with artificial intelligence — in particular, the then-maverick technology of neural networks. Most people thought she was mixing oil and water. “I had a heck of a time getting published,” she recalled. “The neural-network journals would say, ‘What is this quantum mechanics?’ and the physics journals would say, ‘What is this neural-network garbage?’”

Today the mashup of the two seems the most natural thing in the world. Neural networks and other

machine-learning systems have become the most disruptive technology of the 21st century. They out-human humans, beating us not just at tasks most of us were never really good at, such as chess and data-mining, but also at the very types of things our brains evolved for, such as recognizing faces, translating languages and negotiating four-way stops. These systems have been made possible by vast computing power, so it was inevitable that tech companies would seek out computers that were not just bigger, but a new class of machine altogether.

Quantum computers, after decades of research, have nearly enough oomph to perform calculations beyond any other computer on Earth. Their killer app is usually said to be factoring large numbers, which are the key to modern encryption. That's still another decade off, at least. But even today's rudimentary quantum processors are uncannily matched to the needs of machine learning. They manipulate vast arrays of data in a single step, pick out subtle patterns that classical computers are blind to, and don't choke on incomplete or uncertain data. "There is a natural combination between the intrinsic statistical nature of quantum computing ... and machine learning," said [Johannes Otterbach](#), a physicist at Rigetti Computing, a quantum-computer company in Berkeley, California.

If anything, the pendulum has now swung to the other extreme. Google, Microsoft, IBM and other tech giants are pouring money into quantum machine learning, and a startup incubator at the University of Toronto is devoted to it. "'Machine learning' is becoming a buzzword," said [Jacob Biamonte](#), a quantum physicist at the Skolkovo Institute of Science and Technology in Moscow. "When you mix that with 'quantum,' it becomes a mega-buzzword."

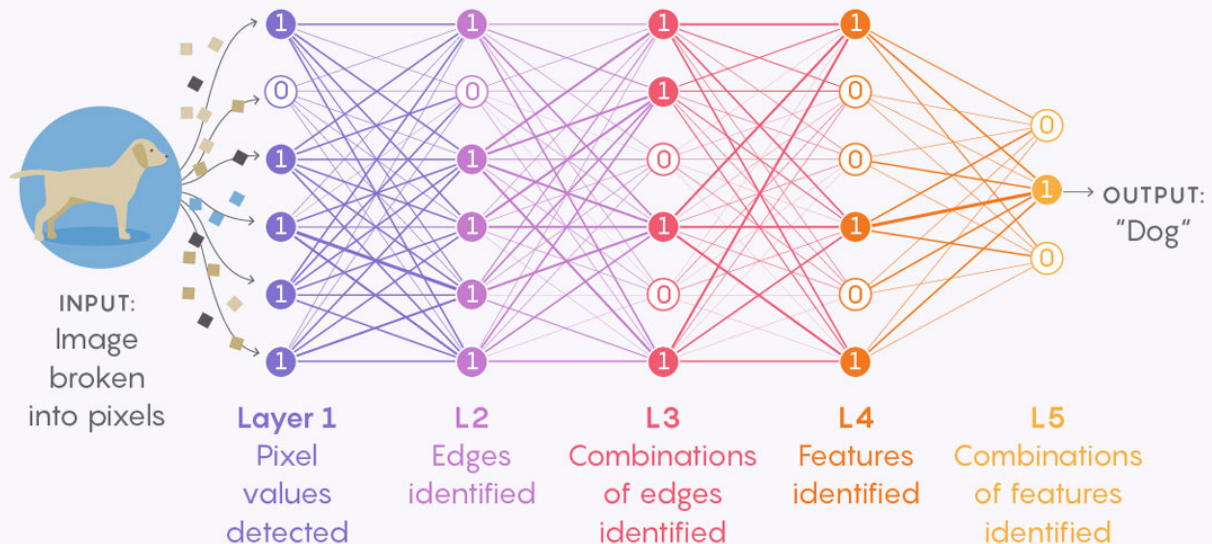
Yet nothing with the word "quantum" in it is ever quite what it seems. Although you might think a quantum machine-learning system should be powerful, it suffers from a kind of locked-in syndrome. It operates on quantum states, not on human-readable data, and translating between the two can negate its apparent advantages. It's like an iPhone X that, for all its impressive specs, ends up being just as slow as your old phone, because your network is as awful as ever. For a few special cases, physicists can overcome this input-output bottleneck, but whether those cases arise in practical machine-learning tasks is still unknown. "We don't have clear answers yet," said [Scott Aaronson](#), a computer scientist at the University of Texas, Austin, who is always the voice of sobriety when it comes to quantum computing. "People have often been very cavalier about whether these algorithms give a speedup."

Quantum Neurons

The main job of a neural network, be it classical or quantum, is to recognize patterns. Inspired by the human brain, it is a grid of basic computing units — the "neurons." Each can be as simple as an on-off device. A neuron monitors the output of multiple other neurons, as if taking a vote, and switches on if enough of them are on. Typically, the neurons are arranged in layers. An initial layer accepts input (such as image pixels), intermediate layers create various combinations of the input (representing structures such as edges and geometric shapes) and a final layer produces output (a high-level description of the image content).

Learning From Experience

Deep neural networks learn by adjusting the strengths of their connections to better convey input signals through multiple layers to neurons associated with the right general concepts.



When data is fed into a network, each artificial neuron that fires (labeled "1") transmits signals to certain neurons in the next layer, which are likely to fire if multiple signals are received. The process filters out noise and retains only the most relevant features.

Lucy Reading-Ikkanda/Quanta Magazine

Crucially, the wiring is not fixed in advance, but adapts in a process of trial and error. The network might be fed images labeled "kitten" or "puppy." For each image, it assigns a label, checks whether it was right, and tweaks the neuronal connections if not. Its guesses are random at first, but get better; after perhaps 10,000 examples, it knows its pets. A serious neural network can have a billion interconnections, all of which need to be tuned.

On a classical computer, all these interconnections are represented by a ginormous matrix of numbers, and running the network means doing matrix algebra. Conventionally, these matrix operations are outsourced to a specialized chip such as a graphics processing unit. But nothing does matrices like a quantum computer. "Manipulation of large matrices and large vectors are exponentially faster on a quantum computer," said [Seth Lloyd](#), a physicist at the Massachusetts Institute of Technology and a quantum-computing pioneer.

For this task, quantum computers are able to take advantage of the exponential nature of a quantum system. The vast bulk of a quantum system's information storage capacity resides not in its

individual data units — its qubits, the quantum counterpart of classical computer bits — but in the collective properties of those qubits. Two qubits have four joint states: both on, both off, on/off, and off/on. Each has a certain weighting, or “amplitude,” that can represent a neuron. If you add a third qubit, you can represent eight neurons; a fourth, 16. The capacity of the machine grows exponentially. In effect, the neurons are smeared out over the entire system. When you act on a state of four qubits, you are processing 16 numbers at a stroke, whereas a classical computer would have to go through those numbers one by one.

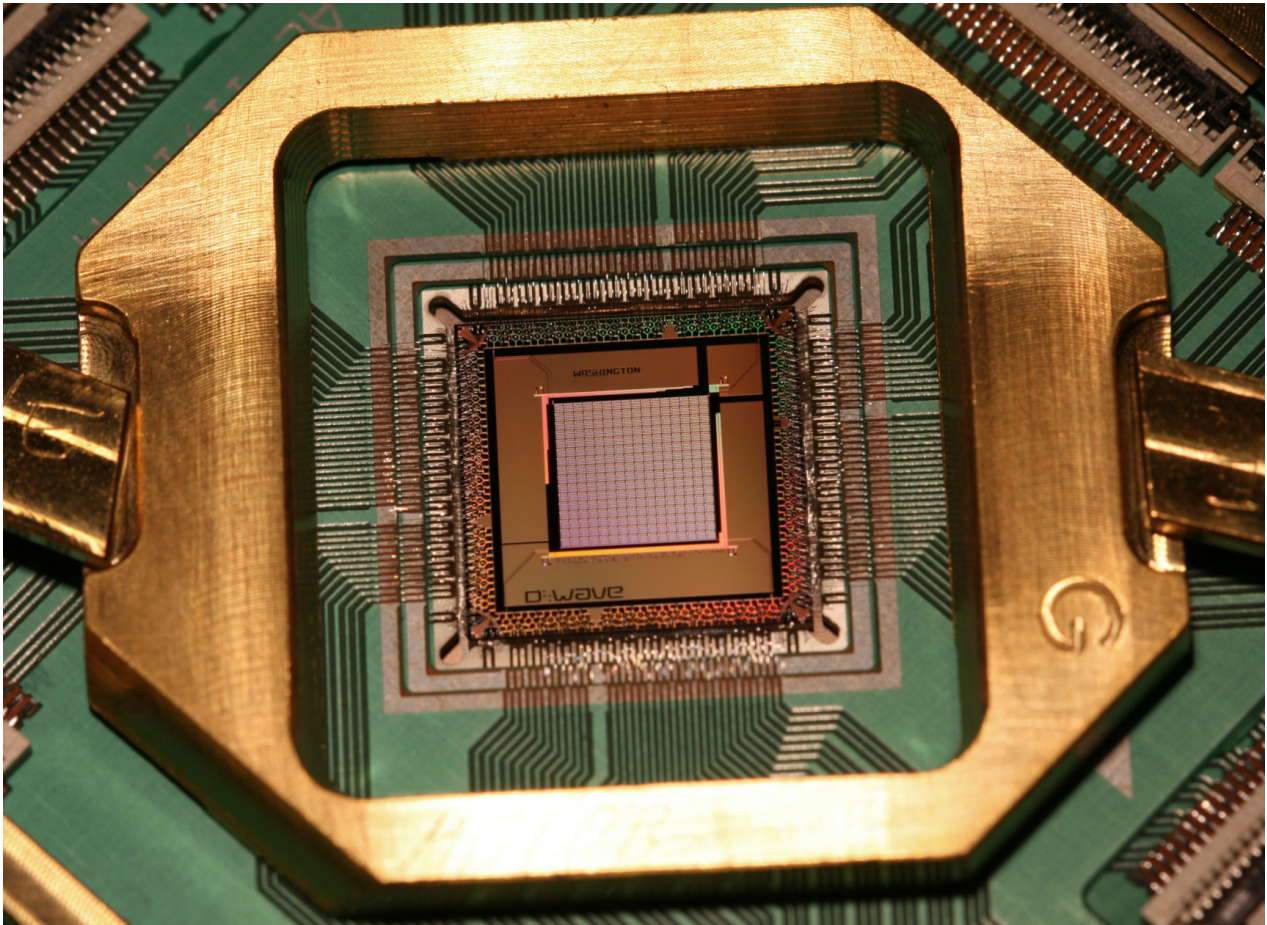
Lloyd estimates that 60 qubits would be enough to encode an amount of data equivalent to that produced by humanity in a year, and 300 could carry the classical information content of the observable universe. (The biggest quantum computers at the moment, built by IBM, Intel and Google, have 50-ish qubits.) And that’s assuming each amplitude is just a single classical bit. In fact, amplitudes are continuous quantities (and, indeed, complex numbers) and, for a plausible experimental precision, one might store as many as 15 bits, Aaronson said.

But a quantum computer’s ability to store information compactly doesn’t make it faster. You need to be able to use those qubits. In 2008, Lloyd, the physicist [Aram Harrow](#) of MIT and [Avinatan Hassidim](#), a computer scientist at Bar-Ilan University in Israel, [showed](#) how to do the crucial algebraic operation of inverting a matrix. They broke it down into a sequence of logic operations that can be executed on a quantum computer. Their algorithm works for a huge variety of machine-learning techniques. And it doesn’t require nearly as many algorithmic steps as, say, factoring a large number does. A computer could zip through a classification task before noise — the big limiting factor with today’s technology — has a chance to foul it up. “You might have a quantum advantage before you have a fully universal, fault-tolerant quantum computer,” said Kristan Temme of IBM’s Thomas J. Watson Research Center.

Let Nature Solve the Problem

So far, though, machine learning based on quantum matrix algebra has been demonstrated only on machines with just four qubits. Most of the experimental successes of quantum machine learning to date have taken a different approach, in which the quantum system does not merely simulate the network; it is the network. Each qubit stands for one neuron. Though lacking the power of exponentiation, a device like this can avail itself of other features of quantum physics.

The largest such device, with some 2,000 qubits, is the quantum processor manufactured by D-Wave Systems, based near Vancouver, British Columbia. It is not what most people think of as a computer. Instead of starting with some input data, executing a series of operations and displaying the output, it works by finding internal consistency. Each of its qubits is a superconducting electric loop that acts as a tiny electromagnet oriented up, down, or up and down — a superposition. Qubits are “wired” together by allowing them to interact magnetically.



[Mwjohnson0](#)

Processors made by D-Wave Systems are being used for machine learning applications.

To run the system, you first impose a horizontal magnetic field, which initializes the qubits to an equal superposition of up and down — the equivalent of a blank slate. There are a couple of ways to enter data. In some cases, you fix a layer of qubits to the desired input values; more often, you incorporate the input into the strength of the interactions. Then you let the qubits interact. Some seek to align in the same direction, some in the opposite direction, and under the influence of the horizontal field, they flip to their preferred orientation. In so doing, they might trigger other qubits to flip. Initially that happens a lot, since so many of them are misaligned. Over time, though, they settle down, and you can turn off the horizontal field to lock them in place. At that point, the qubits are in a pattern of up and down that ensures the output follows from the input.

It's not at all obvious what the final arrangement of qubits will be, and that's the point. The system, just by doing what comes naturally, is solving a problem that an ordinary computer would struggle with. "We don't need an algorithm," explained [Hidetoshi Nishimori](#), a physicist at the Tokyo Institute of Technology who developed the principles on which D-Wave machines operate. "It's completely different from conventional programming. Nature solves the problem."

The qubit-flipping is driven by quantum tunneling, a natural tendency that quantum systems have to seek out their optimal configuration, rather than settle for second best. You could build a classical network that worked on analogous principles, using random jiggling rather than tunneling to get bits

to flip, and in some cases it would actually work better. But, interestingly, for the types of problems that arise in machine learning, the quantum network seems to reach the optimum faster.

The D-Wave machine has had its detractors. It is extremely noisy and, in its current incarnation, can perform only a limited menu of operations. Machine-learning algorithms, though, are noise-tolerant by their very nature. They're useful precisely because they can make sense of a messy reality, sorting kittens from puppies against a backdrop of red herrings. "Neural networks are famously robust to noise," Behrman said.

In 2009 a team led by [Hartmut Neven](#), a computer scientist at Google who pioneered augmented reality — he co-founded the Google Glass project — and then took up quantum information processing, showed how an early D-Wave machine could do a respectable machine-learning task. They used it as, essentially, a single-layer neural network that sorted images into two classes: "car" or "no car" in a library of 20,000 street scenes. The machine had only 52 working qubits, far too few to take in a whole image. (Remember: the D-Wave machine is of a very different type than in the state-of-the-art 50-qubit systems coming online in 2018.) So Neven's team combined the machine with a classical computer, which analyzed various statistical quantities of the images and calculated how sensitive these quantities were to the presence of a car — usually not very, but at least better than a coin flip. Some combination of these quantities could, together, spot a car reliably, but it wasn't obvious which. It was the network's job to find out.

The team assigned a qubit to each quantity. If that qubit settled into a value of 1, it flagged the corresponding quantity as useful; 0 meant don't bother. The qubits' magnetic interactions encoded the demands of the problem, such as including only the most discriminating quantities, so as to keep the final selection as compact as possible. The result was able to spot a car.

Last year a group led by [Maria Spiropulu](#), a particle physicist at the California Institute of Technology, and [Daniel Lidar](#), a physicist at USC, applied the algorithm to a practical physics problem: classifying proton collisions as "Higgs boson" or "no Higgs boson." Limiting their attention to collisions that spat out photons, they used basic particle theory to predict which photon properties might betray the fleeting existence of the Higgs, such as momentum in excess of some threshold. They considered eight such properties and 28 combinations thereof, for a total of 36 candidate signals, and let a late-model D-Wave at the University of Southern California find the optimal selection. It [identified](#) 16 of the variables as useful and three as the absolute best. The quantum machine needed less data than standard procedures to perform an accurate identification. "Provided that the training set was small, then the quantum approach did provide an accuracy advantage over traditional methods used in the high-energy physics community," Lidar said.



Courtesy of Maria Spiropulu

Maria Spiropulu, a physicist at the California Institute of Technology, used quantum machine learning to find Higgs bosons.

In December, Rigetti demonstrated a way to automatically group objects using a general-purpose quantum computer with 19 qubits. The researchers did the equivalent of feeding the machine a list of cities and the distances between them, and asked it to sort the cities into two geographic regions. What makes this problem hard is that the designation of one city depends on the designation of all the others, so you have to solve the whole system at once.

The Rigetti team effectively assigned each city a qubit, indicating which group it was assigned to. Through the interactions of the qubits (which, in Rigetti's system, are electrical rather than magnetic), each pair of qubits sought to take on opposite values — their energy was minimized when they did so. Clearly, for any system with more than two qubits, some pairs of qubits had to consent to be assigned to the same group. Nearby cities assented more readily since the energetic cost for them to be in the same group was lower than for more-distant cities.

To drive the system to its lowest energy, the Rigetti team took an approach similar in some ways to the D-Wave annealer. They initialized the qubits to a superposition of all possible cluster assignments. They allowed qubits to interact briefly, which biased them toward assuming the same or opposite values. Then they applied the analogue of a horizontal magnetic field, allowing the qubits to flip if they were so inclined, pushing the system a little way toward its lowest-energy state. They repeated this two-step process — interact then flip — until the system minimized its energy, thus sorting the cities into two distinct regions.

These classification tasks are useful but straightforward. The real frontier of machine learning is in generative models, which do not simply recognize puppies and kittens, but can generate novel archetypes — animals that never existed, but are every bit as cute as those that did. They might even figure out the categories of “kitten” and “puppy” on their own, or reconstruct images missing a tail or paw. “These techniques are very powerful and very useful in machine learning, but they are very hard,” said [Mohammad Amin](#), the chief scientist at D-Wave. A quantum assist would be most welcome.

D-Wave and other research teams have taken on this challenge. Training such a model means tuning the magnetic or electrical interactions among qubits so the network can reproduce some sample data. To do this, you combine the network with an ordinary computer. The network does the heavy lifting — figuring out what a given choice of interactions means for the final network configuration — and its partner computer uses this information to adjust the interactions. In one demonstration last year, [Alejandro Perdomo-Ortiz](#), a researcher at NASA’s Quantum Artificial Intelligence Lab, and his team exposed a D-Wave system to images of handwritten digits. It discerned that there were 10 categories, matching the digits 0 through 9, and generated its own scrawled numbers.

Bottlenecks Into the Tunnels

Well, that’s the good news. The bad is that it doesn’t much matter how awesome your processor is if you can’t get your data into it. In matrix-algebra algorithms, a single operation may manipulate a matrix of 16 numbers, but it still takes 16 operations to load the matrix. “State preparation — putting classical data into a quantum state — is completely shunned, and I think this is one of the most important parts,” said Maria Schuld, a researcher at the quantum-computing startup Xanadu and one of the first people to receive a doctorate in quantum machine learning. Machine-learning systems that are laid out in physical form face parallel difficulties of how to embed a problem in a network of qubits and get the qubits to interact as they should.

Once you do manage to enter your data, you need to store it in such a way that a quantum system can interact with it without collapsing the ongoing calculation. Lloyd and his colleagues have proposed a quantum RAM that uses photons, but no one has an analogous contraption for superconducting qubits or trapped ions, the technologies found in the leading quantum computers. “That’s an additional huge technological problem beyond the problem of building a quantum computer itself,” Aaronson said. “The impression I get from the experimentalists I talk to is that they are frightened. They have no idea how to begin to build this.”

And finally, how do you get your data out? That means measuring the quantum state of the machine, and not only does a measurement return only a single number at a time, drawn at random, it collapses the whole state, wiping out the rest of the data before you even have a chance to retrieve it. You’d have to run the algorithm over and over again to extract all the information.

Yet all is not lost. For some types of problems, you can exploit quantum interference. That is, you can choreograph the operations so that wrong answers cancel themselves out and right ones reinforce themselves; that way, when you go to measure the quantum state, it won’t give you just any random value, but the desired answer. But only a few algorithms, such as brute-force search, can make good use of interference, and the speedup is usually modest.

In some cases, researchers have found shortcuts to getting data in and out. In 2015 Lloyd, Silvano Garnerone of the University of Waterloo in Canada, and [Paolo Zanardi](#) at USC showed that, for some kinds of statistical analysis, you don’t need to enter or store the entire data set. Likewise, you don’t need to read out all the data when a few key values would suffice. For instance, tech companies use machine learning to suggest shows to watch or things to buy based on a humongous matrix of

consumer habits. “If you’re Netflix or Amazon or whatever, you don’t actually need the matrix written down anywhere,” Aaronson said. “What you really need is just to generate recommendations for a user.”

All this invites the question: If a quantum machine is powerful only in special cases, might a classical machine also be powerful in those cases? This is the major unresolved question of the field. Ordinary computers are, after all, extremely capable. The usual method of choice for handling large data sets — random sampling — is actually very similar in spirit to a quantum computer, which, whatever may go on inside it, ends up returning a random result. Schuld remarked: “I’ve done a lot of algorithms where I felt, ‘This is amazing. We’ve got this speedup,’ and then I actually, just for fun, write a sampling technique for a classical computer, and I realize you can do the same thing with sampling.”

If you look back at the successes that quantum machine learning has had so far, they all come with asterisks. Take the D-Wave machine. When classifying car images and Higgs bosons, it was no faster than a classical machine. “One of the things we do not talk about in this paper is quantum speedup,” said Alex Mott, a computer scientist at Google DeepMind who was a member of the Higgs research team. Matrix-algebra approaches such as the Harrow-Hassidim-Lloyd algorithm show a speedup only if the matrices are sparse — mostly filled with zeroes. “No one ever asks, are sparse data sets actually interesting in machine learning?” Schuld noted.

Quantum Intelligence

On the other hand, even the occasional incremental improvement over existing techniques would make tech companies happy. “These advantages that you end up seeing, they’re modest; they’re not exponential, but they are quadratic,” said [Nathan Wiebe](#), a quantum-computing researcher at Microsoft Research. “Given a big enough and fast enough quantum computer, we could revolutionize many areas of machine learning.” And in the course of using the systems, computer scientists might solve the theoretical puzzle of whether they are inherently faster, and for what.

Schuld also sees scope for innovation on the software side. Machine learning is more than a bunch of calculations. It is a complex of problems that have their own particular structure. “The algorithms that people construct are removed from the things that make machine learning interesting and beautiful,” she said. “This is why I started to work the other way around and think: If I have this quantum computer already — these small-scale ones — what machine-learning model actually can it generally implement? Maybe it is a model that has not been invented yet.” If physicists want to impress machine-learning experts, they’ll need to do more than just make quantum versions of existing models.

Just as many neuroscientists now think that the structure of human thought reflects the requirements of having a body, so, too, are machine-learning systems embodied. The images, language and most other data that flow through them come from the physical world and reflect its qualities. Quantum machine learning is similarly embodied — but in a richer world than ours. The one area where it will undoubtedly shine is in processing data that is already quantum. When the data is not an image, but the product of a physics or chemistry experiment, the quantum machine will be in its element. The input problem goes away, and classical computers are left in the dust.

In a neatly self-referential loop, the first quantum machine-learning systems may help to design their successors. “One way we might actually want to use these systems is to build quantum computers themselves,” Wiebe said. “For some debugging tasks, it’s the only approach that we have.” Maybe they could even debug us. Leaving aside whether [the human brain is a quantum computer](#) — a highly contentious question — it sometimes acts as if it were one. Human behavior is notoriously contextual; our preferences are formed by the choices we are given, in ways that defy logic. In this,

we are like quantum particles. “The way you ask questions and the ordering matters, and that is something that is very typical in quantum data sets,” Perdomo-Ortiz said. So a quantum machine-learning system might be a natural way to study human cognitive biases.

Neural networks and quantum processors have one thing in common: It is amazing they work at all. It was never obvious that you could train a network, and for decades most people doubted it would ever be possible. Likewise, it is not obvious that quantum physics could ever be harnessed for computation, since the distinctive effects of quantum physics are so well hidden from us. And yet both work — not always, but more often than we had any right to expect. On this precedent, it seems likely that their union will also find its place.

This article was reprinted on [Wired.com](#).