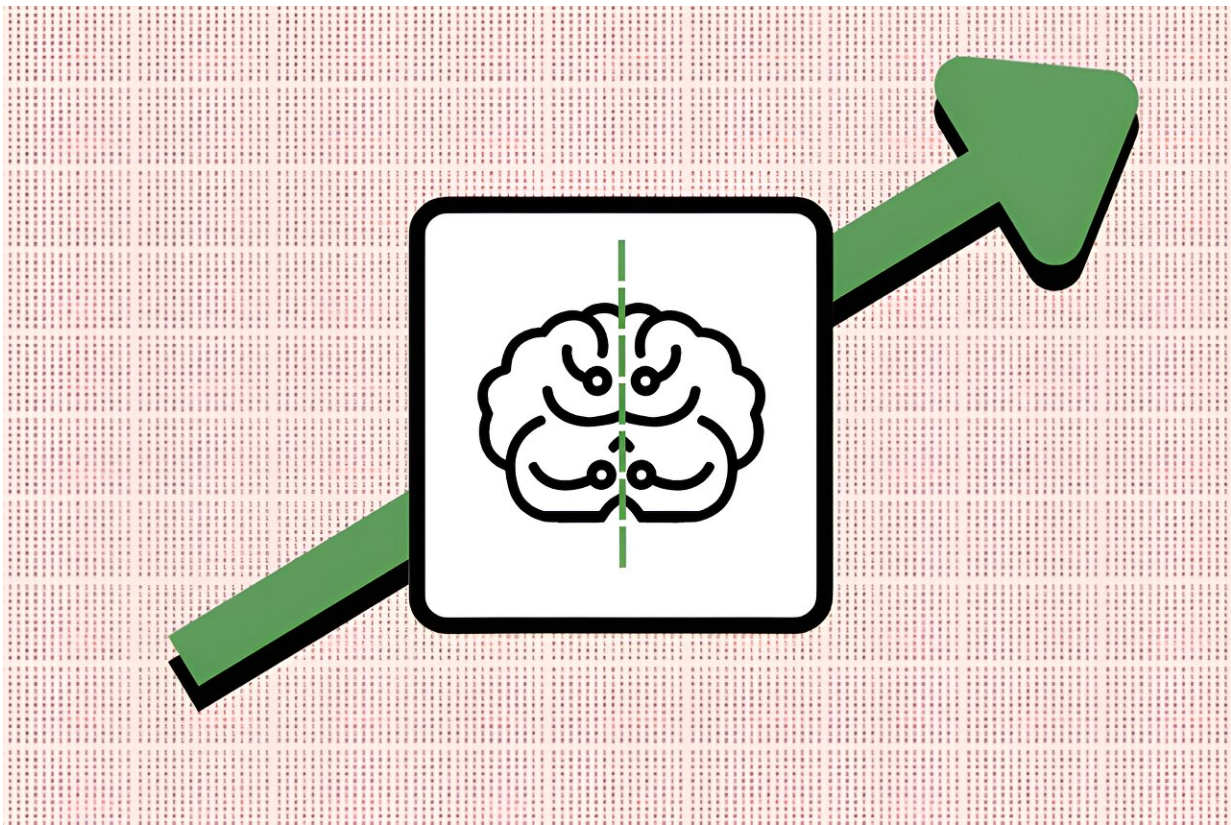# How symmetry can come to the aid of machine learning

February 5 2024, by Steve Nadis

New MIT research provides a theoretical proof for a phenomenon observed in practice: that encoding symmetries in the machine learning model helps the model learn with fewer data. Credit: Alex Shipps/MIT CSAIL

Behrooz Tahmasebi—an MIT Ph.D. student in the Department of

Electrical Engineering and Computer Science (EECS) and an affiliate of the Computer Science and Artificial Intelligence Laboratory (CSAIL)—was taking a mathematics course on differential equations in late 2021 when a glimmer of inspiration struck. In that class, he learned for the first time about Weyl's law, which had been formulated 110 years earlier by the German mathematician Hermann Weyl.

Tahmasebi realized it might have some relevance to the [computer science](#) problem he was then wrestling with, even though the connection appeared—on the surface—to be thin, at best. Weyl's law, he says, provides a formula that measures the complexity of the spectral information, or data, contained within the fundamental frequencies of a drum head or guitar string.

Tahmasebi was, at the same time, thinking about measuring the complexity of the input data to a [neural network](#), wondering whether that complexity could be reduced by taking into account some of the symmetries inherent to the dataset. Such a reduction, in turn, could facilitate—as well as speed up—machine learning processes.

Weyl's law, conceived about a century before the boom in machine learning, had traditionally been applied to very different physical situations—such as those concerning the vibrations of a string or the spectrum of electromagnetic (black-body) radiation given off by a heated object. Nevertheless, Tahmasebi believed that a customized version of that law might help with the machine learning problem he was pursuing. And if the approach panned out, the payoff could be considerable.

He spoke with his advisor, Stefanie Jegelka—an associate professor in EECS and affiliate of CSAIL and the MIT Institute for Data, Systems, and Society—who believed the idea was definitely worth looking into. As Tahmasebi saw it, Weyl's law had to do with gauging the complexity

of data, and so did this project. But Weyl's law, in its original form, said nothing about symmetry.

He and Jegelka have now succeeded in modifying Weyl's law so that symmetry can be factored into the assessment of a dataset's complexity. "To the best of my knowledge," Tahmasebi says, "this is the first time Weyl's law has been used to determine how machine learning can be enhanced by symmetry."

The paper he and Jegelka wrote earned a "Spotlight" designation when it was presented at the December 2023 conference on Neural Information Processing Systems—widely regarded as the world's top conference on machine learning. It is currently [available](#) on the arXiv preprint server.

This work, comments Soledad Villar, an applied mathematician at Johns Hopkins University, "shows that models that satisfy the symmetries of the problem are not only correct but also can produce predictions with smaller errors, using a small amount of training points. [This] is especially important in scientific domains, like computational chemistry, where training data can be scarce."

In their paper, Tahmasebi and Jegelka explored the ways in which symmetries, or so-called "invariances," could benefit machine learning. Suppose, for example, the goal of a particular computer run is to pick out every image that contains the numeral 3. That task can be a lot easier, and go a lot quicker, if the algorithm can identify the 3 regardless of where it is placed in the box—whether it's exactly in the center or off to the side—and whether it is pointed right-side up, upside down, or oriented at a random angle.

An algorithm equipped with the latter capability can take advantage of the symmetries of translation and rotations, meaning that a 3, or any other object, is not changed in itself by altering its position or by rotating

it around an arbitrary axis. It is said to be invariant to those shifts. The same logic can be applied to algorithms charged with identifying dogs or cats. A dog is a dog is a dog, one might say, irrespective of how it is embedded within an image.

The point of the entire exercise, the authors explain, is to exploit a dataset's intrinsic symmetries in order to reduce the complexity of machine learning tasks. That, in turn, can lead to a reduction in the amount of data needed for learning. Concretely, the new work answers the question: How many fewer data are needed to train a machine learning model if the data contain symmetries?

There are two ways of achieving a gain, or benefit, by capitalizing on the symmetries present. The first has to do with the size of the sample to be looked at. Let's imagine that you are charged, for instance, with analyzing an image that has mirror symmetry—the right side being an exact replica, or mirror image, of the left. In that case, you don't have to look at every pixel; you can get all the information you need from half of the image—a factor of two improvement. If, on the other hand, the image can be partitioned into 10 identical parts, you can get a factor of 10 improvement. This kind of boosting effect is linear.

To take another example, imagine you are sifting through a dataset, trying to find sequences of blocks that have seven different colors—black, blue, green, purple, red, white, and yellow. Your job becomes much easier if you don't care about the order in which the blocks are arranged. If the order mattered, there would be 5,040 different combinations to look for. But if all you care about are sequences of blocks in which all seven colors appear, then you have reduced the number of things—or sequences—you are searching for from 5,040 to just one.

Tahmasebi and Jegelka discovered that it is possible to achieve a

different kind of gain—one that is exponential—that can be reaped for symmetries that operate over many dimensions. This advantage is related to the notion that the complexity of a learning task grows exponentially with the dimensionality of the data space. Making use of a multidimensional symmetry can therefore yield a disproportionately large return. "This is a new contribution that is basically telling us that symmetries of higher dimension are more important because they can give us an exponential gain," Tahmasebi says.

The NeurIPS 2023 paper that he wrote with Jegelka contains two theorems that were proved mathematically. "The first theorem shows that an improvement in sample complexity is achievable with the general algorithm we provide," Tahmasebi says. The second theorem complements the first, he added, "showing that this is the best possible gain you can get; nothing else is achievable."

He and Jegelka have provided a formula that predicts the gain one can obtain from a particular symmetry in a given application. A virtue of this formula is its generality, Tahmasebi notes. "It works for any symmetry and any input space." It works not only for symmetries that are known today, but it could also be applied in the future to symmetries that are yet to be discovered. The latter prospect is not too farfetched to consider, given that the search for new symmetries has long been a major thrust in physics. That suggests that, as more symmetries are found, the methodology introduced by Tahmasebi and Jegelka should only get better over time.

According to Haggai Maron, a computer scientist at Technion (the Israel Institute of Technology) and NVIDIA who was not involved in the work, the approach presented in the paper "diverges substantially from related previous works, adopting a geometric perspective and employing tools from differential geometry. This theoretical contribution lends mathematical support to the emerging subfield of 'Geometric Deep

Learning,' which has applications in graph learning, 3D data, and more. The paper helps establish a theoretical basis to guide further developments in this rapidly expanding research area."

More information: Behrooz Tahmasebi et al, The Exact Sample Complexity Gain from Invariances for Kernel Regression, *arXiv* (2023). DOI: 10.48550/arxiv.2303.14269

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: How symmetry can come to the aid of machine learning (2024, February 5) retrieved 8 May 2024 from https://techxplore.com/news/2024-02-symmetry-aid-machine.html