Nobel Prize: Mimicking Human Intelligence with Neural Networks

The 2024 Nobel Prize in Physics honors pioneering work on artificial neural networks, which provided the foundation for many of the artificial intelligence technologies in use today.

10 October 2024: We have replaced our initial short announcement with this full-length news story.

Certain processes in the brain, such as recognition and classification, can be modeled as interactions of artificial neurons, or "nodes," in a highly interconnected network. This physics-inspired approach to human learning has been recognized with the 2024 Nobel Prize in Physics. John Hopfield from Princeton University and Geoffrey Hinton from the University of Toronto share this year's prize for their work on artificial neural networks, which have become the basis of many artificial intelligence (AI) technologies, such as facial recognition systems and chatbots.


The 2024 physics Nobel laureates developed the earliest forms of artificial intelligence (AI) by modeling neurons in the brain. This image was generated by AI.
Credit: Vladlena Demidova/Stock.adobe.com

An artificial neural network is a collection of nodes, each of which has a value that depends on the values of the nodes to which it's connected. In the early 1980s, Hopfield showed that these networks can be imprinted with a kind of memory that can recognize images through an energy-minimization process. Building on that work, Hinton showed how the couplings between nodes could be tuned (or "trained") to perform specific tasks, such as data sorting or classification. Together, the contributions of these physicists set the stage for today's machine learning revolution.

Neurons in the brain communicate with each other through synapses, and the number of synapses connected to any given neuron ranges from a handful to several thousand. Early studies in the 1940s showed that the firing activity of a particular neuron—the electrical pulses it generates—depends on the inputs received from connected neurons [1]. Moreover, connected neurons that fire simultaneously can develop stronger mutual connections, eventually leading to memories that are encoded in the relative synaptic strengths [2]. Many researchers became interested in reproducing this neural behavior in digital networks in which nodes replaced neurons and couplings replaced synapses. But solving real problems with these artificial neural networks proved computationally challenging.

In 1982, Hopfield opened a way forward [3]. He proposed a simple network based on many-body physical systems, such as the atomic spins inside a magnetic material. In analogy with a neural network, each spin (or node) has a specific value based on its orientation, and that spin value can affect nearby spins
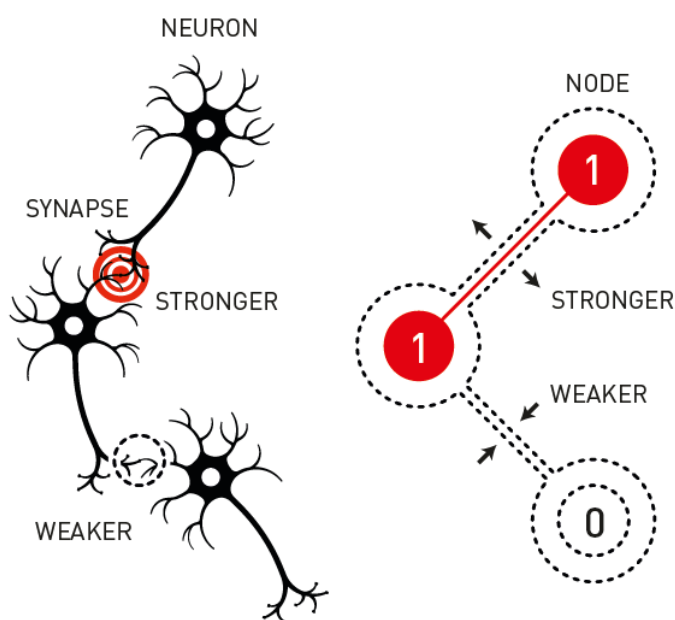
Neurons in the brain are connected through synapses. One popular model of learning is that these connections become stronger (or weaker) depending on the correlated activity of the two connected neurons. Artificial neural networks are built on the principle that the strengths of connections between nodes can be tuned to produce a desired result.
Credit: J. Jarnestad/Royal Swedish Academy of Sciences

through magnetic interactions (or couplings). The spins settle into a stable configuration based on the strengths of those interactions.

Taking spin physics as inspiration, Hopfield set up a network of N nodes connected through weighted couplings. Each node had a value of either 0 or 1, which could be changed (during random updates) depending on the weighted sum of all the other nodes to which it was coupled. He defined an "energy" term based on the relative alignment of the connected nodes and showed that the network evolved toward a low-energy state.

Hopfield then showed that this spin-based neural network could store and retrieve a "memory" in one of the low-energy (stable) states. A real-world example would be memorizing a pattern in an image. Here, the network nodes are associated with the pixels on a screen, and the couplings are tuned so that

the output (or stable state) corresponds to a target image, which might be, for example, the letter "J." If the network is then initialized with a different (input) image—say, that of a highly distorted or poorly written "J"—the node values will naturally evolve the image to the network's stable state. This process illustrates the network associating the input with the stored "J"-pattern memory.

The practical uses of the Hopfield network drew the interest of other researchers, including Hinton. In the mid-1980s, he and his colleagues developed a network called a Boltzmann machine, in which each possible node configuration is assigned a probability based on its energy [4]. The researchers devised an algorithm that adjusted the network's couplings so that the probability distribution matched the statistical distribution in the target data. To make the method more effective, Hinton and colleagues introduced the idea of "visible" layers of nodes that are used for inputs or outputs and separate "hidden" layers that are still part of the network but not connected with the data. A variant on this design, called the restricted Boltzmann machine, became a precursor to deep-learning networks, which are widely used tools in fields such as computer science, immunology, and quantum mechanics.

Giuseppe Carleo, a machine learning specialist from the Swiss Federal Institute of Technology in Lausanne (EPFL), has used restricted Boltzmann machines to find the ground states and time evolution of complex quantum systems [5]. He says that Hinton and Hopfield laid the groundwork for a mechanistic understanding of learning. "The most exciting part for me as a physicist is actually seeing a form of elementary intelligence emerging from first principles, from relatively simple models that can be analyzed with the tools of physics," Carleo says.

Hinton was also instrumental in developing "backpropagation" for the training of neural networks. The method involves computing the difference between a network's output and a set of training data and then tuning the node couplings to minimize that difference. This type of feedback wasn't new, but Hinton and his colleagues showed how adding hidden layers could expand the utility of this process.

"This year's Nobel Prize goes to two well-deserving pioneers," says neural-network expert Stefanie Czischek from the University of Ottawa in Canada. She says the work of Hopfield

and Hinton showed how the principles of spin systems can extend far beyond the realm of many-body physics. "Even though the Hopfield network and the restricted Boltzmann machine have both been replaced by more powerful architectures in most applications, they laid the foundation for state-of-the-art artificial neural networks," Czischek says.

Over the years, neural networks have blossomed into a host of AI algorithms that recognize faces, drive cars, identify cancers, compose music, and carry on conversations. At the Nobel Prize press conference, Hinton was asked about the future impact of AI. "It will have a huge influence, comparable to that of the industrial revolution," he said. He imagines this influence will be welcome in some areas, such as health care, but he also expressed concern over the possibility that AI will exert a controlling influence on our lives.

—Michael Schirber

Michael Schirber is a Corresponding Editor for *Physics Magazine* based in Lyon, France.

REFERENCES

1. W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," **Bull. Math. Biophys. 5, 115 (1943).**
2. D.O. Hebb, **The Organization of Behavior (Wiley & Sons, New York, 1949).**
3. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," **Proc. Natl. Acad. Sci. U.S.A. 79, 2554 (1982).**
4. D. Ackley *et al.*, "A learning algorithm for boltzmann machines," **Cogn. Sci. 9, 147 (1985), https://www.sciencedirect.com/science/article/pii/S0364021385800124.**
5. G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," **Science 355, 602 (2017).**