MAPS OF THE LABYRINTH
A SERIES FROM PHANTOM MAZE — AI & LANGUAGE LAB

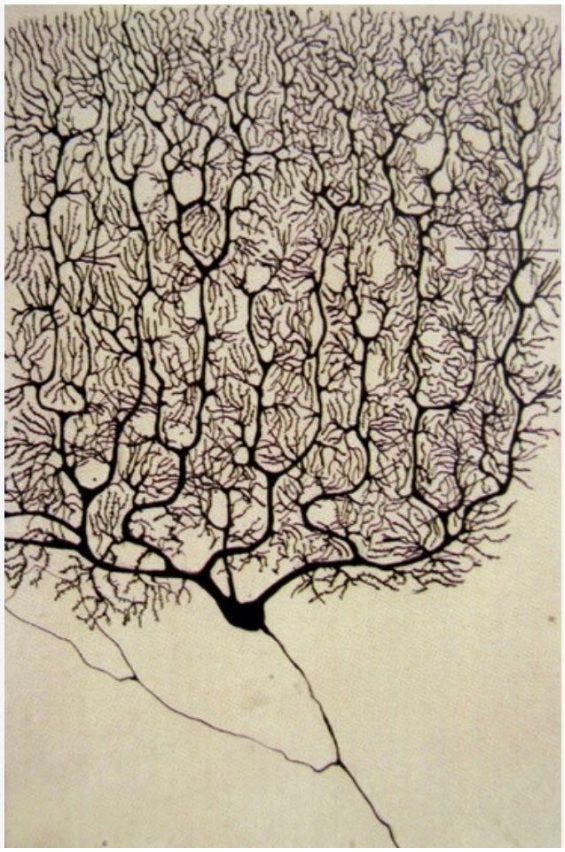
ESSAY 17

Neural Networks: The Spiderweb Without a Spider

Claudia Marsico · Hernán Inverso

DOI 10.5281/zenodo.23575286

What an artificial neuron really does



The neural network is the mathematical structure underlying today's language models, and its name carries the most loaded metaphor in all of artificial intelligence. We repeat "neural network", "neuronal", "neural" so often that it sounds like something obvious, and the strangest part slips away from us. In what sense is it built like a brain of silicon? It is worth looking closely, because this is terrain where technique and poetic licence are devilishly braided together.

The word "net" drags along its history of fishermen and hunters with their mesh of knotted threads, the *rete* of the Romans. From there comes "reticulum", the "little net", and also "retina", because at

the back of the eye something like a net could be seen. Seeing, too, is, at least in its etymology, catching the world in a net. "Neuron" is not far away. The Greek *neuron* had nothing to do with the brain, which they called *enképhalos*, and even less with brain cells, which nobody knew existed. *Neuron* was the sinew, the taut cord of the muscle, and later the nerve. When Wilhelm von Waldeyer synthesized the work in the field in 1891, he chose "neuron" for this unit of the nervous system. Santiago Ramón y Cajal would suggest, years later, that the catchy name had been Waldeyer's entire contribution. That is no small thing, after all.

Meshes of knots speak to us from very far away. In the tradition that formed around the *Flower Ornament Sutra*, the *Avatamsaka Sutra*, there appeared the image of the net of the god Indra, which the patriarch Dushun, in the China of the sixth and seventh centuries, turned into a capital teaching on the interpenetration of all things. Imagine an infinite net stretched over the cosmos with a jewel at every knot. Each jewel reflects all the others, and in each of those reflections all the others are reflected in turn, to infinity. No knot makes sense on its own. Each one is what it is through the whole weave mirrored in it. It is a beautiful intuition about how meaning lives in the relations within a net that repeats the net at every point. An artificial neural network works a little like that, since no loose knot means anything outside the weave of reflections and points. At the same time it is very different, given that Indra's net is a world of jewels contemplating one another while this one is a mesh of sums and multiplications that nobody contemplates.

At the end of the nineteenth century the dominant theory, Camillo Golgi's, understood the nervous system as a continuous reticulum, a single mesh of seamlessly fused fibres in which everything touched everything. Against that idea worked, with a microscope and a draughtsman's talent, the Spaniard Santiago Ramón y Cajal, whom we have already mentioned. Around 1888, after studying stained tissue, he rejected the idea of the continuous mesh and asserted that the brain is a multitude of separate cells, each with its body, its branches and its terminal, which come close without fusing and communicate across gaps that would later be called synapses. In that thicket Cajal saw, in his own words, a forest of trees with their branches spread wide. He drew it thousands of times with a beauty that adorns museums. The discord was so tenacious that, when in 1906 the two were given the same Nobel Prize, Golgi used his lecture to go on defending the continuous net that was losing the game, and Cajal remarked, bitterly, that fate had been cruel to pair him, like a Siamese twin, with a scientific adversary of such contrasting character.

The episode did not end there. Cajal won the anatomical battle, but the "net" survived. The brain was not the continuous reticulum Golgi imagined, but when artificial neural networks appeared half a century later they kept both images at once: separate units, like Cajal's neurons, organized in a weave of connections, except that those "neurons" and those "connections" became mathematical abstractions.

In 1943 the neurophysiologist Warren McCulloch and the logician Walter Pitts proposed an extremely simplified mathematical model of a nerve cell. The basic idea was to reduce the nerve cell to a logical unit: it received signals from other units and, if the excitatory ones reached a certain threshold

and no inhibitory signal intervened, it produced an output; otherwise, it did not. Of the chemical and electrical complexity of a real neuron, barely a small logical device remained.

Today's networks kept the general structure, though they changed the operation a great deal. A unit receives numbers, multiplies them by its weights, adds them up, may shift the result up or down with a parameter called a bias, and finally passes that total through a rule called an activation function. That last operation decides whether the network will be good for anything. One of the simplest and most famous activations, ReLU, does the following: if the number that arrives is negative, it is replaced by zero; if it is positive, it is let through as is. Nothing more. Today's models use smoother variants, but the idea is always to insert into the chain of multiplications and sums a rule that is not proportional, in order to break what mathematicians call linearity. If I double what goes in, what comes out doubles, and the effect of two things together is the sum of their separate effects. Buying twice as many apples costs twice as much. That is linear. The world, however, is full of things that do not work that way. Doubling the dose of a medicine does not double the improvement and may, on the contrary, kill the patient. A little salt improves the stew and twice as much ruins it. Thresholds, saturations, changes of state lie outside the proportional.

From there comes the problem of stacking layers without activations. Think of exchanging currencies with no banks or commissions in between: pesos to dollars, dollars to euros, euros to yen. Each step is a multiplication by a number, and the whole chain, however many links it has, always amounts to a single multiplication, the one that converts pesos to yen directly. Adding intermediaries neither opens up nor blocks anything. The same happens with layers: if each one confined itself to linear or affine transformations, without activations, a hundred layers could be collapsed into one and we would have a hundred-storey network with the capacity of a single storey. The activation prevents that collapse and makes it possible for the second storey to see something the first could not, so that depth is good for something.

Language demands that we do this because the meaning of a sentence is not the sum of the meanings of its words. A "not" placed in front flips everything that follows, and "it is not bad" is not "it is" minus "bad". To say nothing of irony, which means the opposite of what it says. "Unless", "except when", "but" are hinges, not addends. For all these reasons a strictly linear chain cannot represent that kind of interaction and needs thresholds and non-linear turns.

Imagine red and blue dots drawn on a sheet of paper, and the task of separating them with a single straight line. If the reds are on the left and the blues on the right, it is easy, but if the reds form a ring around the blues, no straight line exists that can separate them. One or several curves are needed. Layers with activations allow the network to bend that boundary as many times as it needs. A "deep" network is simply a network in which many of those transformations follow one another, and each can bend a little further what it received from the previous one. The depth of *deep learning* counts those layers, devoid of unfathomable mysteries.

What, then, is a neural network? Recall that a function is a rule that takes an input and returns an output. For example, a temperature conversion table takes degrees Celsius and returns Fahrenheit. A

neural network does something like that, but on a colossal scale. It takes in many numbers, sends them through successive layers of operations and returns other numbers. This function is special, moreover, because it contains an enormous number of adjustable parameters, so that changing those values changes the relation between input and output.

The reach of that flexibility was formalized at the end of the 1980s in what are called the universal approximation theorems. Think of a circle made of three segments. It looks little like a circle, with twenty it looks fairly like one, and with a thousand we can hardly tell the difference. In the case of networks, by combining enough simple pieces they can approximate with precision far more complicated continuous relations. Now, the theorem only guarantees that, under certain conditions, such a configuration exists, not that it is easy to find. Knowing that some combination of millions or billions of parameters produces the behaviour we are after does not tell us which one it is. How do we find the combination of the safe? That is what training is for, which consists in modifying the parameters until a configuration is found that works well.

The history of this idea is a rollercoaster ride. After the McCulloch and Pitts neuron came, in 1958, Rosenblatt's perceptron, that device whose motorized knobs we mentioned when speaking of parameters and weights. It matters here because it added a way of adjusting those weights from examples, which was something new with respect to the 1943 model. At the time there was euphoria, but in 1969 Marvin Minsky and Seymour Papert rigorously showed the limitations of single-layer perceptrons. The example that ended up becoming famous was XOR, the "exclusive or", which is true when one thing or the other happens but not both at once. Its four truth-value combinations cannot be separated by a single linear boundary. The result fed the idea that neural networks were in a dead end, right in the days of the first AI winter, which had many other causes. What is certain is that this line of research was left battered.

The spring of networks returned in the mid-1980s, when David Rumelhart, Geoffrey Hinton and Ronald Williams popularized backpropagation, a procedure that made it possible to use the error observed at the output to work out how the weights located in the inner layers ought to be modified. Multilayer networks could at last be trained in a practical way on problems that the simple perceptron could not represent. The data and the machines were still missing, and they arrived together in 2012, when a deep network called AlexNet, the work of Alex Krizhevsky, Ilya Sutskever and Hinton himself, won a major image recognition competition by a crushing margin and taught the whole world that the old idea, with enough scale, finally worked. From there to today's language models the path becomes recognizable.

Let us pause a moment on what the machine is not. It is not a brain. The artificial "neuron" is to a real neuron what a stick figure is to a human body. It shares the name and little else. A real nerve cell is an electrochemical system of a complexity we still have not finished understanding; the artificial unit is a sum with a curve on top, a mathematical abstraction remotely inspired by a few features of the former. Artificial networks do not reproduce the anatomy, the dynamics or the learning mechanisms of the brain. The kinship is only historical and metaphorical. Taking it literally is the source of a great

many of the confusions circulating about artificial intelligence, such as that adjusting weights is equivalent to the way a brain learns, or that somewhere in the network there must be a neuron that stores a particular idea. Those inferences do not follow from the name.

Nor does the well-worn numerical comparison hold. A human brain has around eighty-six billion neurons and on the order of a hundred trillion synapses; large language models are counted in hundreds of billions of parameters. The proximity of the figures means nothing, because a parameter is not equivalent to a synapse, nor an artificial unit to a nerve cell. The analogy does not improve because the numbers draw closer, though our compulsion to compare them is interesting in itself.

Indra's net was made of jewels that reflected one another. This one is made of numerical operations whose effectiveness also depends on distributed relations, and it ends up recognizing images, translating languages or holding a conversation with no need for a brain hidden in the knots. A natural spiderweb presupposes a spider that wove it and occupies some place in relation to it. In a neural network what the network knows does not live in any isolated piece, nor is there a centre. It is, in that sense, a spiderweb without a spider.

HOW TO CITE

Marsico, Claudia, and Hernán Inverso. 2026. "Neural Networks: The Spiderweb Without a Spider." *Maps of the Labyrinth*, essay 17. Phantom Maze — AI & Language Lab. https://doi.org/10.5281/zenodo.23575286

Maps of the Labyrinth is a series from Phantom Maze, AI & Language Lab. The collection lives at **phantommaze.cloud**; new essays also appear at phantommazelab.substack.com. Subscription is free.

Some of the framing ideas in this piece are developed in *The Egoless Phantom: Mapping the AI Labyrinth*, by Hernán Inverso and Claudia Marsico (Phantom Maze Press, 2026; ISBN 978-987-3729-16-4).

PHANTOM MAZE

Read this essay online · All essays · The book

Sources. Santiago Ramón y Cajal, *Textura del sistema nervioso del hombre y de los vertebrados: estudios sobre el plan estructural y composición histológica de los centros nerviosos adicionados de consideraciones fisiológicas fundadas en los nuevos descubrimientos*, Madrid, Imprenta y Librería de Nicolás Moya, 2 tomes in 3 vols., 1899–1904; Wilhelm von Waldeyer-Hartz, "Ueber einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems", *Deutsche Medizinische Wochenschrift* 17 (1891), pp. 1213–1218, 1244–1246, 1267–1270, 1287–1289, 1331–1332 and 1352–1356; Santiago Ramón y Cajal, *Recuerdos de mi vida*, 3rd ed., Madrid, Imprenta de Juan Pueyo, 1923, especially part two, *Historia de mi labor científica*; Camillo Golgi, "The Neuron Doctrine-Theory and Facts", Nobel Lecture, 11 December 1906; Francis H. Cook, *Hua-yen Buddhism: The Jewel Net of Indra*, University Park, Pennsylvania State University Press, 1977; Warren S. McCulloch and Walter Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", *The Bulletin of Mathematical Biophysics* 5, no. 4 (1943), pp. 115–133; Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review* 65, no. 6 (1958), pp. 386–408; Marvin Minsky and Seymour A. Papert, *Perceptrons: An Introduction to Computational Geometry*, Cambridge, MA, MIT Press, 1969; George Cybenko, "Approximation by Superpositions of a Sigmoidal Function", *Mathematics of Control, Signals, and Systems* 2, no. 4 (1989), pp. 303–314; Kurt Hornik, "Approximation Capabilities of Multilayer Feedforward Networks", *Neural Networks* 4, no. 2 (1991), pp. 251–257; David E. Rumelhart,

Geoffrey E. Hinton and Ronald J. Williams, "Learning Representations by Back-Propagating Errors", *Nature* 323 (1986), pp. 533–536; Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems* 25 (2012), pp. 1097–1105.