MAPAS DEL LABERINTO
UNA SERIE DE PHANTOM MAZE — AI & LANGUAGE LAB

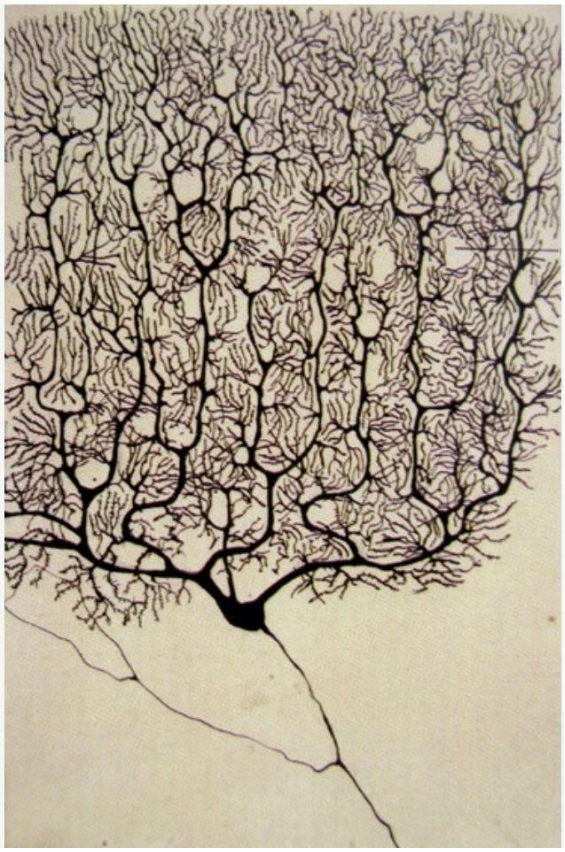
ENSAYO 17

Redes neuronales: la telaraña sin araña

Claudia Marsico · Hernán Inverso

DOI 10.5281/zenodo.23575286

Qué hace de verdad una neurona artificial



La red neuronal es la estructura matemática de fondo de los modelos de lenguaje actuales y su nombre porta la metáfora más cargada de toda la inteligencia artificial. Repetimos tanto “red neuronal”, “neuronal”, “neural” que suena a algo obvio y se nos pierde la parte más extraña. ¿En qué sentido está armada como un cerebro de silicio? Conviene mirar con cuidado, porque es un terreno que trenza endemoniadamente técnica y licencia poética.

El término “red” arrastra su historia de pescadores y cazadores con su malla de hilos anudados. De ahí sale “retículo”, la “red pequeña”, y también “retina”, porque en el fondo del ojo se veía una especie de

red. Ver también es, al menos en la etimología, atrapar el mundo en una red. "Neurona" no está lejos. El griego *neuron* no tenía nada que ver con el cerebro, al que llamaban *enképhalos*, y menos con células del cerebro, que nadie sabía que existían. *Neuron* era el tendón, la cuerda tensa del músculo y más tarde el nervio. Cuando en 1891 Wilhelm von Waldeyer sintetizó trabajos del área, eligió "neurona" para esta unidad del sistema nervioso. Santiago Ramón y Cajal sugeriría años más tarde que el nombre pegadizo fue todo el aporte de Waldeyer. No es poco, después de todo.

Las mallas de nudos nos hablan desde muy lejos. En la tradición formada alrededor del *Sutra del ornamento floral*, el *Avatamsaka Sutra*, apareció la imagen de la red del dios Indra que el patriarca Dushun, en la China de los siglos VI-VII, convirtió en enseñanza capital sobre la interpenetración de todas las cosas. Imaginemos una red infinita tendida sobre el cosmos que en cada nudo tiene una joya. Cada joya refleja a todas las demás y en cada uno de esos reflejos están a su vez reflejadas todas las otras, hasta el infinito. Ningún nudo tiene sentido por sí solo. Cada uno es lo que es por la trama entera que se espeja en él. Es una intuición hermosa sobre la vida del significado en las relaciones en la red que itera la red en cada punto. Una red neuronal artificial funciona un poco así, ya que ningún nudo suelto quiere decir nada fuera de la trama de reflejos y puntos. Al mismo tiempo es muy diferente, dado que la red de Indra es un mundo de joyas que se contemplan mientras esta es una malla de sumas y multiplicaciones que nadie contempla.

A fines del siglo XIX la teoría dominante de Camillo Golgi comprendía el sistema nervioso como un retículo continuo, una única malla de fibras fundidas sin fisuras en el que todo se tocaba con todo. Contra esa idea trabajó, con un microscopio y talento de dibujante, el español Santiago Ramón y Cajal, a quien ya mencionamos. Hacia 1888, después de estudiar tejido teñido, rechazó la idea de la malla continua y afirmó que el cerebro es una multitud de células separadas, cada una con su cuerpo, sus ramas y su terminal, que se aproximan sin fusionarse, comunicándose a través de espacios que más tarde se llamarían sinapsis. Cajal veía en esa espesura, según sus propias palabras, un bosque de árboles con las ramas extendidas. Lo dibujó miles de veces con una belleza que adorna los museos. La discordia fue tan tenaz que, cuando en 1906 les dieron a los dos el mismo premio Nobel, Golgi aprovechó su discurso para seguir defendiendo la red continua que estaba perdiendo la partida y Cajal comentó, con amargura, que el destino era cruel al haberlo atado como a un siamés a un adversario científico de carácter tan opuesto.

El episodio no terminó ahí. Cajal ganó la batalla anatómica, pero la "red" sobrevivió. El cerebro no era el retículo continuo que imaginaba Golgi, pero cuando medio siglo después aparecieron las redes neuronales artificiales, conservaron las dos imágenes a la vez: unidades separadas, como las neuronas de Cajal, organizadas en una trama de conexiones, solo que esas "neuronas" y esas "conexiones" se volvieron abstracciones matemáticas.

En 1943, el neurofisiólogo Warren McCulloch y el lógico Walter Pitts propusieron un modelo matemático extremadamente simplificado de una célula nerviosa. La idea básica era reducir la célula nerviosa a una unidad lógica: recibía señales de otras unidades y, si las excitatorias alcanzaban cierto

umbral y no intervenía una señal inhibitoria, producía una salida; de lo contrario, no. De la complejidad química y eléctrica de una neurona real quedaba apenas un pequeño dispositivo lógico.

Las redes actuales conservaron la estructura general, aunque cambiaron mucho la operación. Una unidad recibe números, los multiplica por sus pesos, los suma, puede mover el resultado hacia arriba o hacia abajo con un parámetro llamado sesgo, y por último hace pasar ese total por una regla que se llama función de activación. Esa última operación decide si la red va a servir para algo. Una de las activaciones más sencillas y famosas, ReLU, hace lo siguiente: si el número que llega es negativo, se lo reemplaza por cero; si es positivo, se lo deja pasar tal cual. Nada más. Los modelos actuales usan variantes más suaves, pero la idea es siempre intercalar en la cadena de multiplicaciones y sumas una regla que no sea proporcional para romper lo que los matemáticos llaman linealidad. Si duplico lo que entra, se duplica lo que sale y el efecto de dos cosas juntas es la suma de sus efectos por separado. Comprar el doble de manzanas cuesta el doble. Eso es lineal. Sin embargo, el mundo está lleno de cosas que no funcionan así. Duplicar la dosis de un remedio no duplica la mejoría y puede, al contrario, matar al paciente. Un poco de sal mejora el guiso y el doble lo arruina. Umbrales, saturaciones, cambios de estado están fuera de lo proporcional.

De ahí viene el problema de apilar capas sin activaciones. Pensemos en cambiar divisas sin bancos ni comisiones de por medio: pesos a dólares, dólares a euros, euros a yenes. Cada paso es una multiplicación por un número, y la cadena completa, por más eslabones que tenga, equivale siempre a una sola multiplicación, la que convierte pesos a yenes directamente. Agregar intermediarios no habilita ni obtura nada. Con las capas pasa lo mismo: si cada una se limitara a transformaciones lineales o afines, sin activaciones, cien capas se podrían colapsar en una y tendríamos una red de cien pisos con la capacidad de uno. La activación impide ese colapso y hace que el segundo piso pueda ver algo que el primero no podía y la profundidad sirva para algo.

El lenguaje pide que hagamos esto porque el sentido de una frase no es la suma de los sentidos de sus palabras. Un “no” delante invierte todo lo que sigue, “no está mal” no es la resta de “está” y “mal”. Ni que hablar de la ironía, que significa lo contrario de lo que dice. “Salvo que”, “a menos que”, “pero” son bisagras, no sumandos. Por todo esto una cadena estrictamente lineal no puede representar esa clase de interacciones y necesita umbrales y vuelcos no lineales.

Imaginemos puntos rojos y azules dibujados en una hoja con la tarea de separarlos con una sola línea recta. Si los rojos están a la izquierda y los azules a la derecha, es fácil, pero si los rojos forman un anillo alrededor de los azules, no existe ninguna recta que los separe. Hace falta una o varias curvas. Las capas con activaciones le permiten a la red doblar esa frontera tantas veces como necesite. Una red “profunda” es simplemente una red en la que muchas de esas transformaciones se suceden unas a otras y cada una puede doblar un poco más lo que recibió de la anterior. La profundidad del *deep learning* cuenta esas capas desprovistas de misterios insondables.

¿Qué es entonces una red neuronal? Recordemos que una función es una regla que toma una entrada y devuelve una salida. Por ejemplo, una tabla de conversión de temperaturas recibe grados Celsius y devuelve Fahrenheit. Una red neuronal hace algo así pero a escala descomunal. Recibe muchos

números, los hace atravesar capas sucesivas de operaciones y devuelve otros números. Esta función es además especial porque contiene una enorme cantidad de parámetros ajustables, de modo que cambiar esos valores cambia la relación entre entrada y salida.

El alcance de esa flexibilidad quedó formalizado a fines de los años ochenta en los llamados teoremas de aproximación universal. Pensemos en una circunferencia hecha con tres segmentos. Se parece poco a un círculo, con veinte se parece bastante y con mil casi no distinguimos la diferencia. En el caso de las redes, combinando suficientes piezas simples pueden aproximar con precisión relaciones continuas mucho más complicadas. Ahora bien, el teorema sólo asegura que, bajo ciertas condiciones, esa configuración existe, no que sea fácil de encontrar. Saber que alguna combinación de millones o miles de millones de parámetros produce el comportamiento que buscamos no nos dice cuál es. ¿Cómo saber cuál es la combinación de la caja fuerte? Para eso está el entrenamiento, que consiste en ir modificando los parámetros hasta encontrar una configuración que funcione bien.

La historia de esta idea es una montaña rusa. Tras la neurona de McCulloch y Pitts vino, en 1958, el perceptrón de Rosenblatt, ese aparato cuyas perillas motorizadas mencionamos a propósito de parámetros y pesos. Interesa aquí porque agregó una manera de ajustar esos pesos a partir de ejemplos, que es algo nuevo respecto del modelo de 1943. En el momento hubo euforia, pero en 1969 Marvin Minsky y Seymour Papert mostraron rigurosamente las limitaciones de los perceptrones de una sola capa. El ejemplo que terminó haciéndose célebre fue XOR, el "o exclusivo", que es verdadero cuando ocurre una cosa o la otra pero no ambas a la vez. Sus cuatro combinaciones veritativas no pueden separarse mediante una única frontera lineal. El resultado alimentó la idea de que las redes neuronales estaban en un callejón sin salida, justo en tiempos del primer invierno de la inteligencia artificial, que tuvo muchas otras causas. Lo cierto es que esta línea de investigación quedó golpeada.

La primavera de las redes volvió a mediados de los ochenta, cuando David Rumelhart, Geoffrey Hinton y Ronald Williams popularizaron la retropropagación, un procedimiento que permitía utilizar el error observado a la salida para calcular cómo convenía modificar pesos situados en las capas internas. Las redes multicapa podían por fin entrenarse de una manera práctica en problemas que el perceptrón simple no podía representar. Faltaban todavía los datos y las máquinas, que llegaron juntos en 2012, cuando una red profunda llamada AlexNet, obra de Alex Krizhevsky, Ilya Sutskever y el mismo Hinton, ganó por una diferencia aplastante un gran concurso de reconocimiento de imágenes y le enseñó al mundo entero que la vieja idea, con suficiente escala, finalmente funcionaba. De ahí a los modelos de lenguaje de hoy el camino se torna reconocible.

Observemos un momento lo que la máquina no es. No es un cerebro. La "neurona" artificial es a una neurona de verdad lo que un monigote de palotes es a un cuerpo humano. Comparte el nombre y poco más. Una célula nerviosa real es un sistema electroquímico de una complejidad que todavía no terminamos de entender; la unidad artificial es una suma con una curva encima, una abstracción matemática inspirada remotamente en algunos rasgos de la primera. Las redes artificiales no reproducen la anatomía, la dinámica ni los mecanismos de aprendizaje del cerebro. El parentesco solo es histórico y metafórico. Tomarla al pie de la letra es la fuente de muchísimas confusiones que

circulan sobre la inteligencia artificial, como que ajustar pesos equivale a la manera en que aprende un cerebro o que debe haber en algún punto de la red una neurona que guarde una idea determinada. Esas inferencias no se siguen del nombre.

Tampoco vale la remanida comparación numérica. Un cerebro humano tiene alrededor de ochenta y seis mil millones de neuronas y del orden de cien billones de sinapsis; los modelos de lenguaje grandes se cuentan en cientos de miles de millones de parámetros. La vecindad de las cifras no significa nada, porque un parámetro no equivale a una sinapsis ni una unidad artificial a una célula nerviosa. La analogía no mejora porque los números se acerquen, aunque resulta interesante nuestra compulsión a compararlos.

La red de Indra estaba hecha de joyas que se reflejaban unas a otras. Esta está hecha de operaciones numéricas cuya eficacia depende también de relaciones distribuidas y termina reconociendo imágenes, traduciendo lenguas o sosteniendo una conversación sin necesidad de un cerebro escondido en los nudos. Una telaraña natural supone una araña que la tejió y ocupa algún lugar respecto de ella. En una red neuronal lo que la red sabe no vive en ninguna pieza aislada ni hay un centro. Es, en ese sentido, una telaraña sin araña.

CÓMO CITAR

Marsico, Claudia y Hernán Inverso. 2026. “Redes neuronales: la telaraña sin araña”. *Mapas del laberinto*, ensayo 17. Phantom Maze — AI & Language Lab. https://doi.org/10.5281/zenodo.23575286

Mapas del laberinto es una serie de Phantom Maze, AI & Language Lab. La colección vive en **phantommaze.cloud**; los ensayos aparecen también en phantommazeailabes.substack.com. La suscripción es gratuita.

Algunas de las ideas que enmarcan este texto se desarrollan en *El fantasma sin ego. Un mapa en el laberinto de la inteligencia artificial*, de Hernán Inverso y Claudia Marsico (Phantom Maze Press, 2026; ISBN 978-987-3729-10-2).

PHANTOM MAZE

Leer este ensayo en línea · Todos los ensayos · El libro

Fuentes. 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 tomos en 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 y 1352–1356; Santiago Ramón y Cajal, *Recuerdos de mi vida*, 3.ª ed., Madrid, Imprenta de Juan Pueyo, 1923, especialmente la segunda parte, *Historia de mi labor científica*; Camillo Golgi, “The Neuron Doctrine-Theory and Facts”, Nobel Lecture, 11 de diciembre de 1906; Francis H. Cook, *Hua-yen Buddhism: The Jewel Net of Indra*, University Park, Pennsylvania State University Press, 1977; Warren S. McCulloch y Walter Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity”, *The Bulletin of Mathematical Biophysics* 5, n.º 4 (1943), pp. 115–133; Frank Rosenblatt, “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain”, *Psychological Review* 65, n.º 6 (1958), pp. 386–408; Marvin Minsky y 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, n.º 4 (1989), pp. 303–314; Kurt Hornik, "Approximation Capabilities of Multilayer Feedforward Networks", *Neural Networks* 4, n.º 2 (1991), pp. 251–257; David E. Rumelhart, Geoffrey E. Hinton y Ronald J. Williams, "Learning Representations by Back-Propagating Errors", *Nature* 323 (1986), pp. 533–536; Alex Krizhevsky, Ilya Sutskever y Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems* 25 (2012), pp. 1097–1105.