

Ensayo/ Assay

El camino hacia una memoria asociativa del conocimiento humano

The road to an associative memory of human knowledge

Un físico y un psicólogo reciben el premio Nobel de Física por sus trabajos pioneros en el área de las redes neuronales artificiales

¿Las máquinas pueden pensar? Esta fue la pregunta que se hizo el matemático Alan Turing en un artículo publicado en la revista *Mind* en 1950 (Turing, 1950). Los computadores comenzaron a desarrollarse con el objetivo inicial de automatizar cálculos repetitivos y aburridos para los humanos. En ese entonces, ser calculadora humana era un trabajo como cualquier otro. Sin embargo, no tardaron en aparecer quienes imaginaron que el computador podría servir para tareas cognitivas, como jugar ajedrez. Claude Shannon, padre de la teoría de la información, publicaría también en 1950 un artículo sobre cómo programar un computador para jugar ajedrez (Shannon, 1950). Curiosamente, en la introducción del artículo Shannon comenta: «...aunque quizás carezca de importancia práctica, esta cuestión podría tener un interés teórico». Tras periodos de euforia, en los que se esperaban resultados exagerados, y momentos de progreso moderado, en la actualidad los computadores no solamente juegan ajedrez, sino que, de cierta manera, pueden inventar historias y chistes, gracias a la progresión de la quizás mal llamada inteligencia artificial.

Los algoritmos más recientes, capaces de procesar el lenguaje natural, escribir códigos para computador y generar imágenes, se basan en las llamadas redes neuronales artificiales, que son modelos simplificados del sistema nervioso. Este año el Comité Nobel de Física de la Real Academia Sueca de Ciencias otorgó el premio a John Hopfield, físico y profesor emérito de la Universidad de Princeton, y a Geoffrey Hinton, psicólogo de formación y profesor en la Universidad de Toronto y en otras instituciones, reconocido como el “padrino de la inteligencia artificial”. Ambos fueron premiados por sus descubrimientos e invenciones fundamentales en el aprendizaje automático mediante redes neuronales artificiales (Nobel Foundation *et al.*, 2024a).

Hopfield primero (1982), y luego Hinton (1986) (Ackley *et al.*, 1985), partieron de la idea de emular el funcionamiento de las neuronas biológicas a través de modelos relativamente sencillos, a partir de los trabajos previos de Warren McCulloch y Walter Pitts,

John Hopfield



John Hopfield. Ill. Niklas Elmehed © Nobel Prize Outreach

Geoffrey Hinton



Geoffrey Hinton. Ill. Niklas Elmehed © Nobel Prize Outreach

Tomado de <https://www.nobelprize.org/all-nobel-prizes-2024/>

quienes propusieron que las neuronas biológicas podrían ser funcionalmente equivalentes a unidades de procesamiento similares a las puertas lógicas que componen los computadores digitales (**McCulloch & Pitts**, 1943); de Donald O. Hebb, quien investigó cómo los mecanismos fundamentales del aprendizaje podrían emularse a través de la adaptación de las conexiones entre las neuronas (**Hebb**, 1949), y de Frank Rosenblatt, quien introdujo el cableado de un modelo de percepción compuesto de múltiples neuronas, así como los mecanismos de adaptación que le permitían aprender (**Rosenblatt**, 1958).

Hopfield se interesó en desarrollar modelos formados por una única capa de neuronas en la que cada neurona está conectada a todas las demás excepto a sí misma. Dichas conexiones son bidireccionales y simétricas. Hopfield se preguntó si esta configuración de neuronas exhibiría capacidades computacionales emergentes (**Tank & Hopfield**, 1987) y demostró en sus investigaciones que dichos modelos no solo permiten memorizar información, sino que pueden “recordar” patrones completos a partir de sus fragmentos o versiones distorsionadas. Esto significa que, cuando se le presenta una entrada incompleta o con ruido, la red puede recuperar el patrón original asociado a esa entrada. En ese sentido decimos que se trata de una “memoria asociativa”, dado que es capaz de “recordar” información asociándola a otra relacionada.

Por otro lado, Hinton desarrolló un modelo distinto de red neuronal conocido como “máquina de Boltzmann” (**Ackley et al.**, 1985). De manera similar a la red de Hopfield, la de Hinton es recurrente, es decir, utiliza conexiones de retroalimentación, lo que le permite evolucionar con el tiempo hasta alcanzar un estado estable, pero se diferencia por su comportamiento estocástico. Sus neuronas se “encienden” y “apagan” de manera aleatoria, buscando configuraciones que reduzcan su energía, tal como algunos sistemas físicos buscan el estado más estable.

Aunque el modelo resultó ser muy ineficiente, sus versiones posteriores tendrían una profunda influencia en el desarrollo de arquitecturas modernas de redes neuronales artificiales, incluidas las redes neuronales profundas convolucionales (**LeCun et al.**, 1989) y los llamados modelos transformadores (**Ashish et al.**, 2017).

Las dinámicas que ambos integraron en sus modelos provienen de la física. Hopfield utilizó, por ejemplo, ecuaciones que describen la dinámica de materiales ferromagnéticos, mientras que Hinton recurrió a principios de termodinámica y física estadística al aplicar un enfoque probabilístico inspirado en el equilibrio energético (**Nobel Foundation et al.**, 2024a).

Por otro lado, Hinton contribuiría también a popularizar el uso del algoritmo de retropropagación (**Rumelhart et al.**, 1986) para entrenar las redes neuronales artificiales. Este algoritmo busca minimizar, de manera iterativa, el error entre las respuestas de la red neuronal y un conjunto de ejemplos de respuestas deseadas, ajustando los parámetros de conexión entre las neuronas de la red neuronal artificial. Pasados casi 40 años, este mecanismo de “aprendizaje” perdura y sigue siendo la base para entrenar los modelos más avanzados en reconocimiento de objetos, procesamiento del lenguaje y generación de imágenes a partir de descripciones en lenguaje natural.

No cabe duda de que los modelos de aprendizaje automático se han convertido en una especie de navaja suiza para la resolución de problemas debido a su adaptabilidad y competencia. Por un lado, han facilitado la detección de patrones en imágenes y en todo tipo de señales, lo cual es la base de los sistemas actuales de reconocimiento de objetos, de diagnóstico médico, y de otras innumerables aplicaciones. Por otra parte, han permitido la creación de modelos predictivos capaces de revelar relaciones complejas y ocultas entre los datos, por ejemplo en el desarrollo de sistemas de predicción meteorológica y en la predicción de la estructura tridimensional de las proteínas a partir de secuencias de aminoácidos. Fue precisamente este último avance, logrado por los investigadores Demis Hassabis y John Jumper, el que fue reconocido con el premio Nobel de química de este mismo año (**Nobel Foundation et al.**, 2024b).

Por último, Hinton desarrolló la idea de utilizar redes neuronales artificiales para generar representaciones distribuidas de los conceptos del lenguaje (Hinton, 1986). Cada concepto se ve asociado a un vector de valores numéricos en un espacio de alta dimensionalidad, de manera que las redes neuronales quedan habilitadas para procesar información textual mediante representaciones que capturan relaciones semánticas.

Por ejemplo, en un modelo de lenguaje el concepto de “curiosidad” puede representarse numéricamente y, mediante operaciones en este espacio vectorial, se pueden derivar relaciones como “Curiosidad - Incertidumbre + Información = Comprensión”. Estas representaciones distribuidas permiten que las máquinas adquieran habilidades bastante sofisticadas de procesamiento simbólico y ofrezcan una capacidad —aún limitada— de razonamiento y deducción (Valmeekam *et al.*, 2022).

En lugar de entrar en mayores detalles del trabajo de los galardonados, prefiero destacar lo que el desarrollo de estas investigaciones podría hacer posible a nivel tecnológico y científico. Por un lado, hoy en día somos testigos de la emergencia de sistemas capaces de aumentar nuestros sentidos que se aplican en campos que van desde el diagnóstico de enfermedades a partir de exámenes médicos (Xue *et al.*, 2024) hasta la detección de exoplanetas en el espacio (Malik *et al.*, 2022). De otra parte, el computador ha dejado de ser una simple máquina de cálculo y se ha convertido en una especie de máquina para razonar, gracias a la capacidad que hoy ha adquirido para descubrir asociaciones conceptuales complejas e interdisciplinarias en el vasto corpus del conocimiento humano.

A pesar de estos grandes avances tecnológicos, no podemos asegurar que las máquinas podrían llegar a pensar. Sin embargo, el computador y la inteligencia artificial prometen convertirse en “catalizadores de innovación” que faciliten descubrimientos y soluciones a partir de la fusión de conocimientos diversos, lo que acelerará la creatividad y la innovación científica.

Andrés Pérez-Uribe

University of Applied Sciences and Arts of Western Switzerland, Yverdon-les-Bains, Suiza

Referencias

- Ackley, D. H., Hinton, G. E., Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive science*, 9(1), 147-169.
- Ashish, V. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, I.
- Hebb, D. O. (1949). *The Organization of Behavior; a Neuropsychological Theory*. Wiley.
- Hinton, G. E. (1986). Learning distributed representations of concepts. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (Vol. 1, p. 12).
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8), 2554-2558.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- Malik, A., Moster, B. P., Obermeier, C. (2022). Exoplanet detection using machine learning. *Monthly Notices of the Royal Astronomical Society*, 513(4), 5505-5516.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- Nobel Foundation, The Nobel Committee for Physics, The Royal Swedish Academy of Sciences (2024a). *Scientific Background to the Nobel Prize in Physics 2024*. <https://www.nobelprize.org/uploads/2024/09/advanced-physicsprize2024.pdf>
- Nobel Foundation, The Nobel Committee for Chemistry, The Royal Swedish Academy of Sciences (2024b). *Scientific Background to the Nobel Prize in Chemistry 2024*. <https://www.nobelprize.org/uploads/2024/10/advanced-chemistryprize2024.pdf>
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386.

- Rumelhart, D. E., Hinton, G. E., Williams, R. J.** (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- Shannon, C. E.** (1950). XXII. Programming a computer for playing chess. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(314), 256-275.
- Tank, D. W. & Hopfield, J. J.** (1987). Collective computation in neuronlike circuits. *Scientific American*, 257(6), 104-115.
- Turing, A. M.** (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.
- Valmeekam, K., Olmo, A., Sreedharan, S., Kambhampati, S.** (2022, November). Large language models still can't plan (a benchmark for LLMs on planning and reasoning about change). In NeurIPS 2022. Workshop: Foundation Models for Decision Making.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I.** (2017). Attention Is All You Need. arXiv.org: 1706.03762v7 PDF.
- Xue, C., Kowshik, S. S., Lteif, D., Puducheri, S., Jasodanand, V. H., Zhou, O. T., Walia, A. S., Guney, O. B., Zhang, J. D., Pham, S. T., Kaliaev, A., Andreu-Arasa, V. C., Dwyer, B. C., Farris, C. W., Hao, H., Kedar, S., Mian, A. Z., Murman, D.L., O'Shea, S. A., ... Kolachalama, V. B.** (2024). AI-based differential diagnosis of dementia etiologies on multimodal data. *Nature Medicine*, 30, 2977-2989.