



CEU

*Universidad
San Pablo*

Apertura Curso Académico 2024-2025

**Inteligencia artificial vs.
inteligencia humana:
Perspectiva actual desde la neurociencia**

José Luis Lavandera Díaz
Catedrático de Fisiología de la Facultad de Medicina
Universidad CEU San Pablo



CEU | *Ediciones*

Inteligencia artificial vs. inteligencia humana: Perspectiva actual desde la neurociencia

José Luis Lavandera Díaz

Catedrático de Fisiología de la Facultad de Medicina
Universidad CEU San Pablo

Universidad CEU San Pablo

Inteligencia artificial vs. inteligencia humana: Perspectiva actual desde la neurociencia

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

© José Luis Lavandera Díaz, 2024

© de la edición, Fundación Universitaria San Pablo CEU, 2024

Maquetación: Andrea Nieto Alonso (CEU *Ediciones*)

CEU *Ediciones*

Julián Romea 18, 28003 Madrid

www.ceuediciones.es

Depósito legal: M-21525-2024

*Estimadas autoridades académicas,
queridos compañeros,
queridos alumnos de la Universidad CEU San Pablo:*

Ante todo, quisiera agradecer a la Junta de Gobierno de la Universidad y al Decano de la Facultad de Medicina, quienes amablemente me brindaron el honor de poder impartir esta Lección Magistral.

En un momento como el actual, en el que existe un enorme debate en la sociedad sobre las bondades y beneficios de la aplicación de la Inteligencia Artificial (IA) en innumerables ámbitos de la vida cotidiana, quisiera compartir con ustedes una visión desde el punto de la Neurociencia, así como una reflexión sobre las incorrectas asunciones a las que podemos llegar cuando se afirma que la Inteligencia Humana (IH) será en un futuro próximo suplantada por la Inteligencia Artificial.

¿Puede la IA sustituir a la inteligencia humana? ¿Sustituirá la inteligencia artificial a los humanos dentro de unos años?

Estas preguntas están alimentadas por el rápido avance que la IA, el cual suscita temores sobre las posibles implicaciones para el futuro del trabajo y el papel de los seres humanos en un mundo cada vez más dirigido por máquinas inteligentes. Aunque la IA ya ha demostrado sus notables capacidades en diversos ámbitos, la pregunta sigue en el aire.

Para tratar de responder estas cuestiones, debemos primero definir, de forma muy escueta, qué se entiende actualmente por Inteligencia Humana y por Inteligencia Artificial y cuáles son sus diferencias.

Definición inteligencia humana

El cerebro humano es el sistema biológico más complejo conocido y el culmen de la evolución biológica en la Tierra. En él almacenamos todos nuestros recuerdos, conocimientos y habilidades. También nos permite interactuar con el entorno que nos rodea de múltiples formas. Además, la neuroplasticidad cerebral nos facilita el aprendizaje de habilidades y la adquisición de nuevos conocimientos. A lo largo de la historia, el cerebro y la mente humanos han sido el objeto de estudio de pensadores, filósofos y científicos. Inicialmente se consideraba que

la mente humana, como un todo (conciencia, emociones, razonamiento y capacidad de elección), tenía un origen divino. El mayor avance en el estudio del cerebro tuvo lugar a principios del siglo xx con los estudios de Ramón y Cajal y su teoría Neuronal (Ramon y Cajal, 1904) que dieron lugar al comienzo de la Neurociencia moderna. El cerebro tiene multitud de capacidades bien definidas, algunas de las más destacadas son el aprendizaje, la memoria, la imaginación, la creatividad, la abstracción, la adaptabilidad y la intuición, englobándose todas ellas en un concepto más complejo de definir cómo es el de «inteligencia». Terminológicamente, «inteligencia» proviene del latín *intelligentia*, que a su vez deriva de *inteligere*, palabra compuesta de dos términos: *intus* («entre») y *legere* («escoger»). Etimológicamente, inteligencia significa «la capacidad de escoger entre varias posibilidades». La evolución en el estudio de la inteligencia ha dado lugar a una serie de definiciones desde el punto de vista de la Antropología, la Filosofía y la Psicología, de las que cabría destacar muy brevemente, por orden cronológico, las dadas por: Francis Galton (1822-1911) como «una aptitud general superior que explica un conjunto de aptitudes especiales». Galton creía que la inteligencia se determinaba fundamentalmente por la herencia. Posteriormente, Alfred Binet (1857-1911) la definió como «la habilidad para tomar y mantener determinada dirección, adaptarse a nuevas situaciones y tener la capacidad para criticar los propios actos». Años más tarde, Lewis Madison Terman (1877-1956) definió la inteligencia como «la capacidad para pensar de manera abstracta». En una forma más descriptiva, David Wechsler (1896-1981) definió la inteligencia como «la capacidad global de actuar con un propósito, de pensar racionalmente y de enfrentarse de manera efectiva con el entorno». Sin embargo, fue Jean Piaget (1896-1980) quien la definió, desde un punto más analítico, como «el conjunto de operaciones lógicas para las que está capacitado el ser humano, yendo desde la percepción, las operaciones de clasificación, substitución, abstracción, etc. hasta el cálculo proporcional». En la segunda década del siglo xx, Robert J. Sternberg, con su Teoría Triárquica de la Inteligencia (1985), la describe como «la capacidad para adaptar el comportamiento a la consecución de un objetivo, incluyendo las capacidades para beneficiarse de la experiencia, resolver problemas y razonar de modo efectivo» y definió tres categorías para describir la inteligencia: Inteligencia componencial-analítica (la capacidad para planificar, ejecutar y el logro del conocimiento), la Inteligencia experiencial-creativa (habilidad fundada en la experiencia para tratamiento de la novedad y la automatización de procesos) y la Inteligencia contextual-práctica (relacionada con la conducta adaptativa al mundo real). Finalmente, Diane E. Papalia, junto a Sally Wendkos Olds, definieron en 1996 la inteligencia como «la interacción entre genética y ambiente que sirve

para adquirir, recordar y utilizar conocimientos, entender conceptos concretos y abstractos, comprender relaciones entre objetos, hechos, ideas, aplicando y utilizando todo ello para resolver problemas cotidianos».

Tipos de inteligencia humana

Sin embargo, desde el punto de vista de la Neurociencia no existe un único tipo de inteligencia, sino que se distinguen diferentes tipos que podrían describirla como «la capacidad mental general que permite razonar, planificar, resolver problemas, pensar de modo abstracto, comprender ideas complejas, asociar, intuir y aprender con rapidez, usando la experiencia, la adaptabilidad y las emociones, a partir de los diferentes tipos y matices de la información proporcionada por los sentidos como inteligencia sensorial», subdividida esta última en:

- **Inteligencia Visual:** habilidad para identificar imágenes, colores, caras y formas.
- **Inteligencia Auditiva:** habilidad para reconocer sonidos, frecuencias.
- **Inteligencia somestésica y cinestésica:** habilidad para usar el cuerpo en la expresión de idea y sentimientos, así como la facilidad en el uso del sentido del tacto. Incluye habilidades de coordinación, destreza, equilibrio, flexibilidad, fuerza y velocidad. Se incluye la capacidad cinestésica y percepción del espacio, orientación, de medidas y volúmenes.
- **Inteligencia Gustativa:** habilidad para identificar y clasificar los sabores o sensaciones que nos producen los alimentos y bebidas.
- **Inteligencia Olfativa:** habilidad para identificar y clasificar los olores.

En definitiva, una inteligencia perceptiva que configura la experiencia necesaria para enfrentarnos y adaptarnos al entorno y aprender de forma continuada en el tiempo.

Definición inteligencia artificial

En realidad, la inteligencia artificial puede definirse como una serie de algoritmos y técnicas que tratan de emular la inteligencia humana.

Se podría decir que la inteligencia artificial comenzó en 1943 con la publicación del artículo «A Logical Calculus of Ideas Immanent in Nervous Activity», de Warren McCulloch y Walter Pitts, en el cual ambos autores presentaron el primer modelo matemático para la creación de una red neuronal.

Posteriormente, en 1950, dos alumnos de Harvard, Marvin Minsky y Dean Edmonds presentaron el primer ordenador de red neuronal, SNARC. Paralelamente, ese mismo año, Alan Turing publicó un artículo histórico denominado «Computing Machinery and Intelligence», en el que trataba la posibilidad de diseñar sistemas computacionales inteligentes como los humanos y que se considera el inicio de la investigación en Inteligencia Artificial (IA).

En 1952, Arthur Samuel creó un *software* capaz de aprender a jugar al ajedrez de forma autónoma. El término Inteligencia Artificial fue utilizado por primera vez en la conferencia «Dartmouth Summer Research Project on Artificial Intelligence», de John McCarthy en 1956, considerada como el verdadero nacimiento de la inteligencia artificial, tal y como se conoce hoy en día.

En 1959, John McCarthy y Marvin Minsky fundaron el MIT Artificial Intelligence Project. Ese mismo año, Arthur Samuel trabajando en IBM, acuñó el término *Machine Learning*. Años después, John McCarthy creó el «AI Lab» en la Universidad de Stanford (1963).

Sin embargo, en los años siguientes, la IA no logró dar en esos primeros momentos los frutos esperados con el nivel de inversiones económicas realizadas. Uno de los proyectos más ambiciosos y de mayor interés para el gobierno norteamericano se realizó en el contexto de la Guerra Fría con el intento del uso de la IA para la traducción simultánea del ruso al inglés. El informe estadounidense ALPAC, en 1966, puso de manifiesto la falta de avances en muchos de los proyectos con IA financiados por el gobierno estadounidense, por lo que fueron cancelados.

Más tarde, en 1973, el informe «Lighthill» del gobierno británico destacó las decepciones de la investigación en IA, por lo que la inversión en IA se redujo considerablemente. Este período hasta 1980 es lo que se denomina el «primer invierno de la IA».

En la década de los 80, Digital Equipment Corporations creó un sistema comercial experto R1 (XCON) diseñado para gestionar y configurar los pedidos de nuevos sistemas informáticos, el cual significó un nuevo auge en las inversiones y proyectos en IA. Grandes empresas de Estados Unidos y Japón hicieron fuertes inversiones en IA para el diseño y producción de sistemas expertos LISP, pero en 1987 surgieron otras alternativas más baratas, por lo que las empresas perdieron su interés en la IA. Esto constituyó el «segundo invierno de la IA».

En 1997, la IA Deep Blue de IBM triunfó sobre el campeón mundial de ajedrez Gary Kasparov, lo que se considera como un hito en la historia de la IA al ser la

primera vez que una máquina derrota a un hombre. Este hecho significó el resurgimiento de la inteligencia artificial apoyado por nuevos desarrollos tecnológicos, entre los que cabría destacar el conseguido por Google en 2008 en el campo del reconocimiento de voz. Andrew Ng, en 2012, entrenó con Deep Learning, una red neuronal alimentada con 10 millones de vídeos de YouTube, usados como datos para reconocer un gato sin que se le hubiese previamente enseñado.

En 2016, otro hito vino de la mano del sistema AlphaGo de Google DeepMind, que triunfó sobre Lee Sedol, el campeón de Go. La IA también logró grandes hitos con DeepMind AlphaStar en Starcraft u OpenAI Five en Dota 2.

Tipos de inteligencia artificial

La IA se divide en IA fuerte, científica o Artificial General Intelligence (AGI), e IA débil o tecnológica (o *narrow*). La AGI trataría de crear sistemas que sean capaces de realizar cualquier trabajo intelectual emulando a la Inteligencia Humana, para lo que sería necesario que dicho sistema de IA pueda igualar las capacidades de la IH y desarrollar una conciencia propia que le permitiese interactuar con el mundo, así como ser capaz de abstraerse para tener una representación interna del propio sistema. A día de hoy, esto es imposible.

En lo que respecta a la IA tecnológica, actualmente la mayoría de sus aplicaciones se centran en el desarrollo de sistemas que realicen una única tarea de manera exitosa. Existen dos variantes de la IA tecnológica, los sistemas expertos y el Machine Learning (ML) o aprendizaje máquina.

En el ML se dice que «el sistema puede aprender» si transforma un cierto mensaje de entrada en uno de salida, de acuerdo con algún principio de transformación, si tal principio está sujeto a cierto criterio de validación de su funcionamiento, y si el método de transformación se ajusta de forma que tienda a mejorar el funcionamiento del sistema de acuerdo con ese criterio. Los sistemas expertos tratan de crear comportamientos inteligentes, en base al conocimiento de un experto en el campo que ayuda a diseñar un programa, o conjunto de reglas, para realizar una determinada tarea (Michalski *et al.*, 1983). El ML clásico utiliza un conjunto de variables seleccionadas manualmente para entrenar un modelo que no haya sido programado explícitamente para esa tarea. Dentro del ML existen modelos que están inspirados en el cerebro y utilizan la neurona como elemento básico de procesamiento de la información, por lo que se les denomina Redes de Neuronas Artificiales (RNA) (I. Goodfellow *et al.*, 2014; Michalski *et al.*, 1983). Este modelo emula el funcionamiento del cerebro, ya que extrae las características de

manera automática a partir de los ejemplos (datos) que se le muestran en la fase de entrenamiento y, por ello, también se le denomina *Representation Learning*, o aprendizaje de representación, ya que, después de entrenar los modelos, estos representan las características de los datos en su estado interno.

Dentro de los modelos de *Representation Learning* existe una división en función del tamaño, o profundidad de la red, los modelos *shallow*, o poco profundos, y los modelos de *Deep Learning*, o de aprendizaje profundo, los cuales son capaces de representar de forma jerárquica las características, al igual que hace el cerebro (I. Goodfellow *et al.*, 2016).

En resumen, el *Machine Learning* (aprendizaje automático) y el *Deep Learning* (aprendizaje profundo) son las dos técnicas principales de IA que se utilizan en la actualidad. El *Machine Learning* es una categoría de IA y el *Deep Learning* es una técnica de *Machine Learning*.

El *Machine Learning* es el proceso en el que una máquina se alimenta con datos, sobre los cuales utilizará técnicas de análisis para aprender a ejecutar una tarea. Para ello no necesita millones de líneas de código de programación específica, por eso se denomina aprendizaje «automático». Si la máquina se alimenta con datos etiquetados, se dice que el aprendizaje es «supervisado», mientras que si, por el contrario, los datos que se le suministran no son etiquetados, se dice que el aprendizaje es «no supervisado».

El *Deep Learning* es un tipo de *Machine Learning* directamente inspirado en la arquitectura de las neuronas del cerebro humano. Una red neuronal artificial está compuesta por múltiples capas, a través de las cuales se procesan los datos. Esto es lo que permite que la máquina «profundice» en su aprendizaje, identificando conexiones y alterando los datos introducidos para conseguir los mejores resultados.

Dentro del *Machine Learning*, tenemos que diferenciar además lo que se denomina IA generativa y la IA predictiva.

La IA generativa utiliza modelos de aprendizaje automático para crear nuevos contenidos a partir de datos existentes. Utiliza datos y algoritmos para aprender nueva información y, basándose en el método de ensayo y error, se ajusta en consecuencia y se adapta hasta el punto de que el sistema operativo no necesita instrucciones específicas. Utilizando muchos datos, puede responder con creaciones aparentemente «nuevas», como vídeos, imágenes, audio y texto, cuando se le proporcionan instrucciones. Uno de los usos de la IA generativa es crear y producir materiales sencillos, pero que requieren mucho tiempo para que los

escriba un humano (ChatGPT). La IA generativa tiene varios inconvenientes potenciales, como la dependencia de la calidad de los datos de entrada y el riesgo de sesgos implícitos, así como riesgos de seguridad.

La IA predictiva se centra principalmente en la previsión, ya sea de patrones, tendencias futuras o acontecimientos. La IA predictiva utiliza también aprendizaje automático para aprender y adaptarse, personalizando la experiencia del cliente y pronosticando eventos futuros para las empresas. Sin embargo, mientras que la IA generativa utiliza el ML para crear nuevos contenidos, la IA predictiva utiliza el ML para realizar análisis estadísticos, con el fin de anticipar e identificar sucesos futuros. En lugar de redes neuronales, la IA predictiva se basa en modelos estadísticos y algoritmos de aprendizaje automático simples para recopilar grandes cantidades de datos (*Big Data*) y proporcionar predicciones basadas en esos datos. Analizando las actividades pasadas de sus empleados o clientes, la IA predictiva puede ayudarle a descubrir qué artículos necesitan reponerse, anticiparse a los comportamientos de los usuarios y tomar decisiones sobre atención médica.

Similitud entre las redes neuronales artificiales y el cerebro humano

Antes de abordar el funcionamiento preciso de las RNA, es necesario establecer un *paralelismo con las neuronas biológicas*.

El cerebro humano, con 1,5 kg, está compuesto por miles de millones de células nerviosas, de las cuales solo en el cerebro tenemos 86.000 millones como neuronas y unos 96.000 millones de otros tipos celulares (astrocitos, células de la microglía, etc.) (Acevedo *et al.*). Es una red de neuronas biológicas en la que cada neurona tiene entre 1.000 y 10.000 interacciones sinápticas con otras neuronas, lo que significa que en momentos de máxima actividad cerebral, tanto sensorial como perceptiva, podrían llegar a tener lugar del orden de entre 100 y 1.000 billones (10^{14} - 10^{15}) de transmisiones sinápticas con una duración del orden de 1 ms.

Desde el punto de vista de la estructura, y para entender la analogía establecida entre las RNA y las neuronas biológicas, estas últimas tienen (Figura 1) dendritas, soma, axón y terminales axónicos (botones sinápticos), siendo estos últimos la parte funcional donde se produce la secreción de los distintos neurotransmisores.

El cuerpo celular (soma) es el centro de control que procesa la información recibida por las dendritas. Las dendritas son los «cables principales» por los que

reciben la información externa. El axón es el «cable conductor» que transmite la señal de salida del cuerpo celular a otras neuronas.

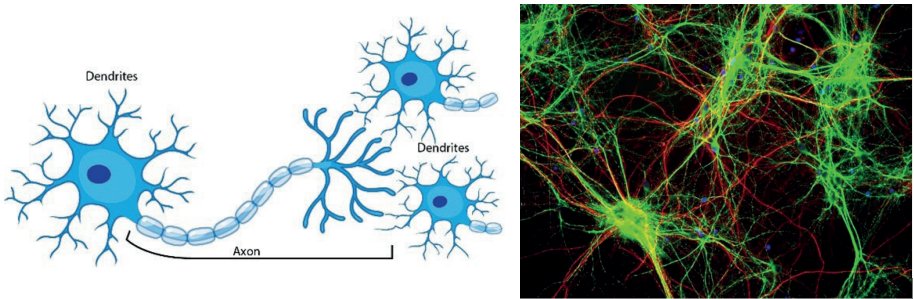


Figura 1. Estructura de las neuronas biológicas y redes que forman.

Por otro lado, los circuitos neuronales que forman las neuronas biológicas son complejos, ejecutan diferentes tipos de procesos (Figura 2) e involucran distintas clases neuronales, aumentando así la versatilidad y la variedad de la señal eléctrica (Hunnicuttt *et al.*).

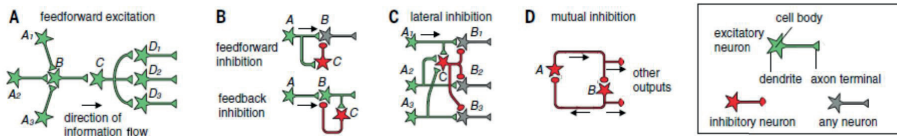


Figura 2. Ejemplos de tipos de circuitos entre neuronas biológicas.

En cuanto a la transmisión de la señal entre una neurona y la siguiente o sinapsis, actúan como enlaces y pesos entre las neuronas y, por tanto, permiten que estas se comuniquen entre sí.

Hay diferentes tipos de neuronas dependiendo de su morfología (unipolares, bipolares, multipolares, pseudopolares), según su función (sensoriales, motoras, de asociación), según el neurotransmisor que secretan (glutamatérgicas, colinérgicas, gabaérgicas, dopaminérgicas, histaminérgicas, glicinérgicas, serotoninérgicas, encefalinérgicas, etc.).

Por otro lado, la mayoría de las neuronas no secretan un único neurotransmisor, sino que pueden sintetizar y secretar dos o más neurotransmisores dependiendo de la frecuencia de la estimulación que reciban, por lo que la presencia de más

de un neurotransmisor confiere una versatilidad considerable a la transmisión sináptica, ya que, además, no es necesario que los neurotransmisores se liberen simultáneamente. Por ejemplo, la estimulación de baja frecuencia suele liberar solo neurotransmisores pequeños, mientras que la estimulación de alta frecuencia es necesaria para liberar neuropéptidos en las mismas terminales presinápticas. Como resultado, las propiedades de señalización química de una sinapsis suelen cambiar según el nivel de actividad presináptica.

A esto último es necesario añadir que en la transmisión sináptica biológica no intervienen solamente neuronas, sino que otras células, como los astrocitos y células de la glía, pueden modificar la transmisión de la señal desde una neurona a la siguiente, tal y como se puede apreciar en la figura siguiente (Figura 3).

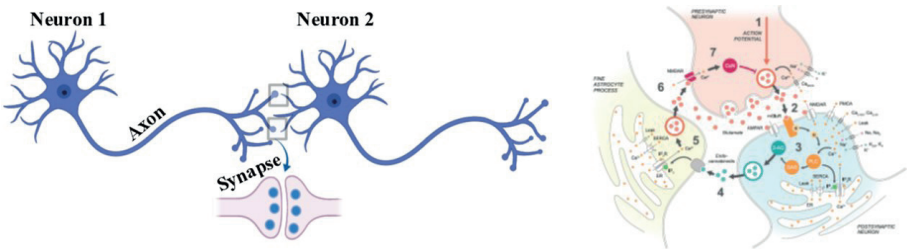


Figura 3. Modelo esquemático de una sinapsis con neurona pre y postsináptica y un astrocito.

¿Cuál es la similitud y diferencia entre las redes de neuronas biológicas y las RNA?

Las neuronas biológicas tienen un centro de control (cuerpo celular o SOMA) que resume la información recogida por las dendritas. El centro de control devuelve entonces una respuesta eléctrica (potencial de acción) según las siguientes reglas:

Si la suma de entrada no supera el umbral de excitación de la neurona, entonces no hay mensaje nervioso a través del axón. Si la suma de entrada supera el umbral de excitación, se envía un mensaje nervioso a través del axón (esa es la idea, pero en realidad es un poco más complicado).

En el caso de las RNA, la neurona artificial se denomina perceptrón y es una unidad neuronal de la RNA (Figura 4). Es un algoritmo que recibe datos (x_i), los pondera (asigna pesos w_i), calcula su suma $f(\sum x_i \cdot w_i)$ y presenta un resultado por medio de una función f no lineal. Una RNA está compuesta por numerosas neuronas

artificiales (perceptrones) constituidas en capas (capa de entrada, capas «escondidas» y capa de salida) como analogía matemática a las redes neuronales biológicas (o del cerebro). Esta elección del algoritmo es lo más parecido al funcionamiento de las neuronas biológicas: Sinapsis/dendritas (x_i); ponderación de cada elemento de entrada w_i ; Cuerpos celulares: aplicación de una función de activación f a la suma de las entradas ponderadas ($\sum x_i w_i$); Axón = salida como resultado del modelo.

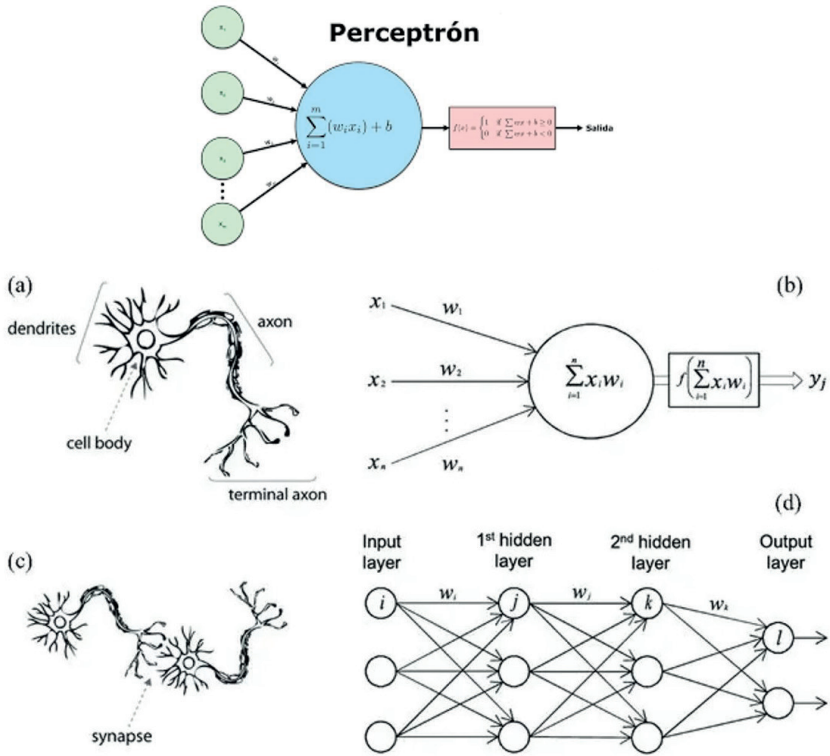


Figura 4. Comparación de una neurona biológica (a) y el perceptrón (b) y de una sinapsis biológica (c) con una RNA (d).

El siguiente ejemplo (Figura 5) ilustra el funcionamiento del reconocimiento de dígitos de una imagen mediante una red neuronal.

1. La capa de entrada recibe los datos como imágenes, texto o números (Figura 8). Cada nodo de entrada representa una característica o una variable de los datos. Por ejemplo, si los datos son una imagen de un dígito escrito a mano, cada nodo de entrada puede representar un valor de píxel de la imagen.
2. La capa de entrada transmite los datos a la capa oculta, formada por neuronas artificiales. Cada neurona artificial tiene asociado un peso y un sesgo, que determinan el grado de influencia de la entrada en la salida.
3. La neurona artificial calcula una suma ponderada de las entradas, añade el sesgo y aplica una función de activación para producir una salida. La función de activación es una función no lineal que determina cómo responde la neurona artificial a la entrada. Por ejemplo, una función de activación común es la función sigmoidea, que asigna cualquier entrada a un valor entre 0 y 1.
4. La salida de la neurona artificial pasa a la siguiente capa, que puede ser otra capa oculta o la capa de salida. La capa de salida produce el resultado final de la red neuronal, como una predicción, una clasificación o una generación. La capa de salida puede tener uno o más nodos en función de la tarea. Por ejemplo, supongamos que la tarea consiste en clasificar el dígito manuscrito en una de 10 clases. En ese caso, la capa de salida puede tener diez nodos, cada uno de los cuales representa una probabilidad de que el dígito pertenezca a esa clase.

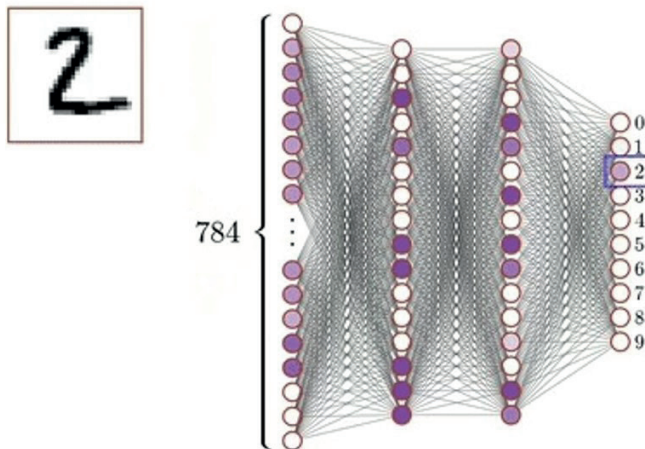


Figura 5. RNA para el reconocimiento de imágenes de dígitos.

5. La RNA se entrena para ajustar los pesos y sesgos de las neuronas artificiales en función de los datos y la salida deseada, lo que se denomina aprendizaje u optimización, empleando diversos algoritmos, como el descenso gradiente, la retropropagación o los algoritmos genéticos. El objetivo del proceso de aprendizaje es minimizar el error o la diferencia entre la salida real y la salida deseada. Una función de pérdida mide el error (p.e. el error cuadrático medio que calcula la media de las diferencias al cuadrado entre la salida real y la deseada) cuantificando el rendimiento de la red neuronal en los datos.

Hasta la fecha, los científicos han desarrollado innumerables modelos de redes neuronales con abrumadora sofisticación mediante nodos multicapa, ampliando la versatilidad de la inteligencia artificial.

Diferencias entre neuronas y RNA

Las redes neuronales artificiales (RNA) son una gran simplificación del cerebro humano y, aun así, son potentes herramientas informáticas. Entre las principales diferencias entre redes neuronales reales y las RNA, cabe destacar:

Las RNA utilizan conexiones ponderadas entre neuronas, mientras que las neuronas biológicas no tienen pesos explícitos, ya que tienen un mayor grado de complejidad y capacidad de aprendizaje y adaptación a través de mecanismos intrínsecos. Los pesos en las redes neuronales artificiales se optimizan mediante entrenamiento para producir la salida deseada. El uso de pesos y algoritmos de aprendizaje es lo que permite entrenar las redes neuronales artificiales para tareas específicas.

Por otro lado, las neuronas artificiales utilizan funciones de activación matemáticas como sigmoide, ReLU, etc. Las neuronas biológicas se activan de una forma no lineal más compleja basada en procesos electroquímicos.

En lo que respecta a la memoria, las neuronas biológicas mantienen su estado interno y tienen memoria gracias a mecanismos como la potenciación a largo plazo. Las neuronas artificiales no suelen mantener el estado entre entradas.

La mayor simplificación es que **las redes neuronales reales son asíncronas, mientras que las RNA son síncronas**. Además, las neuronas reales también tienen un componente temporal que afecta tanto a la actividad de disparo como al aprendizaje. También importa cuándo se producen esas descargas presinápticas. No ocurren simultáneamente como en el caso de una RNA (Figura 6).

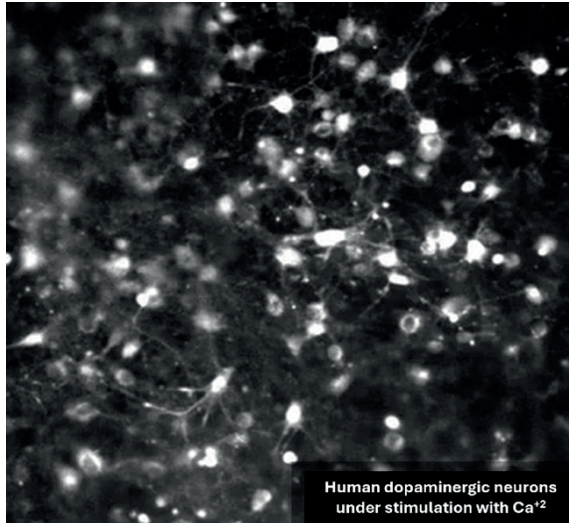


Figura 6. Red de neuronas humanas dopaminérgicas estimuladas con Ca^{+2} .

El tiempo de recorrido de las señales a través de la arborescencia dendrítica varía considerablemente, y lo mismo ocurre en las arborescencias axónicas, ya que algunos axones son mucho más largos que otros.

A lo largo del tiempo, cada neurona real ha adquirido diferentes umbrales de filtrado de señal y de carga general en la dendrita, lo que afecta a las circunstancias precisas en las que se disparará.

En los circuitos cerebrales existen formas a corto y largo plazo de modular las conexiones de las neuronas reales, hay procesos retardados, de acción más prolongada, que realizan una ponderación de las sinapsis, de tal forma que la historia previa de la frecuencia de disparo de una neurona determinada afecta al comportamiento futuro de esa neurona para una entrada determinada. Sin embargo, mientras que en la mayoría de las RNA solo se actualizan los pesos, mediante por ejemplo retropropagación, las propias funciones son estáticas.

En cuanto al arquitectura de red, las redes neuronales artificiales están diseñadas por los investigadores. Sin embargo, las redes neuronales biológicas tienen una arquitectura que surge a través del desarrollo y la evolución.

La disposición de las neuronas artificiales en una RNA es en gran medida *feed-forward*. Hay una entrada (datos) y una salida (*output*).

Los circuitos más interesantes del cerebro son recurrentes, es decir, circulares, recursivos, que se retroalimentan por sí solos. La transmisión de señales de una neurona presináptica a una postsináptica está modulada por cambios en los niveles de neurotransmisores y hormonas, tanto a nivel global como local. No existe un análogo para esto en las RNA.

La gran mayoría de las RNA están organizadas como jerarquías estrictas de cálculo, mientras que en el cerebro tanto las señales excitatorias como las inhibitorias cruzan regularmente las jerarquías en el cerebro, lo que es crítico para su funcionalidad.

Es necesario destacar que en determinados procesos cerebrales una neurona humana suficientemente estimulada puede dispararse a sí misma, enviando un impulso electroquímico por su propio axón, llegando a muchas neuronas distribuidas a larga distancia para proporcionarles estímulo.

A lo largo del tiempo, cada neurona ha adquirido diferentes umbrales de filtrado de señal y de carga general en la dendrita, lo que afecta a las circunstancias precisas en las que se disparará. Este proceso forma parte de la neuroplasticidad y adaptabilidad cerebral.

En lo que respecta a los mecanismos de aprendizaje, las redes neuronales artificiales aprenden mediante algoritmos de entrenamiento y descenso de gradiente. Las neuronas biológicas aprenden y se adaptan mediante mecanismos como el aprendizaje Hebbiano y la plasticidad sináptica dependiente de la experiencia.

En definitiva, las RNA son una interpretación del procesamiento hacia delante de los circuitos visuales del cerebro. No se parecen en nada al procesamiento retrospectivo (quizá más importante) de la visión, ni a los procesos de memoria a corto y largo plazo del hipocampo, ni a los procesos de homeostasis regulados por grupos neuronales de la corteza prefrontal.

Diferencias entre ambos tipos de inteligencia

Conocimiento, abstracción y conciencia

En la actualidad, ni siquiera los sistemas basados en inteligencia artificial más avanzados son capaces de replicar los procesos multisensoriales que realiza habitualmente el cerebro humano, aunque aún no se conoce bien cómo llevan a cabo estos procesos los circuitos neuronales del cerebro. La cognición, entendida como las habilidades de la Inteligencia Humana para recibir, procesar, almacenar

y recuperar información como memoria y aprendizaje, que posteriormente utilizamos para el razonamiento y la toma de decisiones, es, a día de hoy, algo inaccesible para la IA. Así mismo, la abstracción y la introspección son capacidades únicas de la Inteligencia Humana, las cuales nos permiten, a partir de nuestras experiencias, emociones y sentimientos, descubrir nuestras sensibilidades y poder tener una visión subjetiva de nosotros mismos (autoconciencia), a la vez que ser conscientes de la interacción con los demás. Hallazgos recientes arrojan nueva luz sobre los mecanismos neuronales oscilatorios que median en la integración multisensorial en los estados introspectivos y de la percepción consciente (Hense *et al.*).

Aprendizaje y adaptabilidad

La Inteligencia Humana es capaz de ajustar sus perspectivas en respuesta a las condiciones cambiantes de su entorno. Tiene capacidad de adaptarse, aprender y mantener en memoria habilidades, lo que le permitirá recordar información y sobresalir en diversas actividades. Evoluciona constantemente y es capaz de asumir nuevas situaciones y retos en circunstancias cambiantes.

La Inteligencia Artificial no es creativa y no se puede adaptar a entornos cambiantes, porque depende de datos de entrada y reglas fijadas por el ser humano que las crea.

Creatividad y generación de conocimiento

Los humanos tenemos la capacidad de crear y transferir habilidades entre diversas tareas y sintetizar información de varias fuentes. Por otro lado, la capacidad de ejercer un juicio sensato es esencial para la multitarea, como demuestra el hecho de hacer malabarismos con diversos trabajos a la vez. Puede describirse como inventivo o creativo.

La Inteligencia Artificial no puede descubrir formas creativas de resolver diversos problemas crecientes, porque su código no lo admite. Es bien sabido que los robots están codificados para trabajar dentro de su marco. La IA no es capaz de pensar de forma crítica en escenarios complicados, analizar el contexto y desarrollar estrategias complejas.

Empatía

La Inteligencia Humana permite expresar y compartir sentimientos, emociones, experiencia y conocimientos, y «leer» a las personas al instante, lo que posibilita entre otras cosas el trabajo en equipo. Utilizando como vectores de interacción

perceptiva el lenguaje, la voz, la mímica y el tacto, nuestro cerebro puede crear conciencia individual y de grupo, facilitando, una vez más, nuestra interacción y adaptación al entorno social. Son importantes, a este respecto, las neuronas de asociación en el cerebro humano, que permiten integrar todos los estímulos que recibimos del exterior y conservarlos en diferentes formas de memoria.

Actualmente, a pesar de los recientes desarrollos y progresos hechos en IA con Lenguaje Natural Asistido (NPL), capaces de emular el habla humana, es difícil imaginar que la IA pueda mostrar todas las expresiones, contextualizar y leer las emociones de otras personas de forma más favorable que un ser humano.

Intuición, anticipación e instintos viscerales

La Inteligencia Humana está dotada de unas cualidades únicas, como son la intuición y la anticipación. Se trata de una combinación de percepción sensorial, experiencias, conocimientos y procesamiento subconsciente, que nos permite ejecutar acciones rápidas e informadas en ausencia de información completa, simplemente estimuladas por determinadas variables del entorno. Estudios recientes demuestran que hay circuitos cerebrales específicos y complejos encargados de estos procesos y de generar respuestas reflejas conscientes (Bortfeld *et al.*; Martínez *et al.*). En muchos casos, esta intuición y anticipación están acompañadas por instintos viscerales, que son a menudo la fuerza que guía los actos innovadores, las aventuras empresariales de éxito e incluso las decisiones que salvan vidas.

Sin embargo, la IA se basa en algoritmos y razonamientos lógicos, mientras que el ser humano posee una capacidad innata para confiar en su intuición y emitir juicios que van más allá de la pura lógica.

Toma de decisiones éticas

La ética y la moral son parte de la esencia de la sociedad humana e influyen en nuestras elecciones y acciones. Aunque la IA puede seguir normas y directrices programadas, carece del razonamiento ético intrínseco y de la responsabilidad moral que poseen los seres humanos. La inteligencia humana nos permite elegir y tomar decisiones, asumiendo sus consecuencias ante dilemas morales complejos, evaluando desde múltiples perspectivas, matices culturales e impactos sociales a largo plazo. Los humanos llevamos el peso de la responsabilidad de nuestras decisiones, mientras que los sistemas de IA dependen de la intervención y supervisión humanas para garantizar una conducta ética, que en ciertas circunstancias puede ser susceptible de sesgos.

Arquitectura

Finalmente, desde el punto de vista de la «arquitectura de red», es necesario destacar que la IA se basa en el ensamblaje de redes neuronales artificiales más o menos complejas, que funcionan con lenguaje de máquina en un ordenador.

Sin embargo, la Inteligencia Humana se construye sobre una anatomía funcional del cerebro, que se va desarrollando a lo largo de nuestra existencia y que «madura» a medida que «crecemos» biológicamente, y aprendemos y adquirimos experiencia y conocimiento. Tal y como se observa en la Figura 7, cuando se analiza la densidad de la materia gris por técnicas de imagen, a medida que nos desarrollamos, nuestro córtex cerebral experimenta cambios en la densidad de la materia gris, destacando que las cortezas de asociación de orden superior maduran solo después de que se desarrollen las cortezas somatosensoriales y visuales de orden inferior cuyas funciones integran (Gogtay *et al.*). Así pues, se observa que la secuencia de maduración de la corteza coincide con los hitos regionales relevantes del desarrollo cognitivo y funcional. Las partes del cerebro asociadas a las funciones más básicas maduraron pronto: las áreas motoras y sensoriales maduraron primero, seguidas de las áreas implicadas en la orientación espacial, el desarrollo del habla y el lenguaje y la atención (lóbulos parietales superior e inferior). Más tarde, maduran las áreas implicadas en la función ejecutiva, la atención y la coordinación motora (lóbulos frontales) (Drakulich *et al.*; Lu *at al.*: Richards *et al.*).

El sustrato celular de estos cambios corticales puede ser una combinación de mielinización, árbol dendrítico y cambios en la densidad neuronal, glial, vascular y de empaquetamiento de neuritas en diferentes láminas corticales.

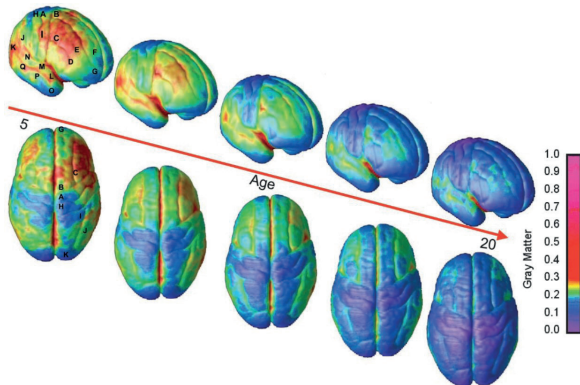


Figura 7. Vistas lateral derecha y superior de la secuencia de maduración de la materia gris sobre la superficie cortical.

Para finalizar, la complejidad de los circuitos que se desarrollan en el cerebro y que sustentan todas las capacidades y cualidades de la Inteligencia Humana se han estudiado en los últimos años por imagen para determinar el «CONECTOMA» del cerebro humano (<https://www.humanconnectome.org/>), y del cual se muestran imágenes en la Figura 8. Se puede observar la macroconectividad entre las diferentes regiones del cerebro de un adulto, como resultado de los centenares de millones de axones que lo cruzan, y que conectan a modo de «finas fibras» los hemisferios y lóbulos cerebrales.

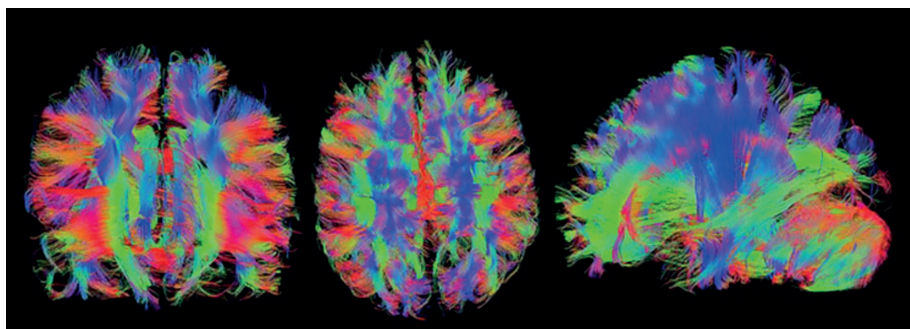


Figura 8. Imágenes del CONECTOMA humano.

Reflexión

Un estudio reciente del MIT (Schaeffer *et al.*) pide cautela al comparar las redes neuronales con el cerebro. Los sistemas informáticos que parecen generar una actividad similar a la del cerebro pueden deberse a que los investigadores los han guiado hacia un resultado concreto. El grupo del MIT que ha realizado el estudio llevó a cabo un análisis de más de 11.000 redes neuronales entrenadas para simular la función de las células reticulares, componentes clave del sistema de navegación del cerebro. Los investigadores descubrieron que las redes neuronales solo producían una actividad similar a la de las células reticulares cuando se les imponían restricciones muy específicas que no se encuentran en los sistemas biológicos.

Lo que esto sugiere es que, para obtener un resultado con células en forma de red, los investigadores que entrenaban los modelos necesitaban «cocinar» esos resultados con opciones de implementación específicas y biológicamente inverosímiles, afirma Rylan Schaeffer, investigador principal del MIT. Sin esas limitaciones,

el equipo del MIT descubrió que muy pocas redes neuronales generaban una actividad similar a la de las células reticulares, lo que implicaría que estos modelos no generan necesariamente predicciones útiles sobre el funcionamiento del cerebro.

Sin embargo, es necesario destacar que la IA es actualmente una herramienta muy importante en todo tipo de industrias y sistemas productivos, así como en el ámbito laboral técnico y en el científico, ya que sin lugar a duda permite automatizar multitud de tareas y simplificar procesos productivos. Es necesario, por ello, verla como una «compañera de viaje», no como una enemiga, y aprovechar todo lo que nos puede ayudar para disponer de recursos y tiempo para formarnos en su manejo, así como en el diseño y análisis de nuevos usos en nuestro campo profesional. Sin duda, me atrevería a sugerir que, independientemente del área de conocimiento, todos los alumnos de todos los grados que impartimos en nuestra Universidad reciban formación en IA, al objeto de que puedan utilizarla de una forma ética para optimizar el acceso al conocimiento y les permita desarrollar un espíritu crítico de liderazgo.

BIBLIOGRAFÍA

- AZEVEDO, F. A.; CARVALHO, L. R.; GRINBERG, L. T.; FARFEL, J. M., FERRETTI, R. E.; LEITE, R. E.; JACOB FILHO, W.; LENT, R.; HERCULANO-HOUZEL, S. (2009 Apr 10). *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain*. *J Comp Neurol.* ; 513(5): 532-41. doi: 10.1002/cne.21974. PMID: 19226510.
- BORTFELD, H. & BUNGE, S. A. (2024). Brain Development. In *Fundamentals of Developmental Cognitive Neuroscience* (pp. 84-113). Chapter, Cambridge: Cambridge University Press.
- DRAKULICH, S.; KARAMA, S. *Structural Brain Imaging of Intelligence*. In: Barbey AK, Karama S, Haier RJ, eds. *The Cambridge Handbook of Intelligence and Cognitive Neuroscience*. Cambridge University Press; 2021:210-234.
- GOODFELLOW, I., et al. *Generative adversarial nets*, in 'Advances in Neural Information Processing Systems 27', Curran Associates. (2014): 2672-2680. Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (nº. 2). MIT press.
- GOGTAY, N.; GIEDD, J. N.; LUSK, L.; HAYASHI, K. M.; GREENSTEIN, D.; VAITUZIS, A. C.; NUGENT, T. E.; HERMAN, D. H.; CLASEN, L. S.; TOGA, A. W.; RAPOPORT, J. L; THOMPSON, P. M. (2004 May 25). *Dynamic mapping of human cortical development during childhood through early adulthood*. *Proc Natl Acad Sci USA*. 101(21):8174-9. doi: 10.1073/pnas.0402680101
- HENSE, A.; PETERS, A.; BRUCHMANN, M.; DELLERT, T.; STRAUBE, T. (2024). *Electrophysiological correlates of sustained conscious perception*. *Sci Rep*. 8;14(1):10593. doi: 10.1038/s41598-024-61281-2. PMID: 38719939; PMCID: PMC11078977
- HUNNICUTT, B.; KRZYWINSKI, M. (2016). *Neural circuit diagrams*. *Nat Methods* 13, 189. <https://doi.org/10.1038/nmeth.3777>
- LENROOT, R. K.; GIEDD, J. N. (2006). *Brain development in children and adolescents: insights from anatomical magnetic resonance imaging*. *Neurosci Biobehav Rev*. 30(6): 718-29. doi: 10.1016/j.neubiorev. 2006.06.001
- LU, L. H.; SOWELL, E. R. *Morphological development of the brain: what has imaging told us?*In: RUMSEY, J. M.; ERNST, M., eds. *Neuroimaging in Developmental Clinical Neuroscience*. Cambridge University Press; 2009:5-21.

- MARTÍNEZ, K., COLOM, R. *Imaging the Intelligence of Humans*. In: BARBEY, A. K., KARAMA, S., HAIER, R. J., eds. *The Cambridge Handbook of Intelligence and Cognitive Neuroscience*. Cambridge University Press; 2021:44-69.
- MCCULLOCH, W. S., PITTS, W. *A logical calculus of the ideas immanent in nervous activity*. *Bulletin of Mathematical Biophysics* 5, 115-133 (1943). <https://doi.org/10.1007/BF02478259>
- MICHALSKI, RYSZARD & CARBONELL, JAIME & MITCHELL, T. (1983). *Machine learning. an Artificial Intelligence approach*. Volume 2. 10.1007/978-3-662-12405-5
- RICHARDS, J. E., CONTE, S. *Brain Development in Infants: Structure and Experience*. In: LOCKMAN, J. J., TAMIS-LEMONDA, C. S., eds. *The Cambridge Handbook of Infant Development: Brain, Behavior, and Cultural Context*. Cambridge Handbooks in Psychology. Cambridge University Press; 2020:94-127.
- RYLAN SCHAEFFER. *Computer Science, Stanford University. Brain and Cognitive Sciences, Massachusetts Institute of Technology. McGovern Institute for Brain Research. 2nd AI Science Workshop at the 39th International Conference on Machine Learning (ICML), 2022.*
- SZEGEDY, C.; ZAREMBA, W.; SUTSKEVER, I.; BRUNA, J.; ERHAN, D.; GOODFELLOW, I. (2014). *Intriguing properties of neural networks*. 2nd International Conference on Learning Representations, ICLR 2014-Banff, Canada.

