

Replicación de principios neurobiológicos en la inteligencia artificial: Aproximación al estado del arte

Aidee Alejandra Velázquez Favela

aavelazquezf@uaemex.mx

Paola Monserrath Velázquez Favela

pmvelazquezf@uaemex.mx

Oscar Salgado Flores

osalgadof773@profesor.uaemex.mx

Universidad Autónoma del Estado de México

Recepción: 02 de junio de 2026

Aceptación: 30 de junio de 2026

Publicación: 15 de julio de 2026

Resumen

El presente artículo desarrolla un análisis documental del estado del arte sobre los principios neurobiológicos que han servido de inspiración para el diseño y funcionamiento de las redes neuronales artificiales (RNA) empleadas en inteligencia artificial (IA). Mediante una revisión sistemática de literatura científica especializada en neurociencia, ciencias cognitivas e inteligencia artificial, se examinan procesos biológicos como la plasticidad sináptica, la poda neuronal, la codificación distribuida y la organización jerárquica de la información, así como su correspondencia con mecanismos computacionales utilizados en modelos contemporáneos de aprendizaje automático y aprendizaje profundo. Asimismo, se analizan las principales similitudes y diferencias entre las redes neuronales humanas y las RNA, considerando aspectos relacionados con el aprendizaje, la representación del conocimiento, la capacidad de generalización, la eficiencia energética y la adaptabilidad. Los resultados muestran que, aunque la IA ha logrado replicar de manera parcial algunos principios funcionales del cerebro humano, persisten diferencias sustanciales asociadas con la plasticidad estructural, el razonamiento contextual, la integración multimodal y la conciencia. Se concluye que las RNA

constituyen modelos inspirados en la neurobiología más que réplicas de los procesos cognitivos humanos. Finalmente, se destaca la importancia de promover una comprensión interdisciplinaria de la relación entre neurociencia e inteligencia artificial, particularmente en el ámbito educativo, con el propósito de favorecer una alfabetización digital crítica y fundamentada sobre el alcance y las limitaciones de estas tecnologías.

Palabras clave: inteligencia artificial, redes neuronales artificiales, neurociencia, plasticidad sináptica, aprendizaje profundo.

Replication of Neurobiological Principles in Artificial Intelligence: An Approach to the State of the Art

Abstract

This article presents a documentary state-of-the-art analysis of the neurobiological principles that have inspired the design and operation of artificial neural networks (ANNs) used in artificial intelligence (AI). Through a systematic review of specialized scientific literature in neuroscience, cognitive sciences, and artificial intelligence, biological processes such as synaptic plasticity, neuronal pruning, distributed coding, and hierarchical information organization are examined, together with their correspondence to computational mechanisms employed in contemporary machine learning and deep learning models. The study also analyzes the main similarities and differences between human neural networks and ANNs, considering aspects related to learning, knowledge representation, generalization capacity, energy efficiency, and adaptability. The findings indicate that, although AI has partially replicated certain functional principles of the human brain, substantial differences remain regarding structural plasticity, contextual reasoning, multimodal integration, and consciousness. It is concluded that ANNs should be understood as neurobiologically inspired models rather than replicas of human cognitive processes. Finally, the article highlights the importance of promoting an interdisciplinary understanding of the relationship between neuroscience and artificial intelligence,

particularly in educational settings, in order to foster critical digital literacy and a well-informed perspective on the scope and limitations of these technologies.

Keywords: artificial intelligence, artificial neural networks, neuroscience, synaptic plasticity, deep learning.

Introducción

Durante las últimas dos décadas, la inteligencia artificial (IA) ha experimentado un crecimiento acelerado que ha transformado significativamente la investigación científica, la industria tecnológica, los sistemas productivos y los procesos educativos. Dentro de este desarrollo, las redes neuronales artificiales (RNA) se han consolidado como uno de los principales fundamentos del aprendizaje automático, al demostrar una notable capacidad para resolver tareas relacionadas con el reconocimiento de patrones, la clasificación de información, la visión computacional, la traducción automática y el procesamiento del lenguaje natural. Estos avances han generado un interés creciente por comprender hasta qué punto los sistemas de IA reproducen mecanismos inspirados en el funcionamiento del cerebro humano.

Las RNA surgieron históricamente como modelos computacionales inspirados en principios generales de organización neuronal. Aunque las neuronas biológicas y los nodos artificiales difieren sustancialmente en su naturaleza y complejidad, ambos sistemas comparten ciertas características funcionales, como el procesamiento distribuido de información, la transmisión de señales mediante conexiones y la capacidad de modificar dichas conexiones a partir de la experiencia. Diversos autores han señalado que conceptos fundamentales de la neurociencia, como la plasticidad sináptica, la organización jerárquica del procesamiento y la representación distribuida del conocimiento, han servido como referencia para el diseño y evolución de las arquitecturas contemporáneas de aprendizaje profundo (Goodfellow et al., 2016; Kandel et al., 2021; Richards et al., 2019).

No obstante, la relación entre inteligencia artificial y neurociencia continúa siendo objeto de debate. Mientras algunos enfoques sostienen que las RNA constituyen únicamente aproximaciones matemáticas simplificadas del aprendizaje biológico, otros consideran que el estudio del cerebro humano sigue proporcionando información valiosa para el desarrollo de sistemas artificiales más eficientes, adaptativos y robustos. Esta discusión ha dado origen a campos emergentes como NeuroAI, orientados a integrar hallazgos neurocientíficos en el diseño de nuevas arquitecturas computacionales y, simultáneamente, utilizar modelos de IA para comprender mejor determinados procesos cognitivos.

En el ámbito educativo, particularmente en el Nivel Medio Superior, comprender la relación entre redes neuronales biológicas y artificiales adquiere una relevancia creciente. La incorporación de herramientas basadas en IA en actividades de enseñanza, aprendizaje, evaluación y producción de conocimiento exige que docentes y estudiantes desarrollen una comprensión crítica acerca de los fundamentos que sustentan estas tecnologías. Más allá de su uso instrumental, resulta necesario conocer los principios científicos que explican sus capacidades, alcances y limitaciones.

En este contexto, el presente artículo tiene como propósito analizar el estado del arte relacionado con los principios neurobiológicos que han servido de inspiración para el desarrollo de las redes neuronales artificiales. Mediante una investigación documental basada en literatura científica especializada en neurociencia, psicología cognitiva e inteligencia artificial, se examinan los mecanismos biológicos que han sido adaptados a modelos computacionales contemporáneos, así como las similitudes, diferencias y desafíos que persisten entre ambos sistemas.

Finalmente, se discuten las implicaciones educativas, científicas, tecnológicas y éticas de estas comparaciones, subrayando la necesidad de un enfoque interdisciplinario que vincule la neurociencia, la psicología, la computación y la

pedagogía para formar ciudadanos críticos y docentes actualizados frente a un mundo en el que la inteligencia artificial forma parte del quehacer cotidiano. En este sentido, diversos organismos internacionales han señalado que la alfabetización en inteligencia artificial constituye una competencia emergente indispensable para la educación contemporánea (UNESCO, 2021).

Desarrollo

Principios neurobiológicos incorporados a las redes neuronales artificiales

El cerebro humano procesa información mediante la interacción simultánea de miles de millones de neuronas organizadas en redes funcionales altamente complejas. Este procesamiento no ocurre de manera lineal, sino a través de la integración dinámica de señales excitatorias e inhibitorias que permiten la percepción, la memoria, el aprendizaje y la toma de decisiones (Purves et al., 2019). Desde mediados del siglo XX, diversos avances en neurociencia han contribuido a comprender cómo las neuronas modifican sus conexiones y generan comportamientos adaptativos, conocimientos que posteriormente sirvieron de inspiración para el desarrollo de las redes neuronales artificiales (RNA).

Entre los principios biológicos más relevantes se encuentran la plasticidad sináptica, la poda neuronal, la codificación distribuida, la organización jerárquica del procesamiento y los mecanismos de atención. Aunque las RNA no reproducen de manera exacta los procesos electroquímicos del sistema nervioso, sí incorporan representaciones matemáticas de varios de estos fenómenos, lo que explica el estrecho vínculo histórico entre neurociencia e inteligencia artificial.

Tabla 1. Principios neurobiológicos y su correspondencia en las redes neuronales artificiales

Principio neurobiológico	Descripción en el cerebro humano	
	humano	Correspondencia en IA
Plasticidad sináptica	Fortalecimiento o debilitamiento de conexiones neuronales a partir de la experiencia.	Ajuste de pesos mediante entrenamiento y retropropagación del error.
Poda sináptica	Eliminación de conexiones poco utilizadas para optimizar recursos neuronales.	Técnicas de pruning para reducir parámetros y mejorar eficiencia.
Codificación distribuida	La información se almacena en patrones de activación distribuidos entre múltiples neuronas.	Representaciones distribuidas en capas ocultas y espacios latentes.
Organización jerárquica	Procesamiento progresivo desde estímulos simples hasta representaciones complejas.	Arquitecturas profundas (Deep Learning) con múltiples capas.
Atención selectiva	Priorización de estímulos relevantes según contexto y objetivos.	Mecanismos Attention y arquitecturas Transformer.
Aprendizaje por experiencia	Adaptación continua mediante interacción con el entorno.	Aprendizaje supervisado, no supervisado y por refuerzo.

Fuente: Elaboración propia con base en Kandel et al. (2021), Purves et al. (2019), Goodfellow et al. (2016), Han et al. (2015) y Vaswani et al. (2017).

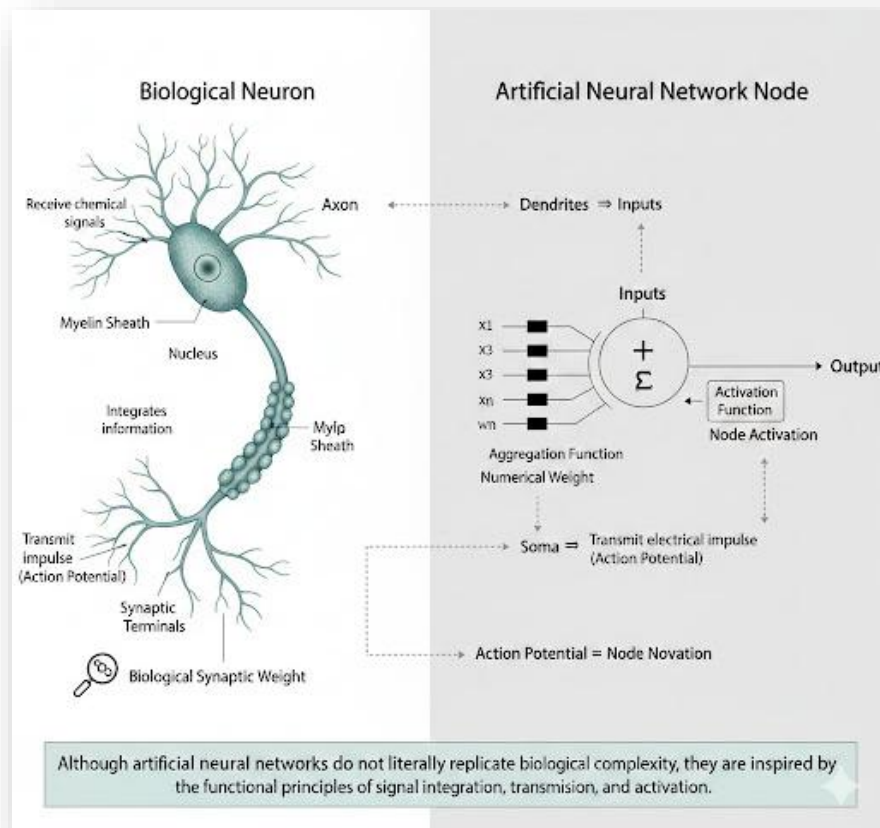
Como puede observarse, los principales avances de las redes neuronales artificiales se han sustentado en la adaptación computacional de mecanismos previamente identificados en el funcionamiento cerebral. Sin embargo, estas adaptaciones representan simplificaciones matemáticas de procesos biológicos considerablemente más complejos.

La plasticidad sináptica constituye uno de los mecanismos fundamentales del aprendizaje humano. Este fenómeno permite fortalecer o debilitar las conexiones neuronales en función de la experiencia acumulada, favoreciendo la adquisición de nuevas habilidades y la consolidación de la memoria. Desde los trabajos pioneros de Hebb (1949), numerosos estudios han demostrado que las modificaciones sinápticas representan la base biológica del aprendizaje y de la adaptación conductual.

Complementariamente, la poda sináptica contribuye a optimizar la eficiencia del sistema nervioso mediante la eliminación selectiva de conexiones poco utilizadas. Este proceso permite reorganizar los circuitos neuronales y conservar aquellas rutas funcionales que resultan más eficaces para responder a las demandas del entorno. Asimismo, la información cerebral no se almacena en neuronas aisladas, sino en patrones distribuidos de actividad que involucran múltiples regiones del sistema nervioso. Este principio, conocido como codificación distribuida, constituye una de las principales bases conceptuales de las redes neuronales artificiales contemporáneas.

Otro elemento relevante es la organización jerárquica del procesamiento cerebral. Diversas áreas corticales procesan la información en niveles progresivamente más complejos, permitiendo pasar de estímulos simples a representaciones abstractas. Esta característica inspiró el diseño de arquitecturas profundas capaces de identificar patrones cada vez más complejos conforme la información avanza a través de múltiples capas de procesamiento.

Figura 1. Estructura comparativa entre una neurona biológica y un nodo de red neuronal artificial



Nota. Figura elaborada mediante inteligencia artificial generativa (Gemini, Google, 2025), con fines ilustrativos.

La Figura 1 ilustra de manera esquemática la analogía funcional que existe entre una neurona biológica y una neurona artificial. Aunque ambos sistemas reciben información, la procesan y generan una respuesta, los mecanismos mediante los cuales realizan estas funciones difieren sustancialmente en términos biológicos y computacionales.

Redes neuronales artificiales y mecanismos de aprendizaje

Las redes neuronales artificiales surgieron como modelos matemáticos diseñados para representar de manera simplificada algunos principios observados en las redes neuronales biológicas. Una RNA está integrada por unidades de procesamiento denominadas neuronas artificiales, organizadas en capas de entrada, capas ocultas y capas de salida. Cada conexión posee un peso numérico que determina la influencia de una señal sobre otra, mientras que las funciones de activación regulan la respuesta producida por cada nodo durante el procesamiento de la información (Goodfellow et al., 2016).

El aprendizaje de una RNA se basa en la modificación iterativa de estos pesos mediante algoritmos de optimización. Entre ellos, la retropropagación del error (backpropagation) constituye uno de los mecanismos más utilizados. Este procedimiento ajusta gradualmente los parámetros internos del modelo con el propósito de minimizar las diferencias entre las respuestas generadas y los resultados esperados. Aunque este proceso difiere sustancialmente de los mecanismos neurobiológicos reales, mantiene una analogía funcional con la plasticidad sináptica observada en el cerebro humano.

El desarrollo del aprendizaje profundo (Deep Learning) permitió ampliar significativamente las capacidades de las RNA. Mediante múltiples capas de procesamiento, estos modelos pueden construir representaciones jerárquicas de los datos, identificar patrones complejos y resolver tareas relacionadas con visión computacional, procesamiento de lenguaje natural, reconocimiento de voz y generación automática de contenidos. Los avances alcanzados durante las últimas dos décadas han consolidado al aprendizaje profundo como uno de los pilares fundamentales de la inteligencia artificial contemporánea.

De manera similar a la poda neuronal observada en el cerebro humano, las RNA incorporan técnicas de optimización conocidas como pruning, cuyo propósito consiste en eliminar conexiones redundantes o poco relevantes sin afectar significativamente el desempeño del modelo. Diversas investigaciones han demostrado que estas estrategias permiten reducir la complejidad computacional y mejorar la eficiencia energética de los sistemas de aprendizaje automático.

A pesar de estas similitudes conceptuales, las diferencias entre ambos sistemas continúan siendo profundas. Las neuronas biológicas transmiten información mediante procesos electroquímicos complejos, mientras que las neuronas artificiales realizan operaciones matemáticas. Asimismo, el cerebro humano posee una capacidad extraordinaria para aprender a partir de pocos ejemplos, adaptarse a contextos cambiantes e integrar simultáneamente información sensorial, emocional y social. En contraste, la mayoría de las RNA requieren grandes volúmenes de datos y presentan limitaciones importantes para generalizar conocimientos fuera de los escenarios para los cuales fueron entrenadas.

Avances recientes en NeuroAI

Durante los últimos años ha emergido una línea de investigación interdisciplinaria conocida como NeuroAI, cuyo objetivo consiste en integrar descubrimientos neurocientíficos dentro del diseño de nuevas arquitecturas de inteligencia artificial. Este campo busca comprender de manera más profunda cómo aprende el cerebro humano para desarrollar sistemas artificiales más eficientes, adaptativos y robustos. Entre las principales áreas de trabajo destacan el estudio de arquitecturas inspiradas en columnas corticales, mecanismos de atención, aprendizaje continuo, plasticidad computacional y representación jerárquica de la información. Estas investigaciones pretenden superar algunas limitaciones actuales de las RNA,

especialmente aquellas relacionadas con la transferencia del aprendizaje, la eficiencia energética y la adaptación a entornos dinámicos.

Un avance particularmente relevante corresponde a los modelos Transformer, cuya arquitectura incorpora mecanismos de atención capaces de priorizar información relevante durante el procesamiento. Diversos autores han señalado que estos sistemas presentan analogías funcionales con ciertos procesos atencionales observados en la memoria de trabajo humana, aunque sin reproducir completamente su complejidad biológica.

En conjunto, los desarrollos recientes muestran que la relación entre neurociencia e inteligencia artificial continúa siendo bidireccional. Por una parte, los conocimientos sobre el cerebro inspiran nuevas tecnologías computacionales; por otra, las herramientas de inteligencia artificial contribuyen al análisis y modelado de procesos neuronales cada vez más complejos. Esta convergencia interdisciplinaria constituye uno de los campos de investigación más prometedores para comprender tanto la cognición humana como la evolución futura de los sistemas inteligentes.

Enfoque Metodológico

La presente investigación se desarrolló bajo un enfoque cualitativo mediante un estudio documental de carácter descriptivo, analítico y comparativo. Su propósito fue examinar el estado actual del conocimiento científico relacionado con los principios neurobiológicos que han servido de inspiración para el diseño y desarrollo de las redes neuronales artificiales empleadas en inteligencia artificial. Este tipo de investigación resulta pertinente cuando se busca comprender, interpretar y sintetizar información proveniente de diversas disciplinas, particularmente en campos donde convergen la neurociencia, la psicología cognitiva, la computación y la inteligencia artificial (Hernández-Sampieri y Mendoza, 2018). El estudio se enmarca en la modalidad de estado del arte, ya que identifica avances, tendencias, convergencias y divergencias presentes en la literatura científica especializada, al tiempo que

adopta una perspectiva comparativa para analizar las semejanzas y diferencias entre los mecanismos de aprendizaje de las redes neuronales humanas y los implementados en las redes neuronales artificiales.

La revisión documental se llevó a cabo mediante una búsqueda sistemática de literatura científica realizada entre enero y marzo de 2026, considerando publicaciones aparecidas principalmente entre 2008 y 2025, además de algunas obras clásicas cuya relevancia histórica y conceptual continúa siendo ampliamente reconocida. La búsqueda se efectuó en las bases de datos SciELO, RedALyC, Scopus, Web of Science, IEEE Xplore, SpringerLink y ScienceDirect, utilizando descriptores en español e inglés relacionados con redes neuronales artificiales, redes neuronales humanas, plasticidad sináptica, aprendizaje profundo, NeuroAI, aprendizaje continuo, mecanismos de atención y biological neural networks, combinados mediante operadores booleanos para ampliar y refinar la recuperación de información.

Como resultado de este proceso se identificó un conjunto inicial de publicaciones que posteriormente fue sometido a un proceso de depuración. Se eliminaron registros duplicados y documentos que no abordaban de manera directa la relación entre neurociencia e inteligencia artificial, conservándose únicamente artículos científicos arbitrados, libros académicos, revisiones sistemáticas y documentos especializados con información bibliográfica verificable y disponibilidad de consulta. Asimismo, se excluyeron publicaciones de divulgación sin respaldo académico, referencias cuya localización no pudo corroborarse y documentos cuyo contenido no guardaba relación directa con los objetivos del estudio. La versión final del corpus documental quedó integrada por las fuentes consideradas de mayor pertinencia, actualidad y solidez metodológica para sustentar el análisis desarrollado.

Una vez seleccionadas las referencias, se realizó un análisis temático orientado a identificar conceptos, procesos y categorías recurrentes relacionados con la plasticidad sináptica, la codificación distribuida, la poda neuronal, la representación

jerárquica de la información, los mecanismos de atención y las estrategias de aprendizaje presentes tanto en sistemas biológicos como artificiales. Posteriormente, estos elementos fueron organizados y contrastados para reconocer convergencias funcionales, diferencias estructurales y tendencias emergentes en el campo de la NeuroAI. La denominada revisión empírica consistió en el análisis crítico de investigaciones experimentales previamente publicadas; por ello, no se desarrollaron experimentos propios, sino que se examinaron evidencias generadas por otros investigadores acerca del funcionamiento del sistema nervioso, los mecanismos de aprendizaje y el desempeño de modelos computacionales inspirados en el cerebro humano. Este procedimiento permitió integrar resultados provenientes de distintas disciplinas y construir una síntesis interpretativa sustentada en evidencia científica verificable (Okoli, 2015).

Con el propósito de fortalecer la confiabilidad del estudio se privilegiaron fuentes publicadas por editoriales académicas, revistas científicas arbitradas e instituciones de reconocido prestigio, además de realizar una triangulación teórica entre investigaciones de neurociencia, psicología cognitiva e inteligencia artificial. Como apoyo complementario durante las etapas de organización de la información, revisión del estilo académico y elaboración de recursos gráficos se utilizaron herramientas de inteligencia artificial generativa. Su participación se limitó exclusivamente a funciones de asistencia en la redacción y representación visual de información, mientras que la selección de las fuentes, el análisis crítico, la interpretación de los hallazgos y la elaboración de las conclusiones fueron realizados íntegramente por los autores. En particular, la Figura 1 fue elaborada con apoyo de Gemini (Google, 2025), utilizándose únicamente como herramienta para la representación gráfica de la comparación entre una neurona biológica y un nodo de red neuronal artificial; el contenido científico representado fue seleccionado, verificado e interpretado por los autores.

Finalmente, dado que la investigación se fundamentó exclusivamente en el análisis de documentos previamente publicados, no involucró participantes humanos ni

animales, por lo que no requirió consentimiento informado ni autorización de un comité de ética. No obstante, durante todo el proceso se observaron los principios de honestidad académica, integridad científica, transparencia en el uso de inteligencia artificial generativa y correcta citación de las fuentes consultadas. Si bien los resultados dependen de la disponibilidad y calidad de la literatura existente, el carácter sistemático de la búsqueda, la depuración de las referencias y el análisis comparativo realizado permiten ofrecer una visión actualizada y metodológicamente consistente sobre la incorporación de principios neurobiológicos en las redes neuronales artificiales.

Resultados

El análisis documental permitió identificar que la evolución de las redes neuronales artificiales (RNA) ha estado fuertemente influenciada por principios derivados del estudio del sistema nervioso humano. A partir de la revisión de la literatura especializada se reconocieron cinco ejes neurobiológicos que han servido como fundamento para el diseño de los modelos contemporáneos de inteligencia artificial: el procesamiento distribuido de la información, la plasticidad sináptica, la organización jerárquica del procesamiento, los mecanismos de atención y la optimización de conexiones mediante procesos análogos a la poda sináptica. Si bien estas correspondencias representan abstracciones computacionales y no reproducciones fisiológicas del cerebro, constituyen la base conceptual de los principales algoritmos de aprendizaje profundo (Goodfellow et al., 2016; LeCun et al., 2015).

En términos metodológicos, la revisión permitió observar que la mayoría de los estudios coinciden en señalar que las RNA replican únicamente funciones específicas del procesamiento neuronal, principalmente aquellas relacionadas con el reconocimiento de patrones, la clasificación de información y el aprendizaje basado en la modificación iterativa de parámetros (Rumelhart et al., 1986). En contraste, capacidades como la reorganización estructural permanente, la integración emocional, la conciencia y el razonamiento contextual continúan siendo

características exclusivas del cerebro humano (Kandel et al., 2021; Purves et al., 2019).

Los hallazgos obtenidos se sintetizan en la Tabla 2, donde se comparan doce procesos cognitivos humanos con sus equivalentes computacionales, y en la Figura 1, que ilustra la correspondencia conceptual entre una neurona biológica y un nodo de una red neuronal artificial. Ambos organizadores permiten apreciar que las semejanzas se concentran en la estructura funcional del procesamiento, mientras que las diferencias se relacionan con la naturaleza biológica, la eficiencia energética, la capacidad de adaptación y la complejidad cognitiva de ambos sistemas.

Tabla 2. Comparación entre procesos cognitivos humanos y mecanismos equivalentes en inteligencia artificial

Categoría	Procesos humanos	cognitivos	Procesamiento en sistemas de IA
Percepción	Integración multisensorial de información visual, auditiva, táctil y contextual.		Procesamiento de datos digitales mediante sensores o entradas estructuradas.
Atención	Selección dinámica de estímulos relevantes según objetivos y contexto.		Mecanismos de atención que asignan diferentes pesos a la información procesada.
Memoria de corto plazo	Retención	temporal de información para tareas inmediatas.	Estados temporales o ventanas de contexto utilizadas durante el procesamiento.
Memoria de largo plazo	Almacenamiento	de conocimientos, experiencias y significados.	Parámetros entrenados que representan patrones estadísticos aprendidos.
Aprendizaje	Modificación de conductas y conocimientos mediante experiencia.		Ajuste de parámetros mediante algoritmos de optimización.

Categoría	Procesos cognitivos humanos	Procesamiento en sistemas de IA
Generalización	Aplicación flexible de conocimientos a situaciones nuevas.	Inferencia estadística dependiente de datos previamente observados.
Razonamiento	Integración de lógica, experiencia y contexto.	Procesamiento algorítmico basado en reglas o patrones aprendidos.
Creatividad	Producción de ideas originales influida por experiencias y emociones.	Generación de nuevas combinaciones a partir de patrones estadísticos.
Emoción	Participación activa en la percepción, memoria y toma de decisiones.	Ausencia de emociones reales; simulación basada en patrones.
Lenguaje	Comunicación simbólica con intención, contexto y significado.	Predicción probabilística de secuencias lingüísticas.
Toma de decisiones	Influida por valores, experiencias y factores sociales.	Optimización de funciones objetivo previamente definidas.
Plasticidad	Reorganización estructural y funcional del sistema nervioso.	Actualización de parámetros dentro de arquitecturas predefinidas.

Fuente: Elaboración propia a partir de la literatura revisada.

Finalmente, el análisis evidenció una tendencia creciente hacia el desarrollo de modelos inspirados de manera más estrecha en la neurociencia, particularmente dentro del campo denominado NeuroAI (Richards et al., 2019). No obstante, la evidencia revisada coincide en que las arquitecturas actuales aún representan aproximaciones parciales al funcionamiento cerebral y mantienen importantes limitaciones para reproducir procesos cognitivos de alto nivel.

Discusión

Los resultados obtenidos confirman que la relación entre las redes neuronales humanas y las redes neuronales artificiales debe entenderse como una inspiración conceptual más que como una equivalencia biológica. Aunque la literatura reconoce que numerosos avances en inteligencia artificial han surgido a partir de principios identificados por la neurociencia, el análisis realizado muestra que la mayor parte de las similitudes corresponden a mecanismos funcionales simplificados cuya implementación responde a modelos matemáticos y no a procesos fisiológicos reales (LeCun et al., 2015; Richards et al., 2019).

Asimismo, la revisión pone de manifiesto que la incorporación de conceptos como la plasticidad, la atención y el procesamiento jerárquico ha permitido desarrollar algoritmos cada vez más eficientes; sin embargo, persisten diferencias sustanciales relacionadas con la capacidad de aprendizaje autónomo, la adaptación estructural, el razonamiento contextual y la eficiencia energética del cerebro humano (Kandel et al., 2021; Strubell et al., 2019). Estos hallazgos coinciden con la postura de diversos autores que advierten que las RNA constituyen modelos inspirados en la biología, pero no representan una reproducción integral de la cognición humana.

Desde la perspectiva educativa, esta distinción resulta especialmente relevante para el Nivel Medio Superior, ya que favorece una comprensión crítica del funcionamiento de la inteligencia artificial y evita interpretaciones que atribuyen a estos sistemas capacidades propias de la inteligencia humana. En consecuencia, el estudio respalda la necesidad de promover una alfabetización digital interdisciplinaria que integre conocimientos de neurociencia, computación y ciencias cognitivas para interpretar con mayor precisión los alcances y limitaciones de la IA. Como ocurre en toda investigación documental, las conclusiones dependen de la calidad y disponibilidad de la literatura analizada. Aunque se privilegiaron fuentes arbitradas y de acceso verificable, el acelerado desarrollo del campo implica que

nuevos hallazgos puedan modificar las interpretaciones actuales. En este sentido, futuras investigaciones podrían complementar esta revisión mediante estudios empíricos que analicen la incorporación de principios neurobiológicos emergentes en nuevas arquitecturas de inteligencia artificial y evalúen sus implicaciones en contextos educativos.

Conclusiones

La revisión documental realizada permitió analizar de manera integral los principios neurobiológicos que han servido de inspiración para el desarrollo de las redes neuronales artificiales y, en consecuencia, para gran parte de los avances contemporáneos en inteligencia artificial. A partir de la evidencia científica examinada, se confirma que la relación entre neurociencia e inteligencia artificial ha sido determinante para la construcción de modelos computacionales capaces de aprender, reconocer patrones y procesar información de manera cada vez más sofisticada.

Los resultados muestran que diversos mecanismos presentes en el cerebro humano, tales como la plasticidad sináptica, la codificación distribuida, la organización jerárquica del procesamiento y la optimización mediante poda de conexiones, han influido directamente en el diseño de las redes neuronales artificiales. Estos principios han sido adaptados al ámbito computacional mediante algoritmos de aprendizaje, ajustes de pesos, arquitecturas profundas y mecanismos de atención que permiten a los sistemas artificiales mejorar su desempeño a partir de la experiencia acumulada.

Sin embargo, el análisis comparativo evidencia que las semejanzas entre ambos sistemas deben interpretarse como analogías funcionales y no como equivalencias estructurales o cognitivas. Mientras que las redes neuronales artificiales operan mediante procedimientos matemáticos definidos, el cerebro humano constituye un sistema biológico altamente dinámico, caracterizado por procesos electroquímicos,

plasticidad estructural continua, integración multisensorial y una estrecha interacción entre cognición, emoción, experiencia y contexto.

Asimismo, se identificó que algunas de las capacidades más distintivas de la inteligencia humana continúan sin ser replicadas por los sistemas de inteligencia artificial actuales. Entre ellas destacan la conciencia, la intencionalidad, la comprensión contextual profunda, la creatividad genuina, la toma de decisiones basada en valores y la capacidad de aprender eficazmente a partir de un número reducido de experiencias. Estas diferencias confirman que, pese a los avances alcanzados durante las últimas décadas, la inteligencia artificial sigue siendo una aproximación parcial y especializada frente a la complejidad de la cognición humana.

La literatura revisada también muestra que el campo emergente de la NeuroAI representa una de las líneas de investigación más prometedoras para el futuro desarrollo tecnológico. La integración de conocimientos provenientes de la neurociencia, la ciencia cognitiva y la inteligencia artificial podría contribuir al diseño de sistemas más eficientes, adaptativos y capaces de aproximarse de manera más realista a ciertos procesos observados en el cerebro humano. No obstante, la evidencia disponible sugiere que aún existe una brecha considerable entre las capacidades de los modelos artificiales y la riqueza funcional de los sistemas biológicos.

Desde la perspectiva educativa, comprender la relación entre redes neuronales humanas y artificiales adquiere una relevancia creciente. En un contexto donde las herramientas basadas en inteligencia artificial forman parte de las actividades académicas cotidianas, resulta indispensable que docentes y estudiantes comprendan tanto sus fundamentos como sus limitaciones. Esta comprensión favorece una alfabetización digital crítica, permite desmitificar ideas erróneas sobre el funcionamiento de la IA y promueve un uso más informado, ético y responsable de estas tecnologías.

Finalmente, el estudio pone de manifiesto la importancia de fortalecer enfoques interdisciplinarios que articulen conocimientos de neurociencia, computación, matemáticas, psicología cognitiva y filosofía de la mente. Solo mediante una comprensión integral de estas disciplinas será posible interpretar con mayor precisión el alcance actual de la inteligencia artificial, valorar críticamente sus aplicaciones y contribuir a la formación de ciudadanos capaces de participar de manera reflexiva en una sociedad cada vez más influida por el desarrollo tecnológico.

Referencias

- Azevedo, F. A. C., Carvalho, L. R. B., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., Jacob Filho, W., Lent, R., & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5), 532–541. <https://doi.org/10.1002/cne.21974>
- Bliss, T. V. P., & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, 232(2), 331–356. <https://doi.org/10.1113/jphysiol.1973.sp010273>
- Citri, A., & Malenka, R. C. (2008). Synaptic plasticity: Multiple forms, functions, and mechanisms. *Neuropsychopharmacology*, 33(1), 18–41. <https://doi.org/10.1038/sj.npp.1301559>
- Frankle, J., & Carbin, M. (2019). The lottery ticket hypothesis: Finding sparse, trainable neural networks. *arXiv*. <https://arxiv.org/abs/1803.03635>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://www.deeplearningbook.org>
- Google. (2025). Gemini [Modelo de inteligencia artificial generativa]. <https://gemini.google.com>

- Han, S., Pool, J., Tran, J., & Dally, W. (2015). Learning both weights and connections for efficient neural networks. arXiv. <https://arxiv.org/abs/1506.02626>
- Hernández-Sampieri, R., & Mendoza, C. P. (2018). Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta. McGraw-Hill.
- Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, 387(2), 167–178. [https://doi.org/10.1002/\(SICI\)1096-9861\(19971020\)387:2<167::AID-CNE1>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z)
- Kandel, E. R., Koester, J. D., Mack, S. H., & Siegelbaum, S. A. (2021). Principles of neural science (6th ed.). McGraw-Hill.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37(43), 879–910. <https://doi.org/10.17705/1CAIS.03743>
- Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A. S., Mooney, R. D., Platt, M. L., & White, L. E. (2019). Neuroscience (6th ed.). Oxford University Press.
- Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R. P., de Berker, A., Ganguli, S., Gillon, C. J., Hafner, D., Kepecs, A., Kriegeskorte, N., Lillicrap, T., Marblestone, A., Pehlevan, C., Roelfsema, P., Sacramento, J., ... Kording, K. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, 22(11), 1761–1770. <https://doi.org/10.1038/s41593-019-0520-2>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. arXiv. <https://arxiv.org/abs/1906.02243>
- UNESCO. (2021). Recomendación sobre la ética de la inteligencia artificial. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. arXiv. <https://arxiv.org/abs/1706.03762>