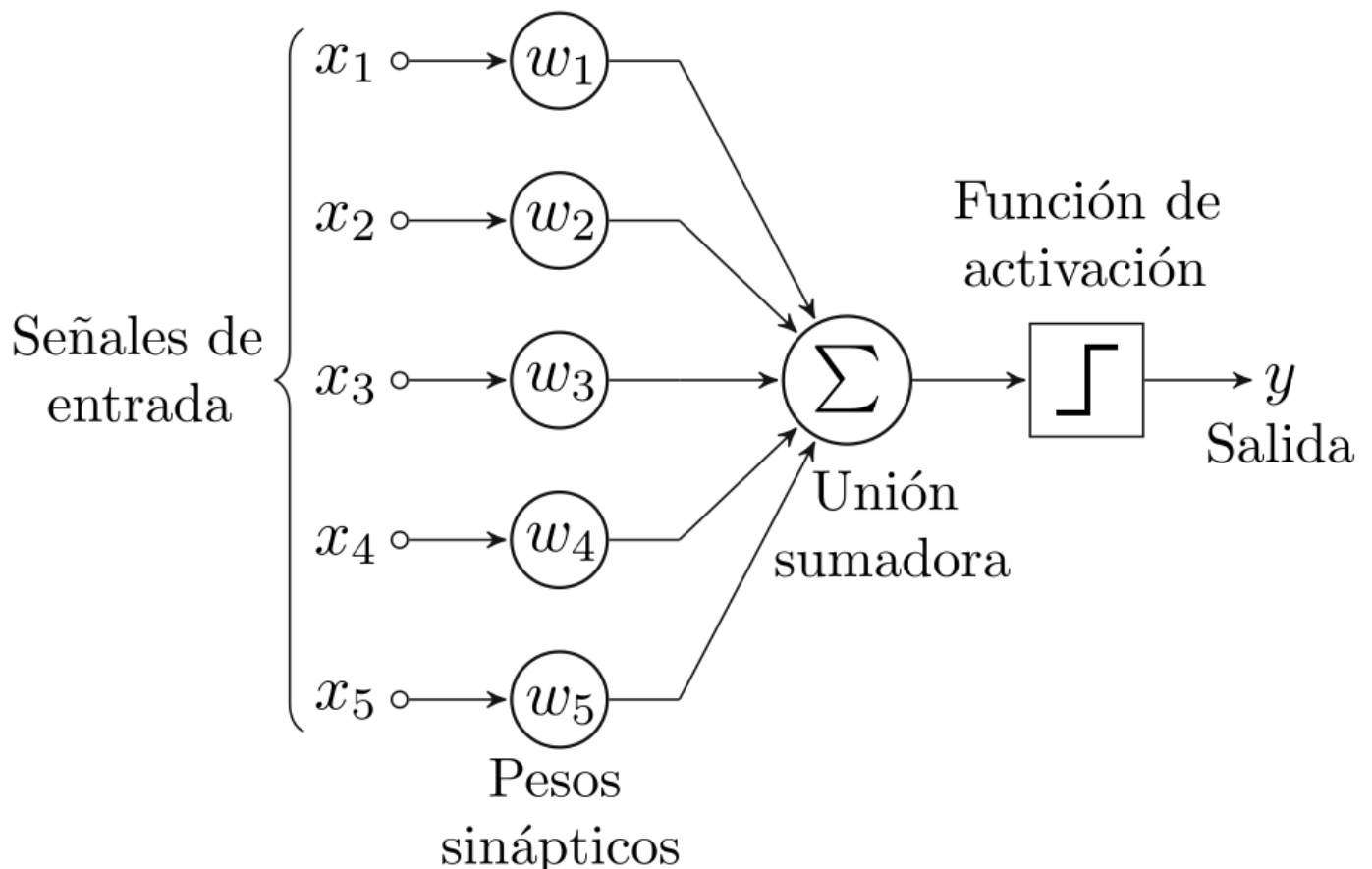


2.1 Una historia de la IA Generativa

Los orígenes, el perceptrón

La idea de que las máquinas puedan aprender no es completamente nueva. Ya en los años 40 y 50 algunos científicos comenzaron a explorar esta posibilidad.

Uno de los primeros modelos fue el **perceptrón**, desarrollado en 1957 por el psicólogo y científico Frank Rosenblatt. El perceptrón estaba inspirado en el funcionamiento de las neuronas del cerebro y trataba de aprender a clasificar información a partir de ejemplos.



Esquema de funcionamiento del perceptrón según Wikipedia

Curiosamente, esta idea tenía una inspiración biológica mucho más antigua. A finales del siglo XIX, el científico español **Santiago Ramón y Cajal** había descrito el funcionamiento de las neuronas y las conexiones entre ellas. Décadas después, muchos investigadores en inteligencia artificial utilizaron estos descubrimientos para intentar construir modelos computacionales que imitaran ese tipo de conexiones.

Durante varios años el avance fue lento, porque los ordenadores no tenían suficiente potencia de cálculo ni existían grandes conjuntos de datos para entrenar los modelos. Sin embargo, a partir de la década de 2010 se produjo un gran salto gracias al aumento de la capacidad computacional y al desarrollo de nuevas arquitecturas de redes neuronales.

De los primeros modelos al auge de la IA generativa

Para entender por qué hoy hablamos de **modelos de lenguaje, IA generativa o agentes inteligentes**, es útil mirar un poco atrás y recorrer brevemente la historia del aprendizaje automático. Lo interesante es que muchas de las ideas fundamentales **no son nuevas**; lo que ha cambiado radicalmente en los últimos años es la **combinación de tres factores clave: datos masivos, potencia de cálculo y nuevos algoritmos**.

Uno de los primeros hitos importantes fue el **perceptrón**, desarrollado a finales de los años cincuenta por **Frank Rosenblatt**. El perceptrón era un modelo matemático inspirado en las neuronas del cerebro que podía aprender a clasificar patrones simples ajustando pesos a partir de ejemplos. Fue uno de los primeros sistemas capaces de **aprender directamente de los datos**, lo que sentó las bases de las redes neuronales modernas.

Sin embargo, estos primeros modelos tenían limitaciones importantes. Un perceptrón simple solo podía aprender relaciones relativamente sencillas entre variables. Durante décadas, el desarrollo de las redes neuronales avanzó lentamente, en parte porque **no existía suficiente capacidad de cálculo ni grandes conjuntos de datos** para entrenar modelos más complejos. Durante los años setenta y parte de los ochenta incluso se produjo lo que se conoce como un **“invierno de la inteligencia artificial”**, un periodo de escepticismo y menor financiación en la investigación.

A finales de los años ochenta y principios de los noventa, investigadores como **Yann LeCun**, **Geoffrey Hinton** y **Yoshua Bengio** empezaron a demostrar que las redes neuronales podían ser mucho más potentes de lo que se pensaba. LeCun, por ejemplo, desarrolló redes neuronales convolucionales capaces de reconocer dígitos escritos a mano, una tecnología que llegó a utilizarse para leer cheques bancarios y códigos postales en sistemas reales.

En paralelo, el mundo comenzó a ver demostraciones públicas del potencial de la inteligencia artificial. Uno de los momentos más simbólicos ocurrió en **1997**, cuando el sistema **Deep Blue** de IBM derrotó al campeón mundial de ajedrez Garry Kasparov. Este evento no implicaba todavía

aprendizaje profundo como el que conocemos hoy, pero mostró hasta qué punto los sistemas informáticos podían superar a los humanos en tareas complejas bien definidas.

A partir de la década de 2010 se produjo un cambio decisivo. La combinación de **grandes cantidades de datos, GPUs capaces de entrenar redes muy grandes y nuevos métodos de aprendizaje profundo** permitió que las redes neuronales alcanzaran niveles de rendimiento nunca vistos. Este periodo es el que conocemos como la **revolución del deep learning**.

Uno de los ejemplos más conocidos llegó en 2016 con **AlphaGo**, desarrollado por DeepMind, que derrotó a campeones humanos en el complejo juego del Go, considerado durante décadas demasiado difícil para las máquinas. Este tipo de sistemas combinaba aprendizaje profundo, aprendizaje por refuerzo y grandes cantidades de datos para desarrollar estrategias propias.

Otro hito muy relevante ocurrió en el ámbito científico con **AlphaFold**, también de DeepMind. Este sistema logró predecir la estructura tridimensional de proteínas, un problema que llevaba más de cincuenta años siendo uno de los grandes desafíos de la biología. En 2020 sus predicciones alcanzaron una precisión comparable a métodos experimentales de laboratorio, lo que supuso un avance enorme para la investigación biomédica.

Mientras tanto, en el campo del lenguaje natural se produjo otro salto fundamental con la aparición de los **transformers**, una arquitectura de redes neuronales presentada en 2017 que permitió procesar grandes secuencias de texto de forma mucho más eficiente que los modelos anteriores. Los *transformers* sustituyeron progresivamente a las redes recurrentes que se utilizaban para modelar secuencias y permitieron entrenar modelos cada vez más grandes capaces de comprender y generar lenguaje con gran coherencia.

Esta arquitectura es la base de los **grandes modelos de lenguaje actuales (LLM)** y de muchas aplicaciones de IA generativa. Gracias a los transformers, a la disponibilidad de enormes colecciones de texto y a infraestructuras de cálculo masivo, hoy es posible entrenar modelos capaces de generar texto, imágenes, código o incluso música.

Mirando esta evolución con perspectiva, podemos ver que la situación actual no es el resultado de un único descubrimiento. Es más bien la convergencia de **ideas que llevan décadas desarrollándose**: desde el perceptrón de Rosenblatt hasta las redes profundas modernas, pasando por el trabajo de investigadores como LeCun, Bengio o Hinton.

En cierto modo, la inteligencia artificial actual es el resultado de una historia larga donde **las ideas estaban ahí desde hace mucho tiempo**, pero **la tecnología, los datos y los algoritmos necesarios para explotarlas plenamente no habían llegado todavía**. Solo cuando estos tres elementos se han combinado —datos masivos, potencia de cálculo y arquitecturas como los *transformers*— hemos podido alcanzar el punto en el que se encuentra hoy la IA generativa.

El estado actual de la inteligencia artificial: una visión de conjunto

Hoy en día, el campo de la **Inteligencia Artificial (IA)** puede entenderse como un sistema de capas o niveles que se apoyan unos sobre otros. En la parte más amplia se encuentra la **IA como disciplina científica**, cuyo objetivo general es desarrollar sistemas capaces de realizar tareas que normalmente asociamos con la inteligencia humana, como aprender, razonar o comprender lenguaje.

Dentro de ese gran campo se encuentra el **Machine Learning (ML)**. En lugar de programar reglas manualmente, el Machine Learning permite que las máquinas **aprendan patrones a partir de datos**. Esto ha permitido construir sistemas capaces de clasificar información, hacer predicciones o detectar relaciones entre variables.

A su vez, dentro del Machine Learning encontramos una subárea muy importante: el **Deep Learning (DL)**. Este enfoque utiliza **redes neuronales con múltiples capas** capaces de descubrir automáticamente patrones complejos en los datos, lo que ha permitido grandes avances en áreas como el reconocimiento de imágenes, voz o texto.

Uno de los campos donde el Deep Learning ha tenido mayor impacto es el **Procesamiento del Lenguaje Natural (NLP)**. Gracias a estas técnicas, los sistemas informáticos pueden analizar, interpretar y generar lenguaje humano. Durante décadas, los modelos de lenguaje fueron relativamente simples, basados en estadísticas de palabras. Sin embargo, los avances recientes han permitido construir modelos mucho más complejos.

El gran salto llegó con la arquitectura **transformer**, presentada en 2017, que permitió procesar grandes secuencias de texto de forma más eficiente y capturar relaciones complejas entre palabras. Esta arquitectura se convirtió en la base de muchos sistemas modernos de procesamiento del lenguaje.

Gracias a estos avances surgieron los **Large Language Models (LLM)**. Estos modelos se entrenan con enormes cantidades de texto y aprenden a **predecir la siguiente palabra en una secuencia**, lo que les permite generar textos coherentes, traducir idiomas o responder preguntas complejas.

A partir de los LLM han aparecido múltiples aplicaciones prácticas. Una de las más visibles son los **chatbots conversacionales**, capaces de interactuar con los usuarios en lenguaje natural y asistir en tareas como la generación de texto, la explicación de conceptos o la resolución de problemas.

En una etapa más reciente han surgido los llamados **agentes de inteligencia artificial**, que representan un paso adicional. Estos sistemas no solo generan respuestas, sino que también pueden **planificar acciones, utilizar herramientas externas y ejecutar tareas para**

alcanzar un objetivo. Cuando varios de estos agentes trabajan de forma coordinada, se habla de **sistemas multiagente**, donde distintos agentes colaboran para resolver problemas más complejos.

En conjunto, el panorama actual puede entenderse como una evolución encadenada:

- **IA** → el campo general que busca crear sistemas inteligentes
- **Machine Learning** → métodos que permiten aprender a partir de datos
- **Deep Learning** → redes neuronales profundas que descubren patrones complejos
- **NLP** → aplicación de estas técnicas al lenguaje humano
- **LLM** → grandes modelos entrenados con enormes corpus de texto
- **Chatbots y herramientas generativas** → aplicaciones que interactúan con usuarios
- **Agentes autónomos y sistemas multiagente** → sistemas capaces de ejecutar tareas y coordinar procesos

Esta evolución muestra cómo una serie de avances científicos y tecnológicos han ido construyendo el ecosistema actual de la inteligencia artificial. Lo que comenzó como modelos estadísticos relativamente simples ha terminado dando lugar a **sistemas capaces de generar texto, imágenes, código o incluso coordinar acciones complejas**, marcando una nueva etapa en la relación entre humanos y máquinas.

Revision #3

Created 2026-03-16 09:11:32 CET by Luis Hueso

Updated 2026-03-16 09:17:11 CET by Luis Hueso