¿Qué es la inteligencia artificial (IA)?



[Artificial Intelligence](https://www.ibm.com/es-es/think/artificial-intelligence)

**9 de agosto de 2024**

* [¿Qué es la IA?](https://www.ibm.com/es-es/think/topics/artificial-intelligence#%C2%BFQu%C3%A9+es+la+IA%3F%C2%A0)
* [Machine learning](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Machine+learning)
* [Deep learning](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Deep+learning)
* [Generative AI](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Generative+AI)
* [Cómo funciona la IA generativa](https://www.ibm.com/es-es/think/topics/artificial-intelligence#C%C3%B3mo+funciona+la+IA+generativa)
* [Beneficios de la IA](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Beneficios+de+la+IA%C2%A0)
* [Casos de uso de la IA](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Casos+de+uso+de+la+IA%C2%A0)
* [Retos y riesgos de la IA](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Retos+y+riesgos+de+la+IA%C2%A0)
* [Ética y gobierno de la IA](https://www.ibm.com/es-es/think/topics/artificial-intelligence#%C3%89tica+y+gobierno+de+la+IA%C2%A0)
* [IA débil vs. IA fuerte](https://www.ibm.com/es-es/think/topics/artificial-intelligence#IA+d%C3%A9bil+vs.+IA+fuerte%C2%A0)
* [Historia de la IA](https://www.ibm.com/es-es/think/topics/artificial-intelligence#Historia+de+la+IA%C2%A0)

**Autores**



**[Cole Stryker](https://www.ibm.com/think/author/cole-stryker)**

Editorial Lead, AI Models

**[Eda Kavlakoglu](https://www.ibm.com/think/author/eda-kavlakoglu)**

Program Manager

¿Qué es la IA?

La inteligencia artificial (IA) es una tecnología que permite a ordenadores y máquinas simular las capacidades humanas de aprendizaje, comprensión, resolución de problemas, toma de decisiones, creatividad y autonomía.

Las aplicaciones y dispositivos equipados con IA pueden ver e identificar objetos. Pueden comprender y responder al lenguaje humano. Pueden aprender de nueva información y experiencia. Pueden hacer recomendaciones detalladas a los usuarios y expertos. Pueden actuar de forma autónoma, por lo que no es necesaria la inteligencia o intervención humana (un ejemplo clásico es un coche autónomo).

Pero en 2024, la mayoría de los investigadores y profesionales de IA (y la mayoría de los titulares relacionados con la IA) se centran en los avances en [IA generativa](https://www.ibm.com/es-es/topics/generative-ai), una tecnología que puede crear textos, imágenes, vídeos y otros contenidos originales. Para comprender completamente la IA generativa, es importante comprender primero las tecnologías sobre las que se basan las herramientas de IA generativa: [machine learning](https://www.ibm.com/es-es/topics/machine-learning) (ML) y [deep learning](https://www.ibm.com/es-es/topics/deep-learning" \t "_self).

Boletín del sector

Las últimas tendencias en IA, presentadas por expertos

Obtenga conocimientos organizados sobre las noticias más importantes e intrigantes de la IA. Suscríbase a nuestro boletín semanal Think. Consulte la [Declaración de privacidad de IBM](https://www.ibm.com/es-es/privacy).

Principio del formulario

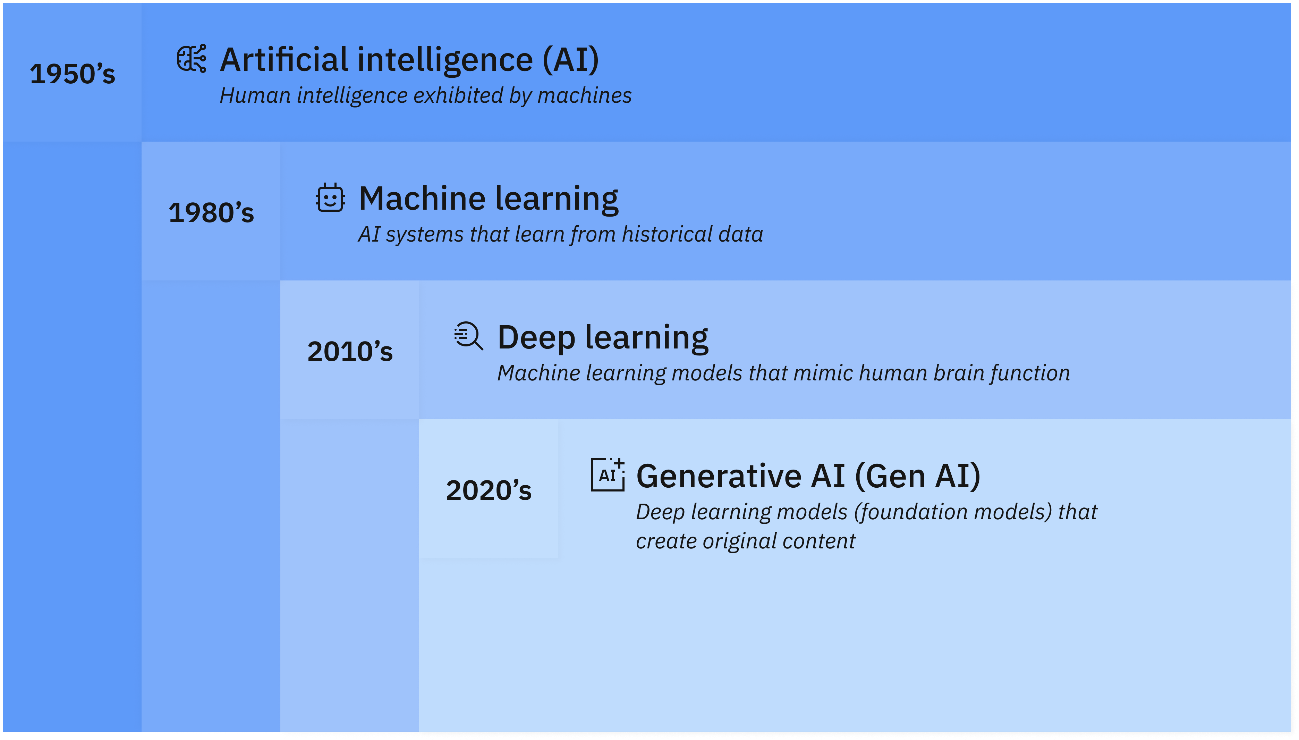
* Correo electrónico comercial

Suscríbase

Final del formulario

Machine learning

Una forma sencilla de entender la IA es como una serie de conceptos anidados o derivados que han ido surgiendo a lo largo de más de 70 años:



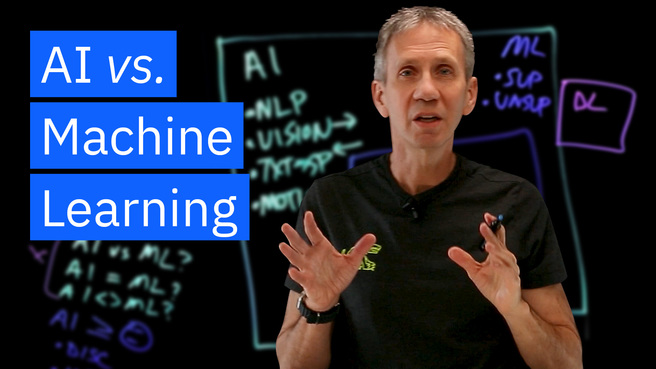
Cómo se relacionan la inteligencia artificial, el machine learning, el deep learning y la IA generativa.

Por debajo de la IA se encuentra el machine learning, que consiste en crear [modelos](https://www.ibm.com/es-es/think/topics/ai-model) mediante el entrenamiento de un algoritmo para que haga predicciones o tome decisiones basadas en datos. Abarca una amplia gama de técnicas que permiten a los ordenadores aprender de los datos y hacer inferencias basadas en ellos, sin necesidad de estar explícitamente programados para tareas específicas.

Existen muchos tipos de técnicas o algoritmos de machine learning, como [regresión lineal](https://www.ibm.com/es-es/think/topics/linear-regression), la [regresión logística](https://www.ibm.com/es-es/think/topics/logistic-regression), los [árboles de decisión](https://www.ibm.com/es-es/think/topics/decision-trees), los [bosques aleatorios](https://www.ibm.com/es-es/think/topics/random-forest), las [máquinas de vectores de soporte](https://www.ibm.com/es-es/think/topics/support-vector-machine) [(SVM)](https://www.ibm.com/es-es/think/topics/support-vector-machine), los [k vecinos más cercanos (KNN)](https://www.ibm.com/es-es/think/topics/knn), el [clustering](https://www.ibm.com/es-es/think/topics/clustering" \t "_self) y otros. Cada uno de estos enfoques es adecuado para distintos tipos de problemas y datos.

Pero uno de los tipos más populares de algoritmo de machine learning se llama [redes neuronales](https://www.ibm.com/es-es/think/topics/neural-networks) (o redes neuronales artificiales). Las redes neuronales se modelan a partir de la estructura y función del cerebro humano. Una red neuronal consta de capas interconectadas de nodos (análogas a las neuronas) que trabajan juntas para procesar y analizar datos complejos. Las redes neuronales son adecuadas para tareas que implican identificar patrones y relaciones complejos en grandes cantidades de datos.

La forma más simple de machine learning es el [aprendizaje supervisado](https://www.ibm.com/es-es/think/topics/supervised-learning), que implica el uso de conjuntos de datos etiquetados para entrenar algoritmos y clasificar datos o predecir resultados con precisión. En el aprendizaje supervisado, los humanos emparejan cada ejemplo de entrenamiento con una etiqueta de output. El objetivo es que el modelo aprenda la asignación entre entradas y outputs en los datos de entrenamiento, de modo que pueda predecir las etiquetas de los datos nuevos e invisibles.



AI vs Machine Learning (5:49 min)

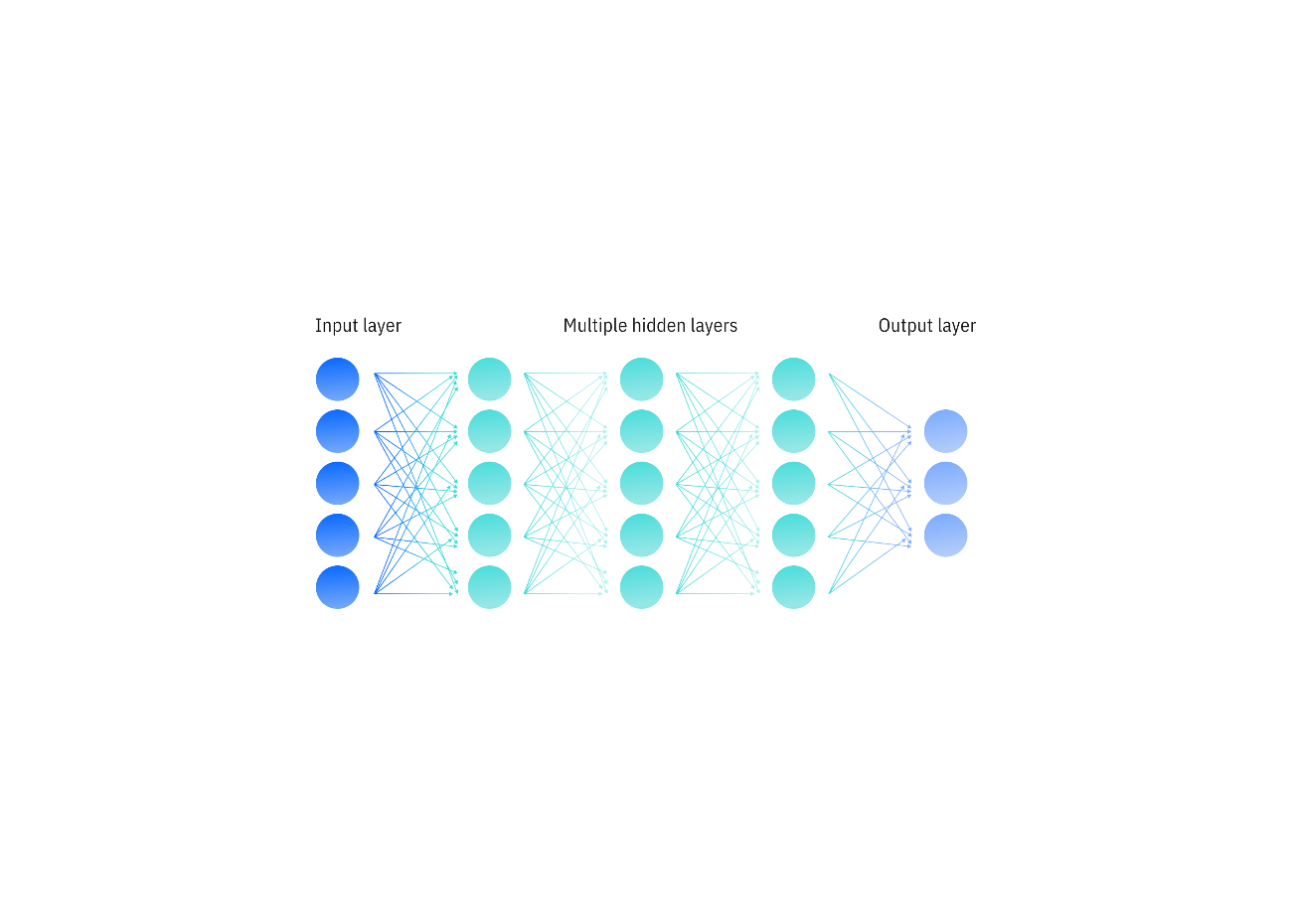
Deep learning

El deep learning es un subconjunto del machine learning que utiliza redes neuronales multicapa, llamadas redes neuronales profundas, que simulan de manera más cercana el complejo poder de toma de decisiones del cerebro humano.

Las redes neuronales profundas incluyen una capa de entrada, al menos tres, pero normalmente cientos de capas ocultas, y una capa de output, a diferencia de las redes neuronales utilizadas en los modelos clásicos de machine learning, que suelen tener solo una o dos capas ocultas.

Estas múltiples capas permiten el [aprendizaje no supervisado](https://www.ibm.com/es-es/think/topics/unsupervised-learning): pueden automatizar la extracción de características de conjuntos de datos grandes, no etiquetados y no estructurados, y hacer sus propias predicciones sobre lo que representan los datos.

Al no requerir intervención humana, el deep learning permite el machine learning a gran escala. Es muy adecuado para el [procesamiento del lenguaje natural (PLN)](https://www.ibm.com/es-es/think/topics/natural-language-processing), la [visión artificial](https://www.ibm.com/es-es/think/topics/computer-vision) y otras tareas que requieren identificar rápidamente y con precisión patrones y relaciones complejas en grandes volúmenes de datos. La mayoría de las aplicaciones de inteligencia artificial (IA) que utilizamos hoy en día se basan en algún tipo de deep learning.



En una red neuronal profunda, múltiples capas de nodos pueden extraer el significado y las relaciones de grandes volúmenes de datos no estructurados ni etiquetados.

El deep learning también permite:

* El [aprendizaje semisupervisado](https://www.ibm.com/es-es/think/topics/semi-supervised-learning), que combina el aprendizaje supervisado y el no supervisado mediante el uso de datos etiquetados y no etiquetados para entrenar modelos de IA para tareas de clasificación y regresión.
* El [aprendizaje autosupervisado](https://www.ibm.com/es-es/think/topics/self-supervised-learning) que genera etiquetas implícitas a partir de datos no estructurados, en lugar de depender de conjuntos de datos etiquetados para las señales de supervisión.
* El [aprendizaje por refuerzo](https://www.ibm.com/es-es/think/topics/reinforcement-learning), que aprende por ensayo-error y funciones de recompensa en lugar de extraer información de patrones ocultos.
* [Aprendizaje por transferencia](https://www.ibm.com/es-es/think/topics/transfer-learning), en el que los conocimientos adquiridos a través de una tarea o conjunto de datos se utilizan para mejorar el rendimiento del modelo en otra tarea o conjunto de datos diferente.

Generative AI

La IA generativa*,*se refiere a modelos de deep learning que pueden crear contenido original complejo, como texto de formato largo, imágenes de alta calidad, vídeo o audio realista y mucho más, en respuesta a la instrucción o solicitud de un usuario.

A un nivel superior, los modelos generativos codifican una representación simplificada de los datos de entrenamiento y se basan en ella para crear un resultado similar, pero no idéntico, a los datos originales.

Los modelos generativos se han utilizado durante años en estadística para analizar datos numéricos. Sin embargo, en la última década han evolucionado para analizar y generar tipos de datos más complejos. Esta evolución coincidió con la aparición de tres sofisticados modelos de deep learning.

* Los [autocodificadores variacionales](https://www.ibm.com/es-es/think/topics/variational-autoencoder), o VAE, que se presentaron en 2013 y permitieron a los modelos generar múltiples variaciones de contenido en respuesta a un mensaje o instrucción.
* Modelos de difusión, vistos por primera vez en 2014, que añaden "ruido" a las imágenes hasta hacerlas irreconocibles, y luego eliminan el ruido para generar imágenes originales en respuesta a las instrucciones.
* Los [transformadores](https://www.ibm.com/es-es/topics/transformer-model?mhsrc=ibmsearch_a&mhq=what%20is%20a%20transformer%20model%26quest%3B) (también llamados modelos transformadores) se entrenan con datos secuenciados para generar secuencias ampliadas de contenido (como palabras en una frase, formas en una imagen, fotogramas de un vídeo o comandos en código de software). Los transformadores son el núcleo de la mayoría de las herramientas actuales de IA generativa, como ChatGPT, GPT-4, Copilot, BERT, Bard y Midjourney.

**Mixture of Experts | 25 de abril, episodio 52**



Descifrar la IA: resumen semanal de noticias

Únase a nuestro panel de ingenieros, investigadores, responsables de producto y otros profesionales de talla mundial que se abren paso entre el bullicio de la IA para ofrecerle las últimas noticias y conocimientos al respecto.

[Ver los últimos episodios del podcast](https://www.ibm.com/es-es/think/podcasts/mixture-of-experts)

Cómo funciona la IA generativa

En general, la IA generativa funciona en tres fases:

1. **Entrenamiento,**para crear un modelo fundacional.
2. **Ajuste,**para adaptar el modelo a una aplicación específica
3. **Generación, evaluación y más ajustes,**para mejorar la precisión.

Entrenamiento

La IA generativa comienza con un "modelo fundacional", un modelo de deep learning que sirve de base para múltiples tipos diferentes de aplicaciones de IA generativa.

Los modelos fundacionales más comunes hoy en día son los [modelos de lenguaje de gran tamaño (LLM)](https://www.ibm.com/es-es/think/topics/large-language-models), creados para aplicaciones de generación de texto. Pero también hay modelos fundacionales para la generación de imagen, vídeo, sonido o música, y modelos fundacionales multimodales que soportan varios tipos de contenido.

Para crear un modelo fundacional, los profesionales entrenan un algoritmo de deep learning con grandes volúmenes de datos relevantes, sin estructurar y sin etiquetar, como terabytes o petabytes de texto de datos, imágenes o videos de Internet. El entrenamiento proporciona una [red neuronal](https://www.ibm.com/es-es/topics/neural-networks) de miles de millones de *parámetros* (representaciones codificadas de las entidades, patrones y relaciones en los datos) que puede generar contenido de forma autónoma en respuesta a las instrucciones. Este es el modelo fundacional.

Este proceso de entrenamiento requiere muchos cálculos, tiempo y dinero: requiere miles de unidades de procesamiento gráfico (GPU) agrupadas y semanas de procesamiento, todo lo cual cuesta millones de dólares. Los proyectos de modelos fundacionales de código abierto, como Llama-2 de Meta, permiten a los desarrolladores de IA generativa evitar este paso y sus costes.

Ajuste

A continuación, el modelo debe ajustarse a una tarea específica de generación de contenido. Esto se puede hacer de varias maneras, entre ellas:

* El ajuste fino, que consiste en alimentar el modelo con datos etiquetados específicos de la aplicación (preguntas o instrucciones que la aplicación probablemente recibirá) y las correspondientes respuestas correctas en el formato deseado.
* El aprendizaje por refuerzo con feedback humano (RLHF), en el que los usuarios humanos evalúan la precisión o relevancia de los outputs del modelo, de modo que éste pueda mejorarse a sí mismo. Esto puede ser tan sencillo como hacer que la gente escriba o responda con correcciones a un chatbot o asistente virtual.

Generación, evaluación y más ajuste

Los desarrolladores y usuarios evalúan periódicamente los outputs de sus aplicaciones de IA generativa y ajustan el modelo, incluso una vez a la semana, para aumentar su precisión o relevancia. En contraste, el propio modelo fundacional se actualiza con mucha menos frecuencia, quizás cada año o cada 18 meses.

Otra opción para mejorar el rendimiento de una aplicación de IA generativa es la generación aumentada por recuperación(RAG), una técnica para ampliar el modelo fundacional y utilizar fuentes relevantes fuera de los datos de entrenamiento para refinar los parámetros y lograr una mayor precisión o relevancia.

Beneficios de la IA

La IA ofrece numerosos beneficios en diversos sectores y aplicaciones. Algunos de los más comúnmente citados incluyen:

* **Automatización de tareas repetitivas.**
* **Más conocimiento y más rápido a partir de los datos**
* **Mejora de la toma de decisiones.**
* **Disminución de los errores humanos.**
* **Disponibilidad 24x7.**
* **Reducción de riesgos físicos.**

Automatización de tareas repetitivas

La IA puede automatizar tareas rutinarias, repetitivas y a menudo tediosas, incluidas tareas digitales como la recopilación, introducción y preprocesamiento de datos, y tareas físicas como la recogida de existencias en almacenes y los procesos de fabricación. Esta automatización libera para trabajar en trabajos de mayor valor y más creativos.

Mejora de la toma de decisiones

Tanto si se utiliza como apoyo a la toma de decisiones como si se trata de una toma de decisiones totalmente automatizada, la IA permite realizar predicciones más rápidas y precisas y tomar [decisiones fiables basadas en datos](https://www.ibm.com/es-es/think/topics/data-driven-decision-making). Combinada con la automatización, la IA permite a las empresas aprovechar las oportunidades y responder a las crisis a medida que surgen, en tiempo real y sin intervención humana.

Disminución de los errores humanos

La IA puede reducir los errores humanos de varias maneras, desde guiar a las personas a través de los pasos adecuados de un proceso, hasta señalar posibles errores antes de que ocurran y automatizar completamente los procesos sin intervención humana. Esto es especialmente importante en sectores como la sanidad, donde, por ejemplo, la robótica quirúrgica guiada por IA permite una precisión constante.

Los algoritmos de machine learning pueden mejorar continuamente su precisión y reducir aún más los errores a medida que están expuestos a más datos y "aprenden" de la experiencia.

Disponibilidad y consistencia las 24 horas del día

La IA está siempre activa, disponible las 24 horas del día y ofrece un rendimiento constante. Herramientas como los chatbots de IA o los asistentes virtuales pueden reducir la necesidad de personal para el servicio de atención al cliente. En otras aplicaciones, como el procesamiento de materiales o las líneas de producción, la IA puede ayudar a mantener la calidad del trabajo y los niveles de producción al automatizar tareas repetitivas o tediosas.

Reducción de los riesgos físicos

Al automatizar los trabajos peligrosos (como el control de animales, la manipulación de explosivos, la realización de tareas en aguas oceánicas profundas, a grandes altitudes o en el espacio exterior), la IA puede eliminar la necesidad de poner a los trabajadores humanos en riesgo de sufrir lesiones o algo peor. Aunque aún no se han perfeccionado, los coches autónomos y otros vehículos ofrecen la posibilidad de reducir el riesgo de lesiones a los pasajeros.

Casos de uso de la IA

Las aplicaciones de la IA en el mundo real son muy numerosas. A continuación se muestra una pequeña selección de casos de uso en diversos sectores para ilustrar su potencial:

Servicio, soporte y experiencia del cliente

Las empresas pueden implementar chatbots y asistentes virtuales con IA para gestionar consultas de clientes, tickets de soporte y más. Estas herramientas utilizan el [procesamiento del lenguaje natural](https://www.ibm.com/es-es/topics/natural-language-processing) (PLN) y las capacidades de IA generativa para entender y responder a preguntas de los clientes sobre el estado del pedido, los detalles del producto y las políticas de devolución.

Los chatbots y los asistentes virtuales permiten prestar soporte de manera continua, proporcionan respuestas más rápidas a las preguntas más frecuentes (FAQ), liberan a los agentes humanos para que puedan centrarse en tareas de mayor nivel y ofrecen a los clientes un servicio más rápido y coherente.

Detección del fraude

Los algoritmos de machine learning y deep learning pueden analizar patrones de transacciones y señalar anomalías, como gastos inusuales o ubicaciones de inicio de sesión, que indiquen transacciones fraudulentas. Esto permite a las organizaciones responder con mayor rapidez a un posible fraude y limitar su impacto, lo que les proporciona a ellas mismas y a sus clientes una mayor tranquilidad.

Marketing personalizado

Los minoristas, los bancos y otras empresas orientadas al cliente pueden utilizar la IA para crear experiencias del cliente personalizadas y campañas de marketing que deleiten a los clientes, mejoren las ventas y eviten la pérdida de clientes. A partir de los datos del historial de compras y los comportamientos de los clientes, los algoritmos de deep learning pueden recomendar productos y servicios que los clientes probablemente deseen, e incluso generar textos personalizados y ofertas especiales para clientes individuales en tiempo real.

Recursos humanos y contratación

Las plataformas de contratación impulsadas por IA pueden racionalizar la contratación mediante la revisión de los currículums, el emparejamiento de candidatos con descripciones de puestos de trabajo e incluso la realización de entrevistas preliminares mediante el análisis de vídeo. Estas y otras herramientas pueden reducir de manera drástica la montaña de papeleo administrativo asociado con la selección de un gran volumen de candidatos. También puede reducir los tiempos de respuesta y el tiempo de contratación, lo que mejora la experiencia de los candidatos, tanto si consiguen el trabajo como si no.

Desarrollo y modernización de aplicaciones

Las herramientas de generación de código de IA generativa y las herramientas de automatización pueden agilizar las tareas de codificación repetitivas asociadas al desarrollo de aplicaciones, y acelerar la migración y modernización (reformateo y cambio de plataforma) de aplicaciones heredadas a escala. Estas herramientas pueden acelerar las tareas, ayudar a garantizar la consistencia del código y reducir errores.

Mantenimiento predictivo

Los modelos de machine learning pueden analizar datos de sensores, dispositivos de Internet de las cosas (IoT) y tecnología operativa (TO) para prever cuándo será necesario el mantenimiento y predecir los fallos de los equipos antes de que ocurran. El mantenimiento preventivo con IA ayuda a evitar el tiempo de inactividad y le permite adelantarse a los problemas de la cadena de suministro antes de que afecten a los resultados finales.

Retos y riesgos de la IA

Las organizaciones se apresuran a beneficiarse de las últimas tecnologías de IA y a sacar partido de sus numerosos beneficios. Esta adopción rápida es necesaria, pero adoptar y mantener flujos de trabajo de IA conlleva retos y riesgos.

Riesgos de datos

Los sistemas de IA se basan en [conjuntos de datos](https://www.ibm.com/es-es/think/topics/data-management) que podrían ser vulnerables al envenenamiento, la manipulación y el sesgo de datos o a los [ciberataques](https://www.ibm.com/es-es/think/topics/cyber-attack) que pueden dar lugar a vulneraciones de datos. Las organizaciones pueden mitigar estos riesgos al proteger la [integridad de los datos](https://www.ibm.com/es-es/think/topics/data-integrity) e implementar la seguridad y la disponibilidad a lo largo de todo el ciclo de vida de la IA, desde el desarrollo hasta el entrenamiento, la implementación y la posimplementación.

Riesgos de modelos

Los [actores de amenazas](https://www.ibm.com/es-es/think/topics/threat-actor) pueden apuntar a modelos de IA para robarlos, realizar ingeniería inversa o manipularlos sin autorización. Los atacantes pueden poner en peligro la integridad de un modelo si alteran su arquitectura, sus ponderaciones o sus parámetros, los componentes básicos que determinan el comportamiento, la precisión y el rendimiento de un modelo.

Riesgos operativos

Al igual que todas las tecnologías, los modelos son susceptibles a [riesgos operativos](https://www.ibm.com/es-es/think/topics/operational-risk), como la deriva del modelo, el sesgo y los fallos en la estructura de gobierno. Si no se abordan, estos riesgos pueden provocar fallos en el sistema y vulnerabilidades de ciberseguridad que los actores de amenazas pueden utilizar.

Ética y riesgos legales

Si las organizaciones no priorizan la seguridad y la ética a la hora de desarrollar e implementar sistemas de IA, corren el riesgo de cometer violaciones de la privacidad y de producir resultados parciales. Por ejemplo, los [datos de entrenamiento parciales](https://www.ibm.com/es-es/think/topics/ai-bias) utilizados para las decisiones de contratación podrían reforzar los estereotipos raciales o de género y crear modelos de IA que favorezcan a determinados grupos demográficos en detrimento de otros.

Ética y gobierno de la IA

La [ética de la IA](https://www.ibm.com/es-es/think/topics/ai-ethics) es un campo multidisciplinar que estudia cómo optimizar el impacto beneficioso de la IA a la vez que se reducen los riesgos y los resultados adversos. Los principios de la ética de la IA se aplican a través de un sistema de [gobierno de la IA](https://www.ibm.com/es-es/think/topics/ai-governance) constituido por barreras que contribuyen a garantizar que las herramientas y los sistemas de IA sigan siendo seguros y éticos.

El gobierno de la IA abarca mecanismos de supervisión que abordan los riesgos. Un planteamiento ético del gobierno de la IA requiere la participación de un amplio abanico de partes interesadas, incluidos desarrolladores, usuarios, responsables políticos y especialistas en ética, que ayuden a garantizar que los sistemas relacionados con la IA se desarrollen y utilicen en consonancia con los valores de la sociedad.

Estos son los valores comunes asociados con la ética de la IA y la [IA responsable](https://www.ibm.com/es-es/think/topics/responsible-ai):

Explicabilidad e interpretabilidad

A medida que la IA avanza, los humanos se enfrentan al reto de comprender y descifrar cómo el algoritmo ha obtenido un resultado determinado. La [IA explicable](https://www.ibm.com/es-es/topics/explainable-ai) es un conjunto de procesos y métodos que permite a los usuarios humanos interpretar, entender y confiar en los resultados y outputs creados por algoritmos.

Equidad e inclusión

Aunque el machine learning, por su propia naturaleza, es una forma de discriminación estadística, la discriminación se vuelve objetable cuando coloca a grupos privilegiados en ventaja sistemática y a ciertos grupos no privilegiados en desventaja sistemática, lo que puede causar diversos daños. Para fomentar la equidad, los profesionales pueden intentar minimizar el sesgo algorítmico en la recopilación de datos y el diseño de modelos, y crear equipos más diversos e inclusivos.

Solidez y seguridad

Una IA sólida maneja con eficacia condiciones excepcionales, como anomalías en la entrada o ataques maliciosos, sin causar daños involuntarios. También está diseñada para resistir las interferencias intencionadas y no intencionadas mediante la protección contra las vulnerabilidades expuestas.

Responsabilidad y transparencia

Las organizaciones deben establecer unas responsabilidades y estructuras de gobierno claras para el desarrollo, la implementación y los resultados de los sistemas de IA. Además, los usuarios deben poder ver cómo funciona un servicio de IA, evaluar su funcionalidad y comprender sus puntos fuertes y débiles. Una mayor transparencia proporciona información a los consumidores de IA para que comprendan mejor cómo se ha creado el modelo o servicio de IA.

Privacidad y cumplimiento

Muchos marcos regulatorios, incluido el RGPD, exigen que las organizaciones cumplan con ciertos principios de privacidad al procesar información personal. Es crucial poder proteger los modelos de IA que puedan contener información personal, controlar qué datos se introducen en el modelo en primer lugar y construir sistemas adaptables que puedan ajustarse a los cambios en la normativa y las actitudes en torno a la ética de la IA.

IA débil vs. IA fuerte

Para contextualizar el uso de la IA en varios niveles de complejidad y sofisticación, los investigadores han definido varios tipos de IA en función de su nivel de sofisticación:

**IA débil**: también conocida como "IA estrecha", define los sistemas de IA diseñados para realizar una tarea específica o un conjunto de tareas.Los ejemplos podrían incluir aplicaciones de asistente de voz "inteligentes" como Alexa de Amazon, Siri de Apple, un chatbot de redes sociales o los vehículos autónomos prometidos por Tesla.

[IA sólida](https://www.ibm.com/es-es/think/topics/strong-ai): también conocida como “inteligencia artificial general” (AGI) o “IA general”, posee la capacidad de comprender, aprender y aplicar conocimientos en una amplia gama de tareas a un nivel igual o [superior a la inteligencia humana](https://www.ibm.com/es-es/think/topics/artificial-superintelligence). Este nivel de IA es actualmente teórico y ningún sistema de IA conocido se acerca a este nivel de sofisticación. Los investigadores sostienen que, si la IA general fuera posible, sería necesario un gran aumento de la potencia de cálculo. A pesar de los recientes avances en el desarrollo de la IA, los sistemas de IA autoconscientes de la ciencia ficción se mantienen firmemente en ese ámbito.

Historia de la IA

La idea de "una máquina que piensa" se remonta a la antigua Grecia. Sin embargo, desde la llegada de la informática electrónica (y en relación con algunos de los temas tratados en este artículo), los acontecimientos clave en la evolución de la IA son los siguientes:

**1950**  
Alan Turing publica [Computing Machinery and Intelligence](https://courses.cs.umbc.edu/471/papers/turing.pdf). En este artículo, Turing, famoso por descifrar el código alemán ENIGMA durante la Segunda Guerra Mundial y a menudo considerado el "padre de la informática", se plantea la siguiente pregunta: "¿Pueden pensar las máquinas?"

A partir de ahí, propone una prueba, ahora conocida como la "Prueba de Turing", en la que un interrogador humano intentaría distinguir entre la respuesta de un ordenador y la de un texto humano. Aunque esta prueba ha sido objeto de mucho escrutinio desde que se publicó, sigue siendo una parte importante de la historia de la IA y un concepto vigente dentro de la filosofía, ya que utiliza ideas en torno a la lingüística.

**1956**  
John McCarthy acuña el término "inteligencia artificial" en la primera conferencia sobre IA en el Dartmouth College. (McCarthy llegó a inventar el lenguaje Lisp). Más tarde, ese mismo año, Allen Newell, J.C. Shaw y Herbert Simon crearon Logic Theorist, el primer programa informático de IA operativo.

**1967**  
Frank Rosenblatt construye el Mark 1 Perceptron, el primer ordenador basado en una red neuronal que "aprendía" por ensayo y error. Justo un año después, Marvin Minsky y Seymour Papert publican un libro titulado Perceptrons, que se convirtió tanto en la obra fundamental sobre redes neuronales como, al menos durante un tiempo, en un argumento en contra de futuras iniciativas de investigación en redes neuronales.

**1980**  
Las redes neuronales, que utilizan un algoritmo de retropropagación para entrenarse, se han generalizado en las aplicaciones de IA.

**1995**  
Stuart Russell y Peter Norvig publican [Artificial Intelligence: A Modern Approach](https://aima.cs.berkeley.edu/), que se convierte en uno de los principales libros de texto en el estudio de la IA. En él, profundizan en cuatro posibles objetivos o definiciones de la IA, que diferencian los sistemas informáticos en función de la racionalidad y el pensamiento frente a la actuación.

**1997**  
Deep Blue de IBM vence al entonces campeón mundial de ajedrez Garry Kasparov, en una partida de ajedrez (y revancha).

**2004**  
John McCarthy escribe un artículo, [¿Qué es la inteligencia artificial?](https://www-formal.stanford.edu/jmc/whatisai.pdf), y propone una definición de IA que se cita con frecuencia. En este momento, la era del big data y el cloud computing ya ha empezado, lo que permite a las organizaciones gestionar conjuntos de datos cada vez mayores, que algún día se utilizarán para entrenar modelos de IA.

**2011**  
IBM Watson supera a los campeones Ken Jennings y Brad Rutter en el concurso televisivo "Jeopardy!". También en esta época, la ciencia de datos comienza a surgir como una disciplina popular.

**2015**  
El superordenador Minwa de Baidu utiliza una red neuronal profunda especial denominada red neuronal convolucional para identificar y categorizar imágenes con una tasa de precisión superior a la del ser humano medio.

**2016**  
El programa AlphaGo de DeepMind, impulsado por una red neuronal profunda, vence a Lee Sodol, campeón del mundo de Go, en un combate a cinco partidas. La victoria es importante dado el enorme número de movimientos posibles a medida que avanza la partida (más de 14,5 billones tras sólo cuatro movimientos). Más tarde, Google compró DeepMind por 400 millones de dólares.

**2022**  
Un aumento de los [modelos de lenguaje de gran tamaño](https://www.ibm.com/es-es/think/topics/large-language-models), o LLM, como ChatGPT, de OpenAI, supone un enorme cambio en el rendimiento de la IA y su potencial para impulsar el valor empresarial. Con estas nuevas prácticas de IA generativa, los modelos de deep learning pueden entrenarse previamente con grandes cantidades de datos.

**2024**  
Las últimas [tendencias en IA](https://www.ibm.com/es-es/think/insights/artificial-intelligence-trends) apuntan a un continuo renacimiento de la IA. Los modelos multimodales que pueden tomar varios tipos de datos como entrada proporcionan experiencias más ricas y sólidas. Estos modelos reúnen las capacidades de reconocimiento de imágenes por [visión artificial](https://www.ibm.com/es-es/think/topics/computer-vision) y de reconocimiento del habla por PLN. Los modelos más pequeños también están avanzando en una era de rendimientos decrecientes con modelos masivos que tienen un gran número de parámetros.