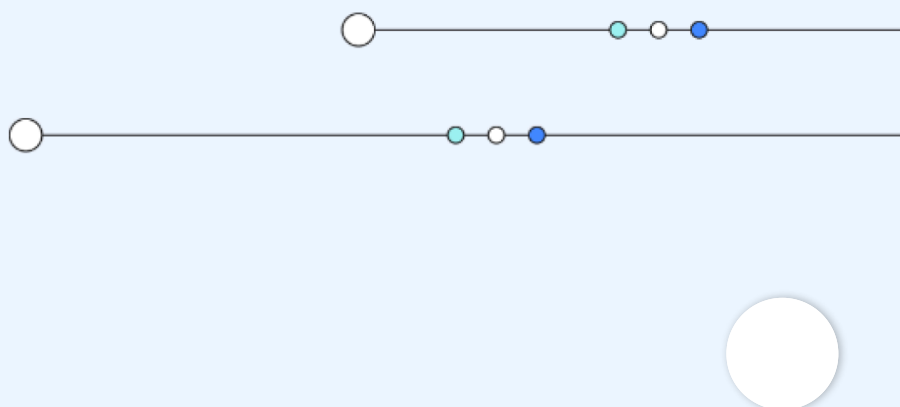


[Inicio](#) / [topics](#) / ¿Qué es el deep learning?

# ¿Qué es el aprendizaje profundo?

**Explorar la solución de aprendizaje profundo de IBM** →

**Suscríbase para recibir actualizaciones sobre IA** →



¿Qué es el aprendizaje profundo?

Cómo funciona el aprendizaje profundo

Tipos de modelos

Casos de uso de aprendizaje profundo

Aplicaciones industriales

Soluciones relacionadas

Recursos

Dé el siguiente paso

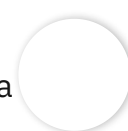
**Actualizado:** 17 de junio de 2024

**Colaboradores:** Jim Holdsworth, Mark Scapicchio

# ¿Qué es el aprendizaje profundo?

El aprendizaje profundo es un subconjunto de [aprendizaje profundo](#) que utiliza [redes neuronales](#) multicapa, llamadas redes neuronales profundas, para simular el complejo poder de toma de decisiones del cerebro humano. Alguna forma de aprendizaje profundo potencia la mayoría de las aplicaciones de [inteligencia artificial \(IA\)](#) en nuestras vidas actuales.

La principal diferencia entre el aprendizaje profundo y el aprendizaje automático es la estructura de la arquitectura de red neuronal subyacente. Los modelos [tradicionales](#)



[de aprendizaje automático](#) “no profundos” emplean redes neuronales simples con una o dos capas computacionales. Los modelos de aprendizaje profundo emplean tres o más capas, pero normalmente cientos o miles de capas, para entrenar los modelos.

Mientras que los modelos de aprendizaje supervisado requieren datos de entrada estructurados y etiquetados para obtener resultados precisos, los modelos de aprendizaje profundo pueden emplear el aprendizaje no supervisado. Con el aprendizaje no supervisado, los modelos de aprendizaje profundo pueden extraer las características, las funciones y las relaciones que necesitan para obtener resultados precisos a partir de datos brutos y no estructurados. Además, estos modelos pueden incluso evaluar y refinar sus resultados para aumentar la precisión.

El aprendizaje profundo es un aspecto de la ciencia de datos que impulsa muchas aplicaciones y servicios que mejoran la [automatización](#), realizando tareas analíticas y físicas sin intervención humana. Esto permite muchos productos y servicios cotidianos, como asistentes digitales, controles remotos de TV habilitados por voz, detección de fraudes con tarjetas de crédito, vehículos autónomos e IA generativa.

Libro electrónico

## Desarrolle flujos de trabajo de IA responsables con gobernanza de IA

Conozca los componentes básicos y las mejores prácticas para ayudar a sus equipos a acelerar la IA responsable.



### Contenido relacionado

[Regístrese para obtener el libro electrónico sobre IA generativa](#)



# Comience su camino hacia la IA

[Aprenda a escalar la IA](#) →

[Conozca AI Academy](#) →

## Cómo funciona el aprendizaje profundo

Las redes neuronales, o redes neuronales artificiales, intentan imitar el cerebro humano a través de una combinación de entradas de datos, pesos y sesgos, todos los cuales actúan como neuronas de silicio. Estos elementos trabajan juntos para reconocer, clasificar y describir con precisión los objetos dentro de los datos.

Las redes neuronales profundas constan de múltiples capas de nodos interconectados, cada uno de los cuales se basa en la capa anterior para refinar y optimizar la predicción o categorización. Esta progresión de cálculos a través de la red se llama propagación hacia adelante. Las capas de entrada y salida de una red neuronal profunda se denominan capas *visibles*. La capa de entrada es donde el modelo de aprendizaje profundo ingiere los datos para su procesamiento y la capa de salida es donde se realiza la predicción o clasificación final.

Otro proceso llamado propagación hacia atrás emplea algoritmos, como el descenso de gradiente, para calcular errores en las predicciones y luego ajusta los pesos y sesgos de la función al moverse hacia atrás a través de las capas para entrenar el modelo. En conjunto, la propagación hacia adelante y la propagación hacia atrás permiten que una red neuronal haga predicciones y corrija cualquier error. Con el tiempo, el algoritmo se vuelve gradualmente más preciso.

El aprendizaje profundo requiere una enorme cantidad de potencia informática. Las [unidades de procesamiento gráfico \(GPU\)](#) de alto rendimiento son ideales porque pueden manejar un gran volumen de cálculos en múltiples núcleos con abundante memoria disponible. La computación en la nube distribuida también podría ayudar. Este nivel de potencia informática es necesario para capacitar algoritmos profundos a través del aprendizaje profundo. Sin embargo, la gestión de varias GPU on premises puede generar una gran demanda de recursos internos y su escalabilidad puede

resultar increíblemente costosa. Para los requisitos de software, la mayoría de las aplicaciones de aprendizaje profundo están codificadas con uno de estos tres marcos de aprendizaje: JAX, PyTorch o TensorFlow.

## Tipos de modelos de aprendizaje profundo

Los algoritmos de aprendizaje profundo son increíblemente complejos y existen diferentes tipos de redes neuronales para abordar problemas o conjuntos de datos específicos. Aquí le mostramos seis. Cada uno tiene sus propios beneficios y se presentan aquí aproximadamente en el orden de su desarrollo: cada modelo sucesivo se ajusta para superar una debilidad en un modelo anterior.

Una posible debilidad en todos ellos es que los modelos de aprendizaje profundo suelen ser "cajas negras", lo que dificulta la comprensión de su funcionamiento interno y plantea desafíos de interpretabilidad. Pero esto se puede equilibrar con los beneficios generales de alta precisión y escalabilidad.

## CNN

Las redes neuronales convolucionales (CNN o ConvNets) se emplean principalmente en aplicaciones de [visión artificial](#) y clasificación de imágenes. Pueden detectar características y patrones dentro de imágenes y videos, lo que permite tareas, como la detección de objetos, el reconocimiento de imágenes, de patrones y facial. Estas



redes aprovechan los principios del álgebra lineal, en particular la multiplicación de matrices, para identificar patrones dentro de una imagen.

Las CNN son un tipo específico de **red neuronal**, que se compone de capas de nodos, que contienen una capa de entrada, una o más capas ocultas y una capa de salida. Cada nodo se conecta a otro, y tiene un peso y umbral asociados. Si la salida de cualquier nodo individual está por encima del valor del umbral especificado, ese nodo se activa y envía datos a la siguiente capa de la red. De lo contrario, no se pasa ningún dato a la siguiente capa de la red.

Al menos tres tipos principales de capas componen una CNN: una capa convolucional, una capa de agrupación y una capa totalmente conectada (FC). Para usos complejos, una CNN puede contener hasta miles de capas, cada una de las cuales se basa en las capas anteriores. Mediante la "convolución" (trabajar y reelaborar la entrada original) se pueden descubrir patrones detallados. Con cada capa, la CNN aumenta su complejidad, identificando mayores porciones de la imagen. Las primeras capas se enfocan en características simples, como colores y bordes. A medida que los datos de la imagen avanzan a través de las capas de CNN, se comienzan a reconocer elementos o formas más grandes del objeto, hasta que finalmente se identifica el objeto previsto.

Las CNN se distinguen de otras redes neuronales por su rendimiento superior con entradas de señales de imagen, voz o audio. Antes de las CNN, se usaban métodos de extracción manual de características, que consumían mucho tiempo, a fin de identificar objetos en las imágenes. Sin embargo, las CNN ofrecen ahora un enfoque más escalable para las tareas de clasificación de imágenes y reconocimiento de objetos, y procesan datos de alta dimensión. Y las CNN pueden intercambiar datos entre capas, para ofrecer un procesamiento de datos más eficiente. Aunque es posible que se pierda información en la capa de agrupación, esto puede verse compensado por los beneficios de las CNN, que pueden ayudar a reducir la complejidad, mejorar la eficacia y limitar el riesgo de sobreajuste.

Las CNN tienen otras desventajas, que son exigentes desde el punto de vista informático: cuestan tiempo y presupuesto, y requieren muchas unidades de procesamiento gráfico (GPU). También requieren expertos altamente calificados con conocimientos de distintos ámbitos y pruebas minuciosas de configuraciones, hiperparámetros y configuraciones.

## RNN



Las **redes neuronales recurrentes (RNN)** se emplean normalmente en aplicaciones de lenguaje natural y **reconocimiento del habla**, ya que emplean datos secuenciales o de series temporales. Las RNN se pueden identificar por sus ciclos de retroalimentación. Estos algoritmos de aprendizaje se utilizan principalmente cuando se emplean datos de series temporales para hacer predicciones sobre resultados futuros. Los casos de uso incluyen predicciones del mercado de valores o pronósticos de ventas, o problemas ordinales o temporales, como traducción de idiomas, **procesamiento de lenguaje natural (PLN)**, reconocimiento del habla y subtítulos de imágenes. Estas funciones a menudo se incorporan a aplicaciones populares, como Siri, búsqueda por voz y Google Translate.

Las RNN emplean su "memoria" a medida que toman información de entradas anteriores para influir en la entrada y salida actuales. Si bien las redes neuronales profundas tradicionales asumen que las entradas y salidas son independientes entre sí, la salida de las RNN depende de los elementos anteriores dentro de la secuencia. Si bien los eventos futuros también serían útiles para determinar la salida de una secuencia determinada, las redes neuronales recurrentes unidireccionales no pueden tener en cuenta estos eventos en sus predicciones.

Las RNN comparten parámetros en cada capa de la red y comparten el mismo parámetro de peso dentro de cada capa de la red, con los pesos ajustados a través de los procesos de propagación hacia atrás y descenso de gradiente para facilitar el aprendizaje por refuerzo.

Las RNN emplean un algoritmo de propagación hacia atrás en el tiempo (BPTT) para determinar los gradientes, que es ligeramente diferente de la propagación hacia atrás tradicional, ya que es específica para los datos secuenciales. Los principios de la BPTT son los mismos que los de la propagación hacia atrás tradicional, en la que el modelo se entrena a sí mismo calculando los errores de su capa de salida a su capa de entrada. La BPTT difiere del enfoque tradicional en que suma errores en cada paso de tiempo, mientras que las redes de prealimentación no necesitan sumar errores, ya que no comparten parámetros en cada capa.

Una ventaja sobre otros tipos de redes neuronales es que los RNN utilizan tanto el procesamiento de datos binarios como la memoria. Los RNN pueden planificar múltiples entradas y producciones de modo que, en lugar de entregar solo un resultado para una sola entrada, los RMM puedan producir salidas de uno a muchos, de muchos a uno o de muchos a muchos.

También hay opciones dentro de los RNN. Por ejemplo, la red de memoria a largo plazo (LSTM) es superior a los RNN simples al aprender y actuar sobre dependencias a largo plazo.

Sin embargo, las RNN tienden a encontrarse con dos problemas básicos, conocidos como gradientes explosivos y gradientes evanescentes. Estos problemas se definen por el tamaño del gradiente, que es la pendiente de la función de pérdida a lo largo de la curva de error.

- Cuando el gradiente *desaparece* y es demasiado pequeño, continúa haciéndose más pequeño, actualizando los parámetros de peso hasta que se vuelven insignificantes, es decir: cero (0). Cuando eso ocurre, el algoritmo ya no aprende.
- Los gradientes *explosivos* ocurren cuando el gradiente es demasiado grande, creando así un modelo inestable. En este caso, los pesos del modelo crecen demasiado y, finalmente, se representarán como NaN (no como un número). Una solución a estos problemas es reducir el número de capas ocultas dentro de la red neuronal, eliminando parte de la complejidad de los modelos RNN.

Algunas desventajas finales: las RNN también pueden requerir un largo tiempo de entrenamiento y ser difíciles de usar en grandes conjuntos de datos. La optimización de las RNN agrega complejidad cuando tienen muchas capas y parámetros.

## Autocodificadores y autocodificadores variacionales

El aprendizaje profundo permitió ir más allá del análisis de datos numéricos al agregar el análisis de imágenes, voz y otros tipos de datos complejos. Entre la primera clase de modelos para lograr esto, se encontraban los [autocodificadores variacionales \(VAEs\)](#). [Fueron los primeros modelos de aprendizaje profundo que se utilizaron ampliamente para generar imágenes y habla realistas, lo que potenció el modelado generativo profundo al hacer que los modelos fueran más fáciles de escalar, que es la piedra angular de lo que concebimos como IA generativa.](#)

Los autocodificadores funcionan codificando datos sin etiquetar en una representación comprimida y luego decodificando los datos a su forma original. Los *autocodificadores* simples se emplearon para una variedad de propósitos, incluida la reconstrucción de imágenes dañadas o borrosas. Los autocodificadores *variacionales* agregaron la capacidad crítica no solo de reconstruir datos, sino también de generar variaciones en los datos originales.

Esta capacidad de generar datos novedosos activó una rápida sucesión de nuevas tecnologías, desde redes generativas adversarias (GAN) hasta modelos de difusión,





capaces de producir imágenes cada vez más realistas, pero falsas. De esta manera, los VAE sentaron las bases para la IA generativa actual.

Los autocodificadores se construyen a partir de bloques de codificadores y decodificadores, una arquitectura que también sustenta los grandes modelos de lenguajes actuales. Los codificadores comprimen un conjunto de datos en una representación densa, organizando puntos de datos similares más juntos en un espacio abstracto. Los decodificadores toman muestras de este espacio para crear algo nuevo y, al mismo tiempo, preservar las características más importantes del conjunto de datos.

El mayor beneficio de los autocodificadores es la capacidad de manejar grandes lotes de datos y mostrar los datos de entrada en forma comprimida, por lo que se destacan los aspectos más significativos, lo que permite tareas de detección y clasificación de anomalías. Esto también acelera la transmisión y reduce los requisitos de almacenamiento. Los autocodificadores se pueden capacitar con datos sin etiquetar, por lo que podrían usarse donde los datos etiquetados no están disponibles. Cuando se emplea el entrenamiento no supervisado, existe el beneficio de ahorrar tiempo: los algoritmos de aprendizaje profundo aprenden automáticamente y ganan precisión sin necesidad de ingeniería manual de características. Además, los VAE pueden generar nuevos datos de muestra para la generación de texto o imágenes.

Los autocodificadores tienen sus desventajas. El entrenamiento de estructuras profundas o complejas puede suponer una pérdida de recursos computacionales. Y durante el entrenamiento no supervisado, el modelo puede pasar por alto las propiedades necesarias y, en su lugar, simplemente replicar los datos de entrada. Los autocodificadores también pueden pasar por alto vínculos de datos complejos en datos estructurados para que no identifiquen correctamente las relaciones complejas.

## GAN

Las **redes generativas adversativas (GAN)** son redes neuronales que se emplean tanto dentro como fuera de la inteligencia artificial (IA) para crear nuevos datos que se parezcan a los datos de entrenamiento originales. Estas pueden incluir imágenes que parecen ser rostros humanos, pero se generan, no se toman de personas reales. La parte "adversativa" del nombre proviene del vaivén entre las dos partes de la GAN: un generador y un discriminador.

- El **generador** crea *algo*: imágenes, video o audio para luego producir una salida con un giro. Por ejemplo, un caballo puede transformarse en cebra con cierto grado de

precisión. El resultado depende de la entrada y de lo bien entrenadas que estén las capas en el modelo generativo para este caso de uso.

- El **discriminador** es el adversario, donde el resultado *generativo* (imagen falsa) se compara con las imágenes *reales* en el conjunto de datos. El discriminador intenta distinguir entre las imágenes, el video o el audio reales y falsos.

Las GAN se entrenan a sí mismas. El generador crea falsificaciones mientras que el discriminador aprende a detectar las diferencias entre las falsificaciones del generador y los ejemplos verdaderos. Cuando el discriminador puede marcar la falsificación, el generador es penalizado. El ciclo de retroalimentación continúa hasta que el generador logra producir una salida que el discriminador no puede distinguir.

El principal beneficio de la GAN es crear resultados realistas que pueden ser difíciles de distinguir de los originales, que, a su vez, se pueden emplear para entrenar aún más los modelos de aprendizaje automático. Configurar una GAN para aprender es sencillo, ya que se capacitan mediante el uso de datos sin etiquetar o con un etiquetado menor. Sin embargo, la desventaja potencial es que el generador y el discriminador pueden ir y venir en competencia durante mucho tiempo, creando un gran drenaje del sistema. Una limitación del entrenamiento es que puede ser necesaria una gran cantidad de datos de entrada para obtener un resultado satisfactorio. Otro problema potencial es el "colapso de modo", cuando el generador produce un conjunto limitado de salidas en lugar de una variedad más amplia.

## Modelos de difusión

Los [modelos de difusión](#) son modelos generativos que se entrenan mediante el proceso de difusión directa e inversa de adición y eliminación de ruido progresivas. Los modelos de difusión generan datos (la mayoría de las veces imágenes) similares a los datos con los que se entrenan, pero luego sobrescriben los datos empleados para entrenarlos. Agregan gradualmente ruido gaussiano a los datos de entrenamiento hasta que es irreconocible, luego aprenden un proceso inverso de "eliminación de ruido" que puede sintetizar la salida (generalmente imágenes) a partir de la entrada de ruido aleatorio.

Un modelo de difusión aprende a minimizar las diferencias de las muestras generadas frente al objetivo deseado. Cualquier discrepancia se cuantifica y los parámetros del modelo se actualizan para minimizar la pérdida, entrenando el modelo para producir muestras que se parezcan mucho a los datos de entrenamiento auténticos.



Más allá de la calidad de la imagen, los modelos de difusión tienen el beneficio de no requerir entrenamiento adversativo, lo que acelera el proceso de aprendizaje y también ofrece un control estricto del proceso. El entrenamiento es más estable que con las GAN y los modelos de difusión no son tan propensos al colapso del modo.

Pero, en comparación con las GAN, los modelos de difusión pueden requerir más recursos informáticos para entrenar, incluido un mayor ajuste. IBM® Research también descubrió que esta forma de IA generativa puede ser secuestrada con [puertas traseras ocultas](#), dando a los atacantes control sobre el proceso de creación de imágenes para que los modelos de difusión de IA puedan ser engañados para generar imágenes manipuladas.

## modelos de transformador

Los [modelos transformadores](#) combinan una arquitectura de codificador-decodificador con un mecanismo de procesamiento de texto y revolucionaron la forma en que se entrenan los modelos de lenguaje. Un codificador convierte el texto sin procesar y sin anotaciones en representaciones conocidas como incrustaciones; el decodificador toma estas incrustaciones junto con las salidas anteriores del modelo y predice sucesivamente cada palabra en una oración.

Mediante el uso de adivinanzas para completar espacios en blanco, el [codificador](#) aprende cómo se relacionan las palabras y las oraciones entre sí, creando una poderosa representación del lenguaje sin tener que etiquetar partes del discurso y otras características gramaticales. De hecho, los transformadores pueden entrenarse previamente desde el principio sin tener en mente una tarea en particular. Una vez que se aprenden estas poderosas representaciones, los modelos pueden especializarse, con muchos menos datos, para realizar una tarea solicitada.

Varias innovaciones lo hacen posible. Los transformadores procesan palabras en una oración simultáneamente, lo que permite el procesamiento de texto en paralelo, acelerando el entrenamiento. Las técnicas anteriores que incluían redes neuronales recurrentes (RNN) procesaban palabras una por una. Los transformadores también aprendieron las posiciones de las palabras y sus relaciones; este contexto les permite inferir significado y desambiguar palabras como “eso” en oraciones largas.

Al eliminar la necesidad de definir una tarea por adelantado, los transformadores hicieron práctico entrenar previamente los modelos de lenguaje en grandes

cantidades de texto sin procesar, lo que les permitió crecer significativamente en tamaño. Anteriormente, los datos etiquetados se recopilaban para entrenar un modelo en una tarea específica. Con los transformadores, un modelo entrenado en una cantidad masiva de datos se puede adaptar a múltiples tareas ajustándolo en una pequeña cantidad de datos etiquetados específicos de la tarea.

Hoy en día, los transformadores de lenguaje se emplean para tareas no generativas, como la clasificación y la extracción de entidades, así como para tareas generativas, como la traducción automática, el resumen y la respuesta a preguntas. Los transformadores sorprendieron a muchas personas con su capacidad para generar diálogos, ensayos y otros contenidos convincentes.

Los transformadores de procesamiento de lenguaje natural (PLN) proporcionan una potencia notable, ya que pueden ejecutarse en paralelo, procesando varias partes de una secuencia de forma simultánea, lo que acelera enormemente el entrenamiento. Los transformadores también rastrean las dependencias a largo plazo en el texto, lo que les permite comprender el contexto general con mayor claridad y crear resultados superiores. Además, los transformadores son más escalables y flexibles para personalizarse por tarea.

En cuanto a las limitaciones, debido a su complejidad, los transformadores requieren enormes recursos computacionales y un largo tiempo de entrenamiento. Además, los datos de entrenamiento deben ser precisos, imparciales y abundantes para producir resultados precisos.

## Casos de uso de aprendizaje profundo

El número de usos para el aprendizaje profundo crece cada día. Estas son solo algunas de las formas en que ahora está ayudando a las empresas a ser más eficientes y atender mejor a sus clientes.

### Modernización de aplicaciones

La IA generativa puede mejorar las capacidades de los desarrolladores y reducir la brecha de habilidades cada vez mayor en los dominios de [la modernización de](#)

[aplicaciones](#) y la automatización de TI. [La IA generativa para la programación](#) es posible gracias a los avances recientes en las tecnologías de modelos de lenguaje grandes (LLM) y el procesamiento de lenguaje natural (PLN). Emplea algoritmos de aprendizaje profundo y grandes redes neuronales entrenadas en vastos conjuntos de datos de código fuente existente. En general, el código de entrenamiento proviene de código disponible públicamente producido por proyectos de código abierto.

Los programadores pueden ingresar [instrucciones de texto sin formato](#) que describen lo que quieren que haga el código. Las herramientas de IA generativa sugieren fragmentos de código o funciones completas, lo que agiliza el proceso de programación al manejar tareas repetitivas y reducir la programación manual. La IA generativa también puede traducir código de un lenguaje a otro, agilizando la conversión de código o los proyectos de modernización, como la actualización de aplicaciones heredadas mediante la traducción de COBOL a Java.

## Visión artificial

[La visión artificial](#) es un campo de la inteligencia artificial (IA) que incluye la clasificación de imágenes, la detección de objetos y la segmentación semántica. Emplea el aprendizaje automático y las redes neuronales para enseñar a las computadoras y a los sistemas de aprendizaje a derivar información significativa de imágenes digitales, videos y otras entradas visuales, y a hacer recomendaciones o tomar medidas cuando el sistema detecta defectos o problemas. Si la IA permite que las computadoras piensen, la visión artificial les permite ver, observar y comprender.

Debido a que un sistema de visión artificial a menudo está entrenado para inspeccionar productos o vigilar los activos de producción, generalmente puede analizar miles de productos o procesos por minuto, notando defectos o problemas imperceptibles. La visión artificial se emplea en industrias que van desde la energía y los servicios públicos hasta la fabricación y la industria automotriz.

La visión artificial necesita muchos datos, y luego ejecuta análisis de esos datos una y otra vez hasta que discierne y finalmente reconoce imágenes. Por ejemplo, para entrenar a una computadora para que reconozca llantas de automóvil, necesita recibir grandes cantidades de imágenes de llantas y elementos relacionados con llantas para aprender las diferencias y reconocer una llanta, especialmente una sin defectos.



La visión artificial emplea modelos algorítmicos para permitir que una computadora se enseñe a sí misma el contexto de los datos visuales. Si se alimentan suficientes datos a través del modelo, la computadora "mirará" los datos y aprenderá a distinguir una imagen de otra. Los algoritmos permiten que la máquina aprenda por sí misma, en lugar de que alguien la programe para reconocer una imagen.

La visión artificial permite a los sistemas obtener información significativa a partir de imágenes digitales, videos y otras entradas visuales y, en función de esas entradas, tomar medidas. Esta capacidad de [proporcionar recomendaciones](#) la distingue de tareas de reconocimiento de imágenes sencillas. Algunas aplicaciones comunes de la visión artificial se pueden apreciar hoy en día en estos sectores:

- **Automotriz:** si bien la era de los automóviles sin conductor aún no ha surgido, la tecnología subyacente ha comenzado a abrirse camino, mejorando la seguridad del conductor y los pasajeros a través de características, como la detección carriles.
- **Atención médica:** la visión artificial se ha incorporado a la tecnología radiológica, lo que permite a los médicos identificar mejor tumores cancerosos en una anatomía saludable.
- **Marketing:** las plataformas de redes sociales ofrecen sugerencias sobre quién podría estar en la fotografía que se ha publicado en un perfil, lo que facilita etiquetar a amigos en álbumes de fotos.
- **Venta minorista:** la búsqueda visual se ha incorporado a algunas plataformas de comercio electrónico, permitiendo a las marcas recomendar artículos que complementarían un guardarropa existente.

## Atención al cliente

La IA está ayudando a [las empresas](#) a comprender y satisfacer mejor las crecientes demandas de los consumidores. Con el auge de las compras en línea altamente personalizadas, los modelos directo al consumidor y los servicios de entrega, la IA generativa puede ayudar a desbloquear aún más una serie de beneficios que pueden mejorar la atención al cliente, la transformación del talento y el rendimiento de las aplicaciones.



La IA permite a las empresas adoptar un enfoque centrado en el cliente aprovechando insights valiosos de la retroalimentación y los hábitos de compra de los clientes. Este enfoque basado en datos puede ayudar a mejorar el diseño y el empaque del producto, así como a impulsar una alta satisfacción del cliente y un aumento de las ventas.

La IA generativa también puede servir como asistente cognitivo para la atención al cliente, proporcionando orientación contextual basada en el historial de conversaciones, el análisis de sentimientos y las transcripciones del centro de llamadas. Además, la IA generativa puede permitir experiencias de compra personalizadas, fomentar la lealtad de los clientes y proporcionar una ventaja competitiva.

## Trabajo digital

Las organizaciones pueden aumentar su fuerza laboral mediante la creación y despliegue de automatización de procesos robóticos (RPA) y [mano de obra digital](#) para colaborar con los humanos y aumentar la productividad, o ayudar cuando sea necesario un respaldo. Por ejemplo, esto puede ayudar a los desarrolladores a acelerar la actualización del software heredado.

La mano de obra digital emplea modelos fundacionales para automatizar y mejorar la productividad de los trabajadores del conocimiento permitiendo la automatización del autoservicio de forma rápida y confiable, sin barreras técnicas. Para automatizar la realización de tareas o la llamada a las API, un modelo de llenado de entidades basado en LLM de nivel empresarial puede identificar información en una conversación y recopilar toda la información necesaria para completar una acción o llamar a una API sin mucho esfuerzo manual.

En lugar de que los expertos técnicos registren y codifiquen los flujos de acción repetitivos para los trabajadores del conocimiento, estos últimos pueden emplear las automatizaciones de mano de obra digital creadas sobre una base de instrucciones y demostraciones conversacionales impulsadas por modelos para la automatización del autoservicio. Por ejemplo, para acelerar la creación de aplicaciones, los [aprendices digitales sin código](#) pueden ayudar a los usuarios finales, que carecen de experiencia en programación, al enseñar, supervisar y validar código de manera efectiva.



# IA generativa

La [IA generativa](#) (también llamada genAI) es una categoría de IA que crea de forma autónoma texto, imágenes, videos, datos u otros contenidos en respuesta a las indicaciones o solicitudes de un usuario.

La IA generativa se basa en modelos de aprendizaje profundo que pueden aprender de patrones en el contenido existente y generar contenido nuevo y similar basado en ese entrenamiento. Tiene aplicaciones en muchos campos, incluido el servicio al cliente, el marketing, el desarrollo de software y la investigación, y ofrece un enorme potencial para optimizar los flujos de trabajo empresariales a través de la creación y el aumento de contenido rápidos y automatizados.

La IA generativa se destaca en el manejo de diversas fuentes de datos, como correos electrónicos, imágenes, videos, archivos de audio y contenido de redes sociales. Estos datos no estructurados forman la estructura para crear modelos y el entrenamiento continuo de la IA generativa, para que pueda seguir siendo eficaz a lo largo del tiempo. El uso de estos datos no estructurados puede mejorar el servicio al cliente a través de chatbots y facilitar un enrutamiento de correo electrónico más eficaz. En la práctica, esto podría significar guiar a los usuarios a los recursos adecuados, ya sea poniéndolos en contacto con el agente adecuado o dirigiéndolos a guías de usuario y preguntas frecuentes.

A pesar de sus tan debatidas limitaciones y riesgos, muchas empresas están avanzando, explorando cautelosamente cómo sus organizaciones pueden aprovechar la IA generativa para mejorar sus flujos de trabajo internos, y mejorar sus productos y servicios. Esta es la nueva frontera: cómo hacer que el lugar de trabajo sea más eficiente sin crear problemas legales o éticos.

[IA generativa para desarrolladores](#) →

## Procesamiento de lenguaje natural y reconocimiento del habla





El **PLN** combina lingüística computación (modelado del lenguaje humano basado en reglas) con modelos estadísticos y de aprendizaje automático para permitir que las computadoras y los dispositivos digitales reconozcan, comprendan y generen texto y voz. El PLN impulsa aplicaciones y dispositivos que pueden traducir texto de un idioma a otro, responder a comandos escritos o hablados, reconocer o autenticar a los usuarios en función de la voz. Ayuda a resumir grandes volúmenes de texto, evaluar la intención o el sentimiento del texto o la voz y generar texto o gráficos u otro contenido bajo demanda.

Un subconjunto de PLN es el PLN estadístico, que combina algoritmos informáticos con modelos de aprendizaje automático y aprendizaje profundo. Este enfoque ayuda a extraer, clasificar y etiquetar automáticamente elementos de texto y datos de voz y luego asignar una probabilidad estadística a cada significado posible de esos elementos. Hoy en día, los modelos de aprendizaje profundo y las técnicas de aprendizaje basadas en RNN habilitan sistemas de PLN que "aprenden" a medida que trabajan y extraen un significado cada vez más preciso de enormes volúmenes de conjuntos de datos de texto y voz sin procesar, sin estructurar y sin etiquetar.

El **reconocimiento del habla**, también conocido como reconocimiento automático del habla (ASR), reconocimiento del habla por computadora o conversión de voz a texto, es una capacidad que permite a un programa procesar el habla humana en un formato escrito.

Mientras que el reconocimiento del habla se confunde comúnmente con el reconocimiento de voz, el primero se centra en la traducción del habla de un formato verbal a uno de texto, mientras que el reconocimiento de voz solo busca identificar la voz de un usuario individual.

## Aplicaciones industriales

Las aplicaciones de aprendizaje profundo del mundo real nos rodean y están tan bien integradas en los productos y servicios, de los cuales los usuarios no son conscientes del complejo procesamiento de datos que se lleva a cabo en segundo plano. Algunos de estos ejemplos incluyen:

### Aprendizaje profundo de la atención al cliente



Muchas organizaciones incorporan tecnología de aprendizaje profundo en sus procesos de [atención al cliente](#). Los [chatbots](#) a menudo se utilizan en diversas aplicaciones, servicios y portales de atención al cliente. Los chatbots tradicionales emplean lenguaje natural e incluso reconocimiento visual, que se encuentran comúnmente en los menús de los centros de atención telefónica. Sin embargo, [las soluciones de chatbot más sofisticadas](#) intentan determinar, a través del aprendizaje, si hay múltiples respuestas a preguntas ambiguas en tiempo real. En función de las respuestas que recibe, el chatbot intenta responder a estas preguntas directamente o enruta la conversación a un usuario humano.

Los asistentes virtuales, como Siri de Apple, Amazon Alexa o Google Assistant, amplían la idea de un chatbot al permitir la funcionalidad de reconocimiento del habla. Esto crea un nuevo método para atraer a los usuarios de forma personalizada.

## Análisis de servicios financieros

[Las instituciones financieras](#) emplean de manera regular [analytics predictivos](#) para impulsar el comercio algorítmico de acciones, evaluar los riesgos comerciales para la aprobación de préstamos, detectar fraudes y ayudar a gestionar las carteras de crédito e inversión de los clientes.

## Mantenimiento de registros de atención médica

La industria de [atención médica](#) se benefició enormemente de las capacidades de aprendizaje profundo desde la digitalización de los registros y las imágenes hospitalarios. Las aplicaciones de reconocimiento de imágenes pueden servir de apoyo a especialistas en imagen médica y radiólogos, ayudándoles a analizar y evaluar más imágenes en menos tiempo.

## Las autoridades emplean el aprendizaje profundo

Los algoritmos de aprendizaje profundo pueden analizar y aprender de los datos transaccionales para identificar patrones peligrosos que indiquen posibles actividades fraudulentas o delictivas. El reconocimiento del habla, la visión artificial y otras aplicaciones de aprendizaje profundo pueden mejorar la eficiencia y efectividad del análisis de investigación al extraer patrones y evidencia de grabaciones de sonido y video, imágenes y documentos. Esta capacidad ayuda a [las autoridades](#) a analizar grandes cantidades de datos de manera más rápida y precisa.



# Soluciones relacionadas

---

## IBM watsonx

IBM watsonx es una cartera de herramientas, aplicaciones y soluciones listas para su empresa diseñadas para reducir los costos y los obstáculos de la adopción de la IA, al tiempo que optimiza los resultados y el uso responsable de la IA.

[Explorar IBM watsonx](#) →

---

## IBM watsonx Assistant: chatbot de IA

IBM watsonx Assistant es el chatbot impulsado por IA para empresas. Esta tecnología de inteligencia artificial empresarial permite a los usuarios crear soluciones de IA conversacional.

[Conozca watsonx Assistant](#) →

---

## IBM Watson Studio

Cree, ejecute y gestione modelos de IA. Prepare datos y cree modelos en cualquier cloud utilizando código abierto o modelado visual. Prediga y optimice sus resultados.

[Explorar watsonx Studio](#) →



# Recursos

Cómo

**Aprendizaje práctico y gratuito para tecnología**

**S**  
Aprenda los conceptos fundamentales

de la IA y la IA

generativa, incluida la

organización de

aviso, los

grandes

modelos

lingüísticos y

los mejores

proyectos de

código

abierto.

[→](#)

Artículo

**Una introducción al aprendizaje**

**profundo**

Explore esta

rama del

aprendizaje

automático

que se

entrena con

grandes

cantidades

de datos y se

ocupa de

desarrollar

aplicaciones

que

trabajan en

conjunto

para realizar

predicciones.

[IBM Developer](#)

Artículo

**Arquitecturas de aprendizaje profundo**

Explore los

fundamentos

del

aprendizaje

automático y

la

arquitectura

de

aprendizaje

profundo, y

descubra sus

aplicaciones

y beneficios

asociados.

[IBM Developer](#)

Artículo

**Compare marcos de aprendizaje profundo**

Elegir el

marco de

aprendizaje

profundo

adecuado en

función de su

carga de

trabajo

individual es

un primer

paso

esencial en

ESG [→](#)

aprendizaje

profundo.

# Dé el siguiente paso

Entrene, valide, ajuste y despliegue IA generativa, modelos fundacionales y capacidades de aprendizaje automático con IBM® watsonx.ai, un estudio empresarial de próxima generación para creadores de IA. Cree aplicaciones de IA en menos tiempo y con menos datos.

**Explore watsonx.ai**



**Reserve una demostración en vivo**

