Módulo 2. Las maquinas no piensan. Fundamentos

Una revisión de la tecnología por dentro, los fundamentos y conceptos que dan soporte a la tecnología actual de IA

- Unidad 2.1. De que estamos hablando: Concepto y Definiciones de IA
- Unidad 2.2. Machine Learning. Como aprender sin instrucciones
- Unidad 2.3 Tipos de aprendizaje
- Unidad 2.4. Imitando al cerebro. Redes neuronales.
- Referencias Módulo 2

Copyright 2026

Unidad 2.1. De que estamos hablando: Concepto y Definiciones de IA

Introducción

> "*Lo siento, Dave. Tengo miedo. No puedo hacer eso.*" - HAL 9000, 2001: Una odisea en el espacio.
>
> "*La verdadera pregunta no es si las máquinas pueden pensar o si pueden tener conciencia, sino si nosotros, como seres humanos, estamos dispuestos a reconocer la inteligencia y la conciencia en formas diferentes a las nuestras.*" - Ray Kurzweil

La Inteligencia Artificial (IA) es un campo de estudio que busca desarrollar sistemas capaces de realizar tareas que normalmente requerirían de la inteligencia humana. Estos sistemas utilizan algoritmos y modelos para procesar información, aprender de ella y tomar decisiones basadas en patrones y reglas.

Dentro de la IA, el *Machine Learning* (Aprendizaje Automático) es una rama fundamental que se centra en desarrollar algoritmos y técnicas que permiten a las máquinas aprender y mejorar su rendimiento a partir de datos, sin ser programadas explícitamente. En lugar de seguir reglas específicas, los algoritmos de Machine Learning se basan en la detección de patrones y en la extracción de conocimientos a partir de grandes conjuntos de datos.

Uno de los enfoques más destacados dentro del *Machine Learning* es el *Deep Learning* (Aprendizaje Profundo), que se inspira en el funcionamiento de las redes neuronales del cerebro humano. Las redes neuronales artificiales utilizadas en el Deep Learning están compuestas por múltiples capas de neuronas interconectadas, lo que les permite procesar información de manera jerárquica y aprender representaciones complejas de los datos. El Deep Learning ha demostrado

Copyright 2026

un rendimiento impresionante en tareas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la generación de contenido.

Estos conceptos de IA, *Machine Learning* y *Deep Learning* han experimentado un rápido crecimiento en las últimas décadas, impulsado por avances en la capacidad de computación, el acceso a grandes volúmenes de datos, algoritmos más potentes y el desarrollo de GPUs (procesadores gráficos). Su aplicabilidad abarca una amplia gama de campos, incluyendo la medicina, la industria, la investigación científica, la conducción autónoma, la traducción automática y mucho más.

A medida que la tecnología continúa evolucionando, se espera que la IA y sus ramas sigan desempeñando un papel fundamental en el desarrollo de soluciones inteligentes y en la mejora de la forma en que interactuamos con el mundo digital.

Conceptos de IA

La Inteligencia Artificial (IA) es un campo de estudio que busca desarrollar sistemas y programas capaces de realizar tareas que normalmente requerirían de la inteligencia humana. Estos sistemas utilizan algoritmos y modelos matemáticos para procesar información, aprender de ella y tomar decisiones o realizar acciones basadas en patrones y reglas.

En primer lugar conviene definir de algún modo ambos términos, Inteligencia y Artificial, cuestión ya comentada al principo del curso pero que creemos conveniente recordar en varias de sus posibles definiciones:

Inteligencia

- **Definición Psicológica:**
  - "*La inteligencia es la capacidad para aprender, razonar, resolver problemas, pensar abstractamente, comprender ideas complejas, adaptarse al ambiente y aprender de la experiencia.*" - Carole Wade y Carol Tavris, psicólogas.
- **Definición Biológica:**
  - "*La inteligencia es un conjunto de habilidades de adaptación al entorno, especialmente habilidades de aprendizaje, memoria, razonamiento y percepción.*" - Stephen Ceci, psicólogo especializado en inteligencia.
- **Definición Operacional:**
  - "*La inteligencia es lo que miden los tests de inteligencia.*" - Alfred Binet, uno de los pioneros en el desarrollo de pruebas de inteligencia.

Artificial

Copyright 2026

- **Definición General:**
  - "*Artificial se refiere a algo que es hecho por el hombre o producido en lugar de ser algo natural.*"
- **Definición Tecnológica:**
  - "*En el contexto de la tecnología y la ciencia, artificial describe algo que ha sido producido mediante un proceso controlado, a menudo imitando o replicando algo que ocurre en la naturaleza.*"
- **Definición Filosófica:**
  - "*Desde una perspectiva filosófica, artificial puede referirse a una imitación o simulación de algo natural, buscando emular sus propiedades o funciones.*"

En el contexto de "Inteligencia Artificial", el término "artificial" indica que la inteligencia que se está describiendo no es una inteligencia biológica o natural, sino una creada y diseñada por humanos mediante algoritmos y sistemas computacionales

En términos generales, se pueden distinguir varios tipos de IA:

IA débil: También conocida como IA estrecha o IA específica, se refiere a sistemas diseñados para realizar tareas específicas de manera eficiente y precisa, pero que carecen de la capacidad de razonamiento o generalización más allá de esas tareas. Ejemplos de IA débil incluyen los asistentes virtuales, sistemas de recomendación y chatbots.

IA fuerte: La IA fuerte se refiere a sistemas que poseen la capacidad de igualar o superar la inteligencia humana en una amplia gama de tareas. Estos sistemas serían capaces de comprender, razonar, aprender y adaptarse en diferentes situaciones, y podrían tener una conciencia de sí mismos. Sin embargo, hasta el momento, la IA fuerte sigue siendo un objetivo teórico y aún no se ha logrado completamente.

IA generalizada: También conocida como IA general, se refiere a sistemas que tienen la capacidad de superar a los humanos en todas las tareas intelectuales, tanto en las tareas específicas como en la capacidad de razonar y generalizar en nuevas situaciones. La IA generalizada es un objetivo a largo plazo y aún no se ha alcanzado.

Es importante tener en cuenta que los avances en IA se han centrado principalmente en la IA débil, desarrollando sistemas especializados para tareas específicas. Sin embargo, los esfuerzos continúan en la búsqueda de lograr una IA más fuerte y generalizada, aunque aún hay muchos desafíos técnicos y éticos por resolver.

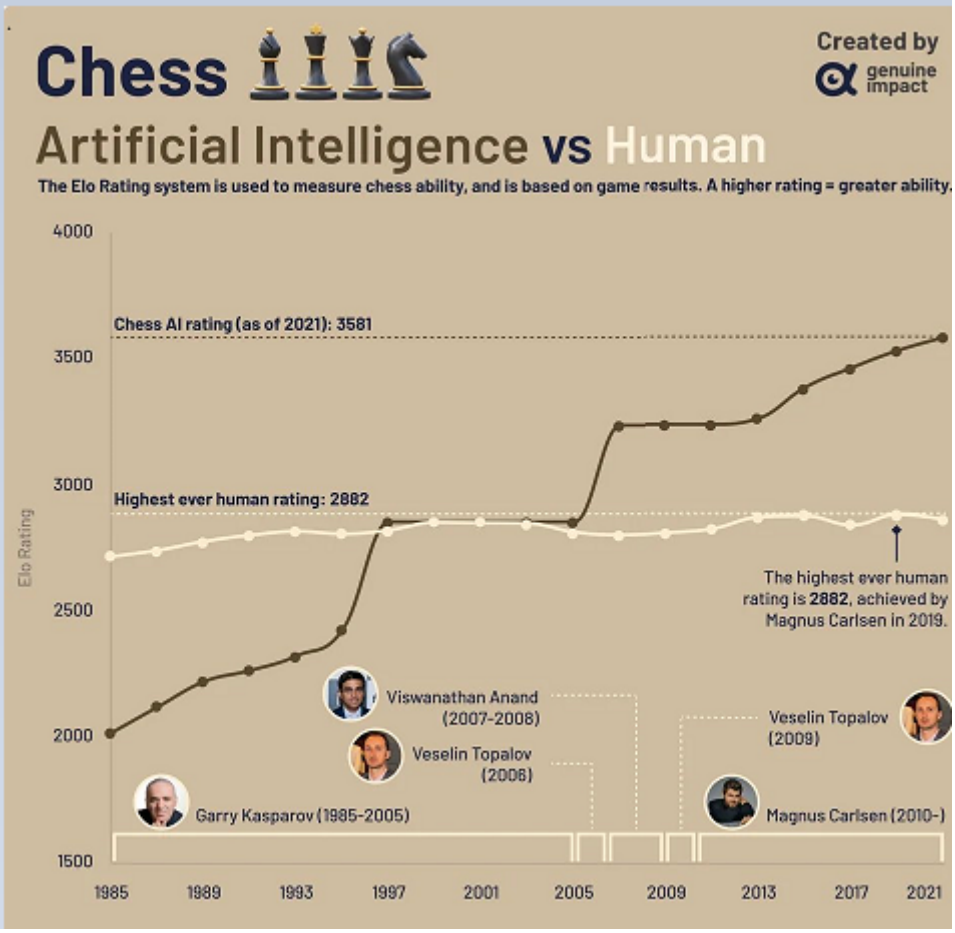
Es fundamental tener en cuenta los límites y las implicaciones éticas de la IA a medida que avanza, y asegurarse de que se utilice de manera responsable y beneficie a la sociedad en su conjunto.

Copyright 2026

> La idea de una IA con conciencia, también conocida como inteligencia artificial consciente o IA consciente, plantea la posibilidad de que las máquinas puedan tener una experiencia subjetiva o conciencia similar a la de los seres humanos. Sin embargo, es importante destacar que actualmente no existe un consenso claro sobre cómo definir o lograr la conciencia en una máquina.

Debe notarse además que la IA es un campo amplio de estudio, que lleva en desarrollo decenas de años y que no solamente abarca lo relacionado con el lenguaje, sino también otros campos como la robótica o la comunicación hombre-máquina

En general tiene que ver con todo lo que signifique imitar el comportamiento humano, especialmente en los campos de la IA conversacional, la síntesis y reconocimiento de voz, la generación de lenguaje natural y la predicción de lenguaje.



Evolución de la inteligencia artificial versus humana en el ajedrez a lo largo del tiempo

En general distinguimos los siguientes campos de estudio:

Campos de Estudio en Inteligencia Artificial

> En vez de intentar producir un programa que simule la mente adulta, ¿por qué no tratar de producir uno que simule la mente del niño? Si ésta se sometiera entonces a un curso educativo adecuado, se obtendría el cerebro de adulto.
>
> **Alan Mathison Turing (1912 - 1954), Matemático Inglés**

Aprendizaje basado en patrones (Machine Learning o ML)

El machine learning es una rama de la inteligencia artificial que se centra en el uso de datos y algoritmos para imitar la forma en la que aprenden los seres humanos. En otras palabras, es un método estadístico para aprender patrones en base a ejemplos.

Por ejemplo, si queremos enseñar a una computadora a reconocer imágenes de perros, le proporcionamos un conjunto de imágenes etiquetadas como “perros” y otro conjunto etiquetado como “no perros”. La computadora utiliza estos datos para aprender a distinguir entre las dos categorías y, con el tiempo, se vuelve más precisa en su capacidad para identificar imágenes de perros

Se utiliza en una amplia variedad de aplicaciones, desde la detección de fraudes hasta la predicción del clima y la recomendación de productos. A medida que la cantidad de datos disponibles sigue creciendo, el llamado ML se está convirtiendo en una herramienta cada vez más importante para ayudar a las empresas a tomar decisiones informadas y mejorar la eficiencia de los procesos productivos.

Dentro del Machine Learning distinguimos especialmente el Aprendizaje Profundo (Deep Learning o DL) que estudia las redes neuronales con múltiples capas (profundas) para procesar grandes conjuntos de datos.

Se basa en el uso de redes neuronales artificiales para imitar la forma en que aprenden los seres humanos. A diferencia del machine learning, que se enfoca en el aprendizaje a partir de datos estructurados, el deep learning se utiliza para aprender patrones a partir de datos no estructurados, como imágenes, audio y texto.

Se utiliza también en una amplia variedad de aplicaciones, desde la detección de fraudes hasta la predicción del clima y la recomendación de productos. A medida que la cantidad de datos

Copyright 2026

disponibles sigue creciendo, el deep learning se está convirtiendo en una herramienta cada vez más importante para ayudar a las empresas a tomar decisiones informadas y mejorar la eficiencia, pero también en una poderosa herramienta para la generación de contenidos de audio texto y vídeo así como chats conversacionales de cualquier temática o dominio de conocimiento.

Procesamiento del Lenguaje Natural (NLP)

El Procesamiento del Lenguaje Natural (NLP, por sus siglas en inglés) es una rama de la inteligencia artificial que se centra en la interacción entre los ordenadores y el lenguaje humano. Su objetivo es permitir que las máquinas comprendan, interpreten y respondan a textos y voces humanas de una manera útil y natural.

En términos simples, el NLP trata de "enseñar" a las computadoras cómo entender el lenguaje humano, incluyendo sus sutilezas, ambigüedades y variaciones. Esto implica procesos como la traducción automática (por ejemplo, convertir texto de un idioma a otro), la respuesta a preguntas (responder preguntas formuladas en lenguaje natural), y la comprensión de sentimientos (detectar emociones en el texto).

Es una tecnología que encontramos en la vida cotidiana en cosas como los asistentes virtuales (Siri, Alexa), los correctores ortográficos automáticos, los sistemas de chatbot para atención al cliente, y las recomendaciones de productos basadas en reseñas. Es un campo en constante evolución que busca mejorar la forma en que las máquinas y los humanos interactúan.

Visión por Computadora

Estudia cómo las máquinas pueden obtener información a partir de imágenes o videos.

La visión por computadora, también conocida como visión artificial, es un campo de la inteligencia artificial que se enfoca en dar a las máquinas la capacidad de "ver" o identificar y entender el contenido visual del mundo. Esto significa procesar y analizar imágenes y videos para identificar objetos, personas, escenas y actividades.

En términos básicos, la visión por computadora permite que las computadoras interpreten y hagan uso de la información visual de la misma manera que lo haría un ser humano, pero a una velocidad y escala mucho mayores. Esto incluye tareas como reconocer rostros en fotografías, identificar objetos en imágenes para sistemas de navegación autónoma (como los utilizados en los coches autónomos), analizar imágenes médicas para diagnosticar enfermedades, y mucho más.

Se utiliza en una amplia gama de aplicaciones prácticas: desde la seguridad y vigilancia hasta la interacción y entretenimiento en redes sociales, pasando por el control de calidad en la fabricación y la agricultura de precisión. Este campo combina técnicas de aprendizaje automático, procesamiento de imágenes y patrones, y análisis de datos para entrenar a las computadoras en el

Copyright 2026

reconocimiento y procesamiento de imágenes.

Robótica

Se centra en el diseño, construcción y operación de robots que pueden interactuar y operar en entornos físicos.

La robótica es un campo de la ingeniería y la ciencia de la computación que se ocupa del diseño, construcción, operación y uso de robots. Los robots son sistemas que pueden moverse, percibir su entorno, procesar información y realizar acciones o tareas en el mundo real, ya sea de manera autónoma o controlada por humanos.

En términos sencillos, la robótica combina elementos de la mecánica (para el diseño y fabricación de las estructuras físicas de los robots), la electrónica (para los sistemas de control y sensores) y la informática (para el procesamiento de datos y la toma de decisiones). Esto permite que los robots realicen una amplia gama de tareas, desde la fabricación industrial y la exploración espacial hasta la asistencia en tareas domésticas y el cuidado de la salud.

La robótica moderna también se entrelaza con la inteligencia artificial, especialmente cuando los robots requieren capacidades avanzadas de percepción, toma de decisiones y aprendizaje. Por ejemplo, los robots que utilizan IA pueden aprender de su entorno y adaptarse a nuevos retos, lo que les permite ser más eficientes y versátiles.

Tiene aplicaciones en numerosos campos, como la manufactura (robots de ensamblaje y soldadura), la medicina (robots quirúrgicos y de rehabilitación), el servicio al cliente (robots de asistencia), el transporte (vehículos autónomos) y muchos más. Es un campo en constante evolución, impulsado por avances tecnológicos y la creciente integración de sistemas inteligentes.

Sistemas Expertos

Un sistema experto es un programa de computadora diseñado para imitar la capacidad de toma de decisiones de un experto humano en un campo específico. Estos sistemas están programados con una gran cantidad de conocimientos y reglas específicas del área en la que se especializan, lo que les permite ofrecer consejos, resolver problemas o tomar decisiones como lo haría un experto real.

Para entenderlo de manera sencilla, piensa en un sistema experto como un "libro de consejos" muy avanzado y específico. Por ejemplo, en medicina, un sistema experto puede diagnosticar enfermedades basándose en los síntomas del paciente, la información médica y las reglas establecidas por médicos reales. Este sistema puede preguntar al usuario (como un médico o paciente) sobre varios síntomas y luego, basándose en su "conocimiento" programado, sugerir posibles diagnósticos o tratamientos.

Copyright 2026

Son útiles en campos donde la toma de decisiones es compleja y requiere un alto nivel de conocimiento especializado, como la medicina, la ingeniería, las finanzas y el derecho. Estos sistemas no reemplazan a los expertos humanos, sino que sirven como herramientas para ampliar su alcance y mejorar la eficiencia en la toma de decisiones.

Aprendizaje por Refuerzo

El aprendizaje por refuerzo es un tipo de aprendizaje automático en el que un agente (un programa de computadora) aprende a tomar decisiones mediante la experimentación y la retroalimentación sobre las acciones que realiza. En lugar de ser enseñado explícitamente qué hacer en cada situación, el agente descubre qué acciones producen los mejores resultados a través de un proceso de prueba y error.

Para entenderlo de manera sencilla, piensa en el aprendizaje por refuerzo como el proceso de enseñar a un niño a andar en bicicleta no solo diciéndole cómo hacerlo, sino dejándolo intentarlo y aprender de sus errores. Cada vez que el niño se inclina demasiado y se cae (un resultado negativo), aprende a evitar esa acción en el futuro. De manera similar, si logra mantener el equilibrio y avanzar (un resultado positivo), aprenderá a repetir y perfeccionar esa acción.

En el aprendizaje por refuerzo, el agente recibe recompensas o penalizaciones basadas en las consecuencias de sus acciones. Su objetivo es maximizar la suma total de recompensas. Este enfoque es particularmente útil en situaciones donde no es factible o posible programar todas las posibles acciones y resultados, como en juegos complejos (como el ajedrez o el Go), la navegación de robots, o la optimización de sistemas.

Este vídeo en el que se enseña a andar a un robot es muy ilustrativo del proceso general

https://www.youtube.com/embed/xAXvfVTgqr0?si=xro-J_KfDqNBWZX1

https://www.youtube.com/watch?v=xAXvfVTgqr0

En definitiva la IA trata de todo lo que tiene que ver con el comportamiento y el razonamiento humano intentando imitaro o reproducirlo basándose en la propia biología humana,.

Se invita al alumno a reflexionar acerca del proceso de aprendizaje que todos hemos seguido en nuestras vidas y de como generamos el propio lenguaje y los recuerdos basados en patrones o similitudes con registros del pasado o de otros tetos o conversaciones.

A día de hoy parece claro que las máquinas no piensan, pero sí aprenden.

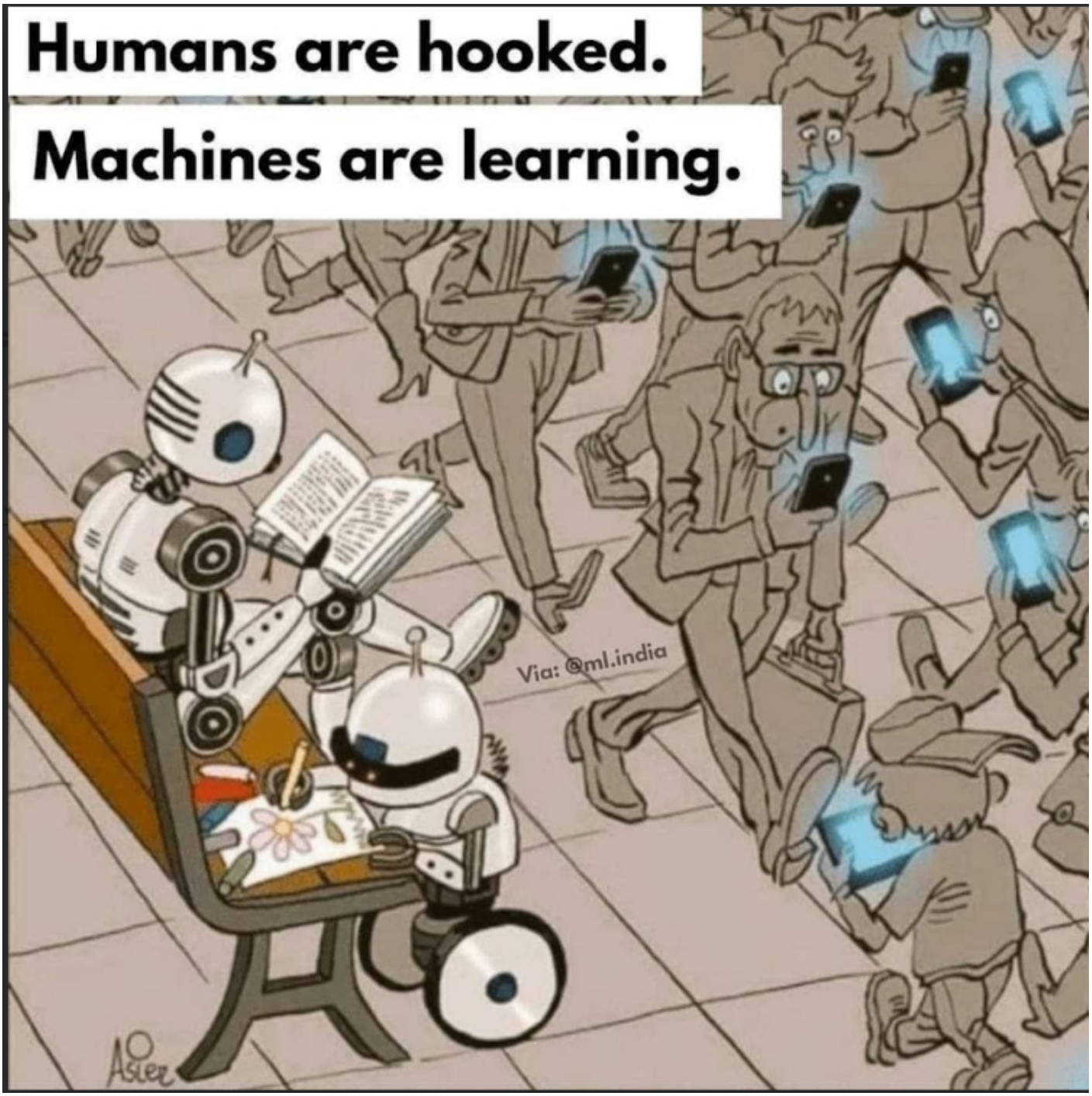
Copyright 2026

Copyright 2026

Unidad 2.2. Machine Learning. Como aprender sin instrucciones

Machine Learning

> **Concepto**
>
> "El aprendizaje automático es el arte y la ciencia de permitir a las máquinas aprender y mejorar automáticamente a partir de los datos, sin necesidad de ser programadas explícitamente para cada tarea específica".



Humanos 'enganchados' mientras las computadoras aprenden

La forma tradicional de hacer que una computadora logre algo es darle instrucciones explícitas (si sucede esto, haz esto, de lo contrario haz aquello, etc.) que se escriben a mano en un lenguaje de programación determinado. Este método de programación de computadoras es extremadamente exitoso y se ha utilizado para desarrollar prácticamente todo el software que se ejecuta en nuestras computadoras, teléfonos e incluso automóviles. Sin embargo, este método no siempre es el más práctico.

Copyright 2026

Por ejemplo, consideremos el desarrollo de un programa para identificar imágenes. Habría que considerar todas las posibles combinaciones de píxeles para que en función de sus características podamos instruir al algoritmo o computadora acerca del objeto de que está 'visualizando'.

Sólo el juego GO ofrece 10 elevado a 700 posibilidades en cuanto a mposibles movivimientos lo que da una idea de lo incorrecto de este enfoque, al menos para el objetivo de que las máquinas aprendan y sean capaces de usar y 'entender' el lenguaje natural, además de distinguir objetos y otras tareas propias de nuestra inteligencia.

Sucede que se nos da muy bien programar cosas que hemos aprendido a hacer como una suma o una multiplicación pero nos resulta extremadamente difiicil enseñar lo que nos es innato, por ejemplo andar, mover un barzo etc...

El aprendizaje automático es un enfoque que permite a la computadora aprender patrones y relaciones en los datos a través de algoritmos y modelos matemáticos. En el caso de la identificación de imágenes, en lugar de programar reglas explícitas, se alimenta a la computadora con un conjunto de imágenes etiquetadas y se le permite aprender por sí misma cómo reconocer diferentes objetos o patrones en las imágenes.

El proceso de aprendizaje automático implica entrenar a la computadora utilizando ejemplos y ajustar sus conexiones internas, de modo que pueda hacer predicciones precisas sobre nuevas imágenes que nunca ha visto antes. Esto se logra mediante el uso de algoritmos de aprendizaje automático, como las redes neuronales, que son capaces de capturar características complejas y extraer información útil de los datos de entrada.

A medida que se alimenta a la computadora con más datos y se perfeccionan los algoritmos de aprendizaje, la precisión y la capacidad de reconocimiento de imágenes de la computadora mejoran con el tiempo. Este enfoque es especialmente útil en casos en los que las reglas explícitas serían demasiado difíciles o tediosas de programar, o cuando los patrones en los datos no son fácilmente discernibles para los humanos.

En resumen, el enfoque tradicional de programación mediante instrucciones explícitas está siendo complementado y en algunos casos reemplazado por el aprendizaje automático, que permite a las computadoras aprender a reconocer patrones y tomar decisiones basadas en datos en lugar de reglas predefinidas. Esta técnica está transformando la forma en que desarrollamos programas y abre un amplio abanico de posibilidades para aplicaciones más inteligentes y adaptativas.

Sería casi imposible escribir un programa así "a mano" debido al gran número de posibilidades a considerar (diferentes orientaciones de objetos, condiciones de iluminación, obstrucciones, etc.). En cambio, sería más fácil utilizar el método de aprendizaje automático, que es un método y una disciplina científica utilizada para programar computadoras utilizando datos en lugar de instrucciones explícitas. En el caso de la identificación de imágenes, el método de aprendizaje

Copyright 2026

automático consistiría en proporcionar a la computadora un conjunto de imágenes etiquetadas según lo que representan y permitir que la computadora descubra por sí misma cómo identificar nuevas imágenes.

Aprendizaje con datos no estructurados

Entendemos por datos no estructurados aquellos que se presentan sin un formato ni estructura específico, como los que provienen del lenguaje natural o en forma de imagen o sonido. Por contra datos estructurados serían aquellos que se presentan con algún tipo de organización como una tabla o un fichero con determinadas características.

En esta sección hablaremos de cómo hacer que las máquinas entiendan datos no estructurados.

Siguiendo con lo comentado al principio de la sección vamos a ver los datos de un modelo para clasificar vehículos distinguiendo entre coches y camiones.

El paradigma de aprendizaje automático permite crear clasificadores que a partir de un conjunto de datos de entrada (en este caso en forma de imágenes etiquetadas) pueda entrenarse y despues al recibir una imagen nueva podrá identificarla correctamente.

Por ejemplo aqui podemos ver un conjunto de imágenes de coches y camiones



COCHE



COCHE



CAMIÓN



CAMIÓN



COCHE



COCHE



CAMIÓN

A la computadora se le han proporcionado varias imágenes de camiones y coches, cada una etiquetada según lo que representan. Estas imágenes etiquetadas se llaman ejemplos (también conocidos como puntos de datos u observaciones) y forman un conjunto de datos. El resultado es un programa llamado modelo que es capaz de identificar nuevos casos y clasificarlos en un tipo de vehículo en este caso.

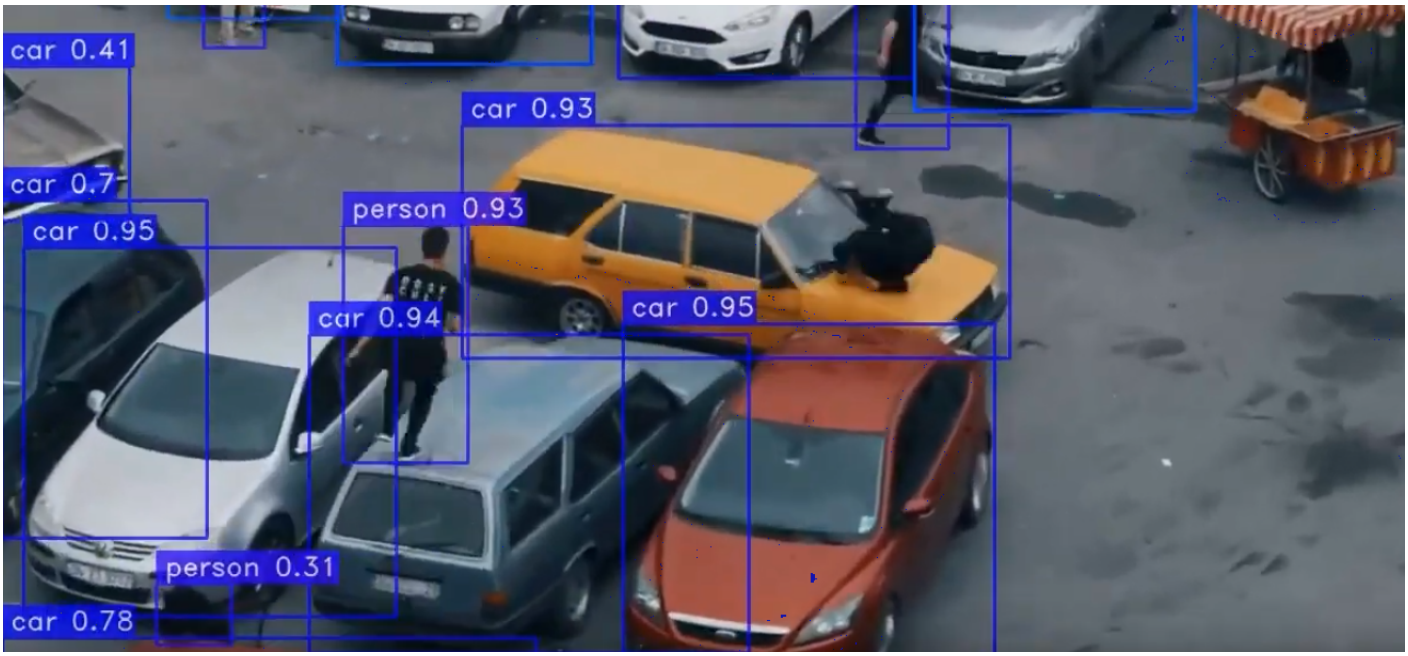
Es preciso señalar la difucltad de este proceso derivada del contexto que rodea a los casos de prueba. en este caso vemos como según que ejemplos no identifican claramente a un coche o camión debido a factores como el ambiente, luces, orientación, distancia, objetos circundantes etc... Es por ello que se reuieren ingentes cantidades de datos para facilitar y mejorar el modelo final de aprendizaje.

En cierto sentido, el aprendizaje automático es programación a través de ejemplos. La computadora (es decir, la máquina) aprende a realizar una tarea a partir de ejemplos de esa tarea. En este contexto, "aprender" significa ampliamente que se utiliza información de ciertos datos para crear el programa.

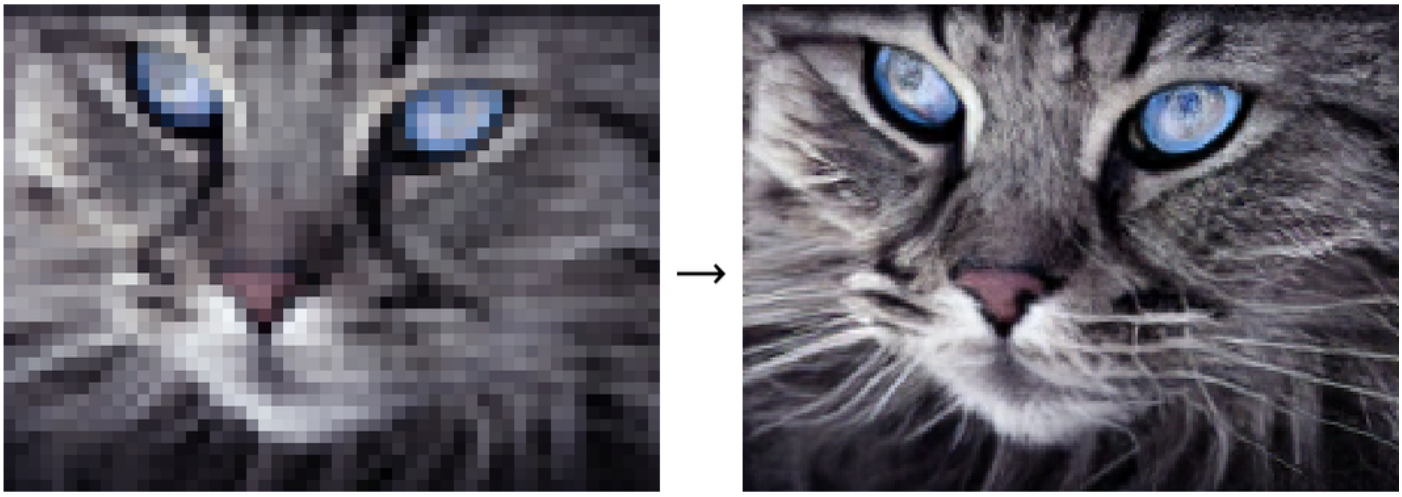
La identificación de imágenes no es la única aplicación en la que el aprendizaje automático es útil; se utiliza en la actualidad para realizar una variedad de tareas, desde identificar correos electrónicos no deseados, hasta predecir los precios de las acciones, o jugar videojuegos. Sin embargo, el aprendizaje automático no es un reemplazo directo de la programación tradicional. Para comprender mejor para qué se puede utilizar el aprendizaje automático, repasemos algunos de sus ámbitos de aplicación actuales.

Un ámbito general de aplicación tiene que ver con la imitación de las habilidades humanas. Esto incluye tareas de percepción, como comprender datos visuales y auditivos; tareas intuitivas, como jugar videojuegos; y la tarea muy importante de comprender texto. A continuación, se presentan ejemplos de aplicaciones que se encuadran en este ámbito.

- **Detectar objetos de un vídeo en tiempo real**

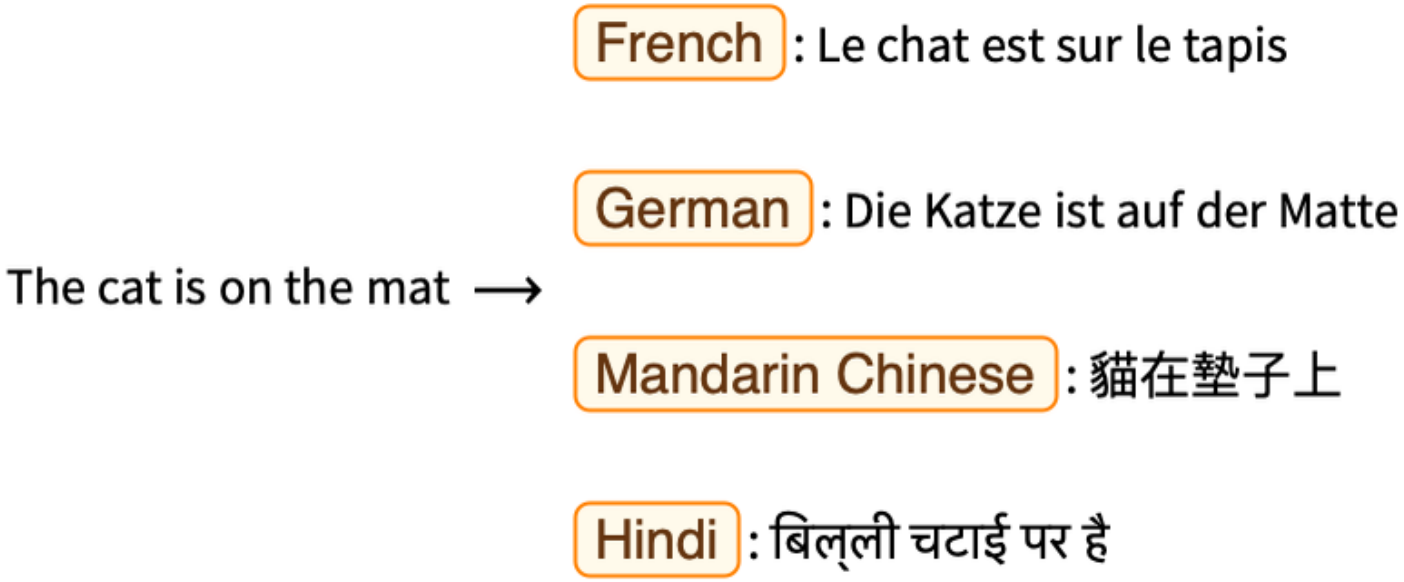


- **Mejorar la calidad imagen**

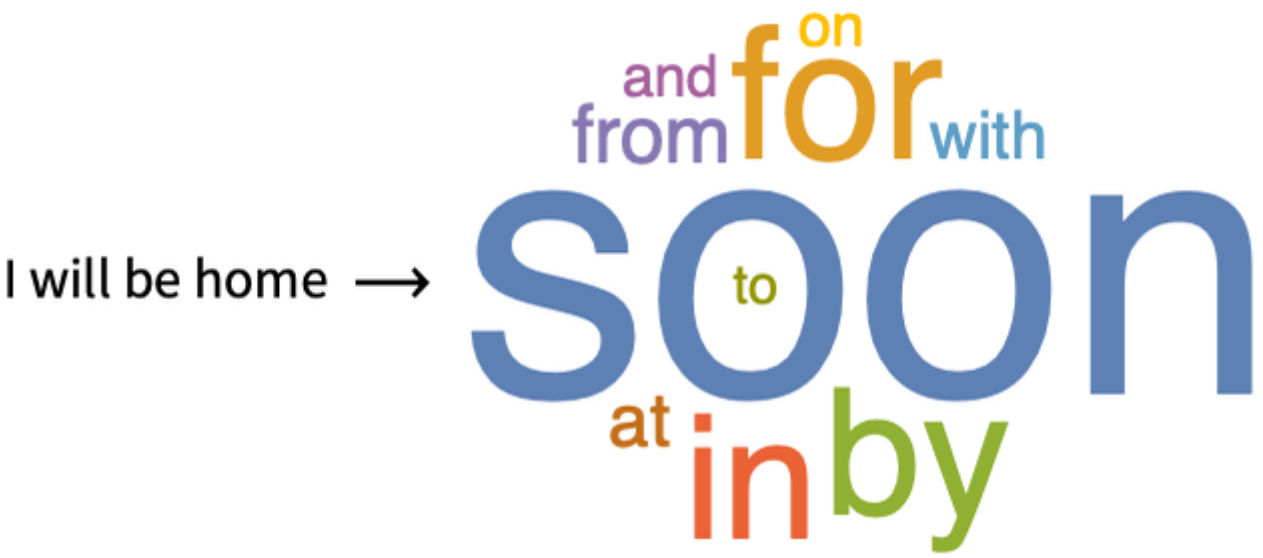


54 × 40 → 216 × 160

- **Traducción**



- **Predicción (generación) de texto**



Son solo algunos ejemplos, que además se van ampliando con la mejora de tecnologías y algoritmos.

En general estamos hablando de convertir información en múltiples formatos (sonido, vídeo, texto e imagen).

Actualmente podemos incluir tareas como conversión de texto a música, texto a imagen, limpiezas de audios o incluso conversión de pensamiento a imagen, algo que de momento se ha conseguido de forma rudimentaria en laboratorio.

Copyright 2026

En general han surgido decenas de aplicaciones llamadas multimodales que permiten la conversión/generación de cualquier tipo de contenido, incluyendo vídeos, imagen y voz.

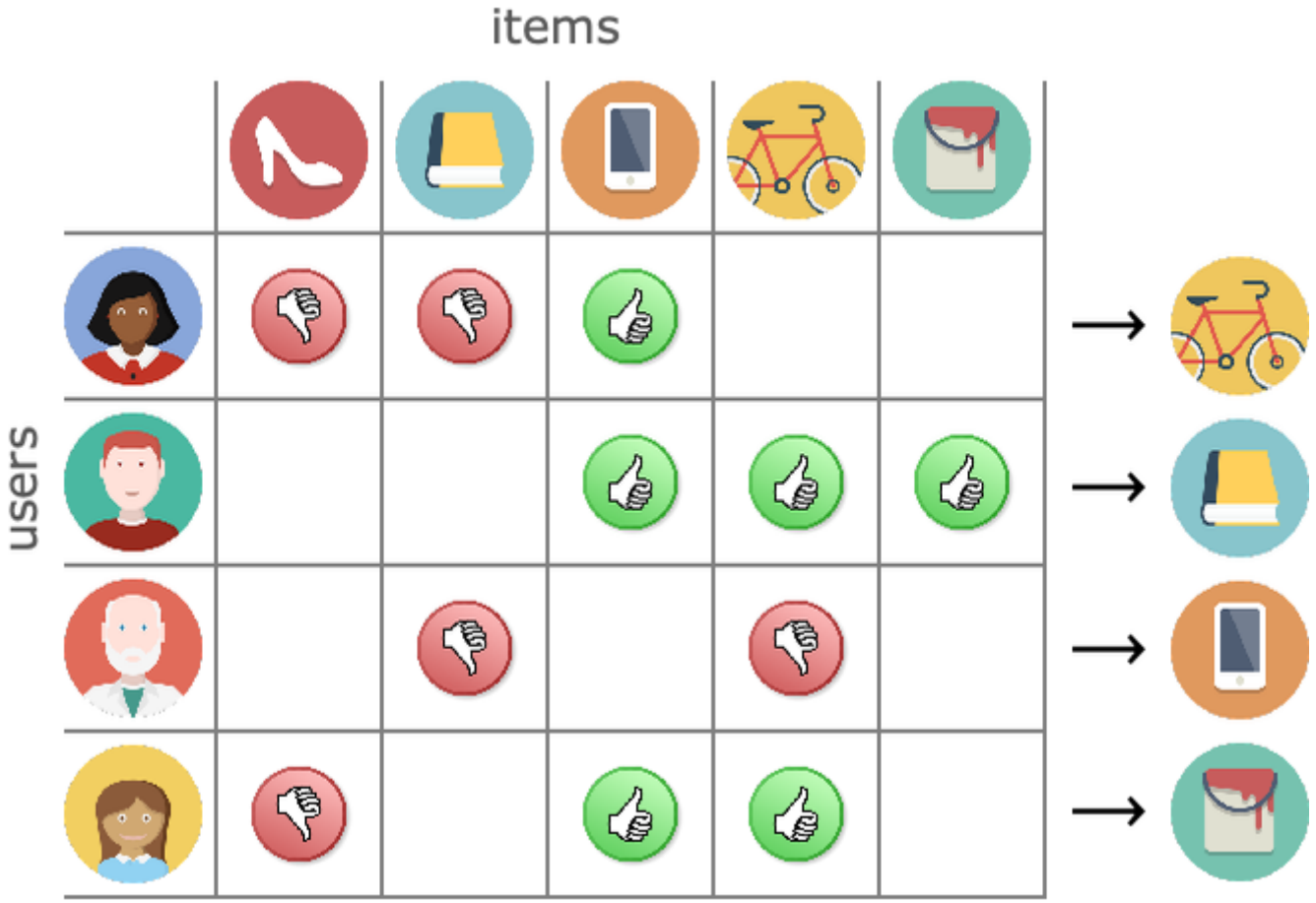
Estas tareas generalmente se consideran parte de la inteligencia artificial (aunque "inteligencia humana" sería más apropiado) y generalmente se abordan utilizando redes neuronales artificiales, un campo conocido como aprendizaje profundo que veremos en la próxima sección. Estas tareas solían ser difíciles o incluso imposibles de resolver en el pasado, pero las cosas han ido cambiando desde la década de 2010 gracias a computadoras más rápidas y al renovado interés en las redes neuronales. En la actualidad, el aprendizaje automático se utiliza ampliamente para resolver estas tareas; por ejemplo, las redes sociales utilizan el aprendizaje automático para analizar grandes cantidades de imágenes y textos con el fin de seleccionar contenido relevante para los usuarios. Cabe destacar que el tipo de datos involucrados en estas tareas (imágenes, audio, texto, etc.) es más complejo y "difuso" que los números organizados en una hoja de cálculo, por lo que se llama datos no estructurados.

Aprendizaje con datos estructurados

Otro ámbito de aplicación importante se refiere a la utilización de grandes cantidades de datos estructurados. Los conjuntos de datos estructurados son los que generalmente se nos vienen a la mente cuando hablamos de datos: números y etiquetas almacenados en hojas de cálculo o bases de datos. Los datos estructurados podrían ser, por ejemplo, datos de ventas recopilados por una empresa minorista: tipo de producto, fechas de venta, precio, etc. La tarea más común al trabajar con datos estructurados es predecir el valor de una variable (también conocida como atributo) de interés, como las futuras cifras de ventas, pero también puede tratarse de comprender los datos, como identificar grupos. A continuación, se presentan ejemplos de tales tareas.

- **Predecir** valores como el precio de una casa o el valor de un stock partiendo de datos históricos
- **Clasificar** los clientes en categorías según sus patrones de comportamiento y su perfil
- **Recomendar** productos en función de los gustos o compras de otros usuarios

Copyright 2026

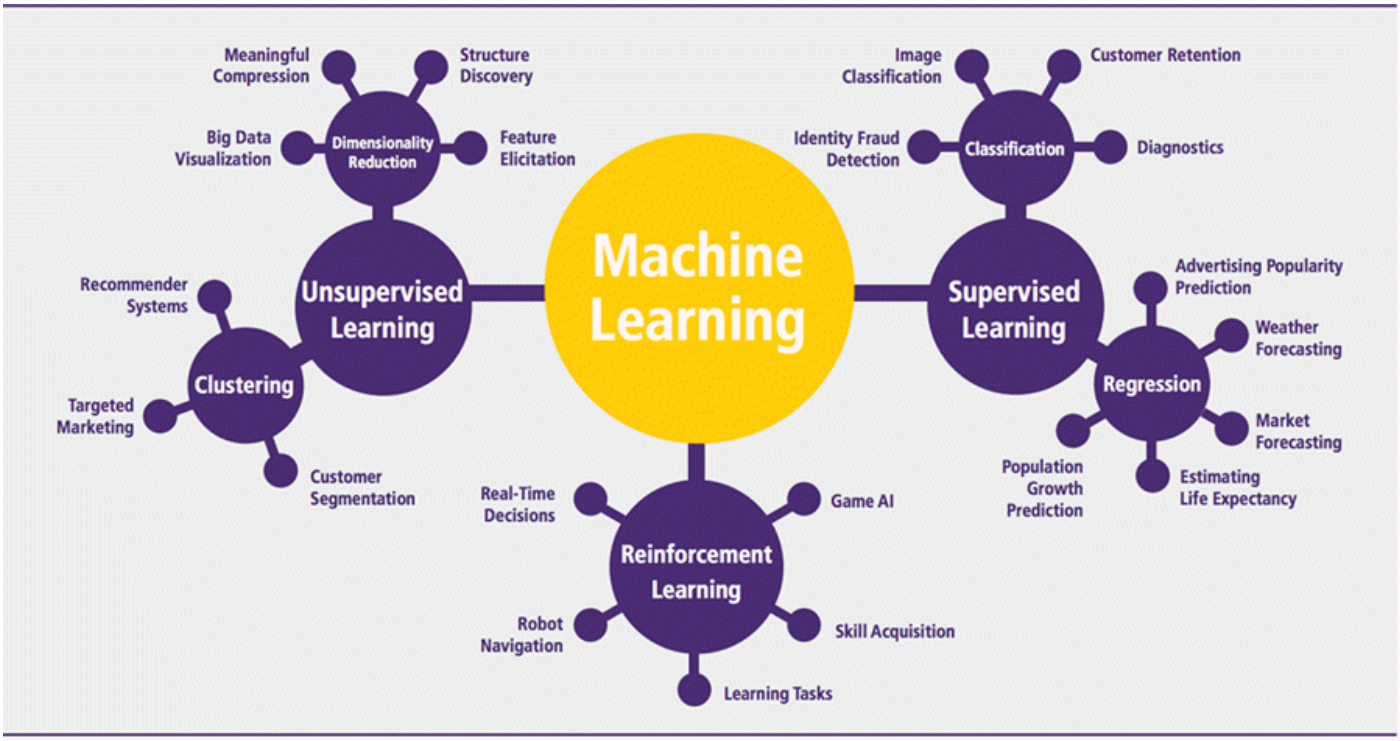


- **Detección de Fraude**

Location	Date	Time	Amount		Fraud probabi
California, United States	Thu 26 Nov 2020	14:09:59	$80.00	→	0.1%
California, United States	Fri 27 Nov 2020	15:29:39	$40.00		0.1%
France	Sat 28 Nov 2020	03:04:15	€1400.00		4%

Para finalizar incluimos una imagen que resume perfectamente las aplicaciones de *Machine Learning*

Copyright 2026



Aclaración El aprendizaje automático a menudo se confunde con la estadística.

Existen muchas similitudes entre estas dos disciplinas, ya que los modelos de aprendizaje automático son modelos estadísticos. En cierto sentido, el aprendizaje automático podría considerarse un subcampo de la estadística. Sin embargo, estas dos áreas difieren en sus objetivos y prácticas. El objetivo del aprendizaje automático generalmente es predecir algo, mientras que el objetivo de la estadística generalmente es comprender algo (por ejemplo, "¿Esta droga ayuda a curar esta enfermedad?"). Como consecuencia, los modelos de aprendizaje automático suelen ser complejos (como conjuntos de árboles, redes neuronales) y "cajas negras", lo que significa que es difícil interpretar lo que hacen. Por otro lado, los modelos en estadística suelen ser simples para poder ser interpretables (como regresión logística, modelos lineales generalizados). Finalmente, el aprendizaje automático a menudo maneja grandes cantidades de datos y de diversos tipos (datos estructurados, imágenes, textos, etc.), mientras que la estadística generalmente trabaja con conjuntos de datos más pequeños y "más simples".

Copyright 2026

Unidad 2.3 Tipos de aprendizaje

> El aprendizaje automático es el nuevo motor eléctrico
>
> **Andrew Ng, British-American computer scientist**

Introducción

La mayoría de humanos funcionamos con patrones, todo lo que hacemos es una copia, mejor o peor imitada, de lo que hemos visto o percibido en el pasado o de lo que nos han dicho. Si vemos repetidamente diferentes versiones de un mismo tipo de animal terminaremos poniendo una etiqueta o un nombre a dicho ser vivo. Cuando vivimos en sociedad dicho nombre nos lo enseñan y así reconocemos distintas variantes que pueden ser nuevas o no vistas anteriormente. Nuestro cerebro se entrena con decenas o cientos de casos para luego poder denotar o saber a qué animal pertenece un nuevo ejemplar.

El aprendizaje automático funciona de manera similar pero de forma computacional, es decir, traduciendo las imágenes (o texto o sonidos) a bits o valores computable por una máquina. Dichos valores permiten entrenar a la máquina igual que lo hace un cerebro humano de manera que sea capaz de agrupar casos o instancias similares (aprendizaje no supervisado) o determinar la categoría (o etiqueta) de un nuevo ejemplar no visto en el proceso de entrenamiento. Es como si partiéramos de un cerebro 'vacío' y empezáramos a entrenarlo generando conexiones entre distintas neuronas.

Así, el aprendizaje automático se basa en la búsqueda de patrones en ingentes cantidades de datos. Dichos patrones de guardan en forma de pesos o parámetros de una red neuronal que nos permitirá estudiar y predecir nuevos casos.

Dicho paradigma tiene tres vertientes o tipos:

1. **Aprendizaje Supervisado**: Es como si le mostrases fotos y le dijeses: “Esto es un gato” o “Esto es un perro”. El robot aprende de los ejemplos que le das para poder decir si la

próxima foto que ve es un gato o un perro.

2. **Aprendizaje No Supervisado**: Es un poco diferente. Aquí, simplemente le das un montón de fotos sin decirle cuál es cuál. El robot tiene que ver las fotos y decir: “Hmm, estas fotos se parecen entre sí, y estas otras fotos se parecen entre sí, así que creo que hay dos grupos diferentes”.
3. **Aprendizaje por Refuerzo**: Es como entrenar a un perro. Cada vez que el robot acierta (por ejemplo, dice correctamente si es un gato o un perro), le das una “golosina virtual” y cuando se equivoca, le dices “inténtalo de nuevo”. Así, el robot está motivado para mejorar y aprender correctamente.

https://www.youtube.com/embed/FtoqQSxJyJA

Evolución de los robots desde 2005

El aprendizaje automático es básicamente enseñarle a una computadora a aprender de los ejemplos y la experiencia, en lugar de seguir instrucciones específicas como en la programación tradicional

Estos paradigmas difieren en las tareas que pueden resolver y en cómo se presenta los datos a la computadora. Por lo general, la tarea y los datos determinan directamente qué paradigma se debe utilizar (y en la mayoría de los casos, es el aprendizaje supervisado). Sin embargo, en algunos casos hay que tomar una decisión. A menudo, estos paradigmas se pueden utilizar juntos para obtener mejores resultados. Este capítulo ofrece una visión general de qué son estos paradigmas de aprendizaje y para qué se pueden utilizar.

Aprendizaje con etiquetas

El aprendizaje supervisado es el paradigma de aprendizaje más común. En el aprendizaje supervisado, la computadora aprende a partir de un conjunto de pares de entrada-salida, que se llaman ejemplos etiquetados:

En este caso identificamos entradas (imágenes de perros y gatos, características de setas, datos del mercado de valores etc...) y salidas (tipo de animal, clase de seta, valor futuro de una acción).

El objetivo del aprendizaje supervisado suele ser entrenar un modelo predictivo a partir de estos pares. Un modelo predictivo es un programa que puede adivinar el valor de salida (también conocido como etiqueta) para una nueva entrada no vista. En pocas palabras, la computadora aprende a predecir utilizando ejemplos reales.

Por ejemplo, consideremos un conjunto de datos de características de animales (ten en cuenta que los conjuntos de datos típicos son mucho más grandes):

EDAD	SEXO	PESO
4 años	hembra	4
6 años	mahco	4.5
5 años y 3 meses	macho	5.1
1 año 3 meses	hembra	1.7

Nuestro objetivo es predecir el peso de un animal a partir de sus otras características, por lo que reescribimos este conjunto de datos como un conjunto de pares de entrada-salida:

Las variables de entrada (en este caso, edad y sexo) generalmente se denominan características, y el conjunto de características que representan un ejemplo se llama vector de características. A partir de este conjunto de datos, podemos generar un modelo que nos permita predecir pesos a partir de nuevos datos (edad y sexo) de entrada.

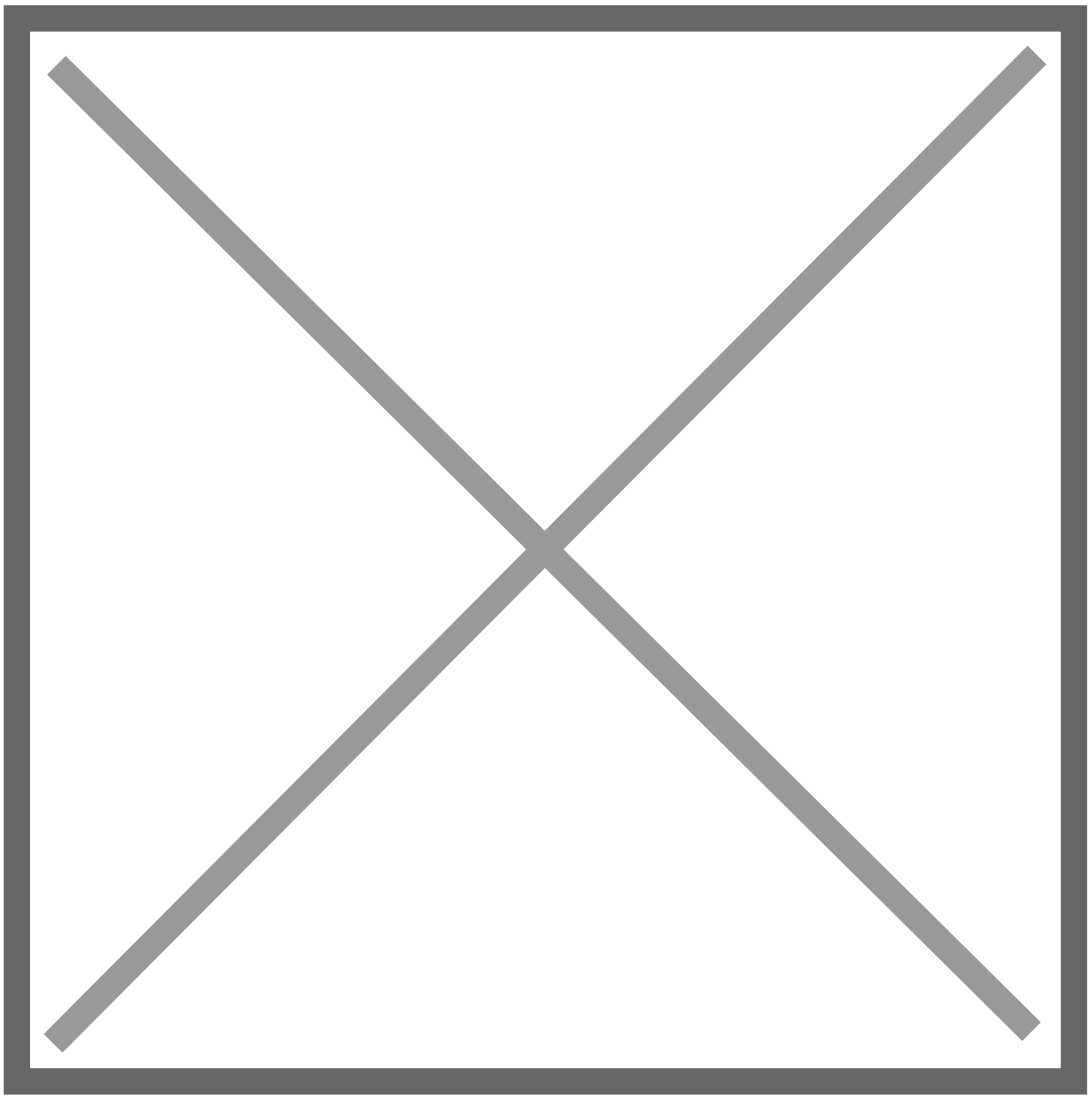
Cuando la variable de salida es numérica hablamos de **regresión**. Si la entrada es texto y la salida es una variable categórica ("gato" o "perro") hablamos de **clasificación**.

Como podemos ver, el aprendizaje supervisado se divide en dos fases: una fase de aprendizaje durante la cual se genera un modelo y una fase de predicción durante la cual se utiliza el modelo. La fase de aprendizaje se llama fase de entrenamiento porque se entrena al modelo para realizar la tarea. La fase de predicción se llama fase de evaluación o fase de inferencia porque la salida se infiere (es decir, se deduce) a partir de la entrada y utilizando el modelo generado.
En esencia el modelo es un conjunto de valores o pesos que se multiplican por los valores de entrada (traducidos a números en el caso de imágenes o texto) para obtener la salida correspondiente.
Es como adivinar los coeficientes en una función matemática:
Si por ejemplo tenemos la ecuación Ax+By+Cz=k podemos construir un modelo que, a partir de distintos valores de x, y y z (muestras) y de k (etiqueta o valor de salida) permita, mediante un proceso iterativo, ajustar los valores de los coeficientes A, B y C de manera que pueda predecir el valor de k para valores de x, y y z no vistos en el proceso de enrtenamiento.
En este caso los coeficientes serían los llamados pesos o parámetros del modelo.
Todo se reduce a hacer modelos que aproximen la realidad de forma que nos permitan predecirla.

Copyright 2026

> "Todos los modelos son incorrectos, pero algunos son útiles
> George E. P. Box, 1919-2013, estadístico británico

La regresión y la clasificación son las principales tareas del aprendizaje supervisado, pero este paradigma va más allá de estas tareas. Por ejemplo, la detección de objetos es una aplicación del aprendizaje supervisado en la que la salida consta de múltiples clases y sus posiciones de caja correspondientes, así el modelo se entrena usando imágenes y sus datos asociados incluyendo las posiciones en forma de rectángulo de cada objeto incluido en la imagen.



Existen numerosas aplicaciones de dicho tipo de aprendizaje, y casi cada día aparecen nuevas, en el siguiente listado hacemos un pqueño resumen de la tipología y posibilidades de las mismas:

Copyright 2026

- **Reconocimiento de Imágenes**: Utilizado en el etiquetado automático de imágenes, reconocimiento facial, y diagnóstico médico a partir de imágenes de radiografía o resonancia magnética.
- **Generación de texto** a partir de un texto anterior usando como entrada frases y palabras de textos reales.
- **Detección de Fraude**: En el sector financiero, para identificar actividades sospechosas en transacciones de tarjetas de crédito o en comportamientos de usuarios en línea.
- **Predicción de Ventas**: En el comercio y la industria para prever tendencias de ventas y gestionar inventarios en base a datos históricos.
- **Diagnósticos Médicos**: Analizando registros médicos para ayudar en el diagnóstico precoz de enfermedades.
- **Reconocimiento de Voz y Procesamiento del Lenguaje Natural (NLP)**: Para convertir el habla en texto, traducción automática, y análisis de sentimiento en textos.
- **Predicción de Riesgos de Crédito**: En el sector bancario, para evaluar la solvencia de los solicitantes de crédito.
- **Personalización de Contenidos**: En plataformas de streaming y servicios en línea, para recomendar películas, música o productos basados en las preferencias del usuario.
- **Análisis de Series Temporales**: Para predecir valores futuros en datos financieros, meteorológicos o de cualquier otro tipo que se registre a lo largo del tiempo.
- **Clasificación de Documentos**: En sistemas de gestión de información, para clasificar y organizar documentos automáticamente.
- **Control de Calidad en Manufactura**: Para detectar defectos en productos en las líneas de ensamblaje a través de imágenes y sensores.

Mientras los datos de entrenamiento consistan en un conjunto de pares de entrada-salida, es una tarea de aprendizaje supervisado.

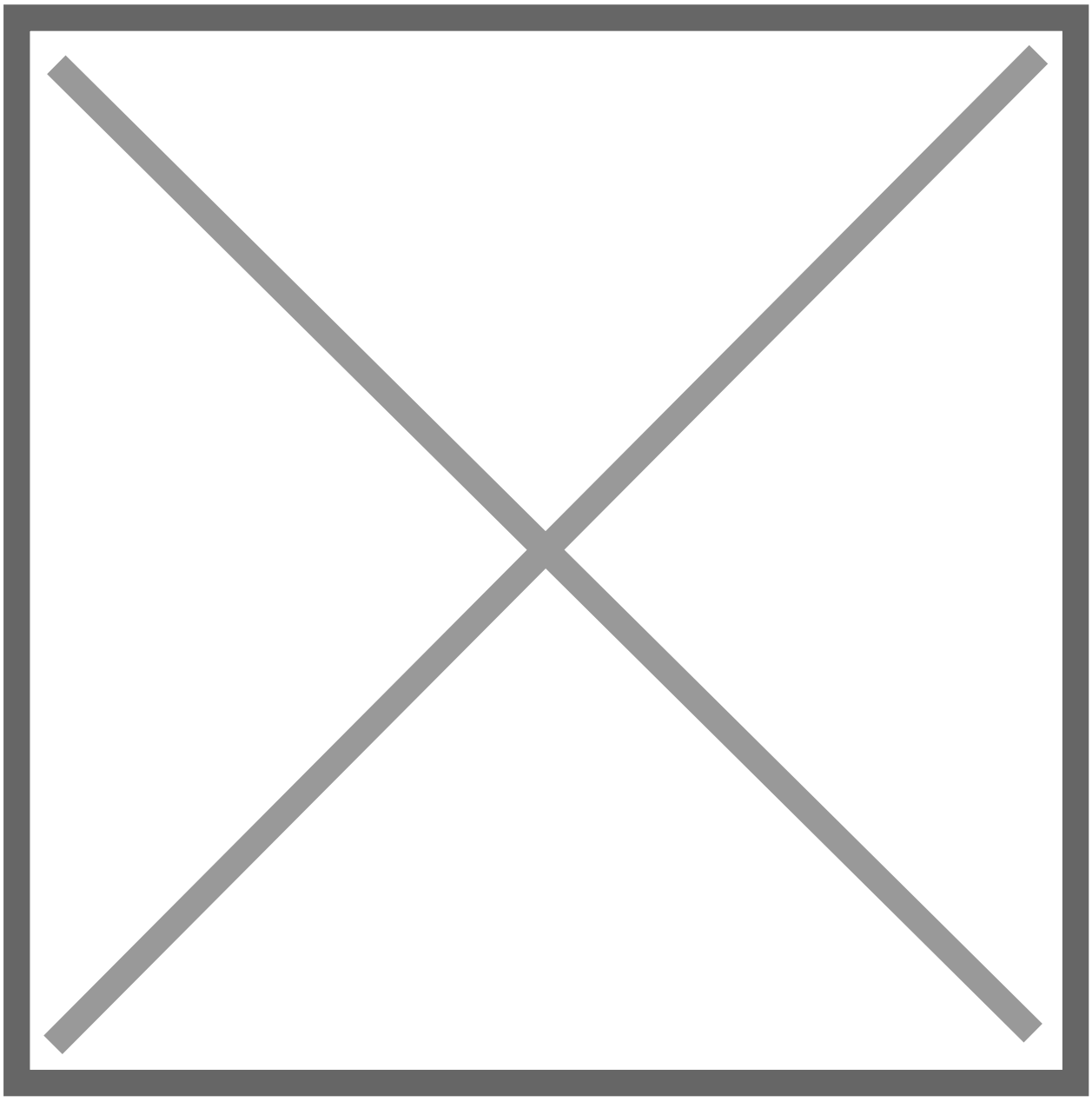
Actualmente, la mayoría de las aplicaciones de aprendizaje automático que se desarrollan utilizan un enfoque de aprendizaje supervisado. Una razón para esto es que las principales tareas supervisadas (clasificación y regresión) son útiles y están bien definidas, y a menudo se pueden abordar utilizando algoritmos simples. Otra razón es que se han desarrollado muchas herramientas para este paradigma. Sin embargo, la principal desventaja del aprendizaje supervisado es que necesitamos tener datos etiquetados, lo cual puede ser difícil de obtener en algunos casos.

Aprendizaje No Supervisado

El aprendizaje no supervisado es el segundo paradigma de aprendizaje más utilizado. No se utiliza tanto como el aprendizaje supervisado, pero es más potente y se asemeja más al funcionamiento del cerebro humano. En el aprendizaje no supervisado, no hay entradas ni salidas, los datos son simplemente un conjunto de ejemplos caracterizados por ciertos parámetros que permiten su agrupación en conjuntos similares de entidades.

Copyright 2026

El aprendizaje no supervisado se puede utilizar para una amplia gama de tareas. Una de ellas se llama agrupamiento o *clustering* y su objetivo es separar los ejemplos de datos en grupos llamados clústeres:



Una aplicación del agrupamiento podría ser separar automáticamente a los clientes de una empresa para crear mejores campañas de marketing (targeted marketing) .

El agrupamiento también se utiliza simplemente como una herramienta de exploración para obtener información sobre los datos y tomar decisiones fundamentadas. También permite detectar casos extraños (outliers) que suelen asociarse a fraudes o errores al ser ejemplos que quedan fuera de cualquier grupo, por ejemplo en un conjunto de transacciones o patrones de

Copyright 2026

comportamiento de clientes de banca.

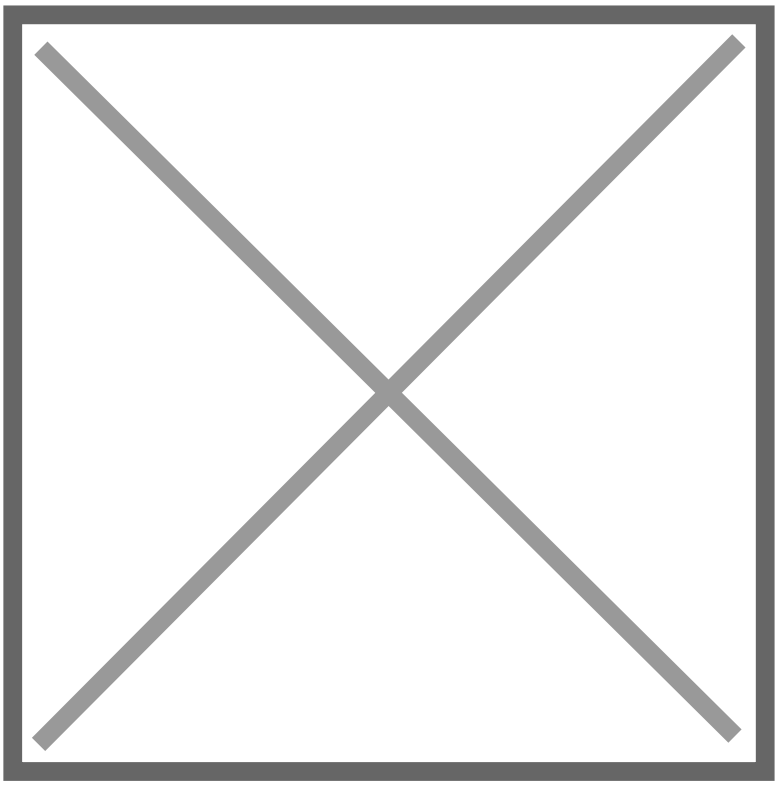
Otra tarea clásica de aprendizaje no supervisado se denomina reducción de dimensionalidad. El objetivo de la reducción de dimensionalidad es reducir el número de variables en un conjunto de datos al tiempo que se intenta preservar algunas propiedades de los datos que son más significativas (por ejemplo el dato 'id de pedido' en una base de datos comercial no aporta información y puede omitirse, o en una base de datos de coches el peso y la potencia suelen estar correlacionados por lo que podemos prescindir de uno de ellos).

Muchos datos incluyen decenas de parámetros que los definen, por ejemplo en una imagen cada pixel es un parámetro o dimensión, usar muchas imágenes con muchos pixeles complica demasiado y sobre todo ralentiza el algoritmo con lo que conviene reducir el número de pixeles eliminando los superfluos o que no aportan mucha información.

Los modelos se entrenan con millones de datos y dimensiones o características así que es importante optimizarlos para reducir el tiempo de entrenamiento y aumentar la eficiencia.

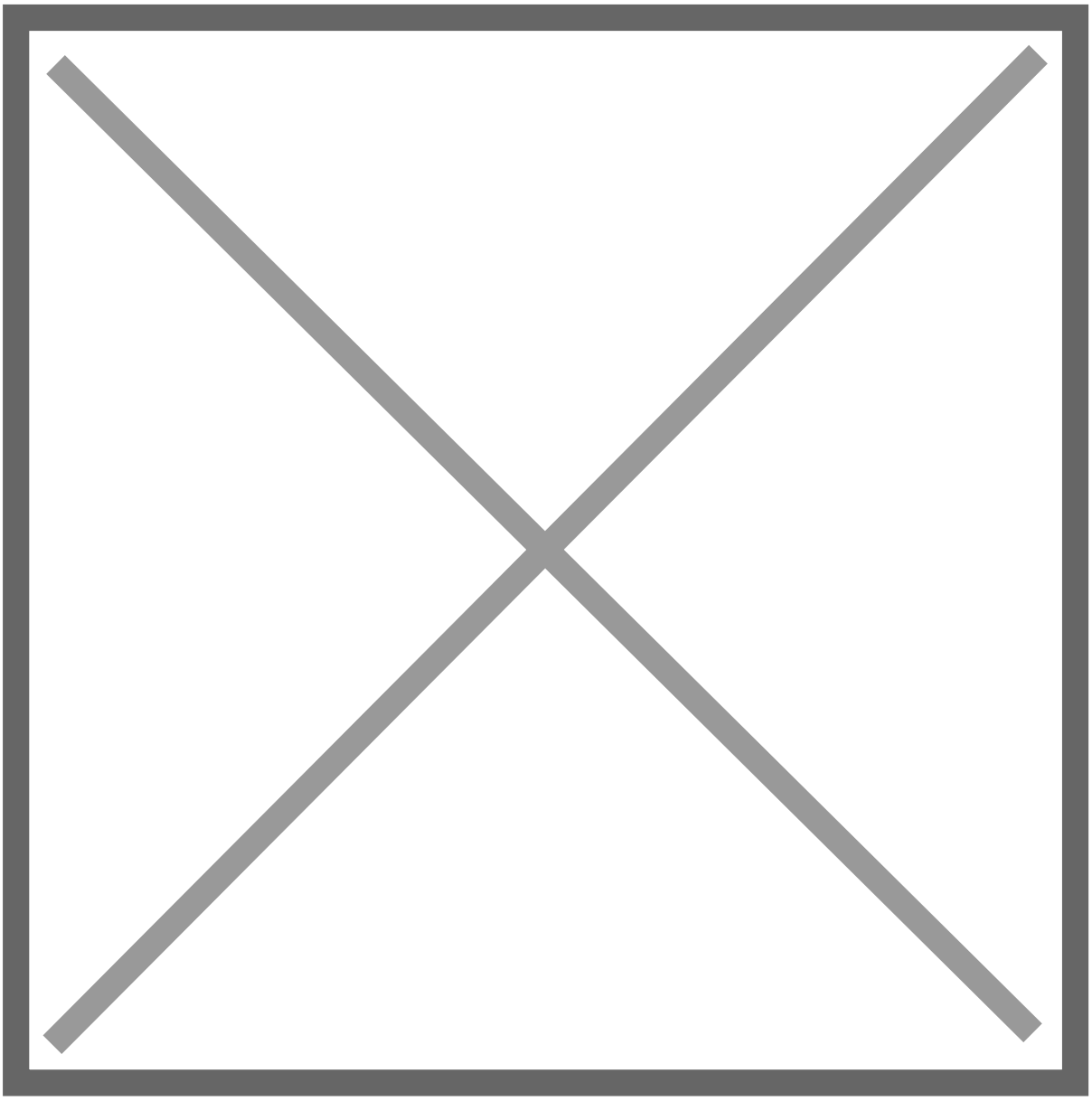
La reducción de dimensionalidad se puede utilizar para una variedad de tareas, como comprimir los datos, aprender con etiquetas faltantes, crear motores de búsqueda o incluso crear sistemas de recomendación. La reducción de dimensionalidad también se puede utilizar como una herramienta de exploración para visualizar un conjunto de datos completo en un espacio reducido.

Copyright 2026



La detección de anomalías es otra tarea que se puede abordar de manera no supervisada. La detección de anomalías se refiere a la identificación de ejemplos que son anómalos, es decir, valores atípicos. Aquí vemos un ejemplo de detección de anomalías realizada en un conjunto de datos numéricos simple:

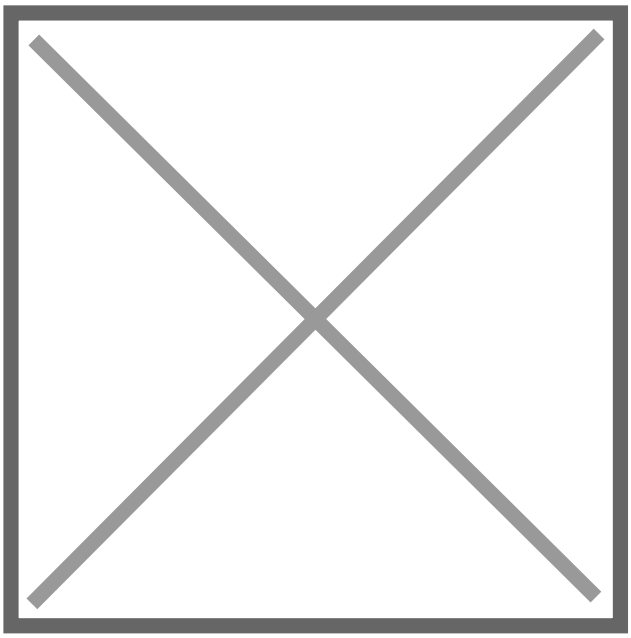
Copyright 2026



Esta tarea podría ser útil para detectar transacciones fraudulentas con tarjetas de crédito, limpiar un conjunto de datos o detectar cuando algo está saliendo mal en un proceso de fabricación.

Otra tarea clásica de aprendizaje no supervisado se denomina imputación de valores faltantes, y el objetivo es completar los valores faltantes en un conjunto de datos:

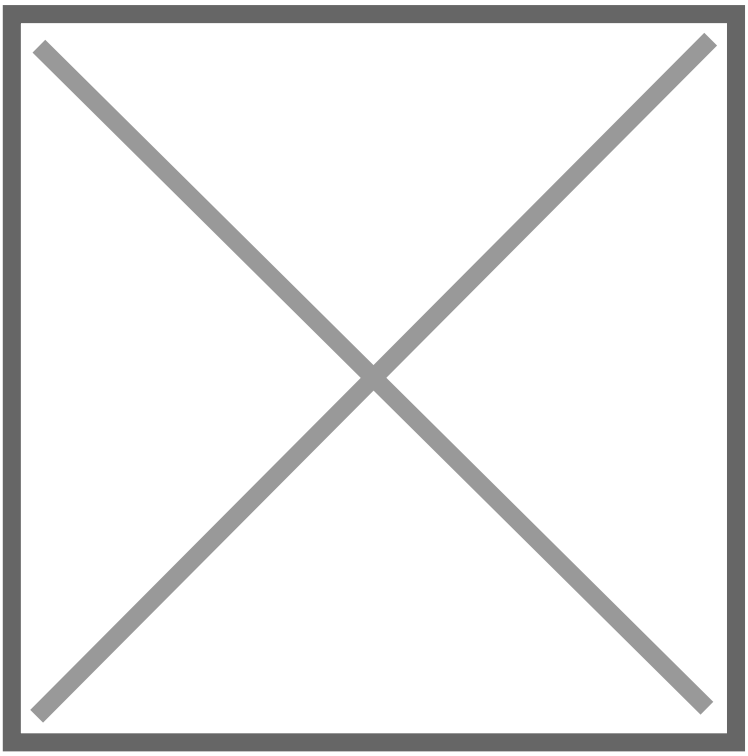
Copyright 2026



Esta tarea es extremadamente útil porque la mayoría de los conjuntos de datos tienen valores faltantes y muchos algoritmos no pueden manejarlos. En algunos casos, las técnicas de imputación de valores faltantes también se pueden utilizar para tareas predictivas, como los motores de recomendación.

Finalmente, la tarea de aprendizaje no supervisado más difícil probablemente sea aprender a generar ejemplos similares a los datos de entrenamiento. Esta tarea se llama modelado generativo y se puede utilizar, por ejemplo, para aprender a generar nuevos rostros a partir de muchos ejemplos de rostros. Aquí vemos un ejemplo de rostros sintéticos generados por una red neuronal a partir de ruido aleatorio usando modelos entrenados con imágenes reales:

Copyright 2026



Técnicas de generación como estas también se pueden utilizar para mejorar la resolución, reducir el ruido o completar valores faltantes.

Como ejemplos genéricos de su uso proponemos los siguientes:

1. **Segmentación de Clientes**: Utilizado en marketing para agrupar clientes en segmentos basados en características similares, lo que ayuda a personalizar estrategias de marketing.
2. **Detección de Anomalías**: Para identificar comportamientos anómalos o inusuales en diversas áreas, como fraudes en tarjetas de crédito, fallos en máquinas o actividad sospechosa en redes informáticas.
3. **Agrupación de Documentos**: Para organizar automáticamente grandes colecciones de documentos, como artículos científicos o noticias, en categorías temáticas.
4. **Recomendación de Productos**: En comercio electrónico y plataformas de streaming, para sugerir productos o contenidos basados en preferencias de usuarios similares.
5. **Análisis de Redes Sociales**: Para identificar comunidades o tipos de mensajes/imágenes
6. **Exploración de Datos Genéticos**: En biología y medicina, para clasificar y estudiar patrones genéticos y biomarcadores.
7. **Análisis de Patrones en Series Temporales**: Como en la detección de tendencias en los mercados financieros o en datos meteorológicos.
8. **Optimización de Procesos Industriales**: Para identificar patrones y mejorar la eficiencia en procesos de manufactura.

Copyright 2026

9. **Visión por Computadora**: En la identificación y clasificación de objetos en imágenes cuando no están etiquetados previamente.

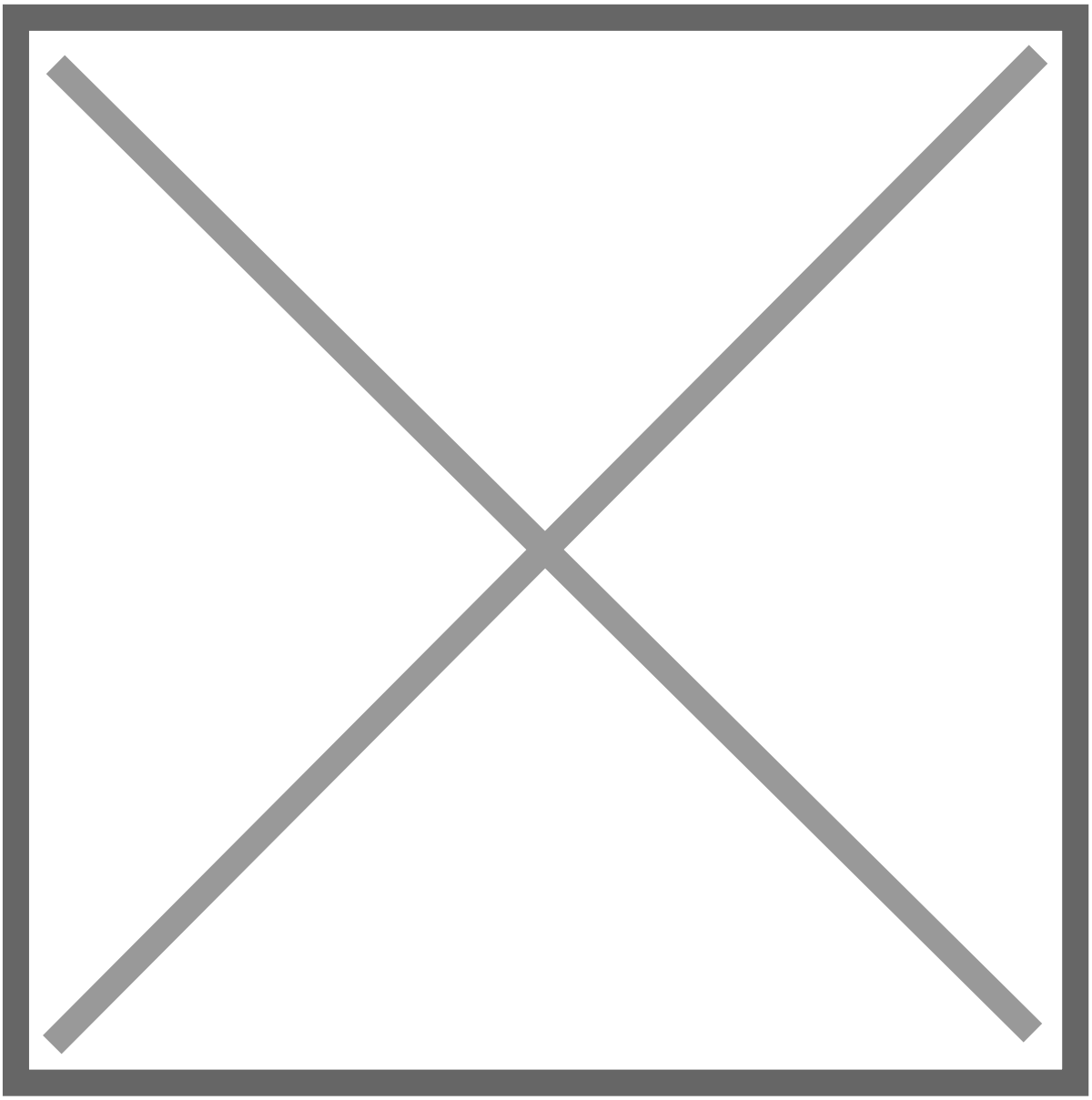
El aprendizaje no supervisado se utiliza un poco menos que el aprendizaje supervisado, principalmente porque las tareas que resuelve son menos comunes y más difíciles de implementar que las tareas predictivas. Sin embargo se puede aplicar a un conjunto más diverso de tareas que el aprendizaje supervisado. En la actualidad, el aprendizaje no supervisado es un elemento clave en muchas aplicaciones de aprendizaje automático y también se utiliza como una herramienta para explorar datos. Además, muchos investigadores creen que el aprendizaje no supervisado es la forma en que los humanos adquieren la mayor parte de sus conocimientos y, por lo tanto, será la clave para desarrollar sistemas artificialmente inteligentes en el futuro.

Aprendizaje de refuerzo

El tercer paradigma de aprendizaje clásico se llama aprendizaje por refuerzo, que es una forma en que los agentes autónomos aprenden. El aprendizaje por refuerzo es fundamentalmente diferente del aprendizaje supervisado y no supervisado en el sentido de que los datos no se proporcionan como un conjunto fijo de ejemplos. En cambio, los datos para aprender se obtienen interactuando con un sistema externo llamado entorno. El nombre "aprendizaje por refuerzo" proviene de la psicología del comportamiento, pero también podría llamarse "aprendizaje interactivo".

El aprendizaje por refuerzo se utiliza a menudo para enseñar a los agentes, como robots, a aprender una tarea determinada. El agente aprende tomando acciones en el entorno y recibiendo observaciones de este entorno:

Copyright 2026



Normalmente, el agente comienza su proceso de aprendizaje actuando al azar en el entorno, y luego el agente aprende gradualmente de su experiencia para realizar mejor la tarea utilizando una especie de estrategia de prueba y error. El aprendizaje generalmente está guiado por una recompensa que se le da al agente según su desempeño. Más precisamente, el agente aprende una política que maximiza esta recompensa. Una política es un modelo que predice qué acción tomar dados los anteriores acciones y observaciones.

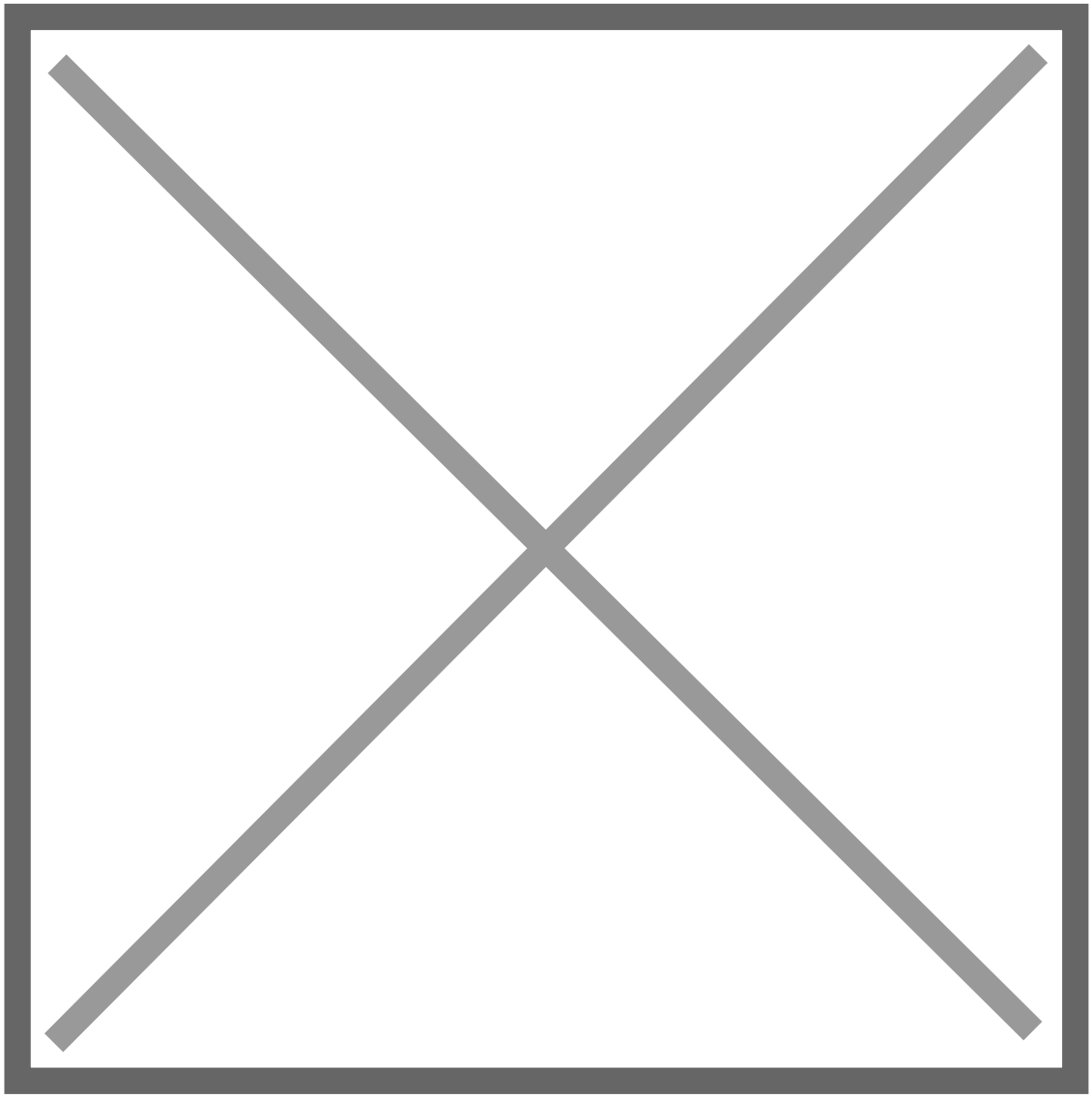
El aprendizaje por refuerzo se puede utilizar, por ejemplo, para que un robot aprenda a caminar en un entorno simulado. Aquí hay una captura de pantalla de un vídeo mostrado anteriormente que ilustra el aprendizaje por refuerzo hecho y grabado por la empresa Boston Dynamics. En él se fija un objetivo como es mantener el equilibrio, y se van ajustando las variables (posición y sincronización de elementos móviles) para poco a poco lograr una estabilidad que contribuya al equilibrio:

Copyright 2026



> Enlace al vídeo

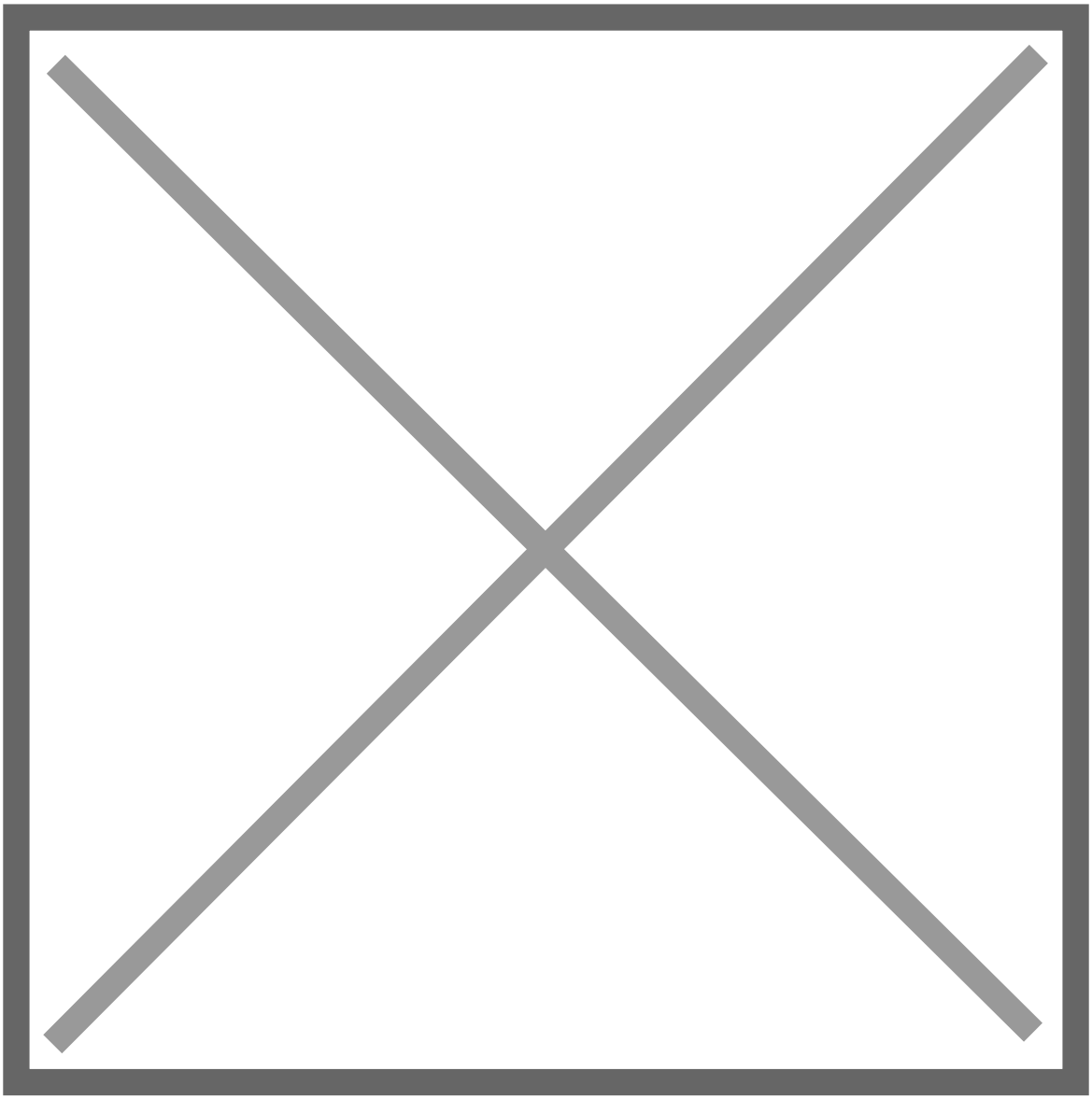
En este caso, las acciones son los valores de torsión aplicados a cada articulación de la pata; las observaciones son los ángulos de las patas, las fuerzas externas, etc.; y la recompensa está unida al equilibrio y la velocidad del robot. El aprendizaje en un entorno simulado como este puede luego utilizarse para ayudar a un robot real a caminar.
Esta transferencia de la simulación a la realidad se ha utilizado, por ejemplo, por *OpenAI* para enseñar a un robot a manipular un cubo de *Rubik*:



También es posible que un robot real aprenda sin un entorno simulado, pero los robots reales son más lentos en comparación con los simulados y los algoritmos actuales tienen dificultades para aprender lo suficientemente rápido. Una estrategia de mitigación consiste en aprender a simular el entorno real, un campo conocido como aprendizaje por refuerzo basado en modelos, que está siendo objeto de investigación activa.

El aprendizaje por refuerzo también se puede utilizar para enseñar a las computadoras a jugar juegos. Ejemplos famosos incluyen *AlphaGo*, que puede vencer a cualquier jugador humano en el juego de mesa *Go*, o *AlphaStar*, que puede hacer lo mismo en el videojuego *StarCraft*:

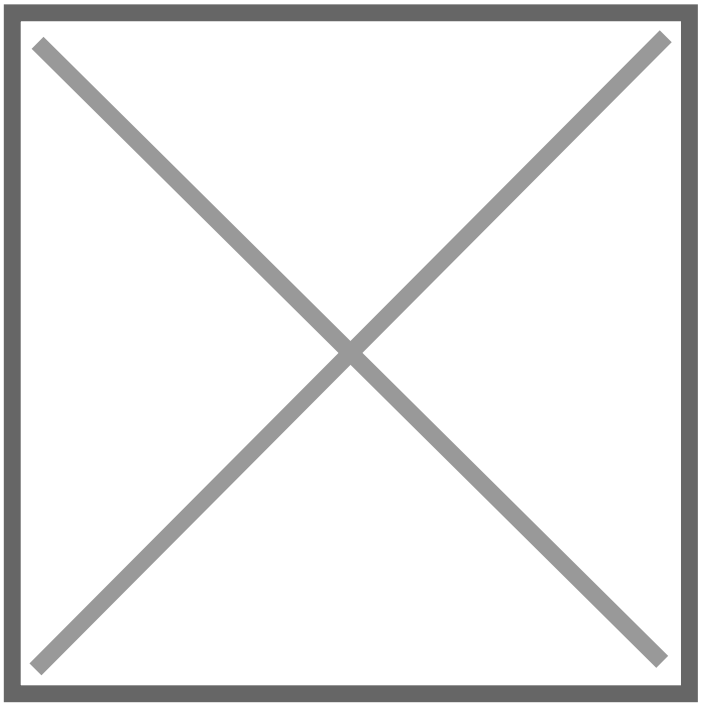
Copyright 2026



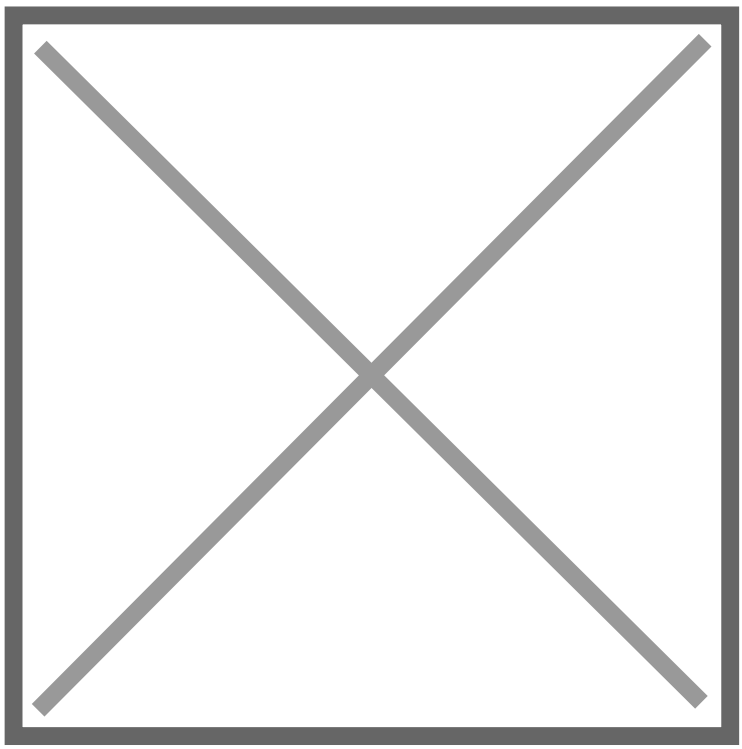
Ambos programas fueron desarrollados utilizando aprendizaje por refuerzo al hacer que el agente juegue contra sí mismo. Cabe destacar que la recompensa en este tipo de problemas solo se otorga al final del juego (ya sea que ganes o pierdas), lo que dificulta aprender qué acciones fueron responsables del resultado.

Otra aplicación importante del aprendizaje por refuerzo se encuentra en el campo de la ingeniería de control. El objetivo aquí es controlar de manera dinámica el comportamiento de un sistema (un motor, un edificio, etc.) para que se comporte de manera óptima. El ejemplo prototípico es controlar un poste que se encuentra sobre un carrito moviendo el carrito hacia la izquierda o hacia la derecha (también conocido como péndulo invertido):

Copyright 2026



En general, se utilizan métodos de control clásicos para este tipo de problemas, pero el aprendizaje por refuerzo está entrando en este campo. Por ejemplo, se ha utilizado el aprendizaje por refuerzo para controlar el sistema de enfriamiento (velocidad del ventilador, flujo de agua, etc.) de los centros de datos de Google de manera más eficiente:



Copyright 2026

> Un problema al aplicar el aprendizaje por refuerzo directamente en un sistema del mundo real es que durante la fase de aprendizaje, el agente podría realizar acciones que podrían dañar el sistema o plantear problemas de seguridad.

El aprendizaje por refuerzo es probablemente el paradigma más emocionante, ya que el agente aprende interactuando, como un ser vivo. Los sistemas activos tienen el potencial de aprender mejor que los sistemas pasivos porque pueden decidir por sí mismos qué explorar para mejorar. Podemos imaginar todo tipo de aplicaciones utilizando este paradigma, desde un robot agrícola que aprende a mejorar la producción de cultivos, hasta un programa que aprende a operar en el mercado de valores, o un *chatbot* que aprende teniendo conversaciones con humanos. Desafortunadamente, los algoritmos actuales requieren una gran cantidad de datos para ser efectivos, por lo que la mayoría de las aplicaciones de aprendizaje por refuerzo utilizan entornos virtuales. Además, los problemas de aprendizaje por refuerzo son generalmente más complicados de manejar que los problemas supervisados y no supervisados. Por estas razones, el aprendizaje por refuerzo se utiliza menos que otros paradigmas en aplicaciones prácticas. A medida que la investigación avanza, es probable que los algoritmos necesiten menos datos para operar y se desarrollen herramientas más simples. El aprendizaje por refuerzo podría convertirse en un paradigma dominante en el futuro.

Copyright 2026

Unidad 2.4. Imitando al cerebro. Redes neuronales.

Métodos de Aprendizaje Profundo

> Este capítulo tiene ciertas cuestiones técnicas que no es necesario entender, se dejan para los lectores que tengan más interés pero en absoluto son un requisito para superar el curso.
>
> Simplemente hemos querido incluirlas por completitud y por dejar clalro que la Inteligencia Artificial no es alg mágico que surge de lo desconocido, sino que nace, en esencia, de un tratamiento estadístico de la información

Dentro del campo del aprendizaje automático surgen las llamadas redes neuronales que, aunque conceptualmente existen desde los años 60, se redefinieron en la década de 2010 y mostraron un rendimiento impresionante en datos de imágenes, texto y audio. Estos métodos se basan principalmente en redes neuronales artificiales, que fueron experimentadas por primera vez en la década de 1950. En aquel momento, las redes neuronales eran principalmente un tema de investigación y no se utilizaban tanto en aplicaciones prácticas. Gracias a la velocidad de las computadoras modernas y a algunas innovaciones algorítmicas, los métodos de aprendizaje profundo se utilizan actualmente de manera intensiva, especialmente en problemas de visión artificial y en gestión del lenguaje natural.

Conviene señalar que hay dos tareas fundamentales en el aprendizaje humano, a saber, la clasificación y la genreación o predicción.

Gran parte de la IA actual y de toda la revolucón que estamos viviendo tiene que ver con ambas, la generación en lo que conocemos como **IA Generativa** y la clasificaicón en tareas de visión aritifcial principalmente.

Copyright 2026

En esta sección nos ocupamos de describir la esencia de las redes neuronales que han permitido el desarrollo de las ténicas de IA genrativa más importantes.

Las Neuronas y el cerebro

El cerebro humano es un órgano increíblemente complejo y fascinante. Está compuesto por miles de millones de células llamadas neuronas, que son las unidades básicas del sistema nervioso. Estas neuronas están interconectadas en una red compleja y trabajan juntas para procesar información y controlar nuestras funciones cognitivas y corporales.

Imaginemos que el cerebro es una gran red de comunicación, donde cada neurona es como un pequeño nodo que envía y recibe mensajes. Estas neuronas se comunican entre sí a través de conexiones especializadas llamadas sinapsis. En estas sinapsis, las neuronas transmiten señales eléctricas y químicas para enviar información de un lugar a otro.

Cuando una neurona recibe una señal de otra neurona a través de sus dendritas, que son como pequeñas ramificaciones que se extienden desde la célula, se genera un impulso eléctrico. Este impulso eléctrico viaja a través del cuerpo de la neurona hacia su axón, que es como un largo cable que lleva la señal hacia las sinapsis.

En las sinapsis, la señal eléctrica se transforma en una señal química. La neurona emisora libera sustancias químicas llamadas neurotransmisores en el espacio entre las células, y estos neurotransmisores se unen a receptores en la neurona receptora, desencadenando un nuevo impulso eléctrico en esa neurona. Este proceso de señalización electroquímica se repite una y otra vez a lo largo de la red neuronal, permitiendo la comunicación y el procesamiento de información.

Es importante destacar que el cerebro humano no funciona de manera lineal, como una cadena de instrucciones paso a paso. En cambio, funciona de manera altamente paralela y distribuida, con múltiples neuronas trabajando simultáneamente en diferentes partes del cerebro. Esta actividad neuronal en paralelo y distribuida es lo que nos permite realizar tareas complejas como pensar, recordar, sentir emociones y realizar acciones.

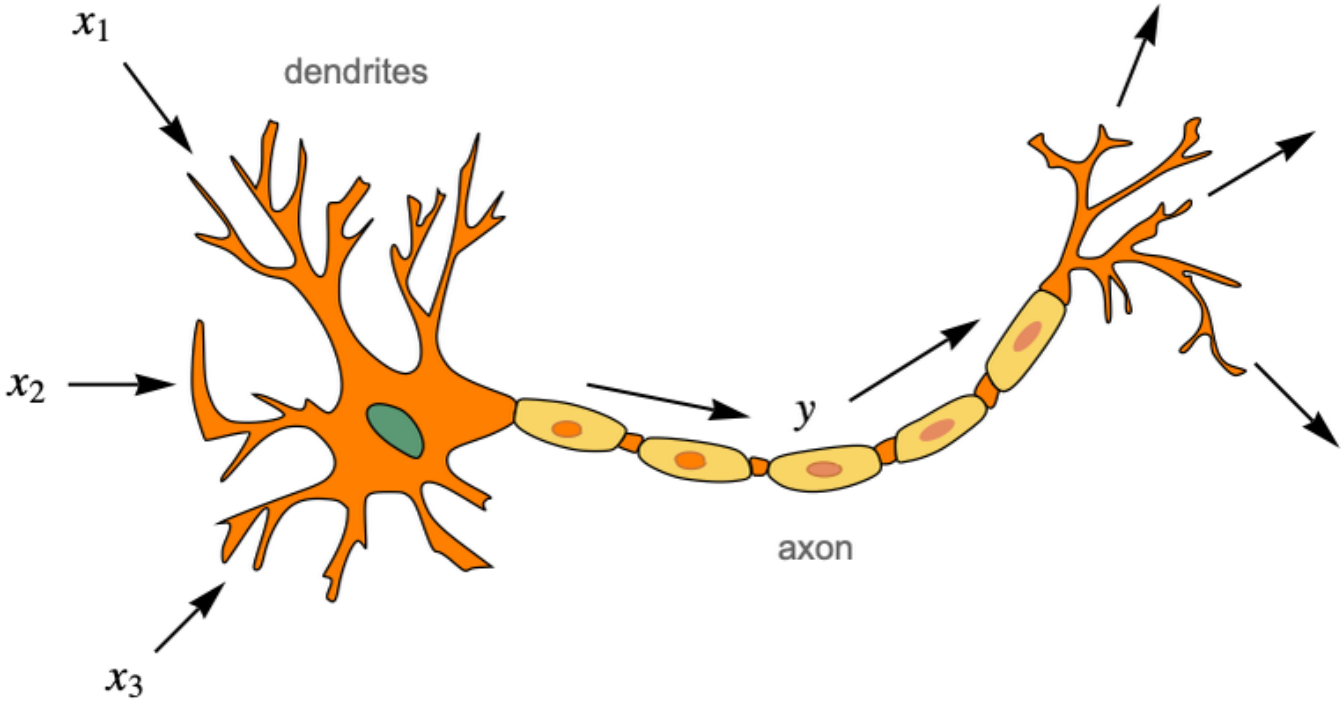
El funcionamiento exacto del cerebro y cómo las neuronas procesan y almacenan información sigue siendo objeto de intensa investigación. Los neurocientíficos continúan estudiando y descubriendo nuevos aspectos sobre el funcionamiento del cerebro y cómo se relaciona con nuestras experiencias y comportamientos.

Neurona Artificial

Las redes neuronales artificiales se inspiran en lo que sabemos sobre el cerebro. En pocas palabras, el cerebro es un sistema de procesamiento de información compuesto por células

Copyright 2026

llamadas neuronas, que están interconectadas en una red. Las neuronas transmiten señales eléctricas a otras neuronas mediante estas conexiones y, juntas, son capaces de realizar cálculos que determinan nuestro comportamiento. Los seres humanos tienen alrededor de cien mil millones de neuronas en sus cerebros y aproximadamente diez mil veces más conexiones. Aquí tienes una representación clásica de lo que es una neurona biológica:



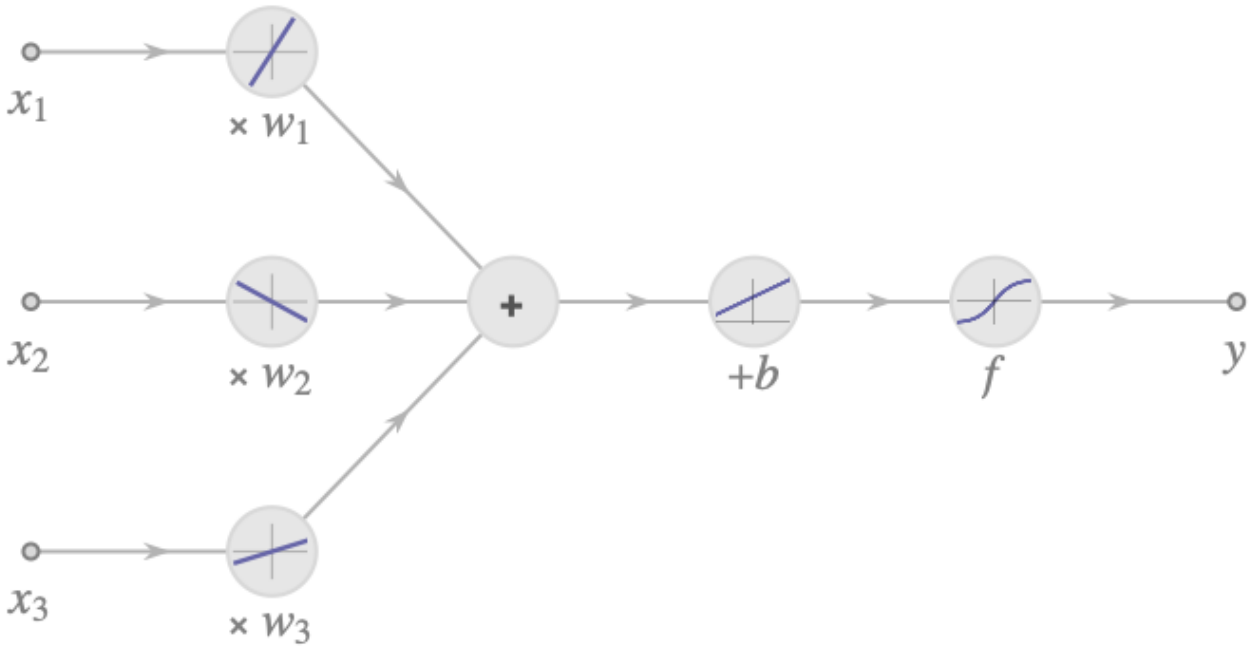
En el lado izquierdo podemos ver las dendritas, que son ramificaciones desde donde la neurona recibe sus entradas eléctricas (mostradas como x1, x2 y x3 aquí). La célula luego "calcula" una salida eléctrica (mostrada como y aquí), que viaja a lo largo del axón y se envía a muchas otras neuronas (potencialmente miles) a través de pequeñas uniones llamadas sinapsis.

Existen una gran variedad de neuronas biológicas y realizan diferentes operaciones. Tienen en común que "disparan" señales eléctricas agudas llamadas picos si se cumplen ciertas condiciones en sus entradas y estados internos. Estos cálculos analógicos son difíciles de simular y, si bien existen muchos modelos de computación de neuronas biológicas, son poco prácticos para el aprendizaje automático.

Las redes neuronales artificiales no intentan imitar exactamente las redes biológicas. En cambio, utilizan los mismos principios subyacentes al tiempo que mantienen las cosas simples y prácticas (de la misma manera que los aviones no imitan a las aves). Las redes neuronales artificiales utilizan neuronas artificiales, que son mucho más simples que sus contrapartes biológicas. Dados

los valores numéricos x1, x2 y x3, la neurona artificial realiza el siguiente cálculo:

Aquí, w1, w2 y w3 son parámetros ajustables llamados pesos (los parámetros en los modelos de lenguaje), que pueden interpretarse como "fortalezas" de conexión entre neuronas. Esto podría corresponder al número de sinapsis entre dos neuronas biológicas. b es otro parámetro ajustable llamado sesgo. En las neuronas biológicas, este valor podría interpretarse como un umbral por encima del cual la neurona dispara. f es una función no lineal llamada función de activación o función de transferencia. Las neuronas biológicas también utilizan algún tipo de función de activación no lineal, ya que o bien disparan o no. Aquí tienes una ilustración del cálculo realizado por esta neurona artificial:



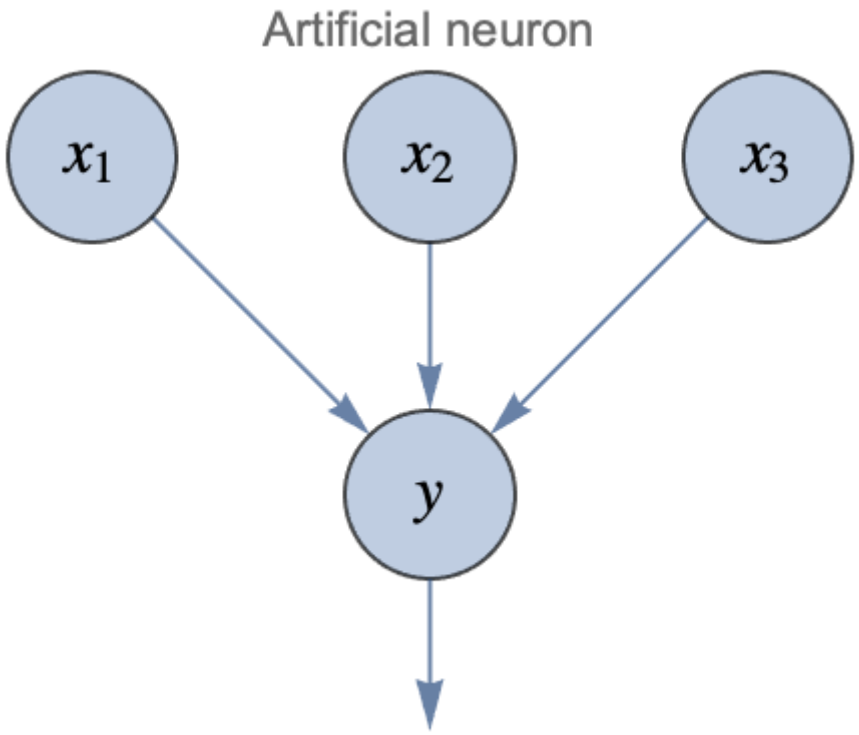
La primera parte es una combinación lineal de las características y luego se aplica una no linealidad. La presencia de esta no linealidad es importante. Permite que las redes neuronales modelen sistemas no lineales (porque la composición de funciones lineales sigue siendo una función lineal). Dado que las neuronas biológicas disparan o no disparan, resulta tentador utilizar algún tipo de función de activación escalón.

Los modelos modernos de *deep learning* tienden a alejarse de las interpretaciones biológicas. Sin embargo, sorprendentemente, siguen utilizando los mismos principios descritos aquí: combinaciones lineales de entradas seguidas de no linealidades.

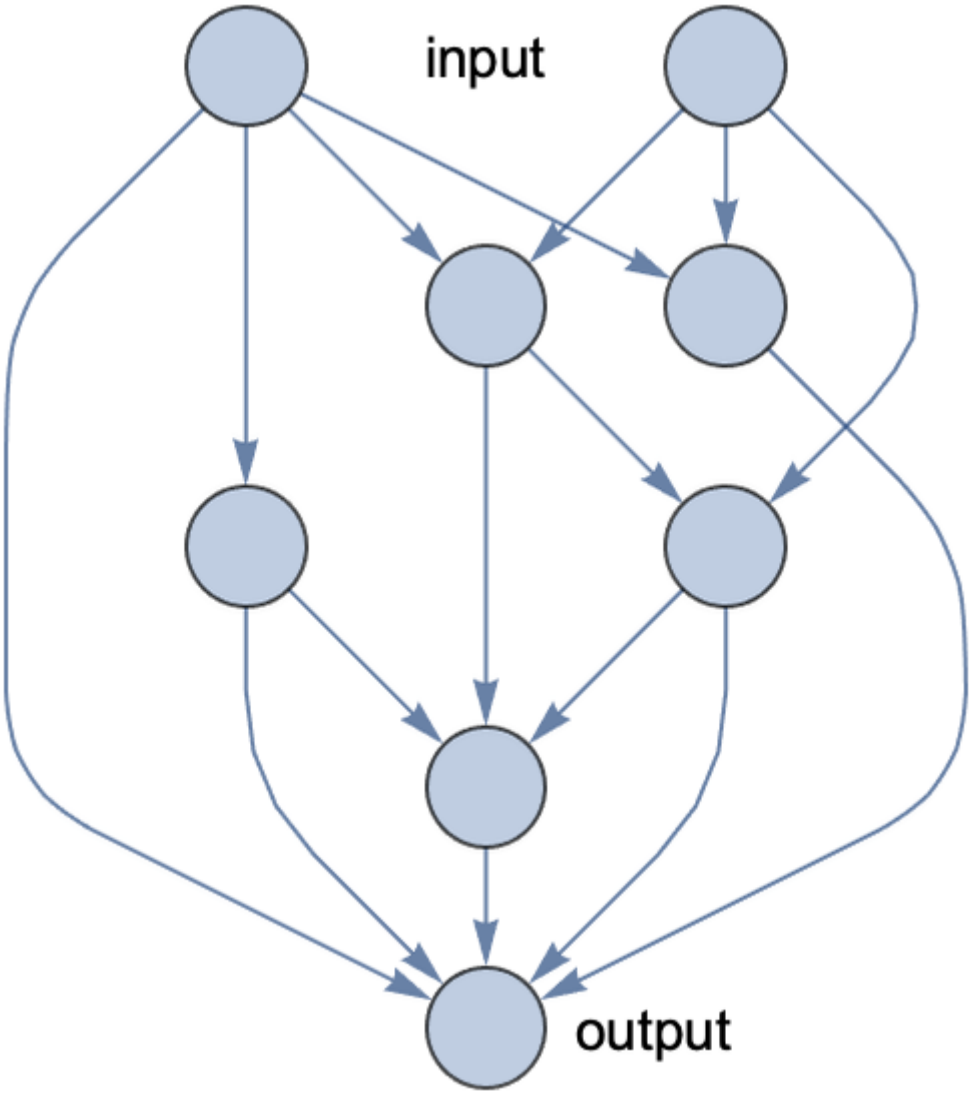
Copyright 2026

Redes Neuronales

Ahora que tenemos una neurona artificial, podemos usarla para crear redes conectando muchas de ellas juntas. Cuando forman parte de una red neuronal, a menudo se les llama unidades y generalmente se representan de la siguiente manera:



Los círculos representan valores numéricos llamados activaciones. Se da por hecho que "y" es una combinación lineal de sus entradas más algún término de sesgo y que el resultado se pasa a través de una no linealidad. Siguiendo esta convención, esto es cómo podría lucir una red neuronal artificial con conexiones aleatorias entre las neuronas:

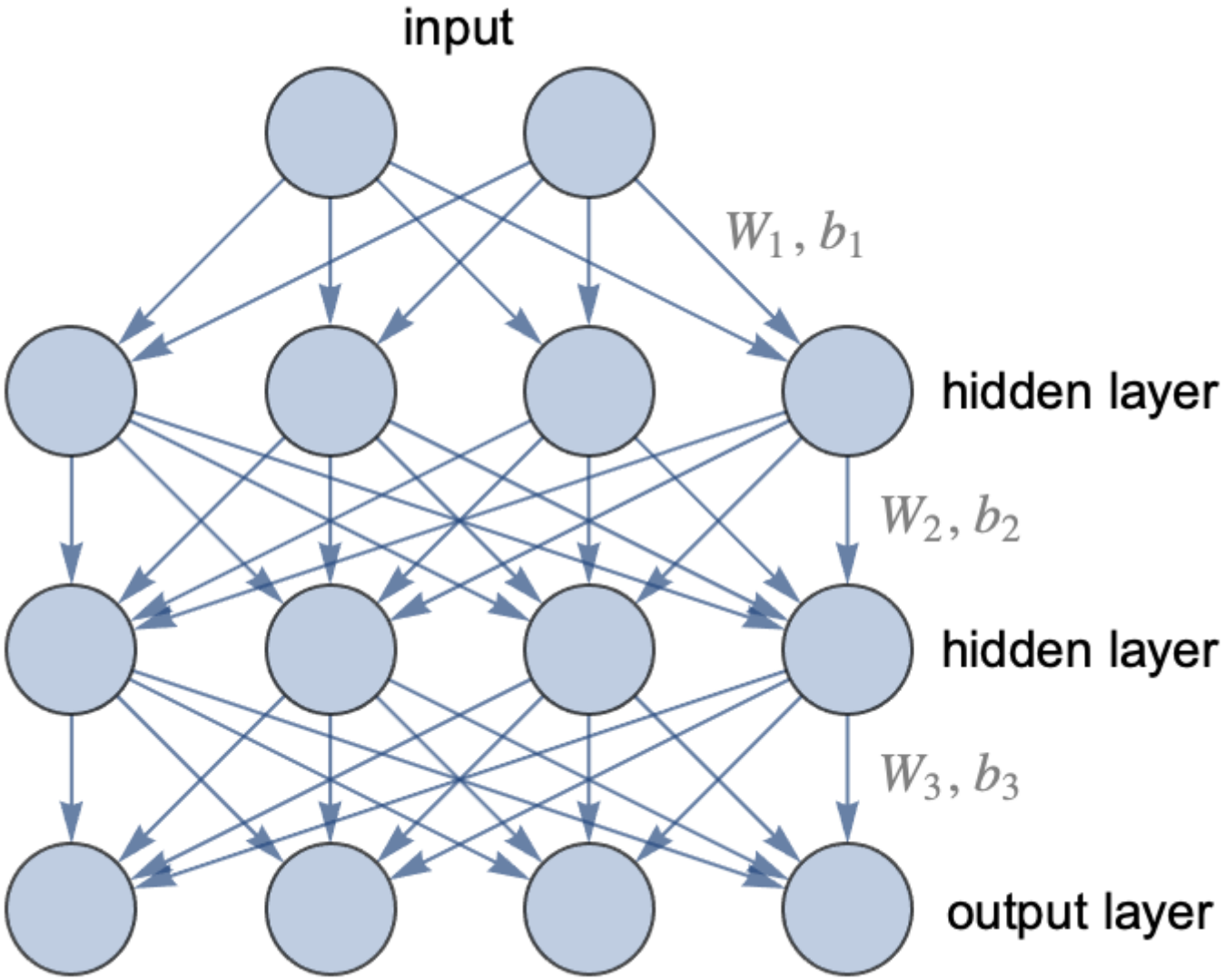


Esta red tiene dos valores de entrada y un valor de salida. Observa que el grafo es dirigido y acíclico, por lo que podemos calcular la salida simplemente siguiendo las aristas.

Esta red es un modelo paramétrico. Hay un parámetro de peso por cada arista y un parámetro de sesgo por cada neurona. Podríamos entrenar esta red de la misma manera que cualquier otro modelo paramétrico: minimizando una función de coste calculada en algunos datos de entrenamiento. A diferencia de lo que ocurre en las redes neuronales biológicas, las aristas no se eliminan ni se añaden durante el proceso de aprendizaje; solo se modifican los parámetros numéricos (pesos y sesgos). También es posible añadir/eliminar aristas, pero es un proceso que consume tiempo, por lo que solo se realiza como un proceso separado para descubrir nuevos tipos de redes neuronales (un proceso conocido como búsqueda de arquitectura neuronal).

Copyright 2026

En la práctica, no utilizamos redes con conexiones aleatorias. En su lugar, utilizamos una arquitectura conocida. Una arquitectura neuronal no es una red exacta, sino una clase de redes neuronales que comparten estructuras similares. La arquitectura más antigua y clásica, inventada en la década de 1960, se llama perceptrón multicapa o red completamente conectada, o a veces red neuronal de propagación hacia adelante. En esta arquitectura, las neuronas se agrupan por capas, y cada neurona de una capa dada envía su salida a cada neurona de la siguiente capa (y solo a ellas). Estas capas completamente conectadas también se llaman capas lineales o capas densas. Aquí tienes un ejemplo de una red de este tipo:



Esta red tiene tres capas (la entrada no es realmente una capa): una capa de salida y dos capas llamadas capas ocultas porque son capas intermedias. Esta red toma dos valores numéricos como entrada y devuelve cuatro valores numéricos como salida. Se podría utilizar para entrenar un clasificador que tenga cuatro posibles clases, y los valores de salida serían las probabilidades de

cada clase. Para una tarea de regresión, solo tendríamos una salida (el valor predicho).

Esta arquitectura en capas permite realizar varios pasos de cálculo, al igual que lo haría un programa clásico, por lo que nos da la capacidad de realizar razonamientos. La mayoría de las redes neuronales artificiales tienen una arquitectura en capas (y las capas también están presentes en las redes neuronales biológicas). En esta ilustración, se incluyeron dos capas ocultas, pero podría haber muchas más. Una red con una capa oculta o menos se llama red superficial, mientras que una red con dos o más capas ocultas se llama red profunda, de ahí el nombre de "aprendizaje profundo". Este nombre resalta la importancia de utilizar modelos que pueden realizar varios pasos de cálculo.

En este grafo, cada flecha representa un peso. Por lo tanto, la primera capa contiene 2x4 = 8 pesos para calcular la activación de la primera capa oculta a partir de las entradas. A estos pesos, debemos agregar un parámetro de sesgo por cada salida. Estos pesos se pueden representar como una matriz.

Cada capa proporciona una serie de valores salida que se transmiten a la siguientes.

En suma el proceso general consiste en ir ajustando los pesos o parámetros de las capas de la red de forma que la salida obtenida se vaya pareciendo lo más posible a los valores reales.

El perceptrón multicapa es la arquitectura original de las redes neuronales, pero nunca logró dominar los métodos clásicos de aprendizaje automático en problemas de datos estructurados. Sin embargo, en 2017 se demostró que los perceptrones multicapa pueden competir con los métodos clásicos de aprendizaje automático en conjuntos de datos estructurados gracias a la arquitectura de auto-normalización.

Sin embargo, el uso de perceptrones multicapa sigue siendo marginal. Las redes neuronales se utilizan principalmente en datos no estructurados (imágenes, texto, sonido, etc.) gracias a arquitecturas como las redes neuronales *convolucionales*, las redes recurrentes o las redes tipo *transformer*. Estas arquitecturas todavía utilizan el concepto de capas, pero su conectividad es bastante diferente a la de los perceptrones multicapa.

Los métodos de aprendizaje profundo destacan en la resolución de problemas con datos de alta dimensionalidad, como las imágenes. En una imagen, cada píxel puede considerarse como una variable, y las interacciones entre los píxeles son cruciales para comprender el contenido general y el significado de la imagen. Un modelo superficial, como la regresión lineal, tendría dificultades para capturar estas interacciones complejas y probablemente fallaría en reconocer objetos con precisión o comprender el contenido de la imagen.

Por otro lado, los modelos de aprendizaje profundo, en particular las redes neuronales convolucionales (CNN), están diseñados específicamente para capturar y aprender patrones e

Copyright 2026

interacciones complejas dentro de la imagen. Las capas de una CNN pueden aprender representaciones jerárquicas, comenzando desde características de bajo nivel como bordes y texturas, hasta características de nivel medio como formas y objetos, y finalmente características de alto nivel que representan conceptos o categorías. Esta capacidad de extraer representaciones jerárquicas permite que los modelos de aprendizaje profundo comprendan la semántica y el contexto de la imagen, lo que les permite realizar tareas como detección de objetos, clasificación de imágenes y generación de imágenes.

La profundidad de la red permite el aprendizaje automático de características relevantes a partir de los datos, sin necesidad de ingeniería de características explícitas. Los modelos de aprendizaje profundo pueden aprender a extraer características en diferentes niveles de abstracción, lo que los hace altamente efectivos en diversas tareas de visión por computadora.

Es importante tener en cuenta que los métodos de aprendizaje profundo requieren grandes cantidades de datos etiquetados para el entrenamiento, así como recursos computacionales significativos para el entrenamiento e inferencia. Sin embargo, con los avances en hardware (*GPUs*) y la disponibilidad de conjuntos de datos a gran escala, el aprendizaje profundo se ha convertido en una herramienta poderosa en el campo de la inteligencia artificial, logrando resultados notables en diversos dominios, incluyendo visión por computadora, procesamiento del lenguaje natural y reconocimiento de voz.

https://www.youtube.com/embed/-P28LKWTzrI

Comparación del poder de una GPU fente a una CPU para pintar una imagen



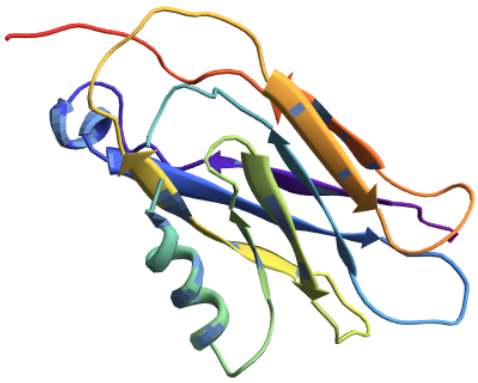
Por sí mismos, los píxeles tienen muy poca información sobre lo que hay en esta imagen. No podríamos identificar este gato simplemente teniendo cada píxel "votando" sobre qué objeto es, al menos no sin que interactúen con otros píxeles. En cambio, es la interacción entre muchos píxeles, formando texturas, formas, etc., lo que proporciona información. Estos patrones no se pueden identificar con un simple cálculo. Los modelos de aprendizaje profundo pueden aprender estos patrones realizando varios cálculos simples. Al principio, detectan cosas como líneas analizando los píxeles. Luego, a partir de estas líneas, detectan formas más complejas. A partir de estas formas, detectan partes de objetos y así sucesivamente hasta que pueden identificar los principales objetos (consulte la sección Redes Convolucionales en este capítulo para una visualización de esta comprensión gradual). Este tipo de proceso de detección escalonado funciona porque la red es profunda y también porque las imágenes tienen una naturaleza algo jerárquica o compositiva. Las imágenes no son el único tipo de datos que es jerárquico/compositivo. El audio, por ejemplo, es bastante similar. El texto también lo es: los caracteres forman palabras, que luego forman oraciones, etc. El significado general surge de interacciones complejas. Las redes neuronales funcionan muy bien para todos estos tipos de datos.

Sin embargo, las redes neuronales no se limitan a imágenes, audio y texto. Por ejemplo, se utilizan para jugar juegos de mesa como el ajedrez o el Go aprendiendo a predecir si una configuración del juego es buena o no. También se utilizan para acelerar simulaciones físicas, predecir si una molécula tiene posibilidades de ser útil como medicamento o predecir cómo se pliega una proteína dada su secuencia de aminoácidos.

Copyright 2026

AHYILNGGTLLGLKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHF-
SISIPVKSDIAPVARLLIYAVLPTGDVIGDSAKYDVENCLANKVDLSFSP-
SQSLPASHAHLRVTAAPQSVCALRAVDQSVLLMKPDAELSASSVYNL-
LPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS-
FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVM-
GRGHARLVHVEEPHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVP-
DTITEWKAGAFCLSEDAGLGISSTASLRAFQPFFVELTMPYSVIRGEA-
FTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEKEQAPHCICANGRQT-
VSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSVPEHGRKDTVIKP-
LLVEPEGLEKETTFNSLLCPSGGEVSEELSLKLPPNVVEESARASVSV-
LGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQL-
TPEIKSKAIGYLNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTA-

→



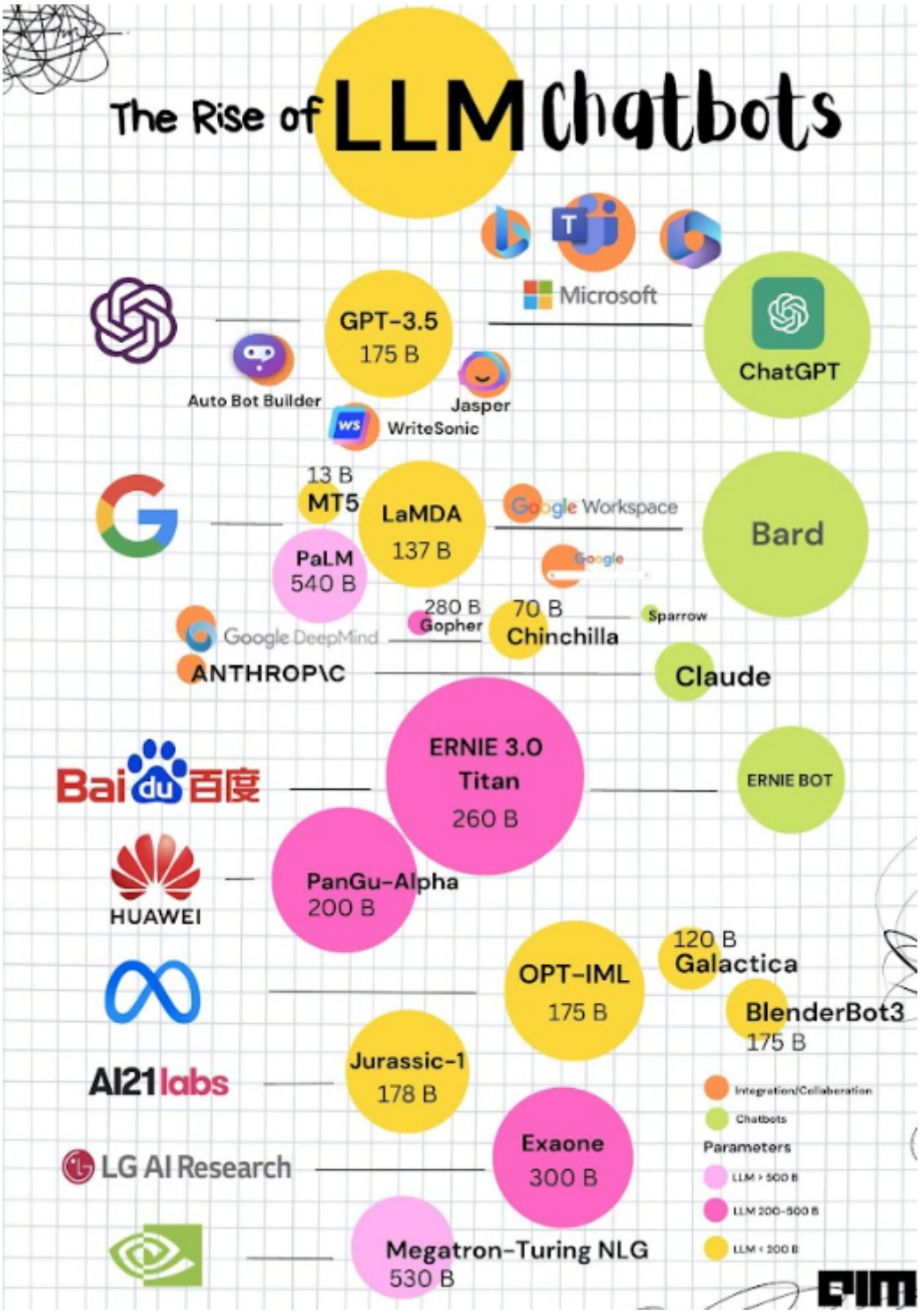
Aunque estas tareas pueden parecer diferentes de las tareas de percepción, también son tareas no estructuradas y tienen en común que los datos son de alta dimensionalidad y a menudo muestran alguna forma de composicionalidad.

Una cosa a tener en cuenta es que las redes neuronales actuales aprenden un tipo particular de programa. Por ejemplo, estas redes solo procesan matrices numéricas y no pueden manipular objetos categóricos, llamados símbolos en este contexto. Además, solo aprenden un conjunto de parámetros continuos y no aprenden construcciones de programación habituales (bucles, declaraciones condicionales, etc.). Dichos programas de red pueden ser muy buenos en algunas tareas pero no en otras. En general, las redes neuronales son bastante buenas para aprender tareas "intuitivas". Una regla general es que si un humano puede realizar una tarea en menos de un segundo, probablemente significa que una red también puede realizar esa tarea. Si le lleva más de un segundo a un humano realizar la tarea, probablemente signifique que los humanos están utilizando algún tipo de razonamiento consciente, que las redes neuronales profundas actualmente no son muy buenas modelando.

Por ejemplo, GPT-3 es un modelo de lenguaje entrenado en 2020 que tiene alrededor de 175 mil millones de parámetros. Estos tamaños son inherentes a las tareas que resuelven: reconocer imágenes requiere mucho conocimiento al igual que generar texto.

Dado que las redes neuronales profundas tienen muchos parámetros, requieren muchos ejemplos de entrenamiento. Las redes de identificación de imágenes suelen entrenarse con decenas de millones a cientos de millones de imágenes. El modelo de lenguaje GPT-3 utilizó un corpus de cientos de miles de millones de palabras. Estos conjuntos de datos contienen más datos de los que cualquier ser humano ha leído, visto u oído. Tal vez algún día descubramos redes con arquitecturas específicas que sean más eficientes en el uso de datos, pero actualmente, los conjuntos de datos grandes son esenciales para entrenar redes neuronales profundas, a menos que comencemos con una red preentrenada.

Copyright 2026



Esta infografía muestra el panorama de modelos grandes de lenguaje y sus parámetros o pesos en octubre de 2023

Actualmente los últimos modelos como chatGPT-4 trabajan con 1.75 trillones de párametros, 1000 veces más que chatGPT-3

En suma las redes usand datos numéricos de entrada para obtener una salida. En el caso de datos de tipo no numérico debemos convertirlos a números mediante diferentes técnicas que no vienne al caso en este curso.

En cualquier caso el proceso o manera en que se usan estas redes para que aprendan es similar y parte siempre de una conversión previa de los datos de entrada en números.

A continuación damos una pincelada de como funciona el proceso.

El proceso de aprendizaje

A continuación describimos el proceso de aprendizaje por pasos, pero sin entrar en detalles técnicos, solamente para entender la idea central.

Podemos dar un resumen de los estadios por lo que pasa cualquier proceso en el aprendizaje automático

1. Definir la pregunta: Es más dificil de lo que parece, y casi lo más importante, saber para que nos puede servir y que queremos resolver
2. Obtención y filtrado de datos: Necesitamos contar con datos de calidad, homogéneos y de fuentes fidedignas
3. Visualización de datos: Esto nos permite detectar errores o casos 'raros' a primera vista
4. Entrenamiento: Aquí ya es donde generamos el modelo
5. Pruebas: En esta parte usamos el grupo de datos de test para hacer las pruebas y evaluar el rendimiento o calidad del modelo
6. Analizar el feedback: Se analizan los resultados y en caso negativo se repite el proceso modificando parámetros como el número de capas y otros valores que afectan al algortitmo utilizado.
7. Usar el modelo obtenido para hacer predicciones.

A continuación describimos brevemente el proceso 4 de entrenamiento y posteriormente describiremos las distintas arquitecturas que sustentan todas las tareas de IA generativa que permiten la creación de modelos de lenguaje para el procesamiento del lenguaje natural.

1. Inicialización

Las ponderaciones y sesgos de la red se inicializan con valores pequeños y aleatorios.

2. Propagación hacia adelante (Forward Pass)

- **Entrada**: Se introduce un conjunto de entradas en la red.
- **Cálculos**: La entrada se propaga a través de las capas de la red. En cada nodo, se realiza una suma ponderada de las entradas, se le añade un sesgo y se aplica una función de activación.
- **Salida**: Se obtiene la predicción de la red para la entrada dada.

3. Cálculo de la Pérdida

- Se calcula la función de pérdida (o coste), que mide la diferencia entre la predicción de la red y la salida deseada.
- Ejemplos de funciones de pérdida incluyen el Error Cuadrático Medio para problemas de regresión y la Entropía Cruzada para problemas de clasificación.

4. Propagación hacia atrás (Backpropagation)

- **Gradiente de la Pérdida**: Se calcula el gradiente de la función de pérdida con respecto a cada uno de los pesos y sesgos en la red, usando la regla de la cadena y derivadas parciales. Esto indica cómo deberían ajustarse los pesos y sesgos para minimizar la pérdida.
- **Actualización de Pesos y Sesgos**: Se ajustan los pesos y sesgos en la dirección opuesta al gradiente para minimizar la pérdida. Esto se hace usando algoritmos de optimización como el Descenso del Gradiente o variantes más avanzadas como Adam.

5. Iteración (*backpropagation*)

Se repiten los pasos 2-4 para un número de iteraciones o épocas, utilizando diferentes conjuntos de entrada, hasta que el modelo alcance un nivel aceptable de rendimiento que se mide mediante las diferencias entre los valores reales y los valores estimados por la red.

6. Evaluación

Finalmente, se evalúa el rendimiento de la red en un conjunto de datos de prueba para asegurarse de que ha aprendido correctamente las relaciones en los datos.

Para la parte de de entrenamiento se utiliza una parte de los datos reales (*training set*) y para la evaluación y validación el resto (*validation set y test set*) .

Este proceso permite que la red neuronal multicapa ajuste sus pesos y sesgos para minimizar la función de pérdida, lo que a su vez mejora la precisión de sus predicciones en tareas de clasificación o regresión.

Este el resumen gráfico de todo el proceso:

Copyright 2026

Fase de aprendizaje



Datos entrenamiento ***Vector características*** ***Algoritmo*** ***Modelo***

Una vez entrenado ya disponemos de un modelo que nos servirá para hacer las predicciones correspondientes en el proceso siguiente:



Datos test ***Vector características*** ***Modelo*** ***Predicción***

> Hemos pretendido dar una pincelada del proceso, en la sección final del módulo encontrarás referencias para estudiarlo en profundidad

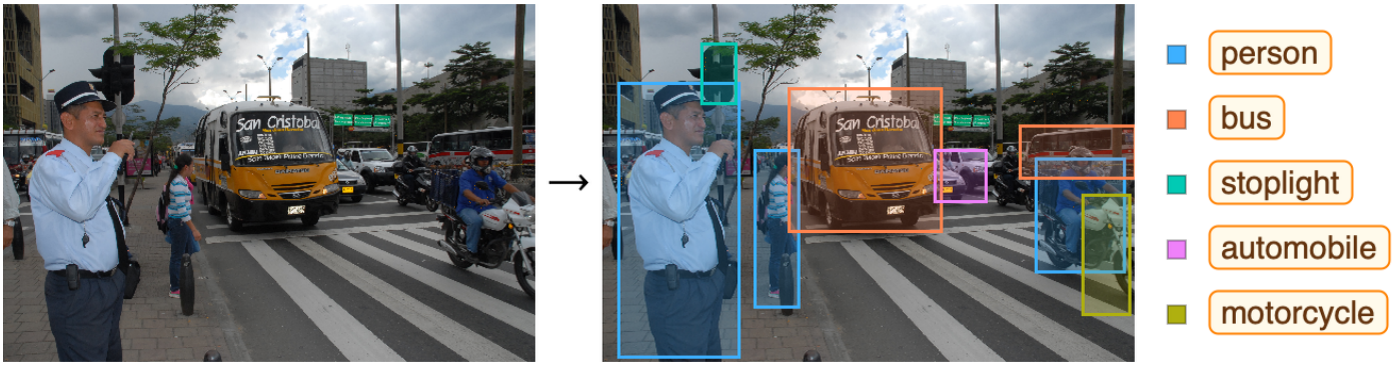
Veremos ahora de forma descriptiva las principales arquitecturas neuronales y sus aplicaciones

Redes convolucionales

Las redes neuronales convolucionales (CNNs por sus siglas en inglés) son un tipo de redes neuronales que han demostrado ser muy efectivas en tareas relacionadas con la visión por computadora y el procesamiento de imágenes. Su arquitectura está inspirada en la forma en que los seres humanos procesamos la información visual.

Una aplicación típica es la detección de objetos como en este ejemplo en el que se detectan los distintos objetos representados por rectángulos.

Copyright 2026



Estructura de una CNN:

Las CNNs están compuestas por diferentes tipos de capas, cada una con una función específica:

1. Convolucional: Aplica diferentes filtros (o kernels) a la imagen de entrada para extraer características importantes como bordes, texturas, etc. Cada filtro se desliza (o convoluciona) sobre la imagen para crear un mapa de características.

2. Activación (ReLU): Introduce no linealidades en el modelo, permitiendo que la red pueda aprender patrones más complejos. La función de activación más común es la unidad lineal rectificada (ReLU).

3. Agrupación (Pooling) Reduce la dimensionalidad de los mapas de características, haciéndolos más manejables y reduciendo el riesgo de sobreajuste. El agrupamiento promedio y el agrupamiento máximo son técnicas comunes de agrupación.

4. Completamente Conectada (Fully Connected): Después de varias capas convolucionales y de agrupación, la información se aplanará en un vector y se alimentará a una o varias capas completamente conectadas, similares a las de un perceptrón multicapa.

5. Salida: La última capa produce la salida de la red, que podría ser una clasificación, una regresión, etc.

Usos Principales:
Son ampliamente utilizadas en una variedad de aplicaciones, incluyendo:

1. Reconocimiento de Imágenes: Clasificar imágenes en categorías.
2. Detección de Objetos: Localizar y clasificar objetos dentro de una imagen.
3. Segmentación Semántica: Clasificar cada píxel de una imagen en una categoría.
4. Reconocimiento Facial Identificar y verificar rostros en imágenes y vídeos.
5. Procesamiento de Lenguaje Natural: Aunque las CNNs se diseñaron originalmente para imágenes, también se han aplicado con éxito a datos de texto y voz.

Copyright 2026

Tipos de CNNs:
Existen varias arquitecturas de CNNs que se han vuelto populares y ampliamente utilizadas:

LeNet-5: Una de las primeras *CNNs*, diseñada para reconocimiento de dígitos.
AlexNet: Ganó el concurso *ImageNet* en 2012 y popularizó las CNNs en la visión por computadora.
VGGNet: Famosa por su arquitectura simple y profunda, utiliza capas convolucionales de pequeño tamaño.
GoogLeNet (*Inception*): Introduce los módulos *inception*, que son bloques de construcción que permiten a la red elegir automáticamente el tamaño del filtro en cada capa.
EfficientNet s una arquitectura de red neuronal convolucional que ha recibido una atención significativa debido a su capacidad para alcanzar un alto rendimiento en tareas de clasificación de imágenes con un número relativamente menor de parámetros en comparación con otras arquitecturas populares. Fue introducido por investigadores de Google en 2019.

Estos son solo algunos ejemplos, y la investigación en este campo está en constante evolución, con nuevas arquitecturas y técnicas que se desarrollan regularmente. Las CNNs son una herramienta poderosa en el campo del aprendizaje profundo y han revolucionado muchas aplicaciones en la visión por computadora y más allá.

Redes Recurrentes

Las Redes Neuronales Recurrentes (RNNs por sus siglas en inglés) son un tipo de red neuronal diseñado para manejar secuencias de datos, capturando la dependencia temporal entre los elementos de la secuencia. A diferencia de las redes neuronales tradicionales, las RNNs tienen conexiones recurrentes que permiten la propagación de información a través de los pasos de tiempo, haciendo posible que la salida en un momento dado dependa de los cálculos anteriores.

Características Clave de las RNNs:

1. Memoria: Las RNNs mantienen un estado interno o memoria que captura información sobre los elementos anteriores de la secuencia.

2. Conexiones Recurrentes: Cada unidad en una RNN recibe entrada no solo del dato actual en la secuencia sino también de su propio estado anterior.

3. Parámetros Compartidos: A través de todos los pasos de tiempo, una RNN utiliza el mismo conjunto de parámetros, lo que reduce significativamente la cantidad de parámetros y facilita el aprendizaje de patrones temporales.

Aplicaciones:

Las RNNs son especialmente útiles para tareas que involucran secuencias de datos, incluyendo:

Copyright 2026

1. Procesamiento de Lenguaje Natural (PLN): Traducción automática, generación de texto, reconocimiento de voz, etc.
2. Series Temporales: Predicción del mercado de valores, análisis de tendencias, etc.
3. Reconocimiento de Secuencias: Como en la escritura a mano o el reconocimiento del habla.
4. Generación de Música: Creación de melodías basadas en patrones aprendidos.

Variantes y Mejoras:

1. LSTM (Long Short-Term Memory): Una mejora con respecto a las RNNs tradicionales que pueden capturar dependencias a largo plazo en los datos gracias a su estructura especial de celda de memoria y puertas de olvido, entrada y salida.
2. GRU (Gated Recurrent Unit): Similar a LSTM pero con una estructura más simplificada, lo que la hace más eficiente en términos de computación.

Desafíos:

- Problema del Gradiente Desvaneciente y Explosivo: En las RNNs, los gradientes pueden volverse extremadamente pequeños o grandes, lo que hace que el entrenamiento sea inestable y difícil. Las LSTMs y GRUs fueron diseñadas para mitigar este problema.

Resumen

Las Redes Neuronales Recurrentes representan un avance crucial en el aprendizaje de patrones temporales y secuenciales, y han habilitado una amplia variedad de aplicaciones en el procesamiento de lenguaje natural, reconocimiento de patrones y más. Su capacidad para mantener un estado interno las hace únicas y poderosas, aunque no están exentas de desafíos, especialmente cuando se trata de secuencias muy largas. Las variantes como LSTM y GRU han demostrado ser soluciones efectivas a muchos de estos desafíos, permitiendo el aprendizaje de dependencias a largo plazo.

Redes Transformers

Las redes Transformer son un tipo de arquitectura de red neuronal introducida en el artículo “*Attention is All You Need*” por Vaswani et al. en 2017. Han revolucionado el campo del procesamiento del lenguaje natural (PLN) y se han convertido en la base para modelos como BERT, GPT, y muchos otros.

Características Clave de las Redes Transformer:

1. **Mecanismo de Atención**: En lugar de depender de recurrencias o convoluciones, las redes Transformer utilizan mecanismos de atención para ponderar la importancia de diferentes partes de la entrada en relación con cada palabra o token. La atención permite

que el modelo se enfoque en las partes relevantes de la entrada para realizar la tarea y sobre todo que tenga en cuenta el contexto en que se encuentra de modo que diferencie el significado por ejemplo de palabras iguales.

2. **Codificadores y Decodificadores**: La arquitectura Transformer original consiste en una serie de bloques de codificadores y decodificadores. Los codificadores procesan la entrada, y los decodificadores generan la salida. En tareas como la generación de texto, solo se utilizan los decodificadores.
3. **Capas de Normalización y Feedforward**: Además del mecanismo de atención, los Transformers también utilizan capas de normalización y capas feedforward para procesar los datos y realizar transformaciones.
4. **Paralelismo**: A diferencia de las RNNs, las redes Transformer permiten el procesamiento paralelo de las secuencias, lo que reduce significativamente los tiempos de entrenamiento.

Aplicaciones:

Las redes Transformer han demostrado ser extremadamente eficaces en una variedad de tareas en PLN y más allá:

1. **Traducción Automática**: Traducir texto de un idioma a otro.
2. **Generación de Texto**: Crear texto coherente y relevante dado un prompt.
3. **Comprensión del Lenguaje**: Entender el significado del texto para responder preguntas, resumir texto, etc.
4. **Reconocimiento de Voz y Generación**: Convertir voz a texto y viceversa.
5. **Clasificación de Imágenes**: Aunque fueron diseñadas inicialmente para texto, las variantes de Transformers también se han aplicado con éxito a la clasificación de imágenes y otras tareas de visión por computadora.
6. **Amplicación de texto, imagne y sonido**

> Veremos ejemplos de estas aplicaciones en el módulo siguiente

Desafíos y Consideraciones:

- **Requerimientos de Computación**: Los Transformers requieren una cantidad significativa de poder de cómputo y memoria, especialmente para grandes conjuntos de datos o modelos grandes.
- **Necesidad de Grandes Cantidades de Datos**: Para aprovechar al máximo su capacidad, los modelos Transformer generalmente requieren grandes cantidades de datos de entrenamiento.

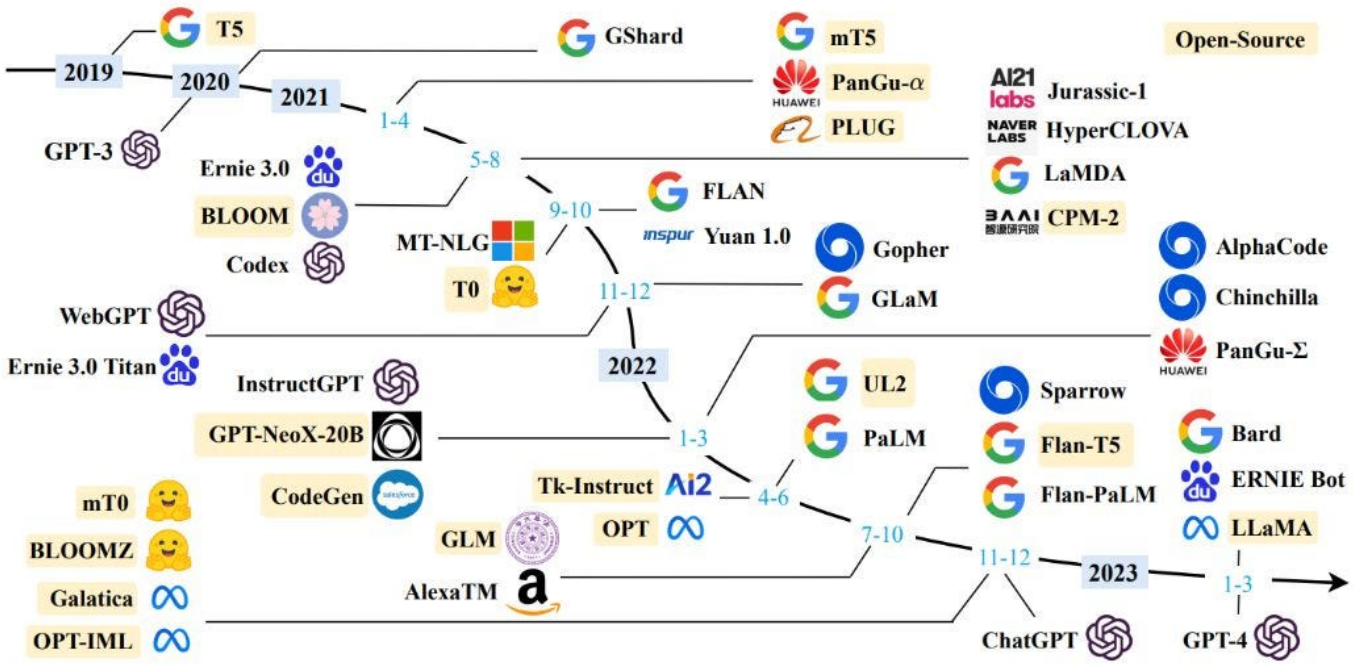
Variantes y Mejoras:

1. **BERT (Bidirectional Encoder Representations from Transformers)**: Utiliza un Transformer para aprender representaciones de palabras en base a su contexto en ambos lados (izquierda y derecha) en un texto.
2. **GPT (Generative Pre-trained Transformer)**: Un modelo de Transformer diseñado para generar texto, preentrenado en grandes cantidades de texto y afinado para tareas específicas.

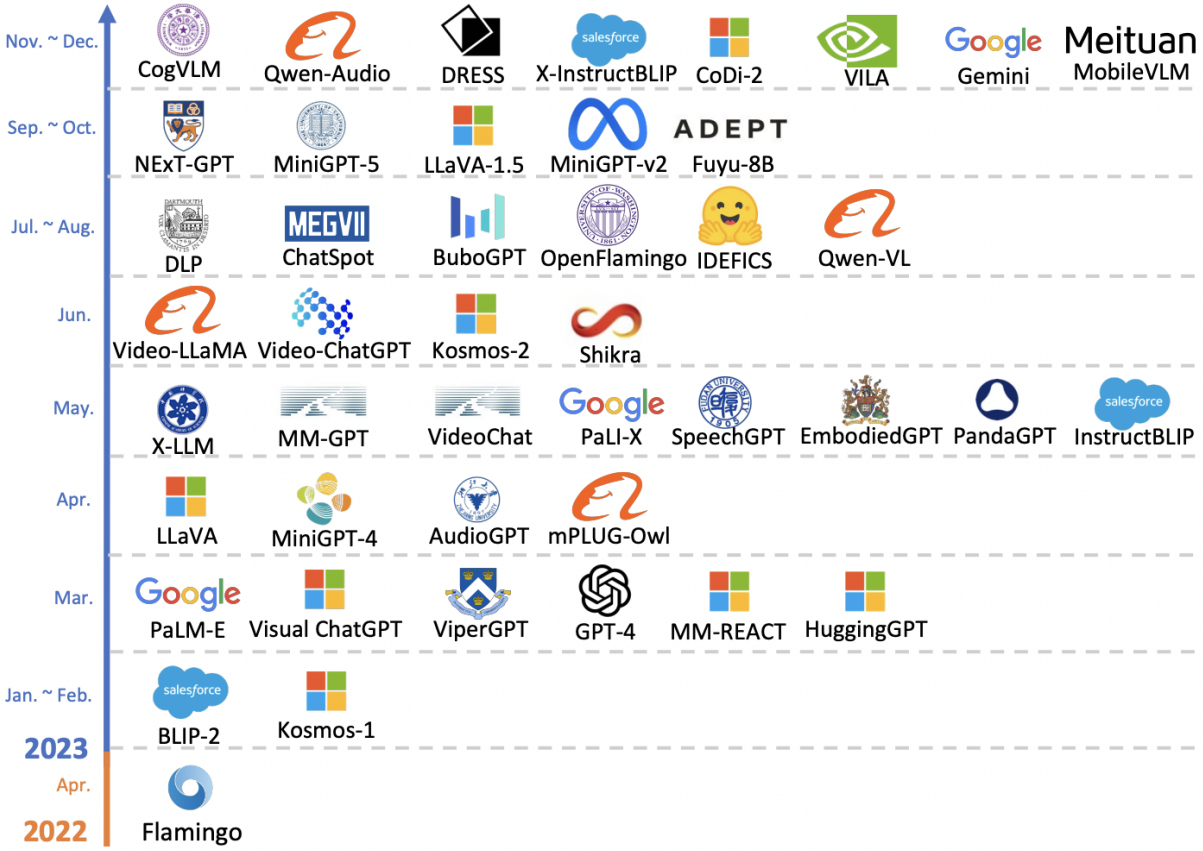
Resumen:

Las redes Transformer han establecido un nuevo estándar en PLN, ofreciendo un rendimiento excepcional en una amplia gama de tareas. Su diseño permite el procesamiento paralelo y el aprendizaje de relaciones complejas en los datos, aunque esto viene con un costo computacional significativo. Las mejoras y variantes continúan evolucionando, ampliando las capacidades y aplicaciones de esta poderosa arquitectura.

Para acabar esta sección dejamos un esquema de la evolución del ecosistema de modelos de tipo *transformer* generados y entrenados hasta el presente.



Fuente: https://www.wolfram.com/language/introduction-machine-learning/deep-learning-methods/



Este es el panorama de chatbots de empresas a finales de 2023

Glosario de términos

Añadimos finalmente un glosario de términos de uso común en el ecosistema de la IA

Término	Descripción
Token	Una unidad básica de procesamiento de texto, como una palabra, un número o un signo de puntuación.
Contexto	El conjunto de tokens o palabras que rodean a otro token específico, importante para determinar su significado.
Modelo de lenguaje	Un sistema de inteligencia artificial diseñado para entender, generar y manipular lenguaje humano.
Aprendizaje profundo	Una técnica de aprendizaje automático que enseña a los ordenadores a aprender realizando tareas.
Red neuronal	Un modelo computacional inspirado en la forma en que funcionan los cerebros humanos.

Copyright 2026

Transformer	Un tipo de arquitectura de red neuronal especializada en procesamiento de lenguaje.
Entrenamiento	El proceso de enseñar a un modelo de lenguaje a entender y generar texto a través de ejemplos.
Generación de texto	La habilidad de un modelo de lenguaje para crear texto nuevo y coherente.
Comprensión del lenguaje	La capacidad de un modelo de lenguaje para entender el significado y la intención detrás del texto.
API (Interfaz de Programación de Aplicaciones)	Una interfaz que permite la interacción con un modelo de lenguaje a través de programas.
Aprendizaje supervisado	Un tipo de aprendizaje automático donde el modelo se entrena con datos etiquetados.
Parámetro de entrenamiento	Valores ajustables en un modelo de aprendizaje automático que se optimizan durante el entrenamiento.
Dimensionalidad	Se refiere al número de atributos o características que tienen los datos de entrada en un modelo.
Aprendizaje no supervisado	Un tipo de aprendizaje automático que utiliza datos no etiquetados para encontrar patrones.
Regularización	Técnicas usadas para reducir el sobreajuste en modelos de aprendizaje automático.
Activación	Función matemática usada en redes neuronales para determinar la salida de un nodo.
Aprendizaje por refuerzo	Un tipo de aprendizaje automático donde un agente aprende a tomar decisiones a través de recompensas.
Sobreajuste (Overfitting)	Cuando un modelo de aprendizaje automático se ajusta demasiado a los datos de entrenamiento.
Subajuste (Underfitting)	Cuando un modelo de aprendizaje automático no puede capturar la estructura subyacente de los datos.

Copyright 2026

Referencias Módulo 2

Referencias

Web: Modelos de lenguaje fundacionales

https://www.promptingguide.ai/models/collection?s=35

Vídeo, 'Qué es y cómo funciona la IA'

Vídeo intuitivo de 8 minutos explicando los principales conceptos de IA

Que es la IA

Libro "*Artificial Intelligence: A Modern Approach*" por Stuart Russell y Peter Norvig.

Este libro es considerado uno de los textos más completos en el campo de la IA. Cubre una amplia gama de técnicas y conceptos de IA, desde los más fundamentales hasta los más avanzados, es un material para dedicarle tiempo.

Libro "*Superinteligencia: Caminos, peligros, estrategias*" por Nick Bostrom.

Bostrom explora el futuro de la inteligencia artificial y los desafíos que enfrentamos cuando las máquinas superen la inteligencia humana.

Web Wolfram

https://www.coursera.org/professional-certificates/ibm-machine-learning

Dell creador del lenguaje Wolfram para trabajo con datos e IA Stephen Wolfram

Copyright 2026

Parte del material de esta sección del curso proviene de su documentación teórica sobre los fundamentos de IA.

Una lecutra imprescindible para profundizar en los conceptos

Vídeo Concepto de ML

Vídeo muy sencillo de 8 minutos sobre el concepto de Machine Learning

Concepto de ML

Libro: *Machine Learning For Dummies*

Clásico libro para aprender sin conocimientos previos

https://www.ibm.com/downloads/cas/GB8ZMQZ3

Libro: "Redes Neuronales: Una Introducción Sencilla". Raúl González Duque.

Curso Aprendizaje Automático

Conjunto de cursos gratuitos de Microsoft para ser científico de datos

https://cloud.google.com/training/machinelearning-ai?hl=es-419

Curso Machine Learning

Curso gratuito de google para desarrolladores

https://developers.google.com/machine-learning/crash-course?hl=es-419

Canal de vídeos relacionados con AI de Stanford

Canal con cientos de vídeos de Yotutube sobre IA y Machine Learning

https://www.youtube.com/@stanfordonline

Curso gratuito de Machine Learning de IBM en la plaforma Coursera

https://www.coursera.org/professional-certificates/ibm-machine-learning