Aprendizaje automático

Introducción al marco conceptual de la inteligencia artificial (IA)

Autor: Zsolt Lengyel, iTStudy Hungary Training and Research Centre

Revisora: Mária Hartyányi, iTStudy Hungary Training and Research Centre

Objetivo de aprendizaje

El rápido auge de la inteligencia artificial ha provocado reacciones extremas, que van desde un entusiasmo desbordante hasta el temor de que pronto nos quite el trabajo. El propósito de este material didáctico es repasar la base teórica necesaria para valorar con realismo qué es realmente la IA y aclarar los conceptos básicos imprescindibles para un uso informado de la IA (como el aprendizaje automático, el conjunto de datos de entrenamiento, el modelo, etc.).

Al terminar este material didáctico, habremos encontrado respuesta a las siguientes preguntas:

- ¿Cómo aprende y se comunica un ordenador?
- ¿Qué entendemos por aprendizaje automático?
- ¿Qué significa que los sistemas de inteligencia artificial no funcionan de forma determinista?
- ¿Es la inteligencia "artificial" una versión mecánica de la inteligencia humana?
- ¿Puede la inteligencia artificial ser más inteligente que el ser humano?

Introducción

El célebre matemático Alan Turing[1] lo expresó así en una conferencia que impartió en 1947 ante la London Mathematical Society:

"Lo que queremos es una máquina que pueda aprender de la experiencia".

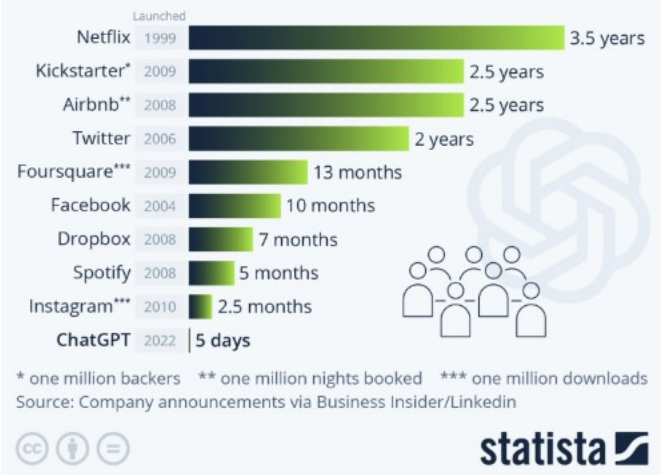
Sin embargo, los ordenadores de la época no permitían crear un programa así, y hicieron falta más de setenta años para acercarse a la visión de Turing: la creación de máquinas capaces de aprender.

Tras la conferencia de Dartmouth de 1956, en la que se convirtió en un campo de investigación independiente, se fueron desarrollando poco a poco los modelos matemáticos y los algoritmos necesarios para el aprendizaje automático. Al mismo tiempo, también se dieron las condiciones tecnológicas: se dispuso de enormes volúmenes de datos digitales, a menudo denominados "big data", y surgieron los ordenadores de alto rendimiento, los servidores, los sistemas de almacenamiento y los procesadores necesarios

[1] Durante la Segunda Guerra Mundial fue un miembro clave del equipo de descifrado de códigos de Bletchley Park, que logró romper Enigma, el famoso sistema de cifrado de los nazis.

para procesarlos. A principios del siglo XXI, la inteligencia artificial ya desempeñaba un papel importante en la industria y la informática, pero fue el lanzamiento público de ChatGPT lo que le dio un amplio reconocimiento social.

El 30 de noviembre de 2022 es una fecha destacada en la historia de la inteligencia artificial. Ese día se lanzó ChatGPT y ocurrió algo sin precedentes: el número de usuarios registrados alcanzó el millón en cinco días, frente, por ejemplo, a Netflix, que necesitó tres años y medio para alcanzar la misma cifra.



¿Cuánto tardaron distintas aplicaciones en alcanzar 1 millón de usuarios? (Fuente: statista.com)

¿Qué despertó este enorme interés?

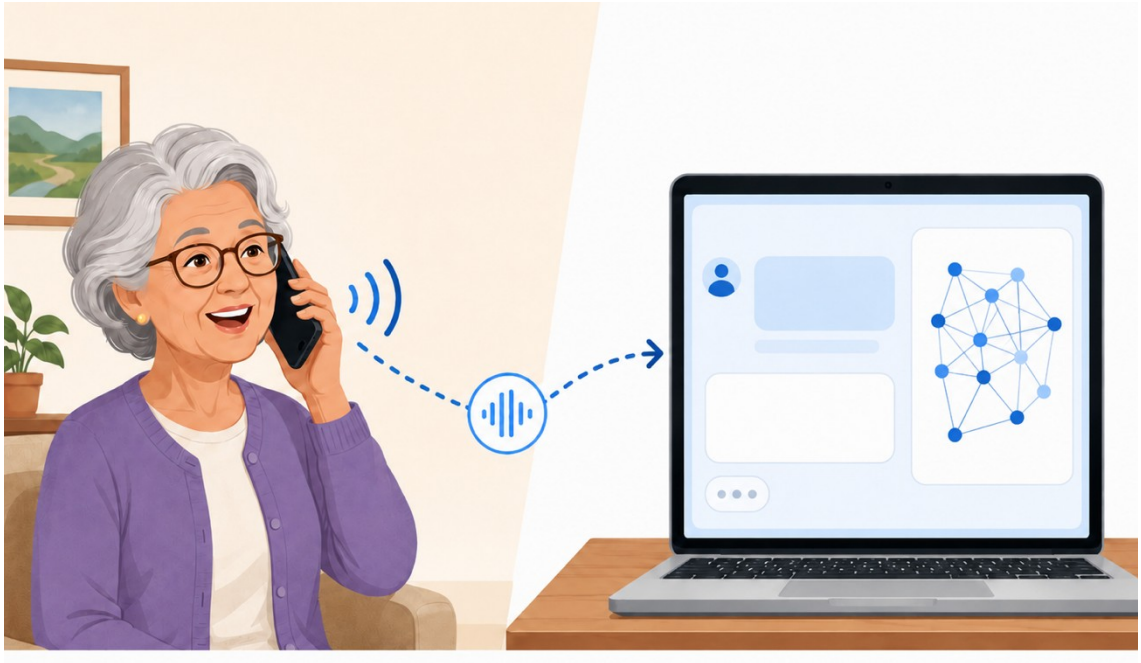
Mientras que antes la IA se consideraba principalmente un campo de investigación de la industria o de los especialistas en informática, la interfaz de ChatGPT permitió a cualquiera comprobar que es posible comunicarse con un ordenador mediante el lenguaje humano natural. ChatGPT respondía a preguntas formuladas en lenguaje natural de forma rápida, coherente y convincente, lo que provocó a la vez entusiasmo, sorpresa y preocupación.



En internet encontramos muchas imágenes parecidas. ¿Qué sugieren estas imágenes? (Fuente: ChatGPT)

En las últimas décadas nos hemos acostumbrado a que los ordenadores formen parte de nuestra vida. Son herramientas indispensables, capaces de ejecutar órdenes. Ni siquiera nos sorprende que los robots de las películas de

ciencia ficción hablen con una extraña voz mecánica. Pero cuando "alguien" nos habla por teléfono con una voz perfectamente humana, todavía hoy nos sorprende descubrir que al otro lado de la línea hay una máquina y no una persona.



La abuela charla con la IA (Fuente: ChatGPT 5.6 Sol)

Han pasado unos años desde la aparición de ChatGPT, pero la confusión y las opiniones extremas que lo rodean siguen presentes. Todavía hoy, muchas personas aceptan sus respuestas de forma incondicional y sin ninguna duda, mientras que otras lo rechazan por completo y no confían en absoluto en su fiabilidad.

¿Aprende realmente la máquina?

Cuando decimos que un ordenador "aprende", desde luego no nos referimos al mismo proceso mediante el cual aprende un ser humano, sobre todo un niño pequeño. La ciencia todavía no es capaz de describir de forma completa y precisa los complejos procesos del aprendizaje humano.

Detrás del aprendizaje automático hay décadas de investigación y desarrollo tecnológico.

> El término que agrupa los algoritmos y métodos que aprenden a partir de datos es **aprendizaje automático** (machine learning), que se ha convertido en uno de los subcampos más importantes de la inteligencia artificial. Mediante los algoritmos de aprendizaje automático, el ordenador reconoce patrones a partir de datos y ejemplos, y los utiliza para generar predicciones y recomendaciones de decisión.

La esencia del aprendizaje automático viene determinada por tres elementos: **los datos, el modelo y el paradigma de aprendizaje**. La "receta" es sencilla: se necesitan **datos** almacenados digitalmente, a partir de los cuales aprende la

IA; **un algoritmo de aprendizaje** que especifica cómo se produce el aprendizaje, y **un modelo** en el que se almacenan las relaciones aprendidas.



Los componentes básicos del aprendizaje automático

Para el entrenamiento se utilizan grandes cantidades de datos; este conjunto de datos se denomina **base de datos de entrenamiento** o **conjunto de datos de entrenamiento**. Como resultado del proceso de aprendizaje se crea un **modelo**, capaz de identificar datos desconocidos a partir de las relaciones que reconoció durante el entrenamiento. Sin embargo, el modelo solo puede extraer conclusiones de lo que ha aprendido de los datos de entrenamiento: no tiene conocimientos ni experiencia propios más allá de ellos.

Los algoritmos de aprendizaje automático (también llamados paradigmas de aprendizaje) se dividen en tres grandes grupos:

- aprendizaje supervisado,
- aprendizaje no supervisado y
- aprendizaje por refuerzo.

Una vez finalizado el proceso de aprendizaje, **el modelo constituirá el núcleo de la aplicación de IA** y se basará en los conocimientos adquiridos.

Tipos de aprendizaje automático (ML)

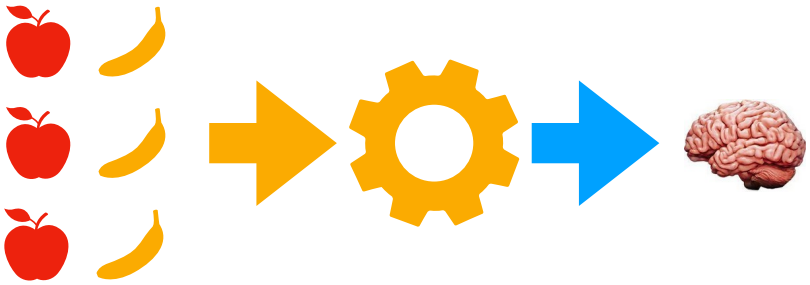
Aprendizaje supervisado

En el **aprendizaje supervisado**, la máquina aprende de ejemplos en los que las respuestas correctas se proporcionan junto con los datos de entrada. Un ejemplo serían millones de fotografías etiquetadas con los nombres de los objetos que aparecen en ellas.



¡Esto es un perro! (Fuente: ChatGPT)

La esencia del algoritmo de aprendizaje se entiende mejor con un ejemplo concreto. Imaginemos que queremos entrenar un modelo para que decida si una imagen muestra una manzana o un plátano.

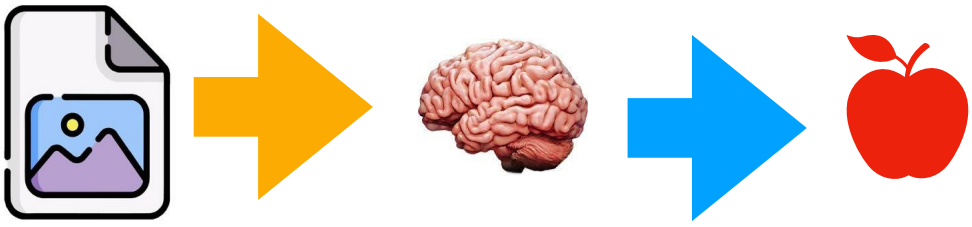


Datos de entrada Algoritmo de aprendizaje Modelo

Entrenar y construir el modelo

Durante el entrenamiento, mostramos al algoritmo miles de imágenes, y para cada una se proporciona la respuesta correcta: “esto es una manzana”, “esto es un plátano”. Estas respuestas correctas predefinidas se llaman etiquetas; de ahí el nombre: datos etiquetados. A partir de estos ejemplos, el modelo aprende qué características visuales (forma, color, etc.) indican cada fruta.

Si el entrenamiento tiene éxito, el modelo también será capaz de decidir, **con una alta probabilidad**, si un objeto desconocido que nunca ha visto antes es una manzana o un plátano, sin que nadie le diga la respuesta.

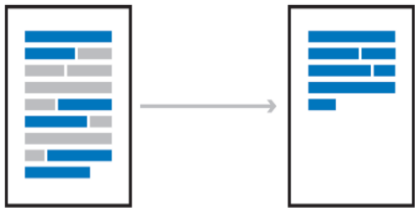


El modelo en funcionamiento (el modelo se representa mediante un diagrama del cerebro humano)

El mismo principio se aplica al reconocimiento de voz, al reconocimiento de escritura a mano y al reconocimiento facial, con la diferencia de que el modelo no se entrena con información visual, sino con otros tipos de datos (por ejemplo, texto o audio).



Reconocimiento de voz



Generación de texto, traducción

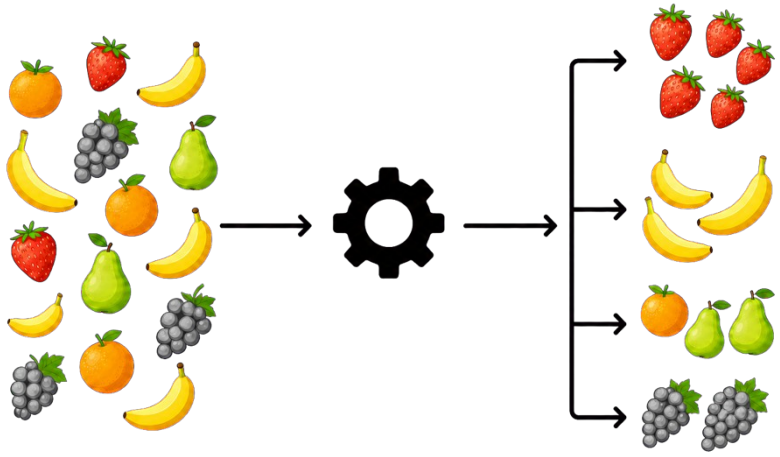
Aprendizaje no supervisado

En el **aprendizaje no supervisado** no existe una respuesta correcta predefinida: a la máquina no se le dan de antemano las respuestas correctas y los objetos observados no tienen nombres predefinidos (etiquetas).



¡Cenicienta lo hizo en un santiamén! (Fuente: ChatGPT)

La tarea del algoritmo es descubrir por sí mismo los patrones ocultos en el conjunto de datos de entrada (agrupaciones, similitudes, diferencias), buscar relaciones en el conjunto de datos y organizar (p. ej., agrupar) los datos.

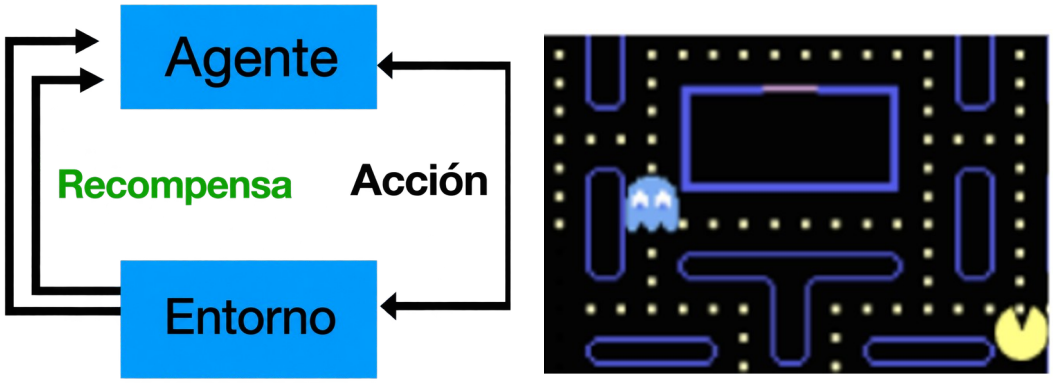


ML no supervisado: por ejemplo, agrupamiento (clustering)

Este tipo de algoritmos se utiliza para construir los modelos de los sistemas de recomendación de contenidos: agrupan los intereses de los usuarios según su comportamiento y recomiendan contenidos en consecuencia. Netflix sugiere películas personalizadas; Spotify, música, y Amazon, productos.

Aprendizaje por refuerzo

En el aprendizaje supervisado, la máquina recibe datos; en el aprendizaje por refuerzo, recibe reglas. El **aprendizaje por refuerzo** es un método de aprendizaje en el que un "agente" virtual interactúa con su entorno e intenta tomar decisiones óptimas. Mediante recompensas o penalizaciones, aprende qué acciones conducen al mejor resultado.



Aprendizaje por refuerzo: el modelo recibe una recompensa o un castigo

El ajedrez, el go y el aprendizaje por refuerzo

En 1997, un ordenador derrotó por primera vez al vigente campeón del mundo de ajedrez, Kaspárov. El programa Deep Blue de IBM era capaz de analizar millones de jugadas muy rápidamente gracias a una enorme potencia de

cálculo y a un sólido sistema de evaluación de posiciones. El programa se creó mediante programación tradicional y no funcionaba con inteligencia artificial.

Sin embargo, el go no se podía enseñar a una máquina con métodos clásicos, porque en el go hay más posiciones y secuencias de juego posibles que átomos en el universo. Ni siquiera hoy existe una máquina lo bastante potente como para almacenar todos los estados posibles del juego. El sistema de DeepMind (AlphaGo), *mediante aprendizaje por refuerzo, practicando contra sí mismo,* aprendió a jugar a un nivel que le permitió derrotar en 2016 a Lee Sedol, uno de los mejores jugadores del mundo.



AlphaGo gana la partida (Fuente: https://artsandculture.google.com/)

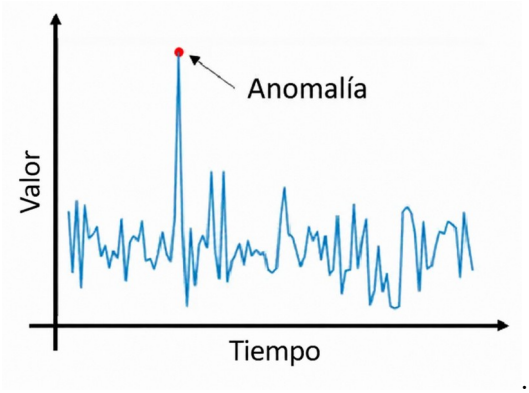
Aprendizaje por transferencia

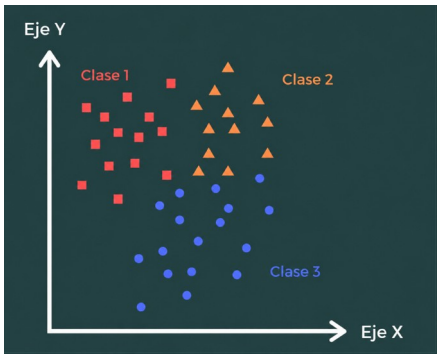
Los tres algoritmos básicos se complementan con un cuarto: el aprendizaje basado en la transferencia de conocimientos, o **aprendizaje por transferencia** (transfer learning). Este algoritmo permite seguir desarrollando un modelo ya terminado y que funciona bien, continuando su entrenamiento en un ámbito más reducido y especializado; por ejemplo, si queremos que el modelo utilice una fuente que le proporcionamos (como un libro de texto) al generar contenidos.

Ámbitos de aplicación de los algoritmos de aprendizaje automático

Los distintos algoritmos de aprendizaje automático se utilizan para tareas diferentes. El aprendizaje supervisado es especialmente adecuado para la clasificación y la predicción, como el reconocimiento de objetos, el filtrado de correo no deseado, la evaluación del riesgo crediticio o el apoyo al diagnóstico médico.

El aprendizaje no supervisado se utiliza para explorar la estructura interna de grandes conjuntos de datos: puede ayudar a identificar grupos de clientes similares, detectar sucesos anómalos o simplificar datos complejos





El aprendizaje por refuerzo se aplica, por ejemplo, en el control de robots, los programas de juegos, la optimización de sistemas de transporte o determinadas tareas de control automatizado, es decir, en situaciones en las que hay que tomar una serie de decisiones sucesivas.

Recursos y costes del aprendizaje automático

Los recursos que requieren los métodos de aprendizaje automático dependen en gran medida de la cantidad de datos, de la complejidad del modelo y del método de aprendizaje elegido. En el aprendizaje supervisado, recopilar y etiquetar los datos de entrenamiento puede suponer un coste importante, mientras que en el aprendizaje no supervisado lo que exige capacidad de cálculo es, sobre todo, el almacenamiento y el procesamiento de grandes conjuntos de datos.

El aprendizaje por refuerzo es especialmente costoso, porque el sistema tiene que adquirir experiencia a través de muchos ensayos, a menudo en un entorno simulado.

¿Por qué es tan caro "entrenar" un modelo?
La empresa de IA *Hugging Face* creó un conjunto de datos de entrenamiento público llamado *FineWeb*, pensado para el preentrenamiento de modelos de lenguaje. El texto se recopiló de varios miles de millones de páginas web y, tras varios meses de trabajo, se filtraron los contenidos defectuosos, los duplicados y los datos personales. *El resultado final son unos 54 terabytes de texto que cualquiera puede utilizar gratuitamente para el entrenamiento*. Los conocimientos del modelo resultante se limitan a lo que ha logrado "aprender" de este conjunto de datos filtrado y procesado.

Ideas clave para docentes

"La inteligencia artificial no piensa, habla" (Vilmos Csányi,[2] 2026).

La cita expresa con gran claridad lo que hemos aprendido sobre el aprendizaje automático en los capítulos anteriores. *Los modelos entrenados, por supuesto, no piensan en el sentido en que pensamos los seres humanos.*

[2] Vilmos Csányi (1935–), etólogo, catedrático, académico y escritor húngaro.

Generan texto mediante métodos matemático-estadísticos: "hablan" sin entender lo que están "diciendo".

¿Qué significa que los sistemas de inteligencia artificial no funcionan de forma determinista?

El software tradicional da el mismo resultado ante la misma entrada. Los sistemas de inteligencia artificial, basados en patrones aprendidos matemáticamente, ofrecen el resultado que el algoritmo considera más probable. Por ello, si varias personas hacen la misma pregunta, pueden obtener respuestas diferentes. La respuesta no es determinista: solo puede considerarse verdadera con cierta probabilidad, y siempre existe alguna probabilidad de error.

¿Significa el adjetivo "artificial" que se ha conseguido una versión mecánica de la inteligencia humana natural?

No. La inteligencia artificial no es una copia mecánica de la inteligencia humana. Los sistemas de IA pueden modelar parcialmente determinadas capacidades asociadas a la inteligencia, como el reconocimiento de patrones, el aprendizaje, las respuestas basadas en el lenguaje o la predicción. Sin embargo, no poseen la comprensión general, la conciencia, la experiencia ni el verdadero entendimiento propios del ser humano.

¿Son los sistemas de IA más inteligentes que el ser humano?

Aunque pueden ser más rápidos y precisos que las personas en determinadas tareas concretas, carecen de la inteligencia general, la experiencia y el criterio que poseen los seres humanos.

Derechos de autor © 2026 Consorcio VITA

Esta obra puede compartirse y adaptarse bajo los términos de la Licencia Internacional Creative Commons Atribución/Reconocimiento-CompartirIgual 4.0 (CC BY-SA 4.0).

Financiado por la Unión Europea. Las opiniones expresadas son únicamente las de los autores y no reflejan necesariamente las de la Unión Europea ni las de la Agencia Nacional.