# 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

Copyright 2026

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).

Copyright 2026

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.

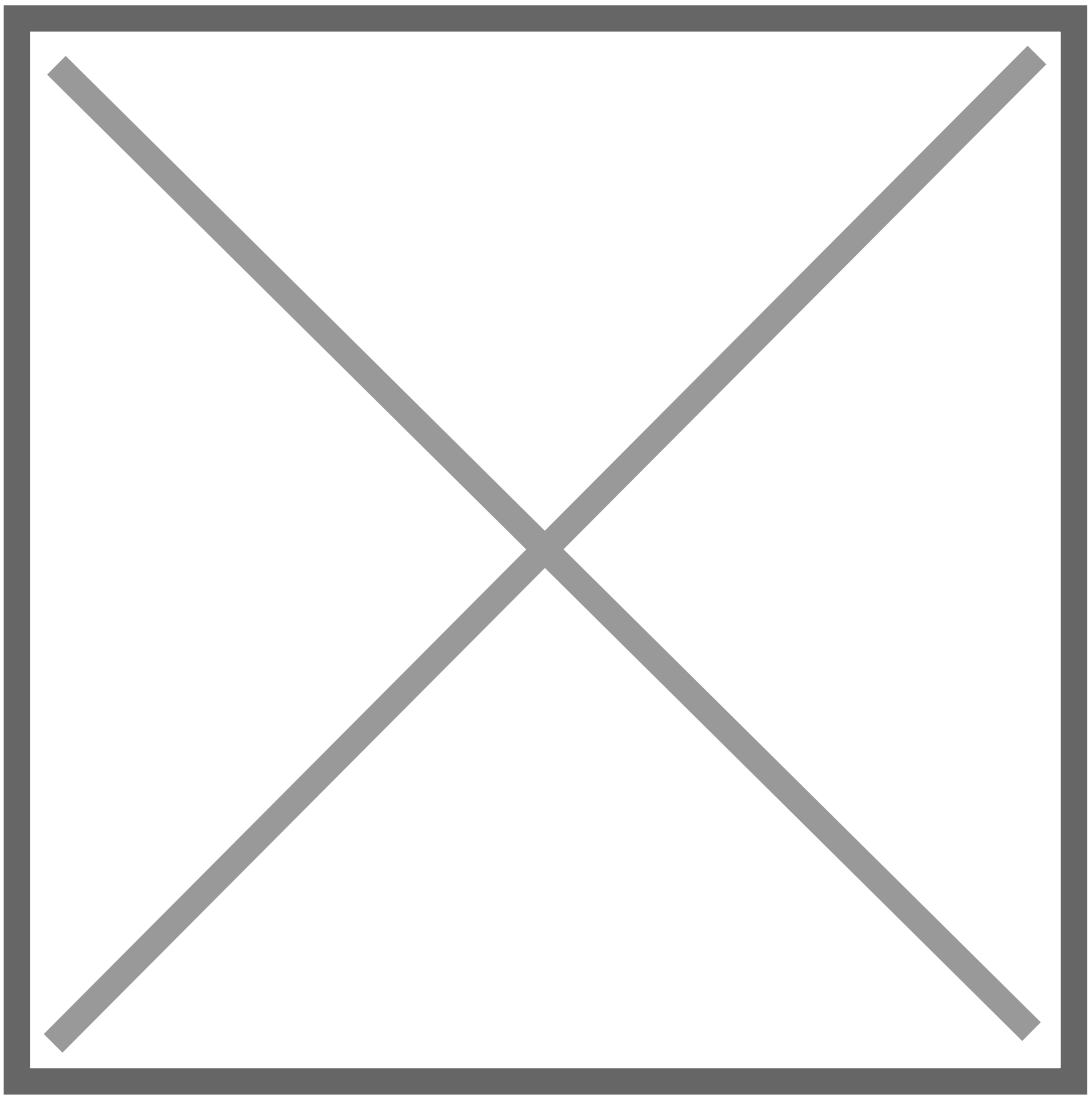
Cuando la variable de salida es numérica hablamos de **regresión**. Si la entrada es texto y la salida es una variable categórica ("gato" o "perro") hablamos de **clasificación**.

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.

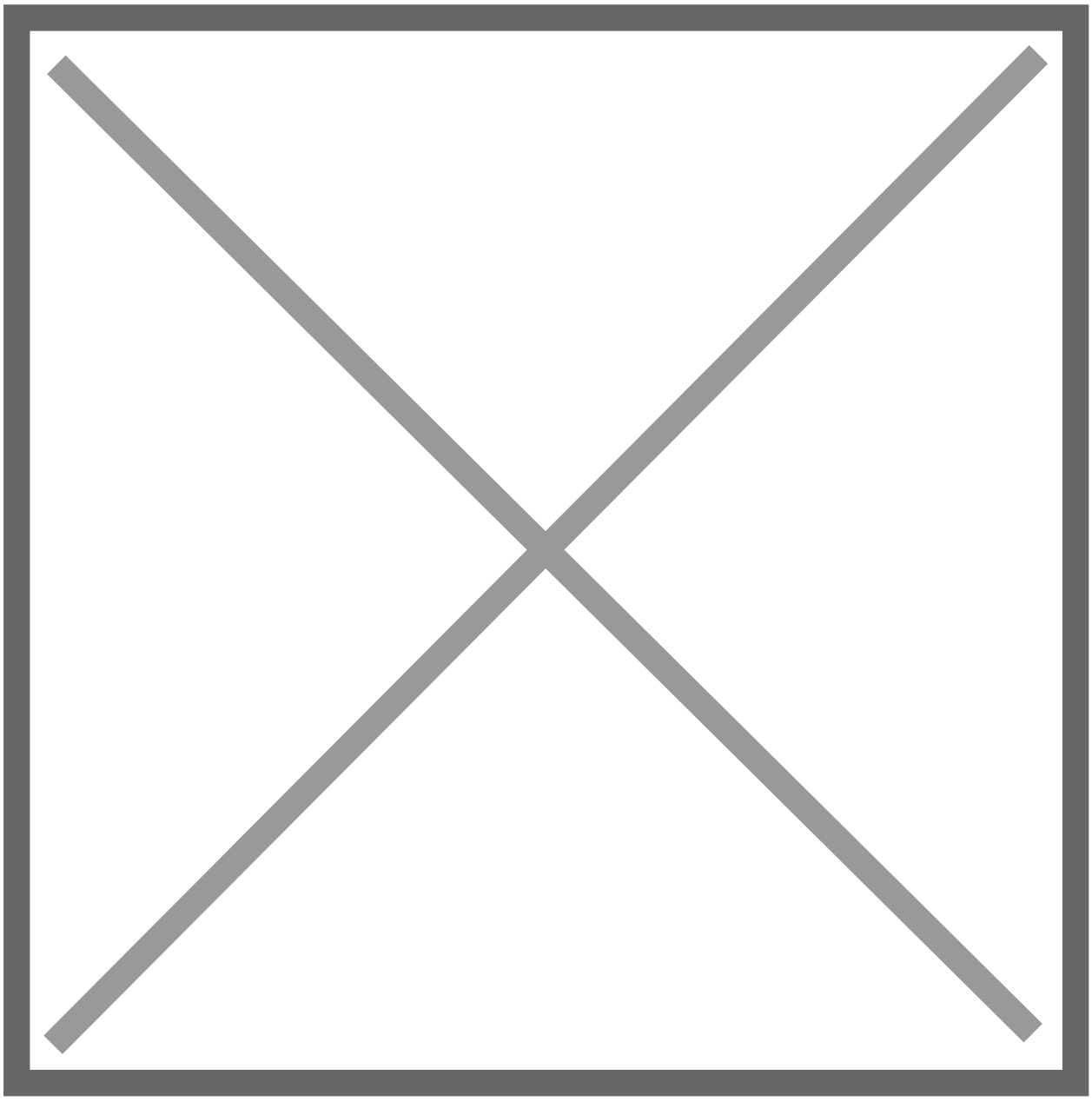
Actualmente, la mayoría de las aplicaciones de aprendizaje automático que se desarrollan utilizan un enfoque de aprendizaje supervisado. Una razón para esto es que las principales tareas supervisadas (clasificación y regresión) son útiles y están bien definidas, y a menudo se pueden abordar utilizando 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 comportamiento de clientes de banca.

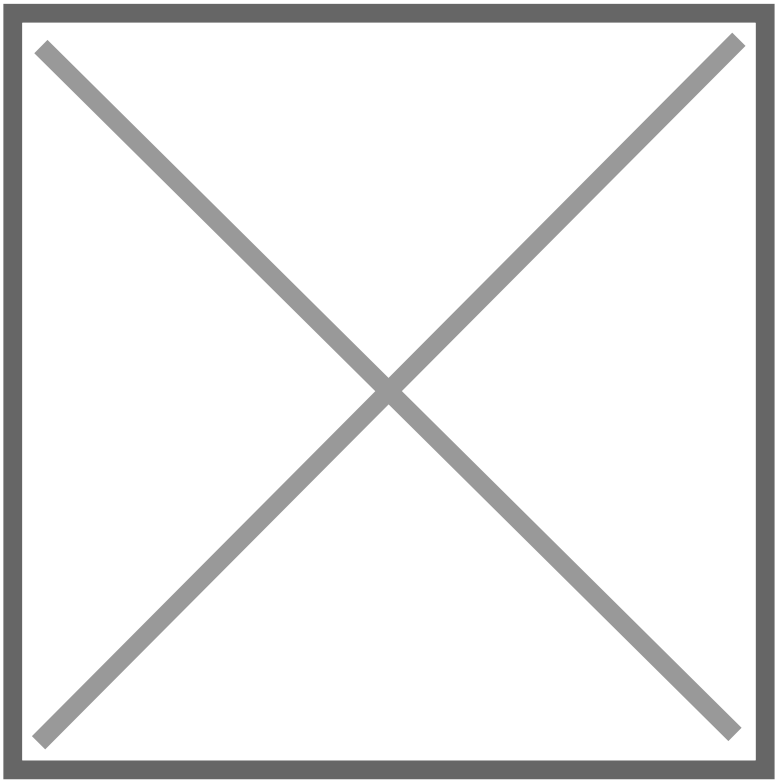
Copyright 2026

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.

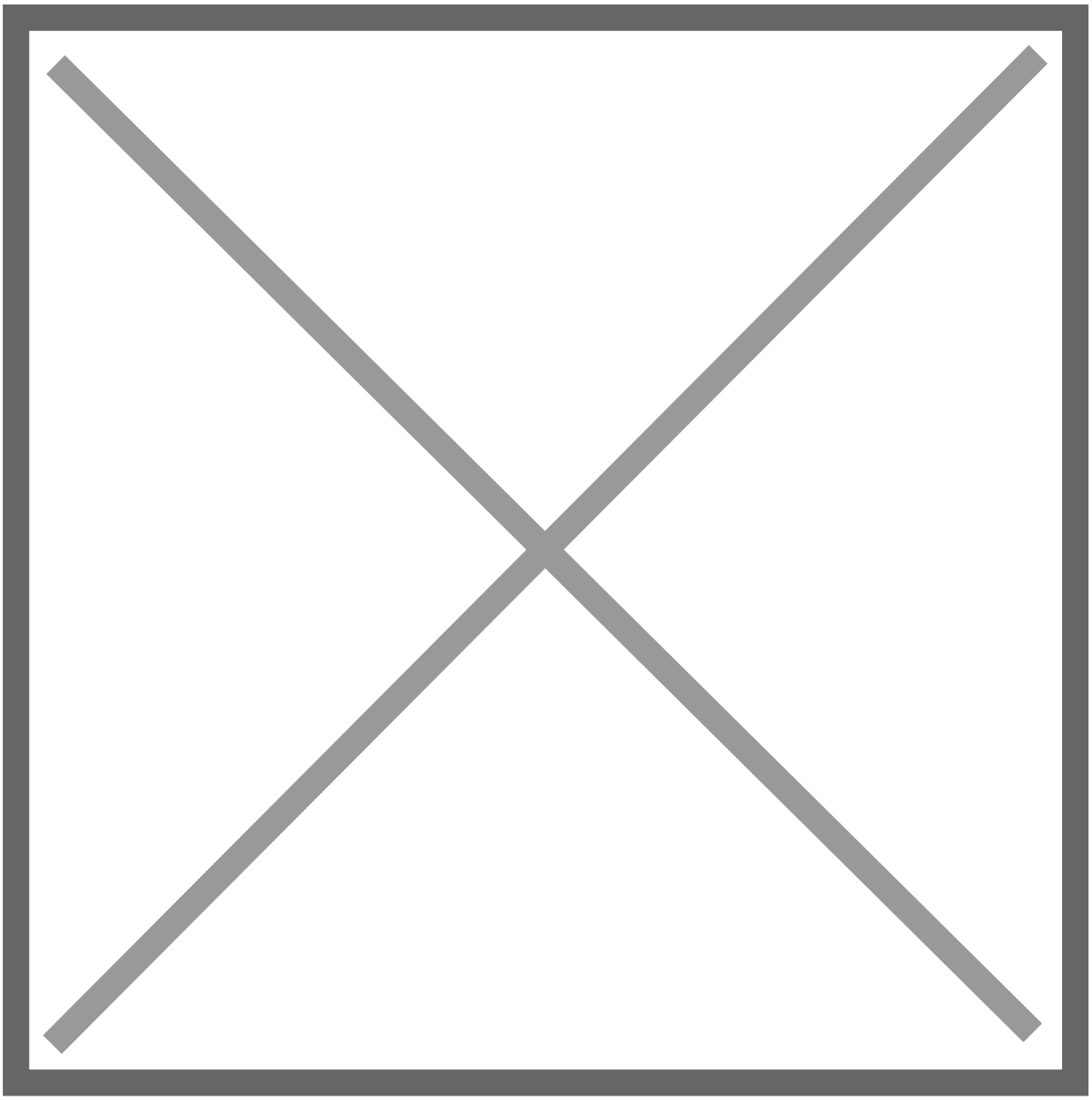
Los modelos se entrenan con millones de datos y dimensiones o características así que es importante optimizarlos para reducir el tiempo de entrenamiento y aumentar la eficiencia.

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

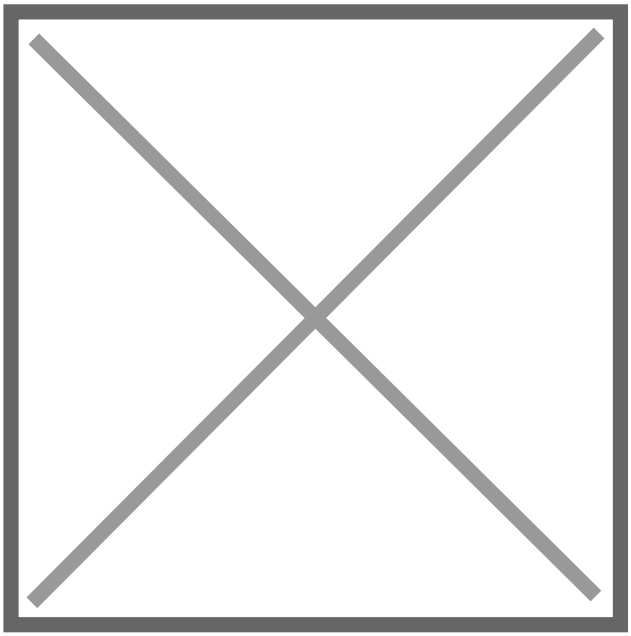
La detección de anomalías es otra tarea que se puede abordar de manera no supervisada. La detección de anomalías se refiere a la identificación de ejemplos que son anómalos, es decir, valores atípicos. Aquí vemos un ejemplo de detección de anomalías realizada en un conjunto de datos numéricos simple:



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.

Copyright 2026

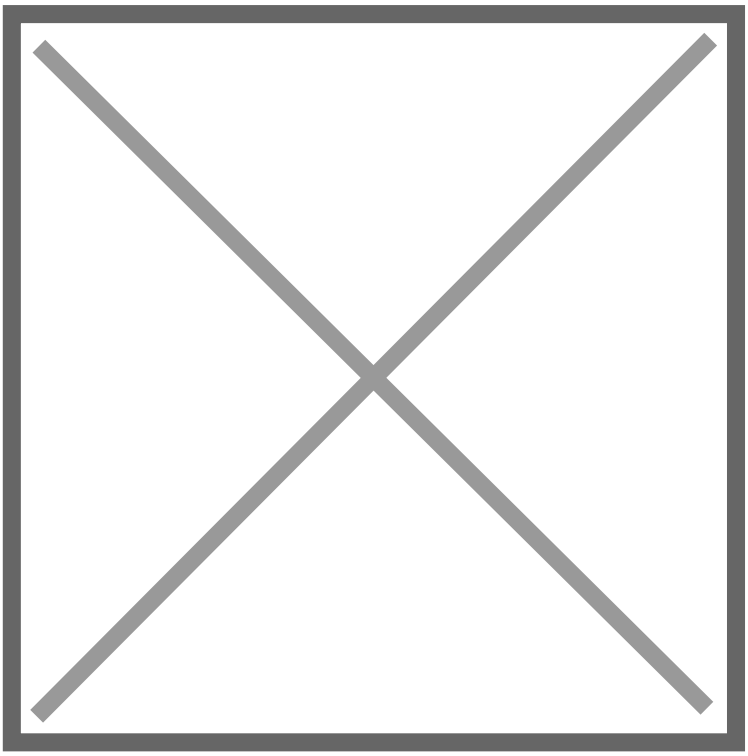
Otra tarea clásica de aprendizaje no supervisado se denomina imputación de valores faltantes, y el objetivo es completar los valores faltantes en un conjunto de datos:



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.

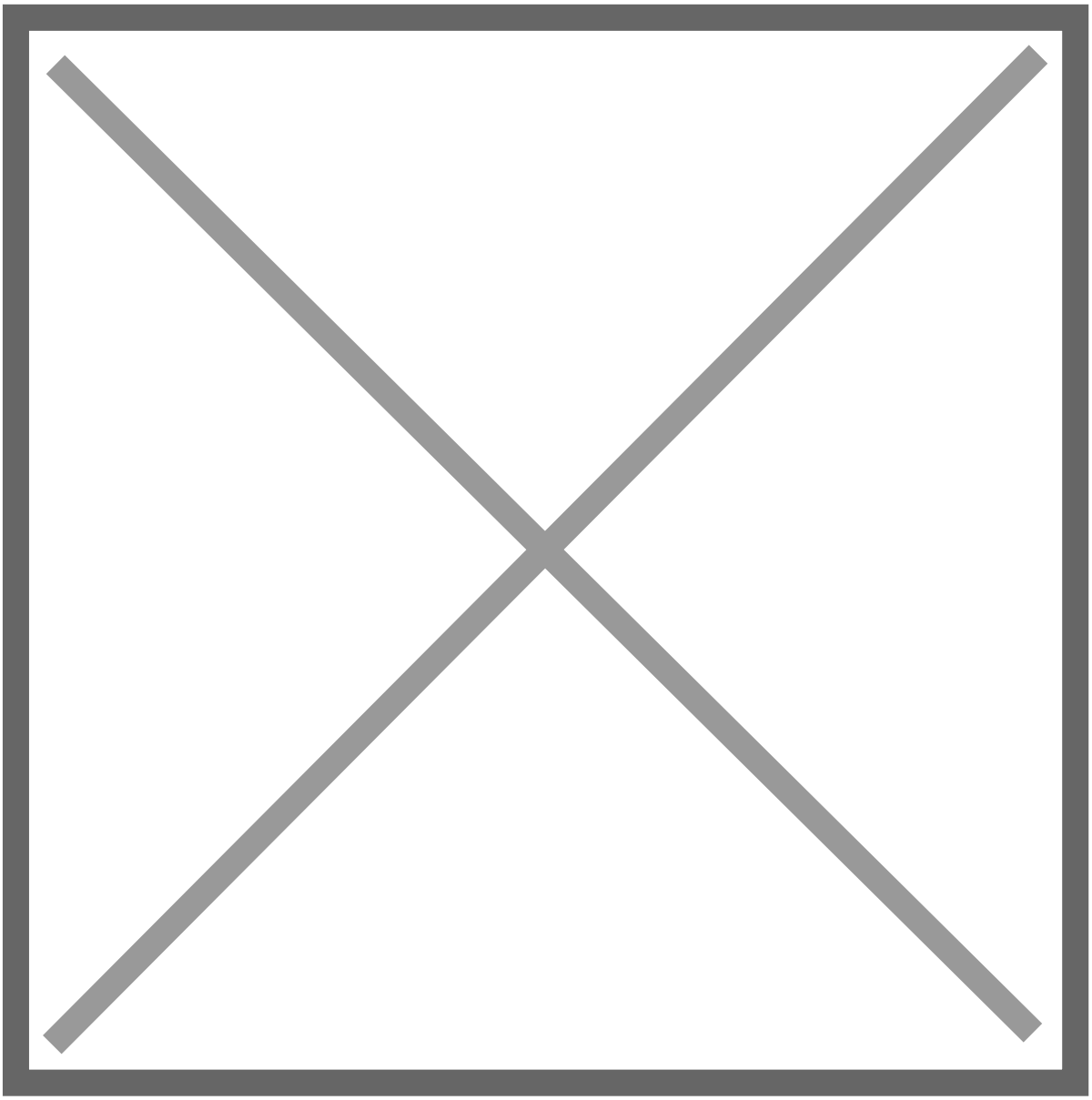
El aprendizaje no supervisado se utiliza un poco menos que el aprendizaje supervisado, principalmente porque las tareas que resuelve son menos comunes y más difíciles de implementar que las tareas predictivas. Sin embargo se puede aplicar a un conjunto más diverso de tareas que el aprendizaje supervisado. 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.

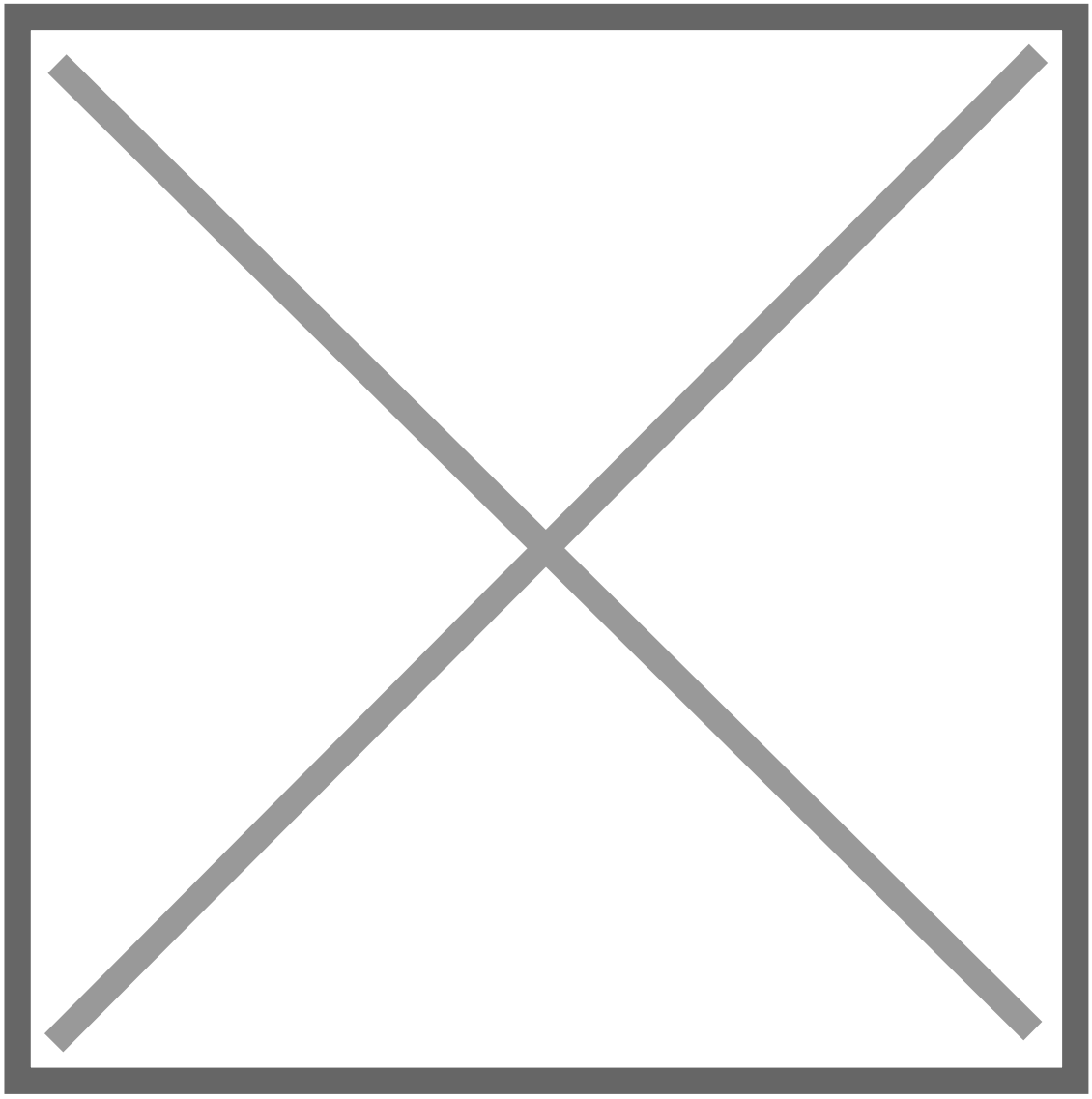
El aprendizaje por refuerzo se puede utilizar, por ejemplo, para que un robot aprenda a caminar en un entorno simulado. Aquí hay una captura de pantalla de un vídeo mostrado anteriormente que ilustra el aprendizaje por refuerzo hecho y grabado por la empresa Boston Dynamics. En él se fija un objetivo 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

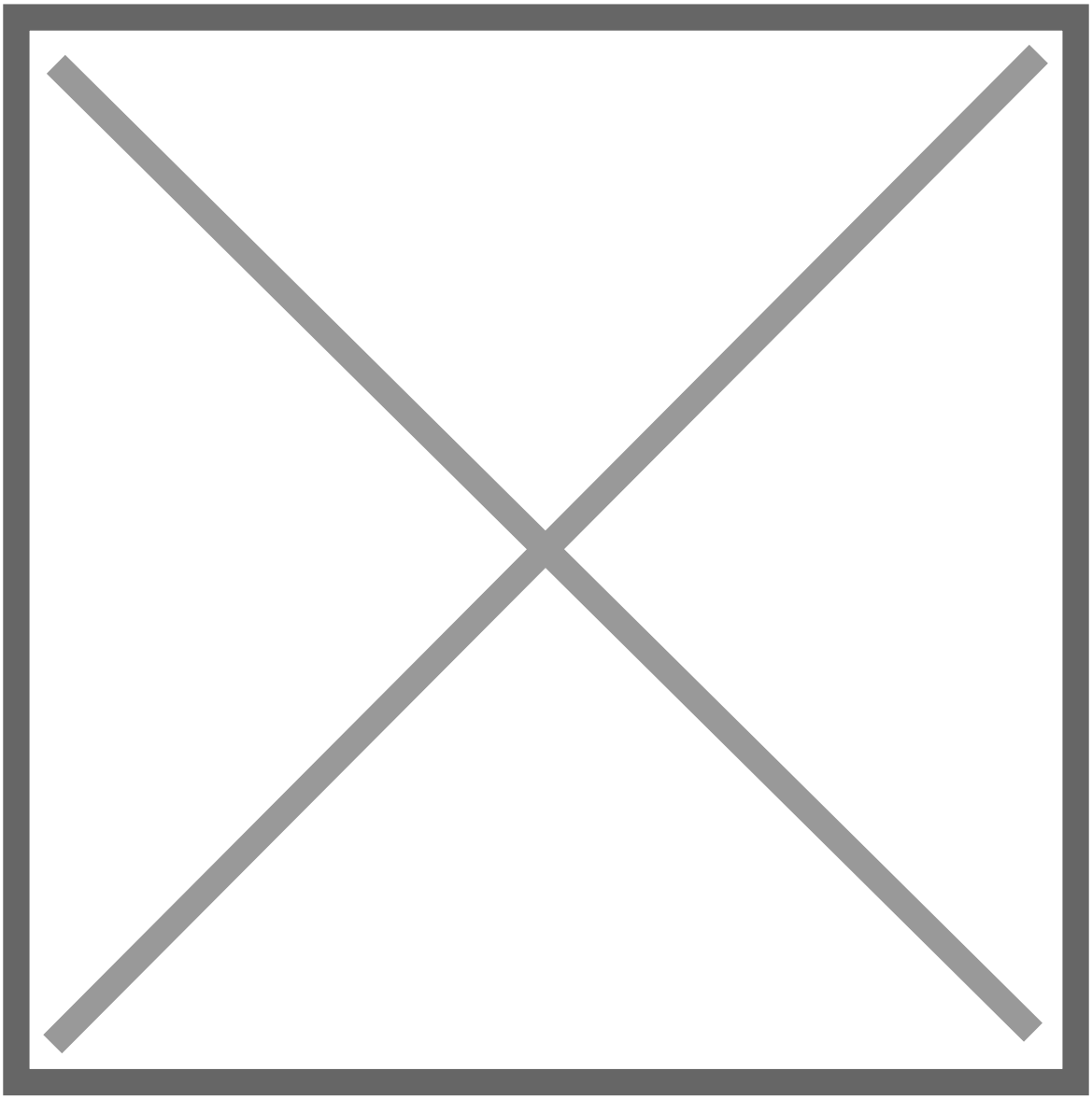
En este caso, las acciones son los valores de torsión aplicados a cada articulación de la pata; las observaciones son los ángulos de las patas, las fuerzas externas, etc.; y la recompensa está unida al equilibrio y la velocidad del robot. El aprendizaje en un entorno simulado como este puede luego utilizarse 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*:

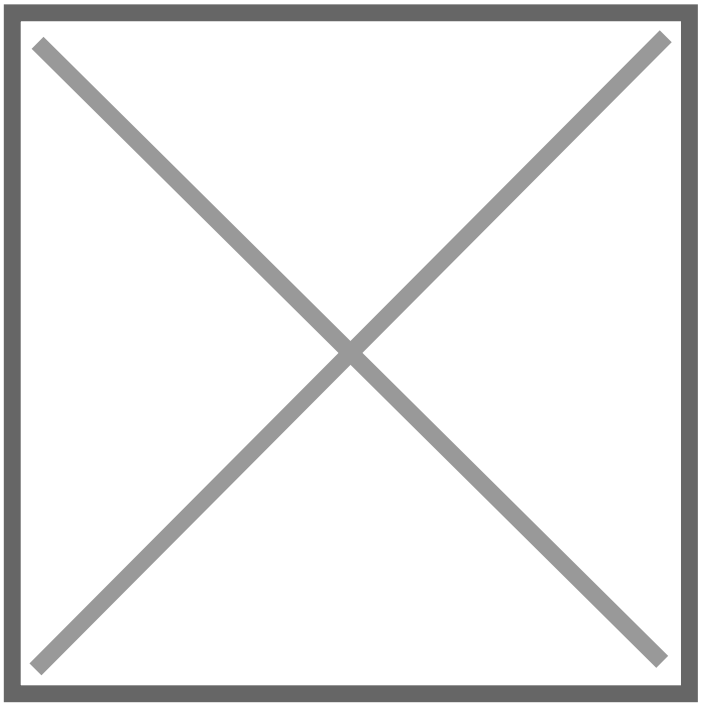
Copyright 2026



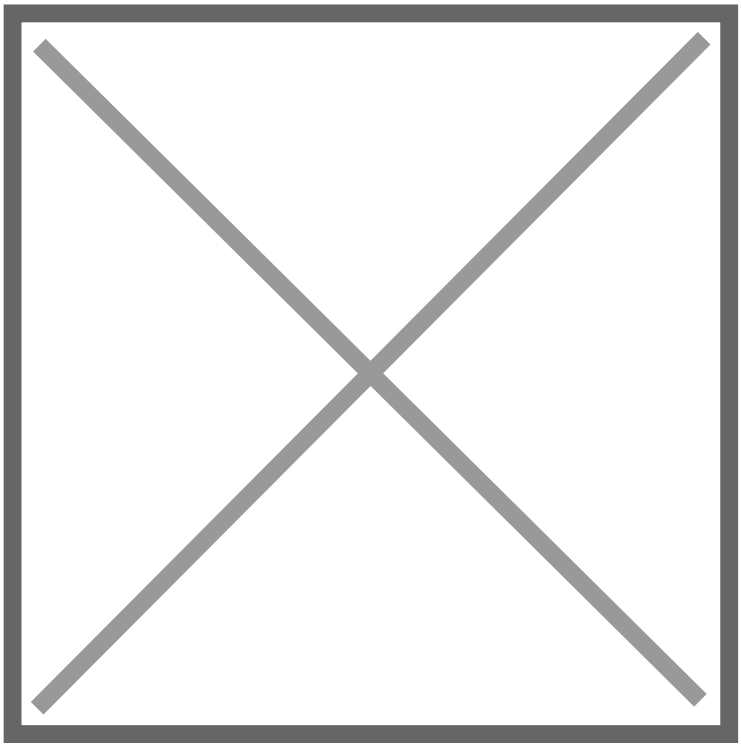
Ambos programas fueron desarrollados utilizando aprendizaje por refuerzo al hacer que el agente juegue contra sí mismo. Cabe destacar que la recompensa en este tipo de problemas solo se otorga al final del juego (ya sea que ganes o pierdas), lo que dificulta aprender qué acciones fueron responsables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

---

Revision #25
Created 2023-06-23 12:03:32 CEST by Luis Hueso
Updated 2024-03-16 12:15:58 CET by Luis Hueso

Copyright 2026