

Aprendizaje supervisado vs no supervisado



Dentro de Machine Learning, existen varias clasificaciones pero dos de ellas por lo general tienden a confundir un poco y son precisamente las relacionadas al aprendizaje supervisado y no supervisado. Estas describen dos formas en que las máquinas puedan entender un conjunto de datos y se espera que aprendan algo útil de ellas.

Aprendizaje supervisado

El aprendizaje supervisado es el más común utilizado entre los dos , incluye algoritmos tales **como regresión lineal y logístico**, clasificación de clases múltiples y **máquinas de vectores de soporte**.

El aprendizaje supervisado se llama así porque el desarrollador actúa como una guía para enseñar al algoritmo las conclusiones a las que debe llegar, es decir la salida del algoritmo ya es conocida. Es similar a la forma en que un niño podría aprender de un maestro.

Requiere que los posibles resultados del algoritmo ya sean conocidos y que los datos utilizados para entrenar el algoritmo ya estén etiquetados con las respuestas correctas. Por ejemplo, un algoritmo de clasificación aprenderá a identificar animales después de haber sido entrenados en un conjunto de datos de imágenes que están apropiadamente etiquetados con las especies del animal y algunas características de identificación.

Aprendizaje supervisado vs no supervisado

APRENDIZAJE SUPERVISADO



Aprendizaje no supervisado

El aprendizaje no supervisado está más estrechamente alineado con la inteligencia artificial, ya que la idea de que una computadora pueda aprender a identificar procesos y patrones complejos sin un humano para proporcionar orientación a lo largo del camino. Algunos ejemplos de algoritmos de aprendizaje no supervisado incluyen **clustering o agrupamiento, k-means y reglas de asociación**.

Acá no existe un conjunto de datos de entrenamiento y los resultados son desconocidos. Esencialmente, se entra en el problema de manera ciega y con solo operaciones lógicas impecables para guiarlo, aunque parezca increíble, el aprendizaje no supervisado es la capacidad de resolver problemas complejos utilizando solo los datos de entrada y los algoritmos lógicos, y en ningún momento se tiene datos de referencias.

Aprendizaje supervisado vs no supervisado

APRENDIZAJE NO SUPERVISADO

Está ligado a la
inteligencia artificial



Incluye clustering o
agrupamiento, k-means
y reglas de asociación



No existe un conjunto de
datos de entrenamiento y
los resultados son
desconocidos, se entra en
el problema de manera
ciega y con solo
operaciones lógicas para
guiarlo

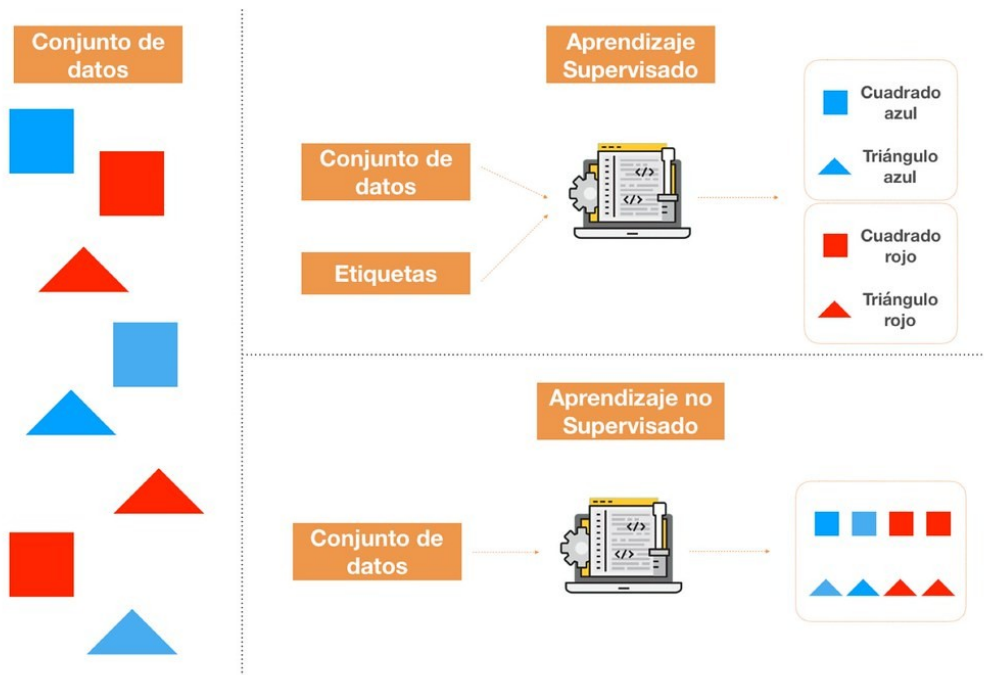


Un ejemplo muy simple de ambos aprendizaje es que digamos que tenemos una imagen digital que muestra una serie de formas geométricas de colores que debemos unir en grupos según su clasificación y color, este es un problema muy común en las aplicaciones de reconocimiento de imágenes de Machine Learning.

Con el aprendizaje supervisado, la solución a este problema es bastante sencillo, simplemente le enseñamos a la computadora que las formas con cuatro lados se conocen como cuadrados y las formas de tres lados se conocen como triángulos. También le indicamos que si la luz emitida por un pixel registra ciertos valores, lo

Aprendizaje supervisado vs no supervisado

clasificamos rojo y otro conjunto de valores como azul. De esta forma muy sencilla nuestro algoritmo ha aprendido a agrupar el conjunto de formas geométricas dado.



Ahora bien con el aprendizaje no supervisado, la solución a este problema se vuelve un poco más complicado. Acá el algoritmo tiene los mismos datos de entrada, imágenes digitales que muestran formas geométricas en diferentes colores y queremos ordenarlos en grupos. Para solucionar este problema, el algoritmo aprende de la información: el problema es de clasificación y que algunas de las formas coinciden con otras, quizás con el mismo número de lados o con marcadores digitales que coinciden con el color. El algoritmo no puede saber que el nombre de los objetos es cuadrado o triángulo, pero si reconocerá otros objetos con más o menos las mismas características y los agrupará asignándoles su propia etiqueta. De esta forma el algoritmo ha resuelto el problema dado.

Técnicamente no hay respuesta correcta o incorrecta, Machine Learning simplemente aprenderá la verdad objetiva de que ciertas formas pertenecen juntas hasta cierto grado de probabilidad. Obviamente, Machine Learning comete errores pero, como nosotros, su fortaleza radica en su capacidad para aprender de sus errores y hacer estimaciones mejor educadas la próxima vez.

La elección de utilizar un algoritmo de Machine Learning supervisado o no supervisado generalmente depende de factores relacionados con la estructura y el volumen de sus datos y del objetivo del problema en cuestión. Un problema complejo por lo general utilizará ambos tipos de algoritmos para construir modelos de datos predictivos que ayuden a tomar decisiones sobre una variedad de desafíos comerciales.