



ORIGINAL

Sayonara, baby. Aprendizaje automático.

Campos Rosas C, Martínez Lorenzo M, Ruiz Portillo L, Bernal Basurto R

Hospital Universitario Torrecárdenas, Almería.

Resumen

El aprendizaje automático, disciplina englobada en la ciencia de datos, implica entrenar sistemas para aprender de los datos y hacer predicciones o decisiones sin ser programados explícitamente para ello.

Introducción



El aprendizaje automático, disciplina englobada en la ciencia de datos, implica entrenar sistemas para aprender de los datos y hacer predicciones o decisiones sin ser programados explícitamente para ello.

Seguro que conocéis la serie de películas sobre el Terminator, del director de cine James Cameron. En esta ficción, un “terminator” es un robot humanoide diseñado por Skynet como soldado prácticamente indestructible y que se dedica a exterminar pobres seres humanos, que están sometidos y en guerra contra las máquinas inteligentes, que se han rebelado y dominan el mundo.

El protagonista, un modelo T-800, bastante malvado en la primera película y transformado en bueno en la segunda, se pelea con el malo de verdad, un T-1000. Cuando termina una de sus batallas, el T-800 se despide con estas

palabras, que han quedado para la posteridad: Sayonara, baby.

No es que la frase sea muy intelectual ni que tenga mucho fondo, pero tampoco había mucho donde elegir en el repertorio del T-800, efectivamente letal como soldado y protector, pero verdaderamente parco en palabras.

Pero el fondo de esta historia, si lo pensamos dos veces, puede ser inquietante. ¿Qué es eso de robots buenos y malos? ¿Hay máquinas inteligentes? ¿Podrían las máquinas rebelarse contra nosotros y llegar a dominarnos e, incluso, asesinarlos? Lo digo porque en los tiempos en los que tanto se habla de inteligencia artificial o de aprendizaje automático, estas ideas se nos pueden pasar por la cabeza.

En efecto, a lo largo de los últimos años han cobrado popularidad términos como inteligencia artificial, aprendizaje automático (*machine learning* para los que saben inglés) y aprendizaje profundo (*deep learning*), que nos pueden hacer pensar que las máquinas puedan ser inteligentes e, incluso, aprender.

Sin embargo, y para tranquilidad de todos (al menos, por el momento), cuando vemos qué significan realmente estos términos, nuestros temores empiezan a disiparse y desaparecen

como lágrimas en la lluvia (palabras de otro robot, un replicante, mucho más instruido que nuestro T-800). Por suerte y, al menos, por el momento, la inteligencia artificial tiene más de artificial que de inteligencia.

Inteligencia artificial

Los inicios de la inteligencia artificial (IA) se sitúan en la década de los 50 del siglo pasado, cuando algunos de los creadores de otra ciencia naciente, la informática, empiezan a plantearse si los ordenadores podrían ser diseñados para “pensar”. Este término tiene tantas implicaciones posibles que, aun hoy día, seguimos dándole vueltas.

Podemos definir la IA como la simulación de la inteligencia humana en máquinas programadas para pensar y aprender como los humanos, de forma que puedan automatizar tareas que normalmente son realizadas por seres humanos. Sin embargo, esta es una definición moderna del término, ya que hasta la década de los 80 no se hace ninguna mención al asunto del aprendizaje.

En la actualidad, los sistemas de IA se pueden entrenar para realizar tareas como reconocer el habla, entender el lenguaje natural, tomar decisiones y jugar juegos. Estos sistemas también se pueden utilizar para mejorar la toma de decisiones y automatizar procesos en diversos sectores. Y, no cabe duda, han conseguido logros sorprendentes, como es el caso de los asistentes virtuales de Google o Amazon, la transcripción y traducción de textos, la personalización de anuncios publicitarios, etc.

Pero esto no debe hacernos temer una futura rebelión de las máquinas: en la actualidad, las máquinas y los sistemas de IA no tienen la capacidad de rebelarse contra los humanos. Los sistemas de IA están diseñados y

programados para realizar tareas específicas y cumplir con ciertos objetivos, pero no tienen la capacidad de tener metas o deseos propios, ni la de tomar decisiones basadas en emociones o motivaciones.

El problema puede estar en el programador, no en la máquina. Es fundamental que el desarrollo de la IA sea supervisado y regulado para evitar posibles problemas éticos y de seguridad.

Aprendizaje automático (*machine learning*)

Podemos definir el aprendizaje automático como un subcampo de la IA que implica entrenar sistemas para aprender de los datos y hacer predicciones o decisiones sin ser programados explícitamente para ello. Veamos qué significa esto.

El primer ordenador de propósito general que se conoce es el *Analytical Engine*, que fue diseñado a mediados del siglo XIX por Charles Babbage. Su amiga, Ada Lovelace, considerada la primera programadora de la historia, decía que una máquina como esta no tenía pretensión alguna de crear nada nuevo, sino simplemente llevar a cabo aquellas tareas que se le ordenase cómo realizar. En resumen, era un asistente para realizar tareas bien comprendidas por los seres humanos.

Casi dos siglos después, esta afirmación puede cuestionarse. ¿Podría una máquina “originar” algo o tendría que limitarse siempre a los procesos comprendidos por los humanos? ¿Podría hacer algo nuevo, mostrarse creativa? ¿Aprender de su experiencia?

La posibilidad del aprendizaje automático implica un cambio en la forma de utilizar los ordenadores en nuestro flujo de trabajo. Con el

planteamiento clásico, como el de Ada Lovelace, se dan al ordenador una serie de instrucciones (un programa informático) que debe seguir de forma generalmente secuencial para transformar los datos en un resultado, la respuesta al problema. Para esto, el programador debe saber con antelación las reglas que gobiernan el comportamiento de los datos, para poder escribir el programa de instrucciones para la máquina.

El aprendizaje automático da la vuelta a este planteamiento: el ordenador “mira” los datos de entrada y la respuesta final correspondiente y, a partir de ambos, averigua las reglas que gobiernan el comportamiento de los datos, que eran desconocidas con antelación.

Un sistema de aprendizaje automático no es programado, sino entrenado: se le proporcionan muestras relevantes a una determinada tarea y encuentra la estructura estadística subyacente a los datos, lo que le permite automatizar la tarea y aplicarla a datos nuevos con los que el sistema no había entrado previamente en contacto.

Por tanto, el aprendizaje automático implica el concepto de aprendizaje a partir de los datos. Ya no se diseña un algoritmo para dar respuesta a un problema concreto, sino que diseñamos un algoritmo de aprendizaje genérico que, a partir de los ejemplos y soluciones que se le proporcionan, es capaz de resolver el problema en otras situaciones.

Esto puede parecer muy abstracto, pero, en realidad, lo hacemos habitualmente al utilizar técnicas de estadística “convencional”. Pensemos es un modelo estadístico sencillo al que estamos más habituados: la regresión lineal simple.

Establecemos el modelo según la siguiente fórmula:

$$y = ax + b + e$$

donde “e” es el componente de error del modelo. Podemos desarrollar un modelo para estimar los coeficientes del modelo lineal sin conocer antes los valores “x” e “y”, pero no podemos presuponer la relación lineal sin mirar antes los datos concretos. Cuando entrenamos el modelo lineal (probamos diferentes coeficientes y vemos cómo se ajusta), estamos “aprendiendo” de los datos.

Con todo lo dicho, podríamos definir el aprendizaje automático como una disciplina que desarrolla y aplica modelos y algoritmos que son capaces de aprender de los datos.

Tipos de aprendizaje automático

Aunque podemos encontrar descrito algún tipo más, en general consideramos dos tipos de aprendizaje automático: supervisado y no supervisado.

En el aprendizaje automático supervisado se proporciona al algoritmo de aprendizaje automático un conjunto de datos con las respuestas correctas ya conocidas para que aprenda a generalizar y hacer predicciones precisas sobre nuevos datos. Se utilizan para predecir el valor desconocido de una variable a partir de una serie de variables conocidas.

Imaginemos que queremos predecir el precio de un coche en función de características como el tamaño, la cilindrada, la autonomía, el tipo de combustible utilizado, etc. A partir de los coches en los que se proporciona el precio y sus características, el modelo trataría de predecir el precio de un coche que no hubiese visto antes si le decimos qué características tiene.

Por otra parte, en el aprendizaje automático no supervisado, el algoritmo no recibe información etiquetada sobre las respuestas correctas. En lugar de eso, se basa en la estructura y relaciones presentes en los datos de entrada para descubrir patrones y características interesantes que son desconocidas para el investigador.

Siguiendo con el ejemplo de nuestros coches, observando todas sus características podría descubrir patrones que permitiesen clasificarlos en distintos grupos: deportivos, de confort, familiares, etc.

Algoritmos de aprendizaje automático supervisado

El aprendizaje supervisado suele dividirse en dos grandes grupos, la regresión y la clasificación. La primera se utiliza cuando la variable de resultado que tratamos de predecir es numérica, mientras que la clasificación se usa para la predicción de variables categóricas.

Como ejemplo de regresión, un modelo puede ser entrenado con datos históricos de precios de viviendas y luego utilizado para predecir el precio de una vivienda específica basándose en su tamaño, ubicación, etc. Los algoritmos de regresión más populares incluyen la regresión lineal, la regresión polinómica, la regresión logística y las redes neuronales.

Finalmente, como ejemplo de clasificación, un algoritmo de clasificación puede ser entrenado con imágenes etiquetadas de animales y luego utilizado para clasificar imágenes de animales no vistos previamente. Los algoritmos de clasificación más populares incluyen la regresión logística, los k-vecinos más cercanos, los árboles de decisión, los bosques aleatorios, las máquinas de vectores de

soporte, el método naive-Bayes y las redes neuronales.

Algoritmos de aprendizaje automático no supervisado

Como ya hemos dicho, el aprendizaje no supervisado no pretende predecir variables a partir de otras variables, sino que se centra en descubrir patrones ocultos que pueden seguir los datos que se proporcionan al algoritmo. Para ello utiliza técnicas de reducción de la dimensionalidad.

La reducción de la dimensionalidad consiste en reducir la cantidad de características o dimensiones de un conjunto de datos, mientras se intenta preservar la información importante. Esto se hace para resolver los problemas que pueden encontrar los algoritmos de aprendizaje para manejar conjuntos de datos con muchas características, y también para simplificar la visualización de los datos.

Como ejemplo, estas técnicas son muy útiles en los estudios sobre genética, en los que pueden estudiarse miles o millones de polimorfismos en cada participante. Resulta útil encontrar un conjunto reducido de variables que explique la mayor parte de la variabilidad genética de cada individuo y hacer el análisis de los datos solo con ese conjunto reducido.

Las dos técnicas más utilizadas para reducción de dimensionalidad son el estudio de componentes principales y los métodos de agrupamiento (conglomerados o *clustering*).

Aprendizaje profundo (*deep learning*).

Todos los algoritmos que hemos visto hasta ahora hacen su trabajo de forma más o menos visible. Como ejemplo tenemos las redes neuronales simples,

que ya hemos visto que pueden servir tanto para regresión como para clasificación.

Escalando un paso en la complejidad de estos algoritmos, podemos crear redes combinando capas de forma sucesiva. Cada capa se encarga de tomar los datos de la capa anterior, hacer alguna transformación en los datos y pasarla a la siguiente. Así, utilizamos estas redes neuronales con varias capas ocultas para modelar patrones y relaciones complejas en los datos. Es lo que se conoce como aprendizaje profundo (*deep learning*, para nuestros amigos angloparlantes).

Estas capas internas pueden aprender características no lineales y combinarlas para realizar tareas complejas como la clasificación de imágenes, el procesamiento del lenguaje natural y la generación de imágenes y audio.

No vamos a entrar en el funcionamiento de las redes neuronales, ya que se merece una entrada (o más) para explicarlo.

¿Qué es Big Data?

Ya podemos imaginar que todas estas técnicas de aprendizaje automático emplean conjuntos de datos más extensos y complejos que los que suelen utilizarse en las técnicas estadísticas más convencionales. Esto suele llevarse a límites insospechados cuando se combinan la multitud de datos que se generan diariamente y la potencia de computación que se ha ido desarrollando durante las últimas décadas.

Cuando esto se sale totalmente de madre, tenemos que inventarnos un concepto nuevo: Big Data (y permitidme que utilice el término en inglés). De una manera sencilla, podemos hablar de Big Data cuando los

datos que estamos utilizando no pueden manejarse en ordenadores convencionales. Y no estoy hablando de gigabytes ni de terabytes, sino de petabytes (1000 Tb) e, incluso, exabytes (1000 Pb) de datos.

Esto hace que los sistemas informáticos capaces de trabajar con Big Data no estén al alcance de cualquiera. En la práctica, solo algunas grandes corporaciones como Google, Facebook o Amazon trabajan realmente con Big Data.

Un cambio de paradigma

Todo lo que hemos hablado en la entrada de hoy puede englobarse bajo el término de Ciencia de Datos, que puede definirse de forma simplificada como la disciplina encargada del estudio de métodos, técnicas y herramientas para el almacenamiento, recuperación, análisis y visualización de datos.

La Ciencia de Datos se ha desarrollado de forma paralela a nuestra capacidad de computación y de generar cada vez mayores volúmenes de datos. En los últimos años se ha ido popularizando de manera progresiva y creo, además, que va a cambiar la forma de orientar nuestros estudios estadísticos.

En el pasado, con volúmenes de datos no muy grandes y con una capacidad de computación reducida (imaginad antes de disponer de ordenadores), el esfuerzo se dirigió en desarrollar técnicas estadísticas de inferencia capaces de funcionar con los medios disponibles.

Sin embargo, en la actualidad disponemos de un volumen de datos mucho mayor y de una capacidad de almacenamiento y de cómputo infinitamente más grande. No cabe duda de que los algoritmos de aprendizaje automático irán ocupando el hueco que les ha correspondido durante tanto

tiempo las técnicas estadísticas más clásicas.

Creo que, sin que nos demos cuenta, la sección de Métodos de nuestros trabajos científicos se verá invadida por redes neuronales, bosques aleatorios o máquinas con soporte de vectores, por decir algunos. Seguid mi consejo: estudiad técnicas de Ciencia de Datos.

Nos vamos...

Y aquí vamos a finalizar esta larga y espesa (aunque espero que provechosa) entrada de hoy.

Prácticamente no hemos hablado nada de en qué consisten los diferentes algoritmos que hemos mencionado. Podríamos haber hablado también un poco de cómo funciona una red neuronal, cómo se entrena y cómo aprende. No os preocupéis, volveremos

sobre estos temas. Pero esa es otra historia...

Bibliografía

– Chollet F, Kalinowski WT, Allaire JJ. Deep learning with R, 2nd ed. New York: Manning Publications Co (2022).

– Baumer BS, Kaplan DT, Horton NJ. Modern Data Science with R, 2nd ed. Boca Ratón, Florida: Taylor & Francis Group, LLC (2021). ([HTML](#))

– Mailund T. Beginning Data Science in R4. Data análisis, visualization, and modelling for the data scientist, 2nd ed. Aarhus, Dinamarca: Apress (2022).

Correspondencia al autor

Manuel Molina

mma1961@gmail.com

Servicio de Gastroenterología.

Hospital Infantil Universitario La Paz.

Madrid. España.

Aceptado para el blog en febrero de 2023

