



En qué consiste el aprendizaje automático (machine learning) y qué está aportando a la Neurociencia Cognitiva

Carlos González-García

Dept. of Experimental Psychology, Ghent University, Bélgica

Tipo de artículo: Actualidad.

Disciplinas: Psicología, Neurociencia.

Etiquetas: cerebro, cognición, machine learning, inteligencia artificial, aprendizaje automático.

El aprendizaje automático permite una comprensión hasta ahora inaudita de conjuntos complejos de datos, lo que le está otorgando un creciente protagonismo en nuestra sociedad en general y en la neurociencia cognitiva en particular. Estas aplicaciones han supuesto un ilusionante avance en el estudio de cuestiones básicas acerca de nuestro sistema cognitivo, al igual que en el diagnóstico de algunas importantes enfermedades que afectan a este sistema. Pese a la novedad de estos trabajos, la flexibilidad del aprendizaje automático permite pronosticar que las aportaciones más relevantes del aprendizaje automático están aún por llegar.



El aprendizaje automático (“machine learning”, en inglés) hace referencia al subcampo dentro de las ciencias de la computación especializado en el reconocimiento de patrones complejos en conjuntos de datos. A diferencia de la programación clásica, en la que un programa ejecuta una y otra vez la misma (más o menos compleja) operación, la principal característica del aprendizaje automático es que sus programas consiguen extraer de forma autónoma (es decir, sin ser programados específicamente para ello) información relevante en los datos que están siendo procesados. Esta información permite que el programa “aprenda”, es decir, que mejore en su ejecución de la tarea para la que había sido programado (Turing,

1950). Mediante el desarrollo de algoritmos sofisticados (los cuales pueden ser entendidos como "modelos"), estos acercamientos permiten identificar relaciones invisibles para el ojo humano. Este tipo de algoritmos

(cc0) John Hain.

González-García, C. (2018). En qué consiste el aprendizaje automático (machine learning) y qué está aportando a la Neurociencia Cognitiva. *Ciencia Cognitiva*, 12:2, 48-50.



interaccionan con nosotros en nuestro día a día cuando, por ejemplo, la cámara de fotos de nuestro móvil reconoce una cara o cuando utilizamos una aplicación de traducción automática. En este sentido, parte del éxito de estas herramientas se debe a su extenso campo de acción: desde sistemas que detectan mutaciones en nuestro ADN (Libbrecht y Noble, 2015) hasta el "big data", el cual identifica patrones en enormes conjuntos de datos acerca de, por ejemplo, diferentes segmentos de nuestra sociedad (Boyd y Crawford, 2012). Como es esperable, las ciencias cognitivas no han sido inmunes al desarrollo de estas herramientas. Una simple consulta de los términos "machine learning" y "brain" en un repositorio de artículos (PubMed; fecha de consulta: julio 2017) permite observar que, mientras en 1990 únicamente un artículo contenía estas etiquetas, 298 trabajos publicados en 2016 coinciden con nuestra búsqueda.

Pero, ¿cómo funciona en la práctica el aprendizaje automático? Para responder a esta pregunta quizá sea beneficioso entender qué diferencia conceptual suponen, dentro de la neurociencia cognitiva, estos análisis respecto a acercamientos más clásicos (véase la Figura 1, panel A).

Imaginemos que el objetivo de nuestro estudio es comparar el procesamiento de imágenes de perros y gatos en dos regiones cerebrales (región X y región Y). Para cada una de estas regiones obtenemos un patrón de activación por estímulo presentado (en la figura, cada cuadrado del patrón representa la activación [negro = activado; blanco = desactivado] en un vóxel [unidad volumétrica utilizada en las imágenes de resonancia magnética]). Siguiendo la lógica de los análisis clásicos, procederíamos a promediar la actividad a lo largo de estos vóxeles. En todos los patrones presentados, hay 3 vóxeles activados y 3 desactivados. El promedio, por tanto, sería el mismo para ambos animales en las dos regiones, sugiriendo que éstas están implicadas de manera similar en el procesamiento de perros y gatos. Sin embargo, podemos observar que, mientras en la región X esta inferencia parece ser acertada, los patrones en la región Y son diferentes, indicando una potencial diferencia en la representación de perros y gatos en esta región.

Pese a que en este ejemplo la diferencia salta a simple vista, este tipo de observaciones se antojan complicadas cuando lo que tenemos delante es un conjunto de datos más amplio. Es aquí donde los investigadores han utilizado técnicas de aprendizaje automático para detectar, entre toda la complejidad de nuestros datos, patrones asociados a diferentes representaciones. Básicamente, el funcionamiento de estas técnicas (Figura 1, panel B) consiste en el entrenamiento de un algoritmo para diferenciar dos clases (en

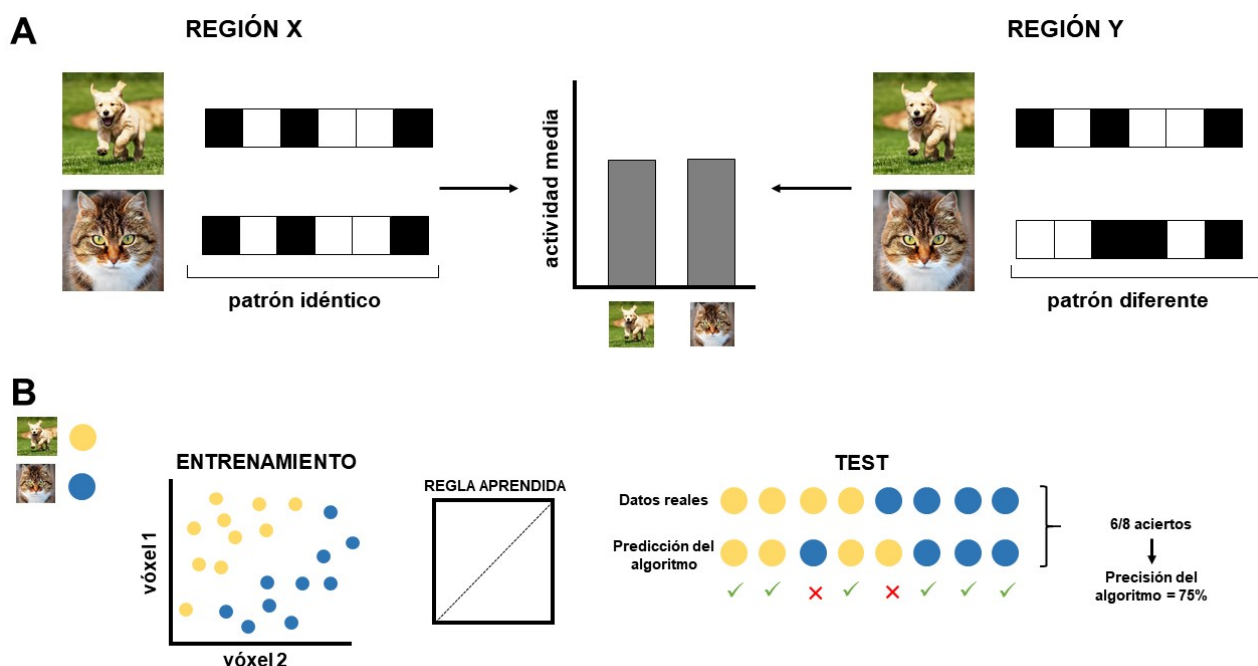


Figura 1.- Ilustración del funcionamiento de la aproximación clásica y la basada en el aprendizaje automático a la localización de funciones en el cerebro.

nuestro caso, perros y gatos) mediante la presentación de patrones asociados a ambas categorías (puntos amarillos y azules en la figura). De esta manera, el algoritmo aprende qué regla utilizar para separar ejemplares de cada clase (en este ejemplo, una recta diagonal). El paso clave (“test”) consiste en la presentación de ejemplares nuevos sin etiquetar, con el objetivo de comprobar la precisión del algoritmo a la hora de asignar cada ejemplar a la clase que corresponde. De esta manera, cuando la actividad de una zona cerebral permite clasificar los ejemplares de dos categorías diferentes por encima del nivel del azar, asumiremos que esa zona está representando de forma diferencial dichas categorías.

¿Qué importancia tienen estas suposiciones en neurociencia cognitiva? Mientras que las técnicas clásicas nos permitían detectar qué zonas parecían estar implicadas en ciertos procesos, no era posible entender qué representaciones estaban siendo codificadas en esas regiones. Así, por ejemplo, ante la activación de la corteza prefrontal durante una tarea de atención, el aprendizaje automático nos permite estudiar si esta activación subyace a procesos inespecíficos de control (en los cuales diferentes categorías serían representadas de forma similar) o, por el contrario, a la codificación diferencial de contenidos relevantes para la tarea (Haynes, 2015).

La aplicación del aprendizaje automático en contextos de corte más clínico es especialmente ilusionante. Por ejemplo, diversos estudios han demostrado cómo, a partir de imágenes estructurales del cerebro, es posible detectar si una persona con trastorno cognitivo leve (TCL) desarrollará en el futuro una demencia tipo Alzheimer (Moradi, Pepe, Gaser, Huttunen y Tohka, 2015). En este estudio, los investigadores sacaron partida de una de las comentadas cualidades del aprendizaje automático, la extracción de regularidades significativas en complejos conjuntos de datos. Concretamente, “alimentaron” al clasificador con tres fuentes de información: imágenes del estado estructural del cerebro, las puntuaciones en diferentes cuestionarios de habilidades cognitivas y la edad de cada participante, todo ello para pacientes con TCL que posteriormente desarrollaron Alzheimer, así como para pacientes con TCL que no desarrollaron la enfermedad (en nuestro ejemplo anterior, esto correspondería a los diferentes puntos azules y amarillos). Los resultados de este estudio demostraron que el algoritmo podía diferenciar con un 80% de precisión los dos grupos de pacientes sólo utilizando las imágenes cerebrales y que este porcentaje aumentaba al 90% cuando se combinaban todas las fuentes de información. Crucialmente, este tipo de resultados permite adelantar el diagnóstico entre 1 y 3 años en comparación con otras herramientas disponibles, lo cual supone una ventaja crucial de cara al tratamiento.

En cualquier caso, el aspecto más interesante del aprendizaje automático es su flexibilidad. Aunque sea de manera introductoria, este artículo permite vislumbrar que estas herramientas son moldeables y aplicables a multitud de problemáticas diferentes. En un mundo como el actual, en el que la cantidad de datos generados en cada instante es ingente, el potencial del aprendizaje automático es innegable. El hecho de que los recursos provenientes del aprendizaje automático utilizados en neurociencia cognitiva sean aún limitados sugiere que su potencial es aún mayor en esta disciplina. Es de esperar, por tanto, que las aportaciones más relevantes estén aún por llegar.

Referencias

- Boyd, D., y Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Communication & Society*, 15, 545–662.
- Haynes, J.-D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron*, 87, 257–270.
- Libbrecht, M. W., y Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, 16, 321–332.
- Moradi, E., Pepe, A., Gaser, C., Huttunen, H., y Tohka, J. (2015). Machine learning framework for early MRI-based Alzheimer’s conversion prediction in MCI subjects. *NeuroImage*, 104, 398–412.
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59, 433.

Manuscrito recibido el 5 de agosto de 2017. Aceptado el 12 de junio de 2018.