

TÉCNICAS DE PROGRAMACIÓN
«DEEP LEARNING»:
¿Simulacro o realización artificial de la inteligencia?

Juan J. Padial
Universidad de Málaga

Resumen: Se expone y critica el concepto de «aprendizaje máquina», dilucidando la naturaleza, alcance y límites del concepto de aprendizaje aplicado a los algoritmos. La actividad que realizan estos se entiende orteguianamente como «alteración» y al aprendizaje humano se lo entiende necesitado de la facultad de «ensimismamiento». Como la máquina carece de esta facultad, la estupidez es siempre co-programada con el algoritmo.

Palabras clave: Aprendizaje automático, aprendizaje profundo, alteración, ensimismamiento, actividad intelectual.

«Deep learning»: Simulation or artificial realization of intelligence?

Abstract: The concept of «machine learning» is exposed and criticized, elucidating the nature, scope and limits of the concept of learning applied to algorithms. The activity carried out by these machines is understood following Ortega y Gasset as «alteration» and human learning is understood as needing the faculty of «self-absorption». Since the machine lacks this ability, stupidity is always co-programmed with the algorithm.

Keywords: Machine learning, Deep learning, alteration, self-absorption, intellectual activity.

Recibido: 29/01/2019 **Aprobado:** 20/02/2019

Entre las ciencias de la computación y de la inteligencia artificial, algunos de los campos más prometedores son los del *aprendizaje automático* — *Machine Learning*— y el subcampo de las técnicas *Deep learning* —aprendizaje profundo—. Es este aspecto sobre el que hoy quisiera concentrar la atención. Se trata de métodos importantísimos en inteligencia artificial para

procesar datos algorítmicamente de modo que el algoritmo pueda «reconocer» en esos datos determinadas pautas. Éste «reconocimiento» permite reacciones y adaptaciones del ordenador al entorno.

¿Se puede hablar de un aprendizaje máquina, dada esta adaptación y reacción a diversos ambientes? ¿No sucederá a los informáticos, matemáticos e ingenieros lo que le pasó a Ortega y Gasset con la sociología de su tiempo cuando

...acudí lleno de ilusión, desplegadas todas las velas de la esperanza, a los libros de sociología, y me encontré con una cosa increíble, a saber: que los libros de sociología no nos dicen nada claro sobre qué es lo social, sobre qué es la sociedad.¹

¿Nos dicen algo claro los libros de informática y computación sobre lo que es la inteligencia y el aprendizaje, o hemos de buscarlo en los libros de teoría del conocimiento y antropología filosófica? Desde el mismo texto que acabo de citar, Ortega respondería que esas reacciones y adaptaciones del ordenador al entorno podrían considerarse «alteraciones» como las que caracterizan esencialmente el vivir animal. Y que al igual que estos viven siempre alterados, volcados hacia los datos externos, enajenados, y nunca consiguen recogerse en sí mismos, ensimismarse o meditar, tampoco lo conseguirían las máquinas. Estas páginas pretenden ser un acercamiento aclaratorio a estas cuestiones de la matemática y la informática desde la antropología filosófica.

Jamás la humanidad había generado tantos datos como lo hace en la última década. Año tras año la detección, elaboración y procesamiento de datos crece exponencialmente, en cantidad, heterogeneidad y precisión. Respectivamente al conocimiento, estamos, con justicia, en los tiempos de

¹ J. Ortega y Gasset, «El hombre y la gente» en *Obras completas*, 1983, Vol. VII: 80.

los *Big Data*,² del *data mining* y de los *analytics*, y esto es algo que debemos pensar, y no meramente constatar. Analistas, programadores de algoritmos y mineros de datos trabajan estrechamente, porque los trillones de datos que generan y registran —y *alteran* (Ortega) a— los algoritmos incesantemente no son todos ellos relevantes. Los nuevos mineros —matemáticos e ingenieros— han de extraer los datos significativos. Y para que la información no sea superflua el algoritmo ha de ajustarse a la pregunta que realiza el analista, y que guía el proceder del programador. Así, parece que el comportamiento inteligente está fuera del algoritmo: en quiénes diseñan —ensimismados— las preguntas pertinentes, programan algoritmos para generar datos sobre esas preguntas, y buscan los datos adecuados. Es indudable que podemos reconocer procesos de aprendizaje en las interacciones y las meditaciones que se dan entre estos tres tipos de especialistas.

Comparada con las de ellos, la inteligencia y el aprendizaje que observamos en máquinas parece aún débil, por astronómicas que nos parezcan las cifras de los datos generados. Pero esta debilidad lejos de ser una objeción puede ser indicio de una estrategia prometedora, más aún el acercamiento más adecuado al tema de la inteligencia artificial. Así lo pensaron Alan Turing o Marvin Minski.

El primero de ellos, en un famosísimo artículo de 1950 sugiere la estrategia de diseñar algoritmos que puedan simular el proceso humano de aprendizaje desde la infancia:

En vez de tratar de producir un programa similar a la mente adulta, ¿por qué no tratar en su lugar de producir uno que simule la mente de un niño? Si ésta fuera luego sujeta al curso apropiado de educación, uno obtendría el cerebro adulto. Presumiblemente, el cerebro de un niño es algo así como un cuaderno como el

² Los aceleradores de partículas del CERN generan en la actualidad un petabyte (10^{15} bytes) de datos por segundo. Sin las técnicas de Inteligencia Artificial no se podrían analizar estos datos, ni discernir cuáles son relevantes, de la infinidad de ellos que carecen de cualquier valor para determinar la estructura física del universo. Cfr.: <https://plato.stanford.edu/entries/artificial-intelligence/#BlooMachLear>

que se compra en una tienda. Un mecanismo más bien simple, con muchas hojas en blanco. (Mecanismo y escritura son casi sinónimos para nuestro punto de vista.) Nuestra esperanza es que haya tan poco mecanismo en el cerebro del niño, que algo así pueda ser programado fácilmente. Podemos asumir en una primera aproximación que la cantidad de trabajo en la educación sería muy similar a la de un niño humano.

Hemos por tanto dividido nuestro problema en dos partes: el programa-niño y el proceso educativo. Éstos permanecen estrechamente relacionados. No podemos esperar encontrar un buen niño-máquina al primer intento. Uno debe experimentar enseñando a una máquina así y ver cómo de bien aprende. Luego, se puede intentar con otra y ver si es mejor o peor.³

Desde luego, muchas de las tareas que aprende el niño no exigen su autoconciencia, aunque ésta se genere y consolide por mediación de estas actividades. Dominar la percepción en tres dimensiones, lograr manipular objetos en esos espacios, etc., son aprendizajes que afinan la conciencia perceptiva y la acción sobre el mundo. Lo mismo sucede con la conquista de los hábitos motores, verbales y técnicos. Sabemos que estos aprendizajes son los de una mente encarnada y en el mundo. Y tenemos excelentes descripciones psicológicas y fenomenológicas de estos saberes. El objeto de estas páginas es dilucidar si el «programa-niño» del que habla Turing, y que se está consiguiendo implementar en muchos de los algoritmos que usamos actualmente, son también mentes. En este caso el acercamiento débil a la inteligencia artificial habría logrado una inteligencia artificial, paradójicamente, fuerte, puesto que el ordenador así programado, sería él mismo una mente.

El «programa-niño» de Turing sería una mente cuya única forma de saber sería probabilística. Una mente que reacciona y se adapta al entorno al encontrar determinadas pautas y correlaciones. Eso no es encontrar causas

3 A. Turing, «Maquinaria computacional e inteligencia», traducido por Cristóbal Fuentes, 2010, versión digital en <http://xamanek.izt.uam.mx/map/cursos/Turing-Pensar.pdf> p. 19. Traducción levemente modificada.

—tener ciencia, saber apodíctico. El problema es que el algoritmo reacciona automáticamente, sin saber del mundo, ni de sí mismo. En cualquier caso, las bases matemáticas de estos tipos de inferencia estadística se los debemos al reverendo Thomas Bayes, quien en su artículo publicado póstumamente en 1763: «*An Essay towards Solving a Problem in the Doctrine of Chances*» puso las bases lógicas y matemáticas del aprendizaje de la experiencia, es decir, del aprendizaje que podemos obtener a partir de hechos generados, determinados y medidos continuamente. A mayor acopio de hechos y mayor precisión en su determinación, se incrementa nuestra capacidad de determinar la causa de estos. Un suceso está correlacionado con otro. Y a partir de los detectados, podemos inferir la probabilidad de los sucesos con él relacionados. La inferencia bayesiana puede ser entendida como la forma, por antonomasia, de aprendizaje e inducción. A partir de lo objetivamente dado, Bayes averiguó el modo de obtener conclusiones generales sobre los sucesos correlacionados y el grado de asentimiento que podemos prestar a dichas conclusiones.

El problema de los cálculos bayesianos era su enorme complejidad cuando comenzaban a multiplicarse los factores a ser tenidos en cuenta. Y por eso, esta rama de la matemática comenzó a desarrollarse cuando tuvimos ordenadores con una capacidad enorme de cálculo. En 1980 Judea Pearl desarrolló el sistema *Bayesian networks*, que permitía simplificar gráficamente los cálculos bayesianos. Las redes bayesianas permiten así deducir fácilmente variables no observadas, y aprender parámetros y estructuras entre datos. Se llama *Deep learning* a aquel aprendizaje que una máquina puede realizar de una determinada función (x), desde datos que se presentan en niveles «ocultos» o profundos, de la red neuronal. El concepto de «red neuronal» está conectado con el de «aprendizaje profundo» pues una red neuronal simula el procesamiento de las variables que en un momento se le presentan para decidirse por una alternativa u otra. Así, uno decide una ruta u otra de tráfico, una comida u otra, etcétera, en función de datos tan variables como el tráfico, el tiempo disponible, la hora, etcétera. Esos datos son procesados a nivel profundo en las redes neuronales de nuestro cerebro, y

allí se logra una u otra predicción. Los algoritmos «*Deep learning*» intentan simular o formalizar ese procesamiento de información neuronal de los datos, buscando la serie de operaciones lineales y las funciones a través de las que hay que pasar los datos para obtener una predicción escalar única. Por ejemplo, las respuestas «inteligentes» alternativas que *Gmail* propone para nuestros correos electrónicos están basadas en este tipo de cálculos probabilísticos. Aparecen así las técnicas de *Deep learning* como modos de operar con probabilidades inversas como las que motivaron las investigaciones del reverendo Bayes, o con otras técnicas y modelos estadísticos como los Modelos ocultos de Andréi Márkov, o los algoritmos de Baum-Welch, por ejemplo.

Desde entonces se han desarrollado innumerables aplicaciones de sistemas de aprendizaje de inteligencia artificial. La lógica bayesiana subyace, por ejemplo, a las inferencias de los algoritmos de Google para decidir si un determinado correo es *spam* o no. Cada día ven la luz novedosas aplicaciones del cálculo bayesiano al derecho, la medicina, la farmacología, la ingeniería, y en general a todas las ramas del saber y la técnica humana.

Aquí surgen algunas preguntas filosóficas. Las hay muchas. Generalmente las disciplinas implicadas son la filosofía de la ciencia y la ética. Hoy quisiera acercarme desde una disciplina diversa: la antropología filosófica. Sólo desde ella es posible responder a preguntas tan fundamentales como las que se hacen por la inferencia o las repercusiones éticas. Estas otras preguntas a las que quisiera intentar responder son: ¿Es justo calificar de aprendizaje al procesamiento algorítmico de datos, o esto es una mera metáfora? ¿Cuál es la naturaleza del aprendizaje que puede realizar un sistema formal artificial? ¿Qué quiere decir que unos algoritmos aprenden?⁴ Aunque no comparto con Dennett que los estudios sobre inteligencia artificial son las «investigaciones más abstractas sobre la posibilidad de la inteligencia o

⁴ Los defensores de la Inteligencia Artificial fuerte, como Jack Copeland, reducen el busilis del problema de la inteligencia artificial en conjunto al algoritmo

del conocimiento»,⁵ si pienso que el enfoque *top-down* es el correcto y que la reflexión sobre la Inteligencia Artificial puede iluminar en mucho qué signifique el aprendizaje.

Por razones temporales abordaré esta cuestión como un artículo de *Summa*. Y por cuestiones de complejidad del asunto, me centraré en aprendizajes relativamente simples que realizamos animales y seres humanos. Intentaré dilucidar si dichos aprendizajes pueden ser realizados también por máquinas. Entre estos aprendizajes relativamente simples, están el reconocimiento visual, el auditivo y la adquisición del lenguaje (para los seres humanos).

Cuestión I. Sobre si las máquinas pueden aprender

Artículo 1. Decir de una máquina que aprende ¿es una simple metáfora?

Objeciones por las que parece que parece que las máquinas pueden aprender

1. El aprendizaje exige un procesamiento de información. Pero el procesamiento de información no tiene por qué realizarse conscientemente. Procesar información en cuanto información es conocerla, ser al menos consciente de la información. Al procesar información también se puede procesar la energía que transporta dicha información. Eso es lo propio del conocimiento sensible, donde no es posible conocer sin el procesamiento por parte del cerebro de energía eléctrica y química. En las operaciones vitales orgánicas el procesamiento es ante todo de energía con un mínimo de información. Nuestro cerebro también procesa la información de las diferentes variables que influyen en una decisión u otra computando energía eléctrica y/o química. El computador también lo hace así. Y el procesamiento artificial de información puede verse «mejorado» si se lo «entrena» con resultados ya conocidos y correctos. Entonces el sistema ha aprendido

⁵ D. Dennett, «Artificial Intelligence as Philosophy and as Psychology», en *Philosophical Perspectives in Artificial Intelligence*, 1979: 64.

qué *outputs* son más probables en dependencia de qué *inputs* de información se introduzcan. Si la inferencia normal procede de la premisa a la conclusión, la probabilidad inversa procede teniendo en cuenta las correlaciones entre conclusiones y premisas. Y esa es la clave del «aprendizaje máquina», que se diferenciaría únicamente del aprendizaje realizado por el cerebro por el número de centros —neuronas o parámetros del algoritmo respectivamente— que intervienen en el procesamiento de los datos.

2. Según Roger Penrose, la Inteligencia Artificial busca la «imitación por medio de máquinas, normalmente electrónicas, de tantas actividades mentales como sea posible, y quizá llegar a mejorar las capacidades humanas en estos aspectos».⁶ Si se imita una actividad mental entonces se la estaría realizando, y la imitación de un aprendizaje sería indiscernible de un aprendizaje real. Por eso Marvin Minsky pudo definir la inteligencia artificial como «la realización de sistemas informáticos con un comportamiento que en el ser humano calificamos como inteligente».⁷ Además, el aprendizaje está esencialmente relacionado con la mejora o crecimiento. La adquisición de *nuevos* conocimientos es un tipo de aprendizaje. Pero también es una forma de aprendizaje la mejora en una técnica o en una tarea. Los sistemas de aprendizaje máquina mejoran su realización de las tareas para las que están programados, o con la reiteración de la tarea, o con ejemplos de una realización arquetípica de dicha tarea.

3. Según los psicólogos de la *Gestalt*, se denomina forma gestáltica a aquella configuración o estructuración que hace reconocible y significativa un conjunto informe y caótico de sensaciones. Sin esa forma sólo tendríamos datos sensibles, pero nunca percepciones. Lo mismo sucede con las técnicas de aprendizaje de los algoritmos. Éstos han de trabajar con una infinidad creciente y heterogénea de datos, que han de categorizar. Así, llamaríamos propiamente visión computerizada a la que fuese capaz de encontrar entre

⁶ R. Penrose, *La nueva mente del emperador*, Mondadori, Madrid 1991: 33.

⁷ https://encyclopaedia.herdereditorial.com/wiki/Inteligencia_artificial

millones de rostros fotografiados aquellos elementos que permiten identificarlos y clasificarlos, es decir de encontrar formas gestálticas en ellos. Lo mismo sucede si el algoritmo ha de reconocer los patrones de compra de un usuario y detectar si hay un uso fraudulento de su tarjeta bancaria. Lo mismo ha de ser realizado en los programas de reconocimiento auditivo y en los asistentes por voz. La máquina debe discernir configuraciones y estructuraciones en los tonos de voz y en la secuencia de fonemas emitidos. Estos tonos, y las secuencias novedosas entre ellas son las que deben ser propiamente aprendidas.

4. Entre los seres humanos una de las primeras formas de aprendizaje, y de las más importantes, es la adquisición del lenguaje. En ella se aprende a conectar un objeto o evento con un término, que adquiere su significación en las relaciones (sintagmáticas y paradigmáticas, sincrónicas y diacrónicas) que tiene con el resto de los términos significativos. Parece pues que la significación depende de la sintaxis. Ésta es la que fundamentaría la semántica y el uso. Pero los algoritmos son estructuras formales, y por ello sintácticas. Luego la conexión de un dato con una categoría significativa sería cuestión sintáctica. Además, que los algoritmos sean capaces de establecer relaciones con un número infinitamente superior de otras estructuras sintácticas que el que puede realizar la mente humana, permitiría acceder al universo de la significación. Esta es la tradición filosófica de la *mathesis* como origen de la inteligencia.

En cambio está lo que dice Douglas Hofstadter «la analogía es más que un pequeño problema; más bien es el azul que llena todo el cielo de la cognición; la analogía es todo, o casi, en mi opinión»⁸

Solución: Entre los animales parece que el aprendizaje requiere de la conciencia. No obstante, en ningún otro periodo aprenden más los seres vivos que en la infancia, cuando la autoconciencia aún no ha emergido. Conforme ésta va emergiendo, la criatura humana va aprendiendo infinidad de cosas que requieren de su autoconciencia, como el reconocimiento de su rostro,

⁸ Hofstadter, 2001: 499.

de su cuerpo en general, de su identidad al ser interpelado, etcétera. Pero hay muchos otros aprendizajes que previa o simultáneamente han de ser realizados por el niño y que no requieren de la autoconciencia. Basta con la conciencia. Así ha de aprender a enfocar sus ojos con corrección, a discriminar objetos visualmente, a erguirse sobre sus extremidades inferiores, etc. Estos aprendizajes serían imposibles sin la percepción visual, y sin la sensibilidad propioceptiva, cinestésica, cenestésica, etcétera. Estos aprendizajes dan lugar a hábitos biopsicológicos, entre los cuales, los corporales son los más claros. Ellos permiten la toma de posesión del propio cuerpo, y por ello son condiciones de posibilidad de la emergencia de la autoconciencia.

La adquisición de hábitos corporales, verbales o técnicos es además condición de la emergencia de la conciencia y la autoconciencia. En los primeros años de vida, el individuo va logrando un saber de sus sentimientos y tendencias que se consolida en diversos modos, característicos e idiosincráticos, de «saber-hacer o saber-comportarse» respecto a esos impulsos y sentimientos. Esto son los diversos hábitos corporales y técnicos, que por lo tanto tienen un papel individualizante de ese individuo. Constituyen así sus costumbres y su carácter. Y el individuo aprende también a reconocerse en ese modo peculiarísimo de ser. Tanto la adquisición de los hábitos como su reconocimiento como individualizador de la subjetividad, son saberes inobjetivos; por un lado, porque se aprenden inconscientemente; y por otro, porque el conocimiento de la subjetividad no puede ser objetivo.

El problema entonces radica en que el aprendizaje humano y animal es o bien objetivo y principiado por la inteligencia, o bien inconsciente mediante la adquisición de hábitos y costumbres. Así pues, para que el aprendizaje máquina fuese algo más que una metáfora sumamente atractiva, las máquinas deberían ser susceptibles de adquirir hábitos o tener autoconciencia intelectual. Lo segundo desde luego no lo han alcanzado las máquinas, tampoco las que manejan algoritmos bayesianos para sus procesos de aprendizaje. Y lo primero, que las máquinas sean susceptibles de hábitos requeriría que tuviesen un cuerpo del que tomar posesión, y un inconsciente que mediara entre ese cuerpo y una autoconciencia intelectual

o reflexiva, y que pudiera comparecer en esta última. Pero dado que no hay ninguna máquina que haya conseguido autoconciencia intelectual, ningún aprendizaje máquina puede realizarse inconscientemente. Es decir, dado que es falsa la primera condición, no puede darse la segunda, puesto que aquella es requisito necesario para la adquisición de hábitos. El inconsciente precisa de la conciencia reflexiva.

Si esto es así, estaríamos muy lejos de poder afirmar que una máquina pueda aprender. Al menos que pueda aprender como lo hacemos los seres humanos. Pero quizá haya otros aprendizajes que no requieran ni de la conciencia ni de la autoconciencia, y sin los cuales la vida no sería posible. Por ejemplo, cada vez hay mayor evidencia de lo que los inmunólogos denominan *trained immunity*, que se adquiere mediante un *aprendizaje innato*. Así la respuesta de las células innatas del sistema inmune es alterada por vacunaciones o por infecciones previas. El sistema inmune puede adaptarse muy específicamente a un determinado tipo de invasor recurrente, protegiendo así al organismo. Por eso los inmunólogos hablan de una *respuesta innata mejorada* —*improved*—. Pero guardar memoria de un patógeno que invade el organismo, para que en invasiones sucesivas sea combatido y no logre presentarse por ejemplo una inflamación anormal de tejidos o una parálisis del sistema inmune, no significa haber aprendido. Los linfocitos T y B y algunas células del sistema inmune innato pueden archivar datos de la estructura del ARN de los patógenos encontrados y de la respuesta inmunológica adecuada a ellos.⁹ Pero la memoria inmunológica no exige un aprendizaje en el sentido estricto de la palabra, sino una reprogramación epigenética del sistema inmunológico que mejora la respuesta tras una primera infección. Un sistema inmune reprogramado responde antes y mejor al patógeno. Pero eso no significa ni que sepa de él ni que tome posesión de él haciéndolo propio. Y estas son las dos formas propias del aprendizaje. Si ahora trasladamos al campo de la informática el tema de la memoria, se observa la distancia que

⁹ Netea, *et alia*, 2016.

separa a máquinas como *Deep Blue* de *Alpha Go*. La primera tenía almacenadas en su disco duro millones de jugadas. No era un sistema informático con capacidad de aprendizaje y reconocimiento de pautas y estrategias «significativas» en su propia «experiencia».

Respuesta a las objeciones:

1. Procesar energía no implica aprendizaje, por lo mismo que una condición necesaria de algo no implica que dada esa condición entonces se produzca ese algo. Una condición necesaria no tiene por qué ser suficiente. Y que todo aprendizaje requiera procesamiento de información y energía no implica que esa condición sea suficiente para que un sistema pueda aprender.

2. La clave del aprendizaje máquina estriba en la reducción de incertidumbre, o lo que es igual, en el cálculo de la mayor probabilidad. Evidentemente hay factores que limitan el aprendizaje y las opciones entre los seres humanos. Por ejemplo, los prejuicios o lo que resulta evidente y natural en un determinado sistema sociocultural una vez que el individuo ha quedado inculturizado gracias a los procesos de socialización primaria y secundaria. Pero los seres humanos saltan habitualmente por encima de estas constricciones gracias a su reflexión, que erosiona la fuerza de las normas imperantes en su cultura, y a nuevas oleadas de creación e invención, por lo que el aprendizaje real es y puede ser irrestricto. Esto introduce una diferencia esencial entre el «aprendizaje» automático y el humano, puesto que los métodos para procesar los datos son definidos por el programador del algoritmo. El algoritmo puede crear una base de datos en que se relacionen la probabilidad de una consecuencia exitosa si se dan una serie de variables. Ese fue el proceder de *Alphago Zero*, que previamente a jugar con un ser humano, jugó millones de partidas consigo mismo, a fin registrar qué jugadas ganadoras eran más probables dadas unas determinadas posiciones. Pero lo que nunca podría realizar el algoritmo es modificar la estructura de parámetros y métodos de cálculo de su propio algoritmo. Es decir, tomar posesión de su propia estructura, que es precisamente la clave de los aprendizajes humanos más básicos, como la creación de hábitos corporales,

verbales y técnicos, como se ha visto en la solución a este artículo. Una cosa es que el algoritmo encuentre pautas en un conjunto gigantesco de datos, y pautas que no estaban especificadas previamente en el algoritmo, y otra que modifiquen su búsqueda y reconocimiento de pautas.

3. La inteligencia humana puede encontrar lo que en sí y de suyo es significativo. Hay datos que son significativos por sí mismos y no en función de una programación antecedente. Las técnicas «Deep learning» procesan un número ingente de datos, aunque los parámetros de los algoritmos son muy limitados. Así, por ejemplo, algoritmos, como los que controlan las pulseras de actividad física, continuamente monitorizan nuestras constantes vitales, el medio ambiente en que vivimos y toda nuestra actividad consciente e inconsciente (horas de sueño, etcétera). Y pueden «aconsejar» a los diversos usuarios sobre su necesidad de mayor o menor actividad física teniendo en cuenta un número muy reducido de parámetros y sobre todo los *outputs* que realiza el usuario en función de las variables dadas. La inteligencia no está preprogramada, ni enseñar es programar a reconocer patrones. Entender es siempre entender universalmente, y no reconocer lo que es significativo para el algoritmo.

4. A la cuarta objeción respondo diciendo que la significatividad de los términos, y por lo tanto el carácter significativo de objetos y eventos, depende ante todo del universo de discurso o contexto. Por lo tanto, y como Wittgenstein señaló, la pragmática tiene la primacía en el aprendizaje.

Artículo 2. ¿El aprendizaje es programable? ¿Y la estupidez?

Objeciones por las que parece que el aprendizaje es programable y la estupidez no lo es.

1. El aprendizaje animal se da en el marco de una conexión indefectible entre la información percibida y el comportamiento posible a partir de dicha percepción. Este ciclo está cerrado (por ello el comportamiento es instintivo), pero dotado en los animales más evolucionados de cierto grado de

plasticidad, por lo que se desarrolla y consolida merced al aprendizaje (fenómeno del *imprinting*). Pues bien, las técnicas de aprendizaje en máquinas trabajan sobre este esquema de conexión entre información suministrada y conducta. Tratan así de desarrollar algoritmos que a partir de una información generalicen un tipo de comportamiento que no estaba preprogramado. Así aprenden al realizar una inducción a partir de los datos suministrados. Consiguientemente desarrollan una programación previa de análisis y consolidan una conducta subsiguiente, y por eso simulan el aprendizaje de un animal.

2. Si los animales pueden aprender, entonces las máquinas también, porque el ciclo funcional de las vivencias entre los animales no se diferenciaría de los módulos necesarios para que haya mente, según la teoría modular de la mente de Jerry Fodor. Y es que animales y máquinas comparten un sistema de entradas de información y un sistema de procesamiento de dicha información, a la que se deben los procesos cognitivos. Según el mismo Fodor, toda actividad mental es reductible a actividad computante, luego el aprendizaje también lo sería.

3. El círculo funcional de la vivencia humana, según Üexkull, se distingue de los círculos animales, pues está roto en tres puntos, lo que impide esa equivalencia entre sentido y conducta. Lo que es percibido como significativo puede ser algo novedoso. A su vez, la conducta a realizar no está preprogramada filogenéticamente en la mayor parte de los casos, sino que ha de ser inventada. Llamamos aprendizaje exitoso a aquel que es capaz de encontrar un nuevo dato significativo, de encontrar un nuevo modo de resolver un problema, o de resolver otra serie de problemas, con otros planteamientos y con otros datos. Pero parece que eso es lo que intentan las técnicas de *machine learning*, encontrar nuevos patrones significativos y realizar predicciones novedosas, sin estar programados explícitamente para ello. Así pueden realizar predicciones, por ejemplo, sobre la demanda de un producto, teniendo en cuenta un conjunto de millones de datos. Los algoritmos por ejemplo de *Amazon* pueden realizar predicciones sobre millones de productos con 18 meses de antelación.

En cambio está lo que ha declarado Gurdeep Singh Pall, vicepresidente corporativo de Microsoft y de su rama sobre Inteligencia Artificial aplicada a los negocios, quien calificó hace poco a los actuales sistemas de Inteligencia Artificial como «*idiot savants*»¹⁰, idiotas con erudición o instruidos.

Solución: La evolución de los vivientes orgánicos, a pesar de tener una estructura arbórea y no lineal, si la consideramos desde el punto de vista genético, es la historia de sucesivas programaciones filogenéticas que permiten a los seres vivos más actividades diferenciadas. Nadar, volar, ver, tener una conducta maternal, implican siempre una evolución de las capacidades biopsicológicas, y no meramente de las cognitivas, pues éstas últimas se encuentran siempre en correlación, con un mayor grado de encefalización, de complejidad del sistema endocrino, del locomotor, del periodo de gestación, etcétera. Hay programaciones filogenéticas de un grado tal de complejidad que permiten al animal cierto grado de aprendizaje. Y sin ese grado, el animal no puede consolidar la tendencia instintiva. Así a los grandes predadores no les basta el impulso instintivo a la caza, ni la percepción de una pieza como comestible o peligrosa. Les es necesario *aprender* a cazar. Y eso implica que el animal no es sólo cazador, sino sagaz.¹¹ El aprendizaje pues enlaza con la programación filogenética afinándola a un nicho ecológico donde las variables (del ambiente que han de interiorizar, de la complejidad biopsicológica que han de coordinar, y de las respuestas conductuales que pueden realizar) tienen una complejidad mucho mayor que las que tienen animales con un ciclo funcional de vivencias inferior. Así pues, el aprendizaje es necesario para estos animales, entre los que también nos encontramos nosotros, como un mecanismo extrauterino que complementa la adaptación filogenética heredada.

Así pues, en los seres vivos se dan dos tipos de aprendizaje: el que tienen como especie o filogenético, y el que cada individuo alcanza a lo largo de su desarrollo vital. Así hay un aprendizaje en cuanto especie y otro en cuanto

¹⁰ Citado por *The Economist*: «GrAI expectations Artificial intelligence in business»

¹¹ Cfr.: Choza, 1988: 72.

individuo. El aprendizaje en cuanto especie es el resultado de una larguísima evolución de las capacidades biopsicológicas y da lugar a la preprogramación filogenética de cada especie. El aprendizaje individual tiene como centro el subsistema cognoscitivo y no el genético; y como órgano el cerebro, y no la dotación genética heredada. Pues bien, en las máquinas podemos encontrar un correlato del aprendizaje en cuanto especie. En efecto, gracias a las técnicas de Inteligencia Artificial de aprendizaje, las máquinas logran adaptar su programación a la resolución de problemas para las que no estaban inicialmente pre-programadas. Esto no quiere decir que la resolución del problema novedoso desplace o anule el programa inicial, pues la máquina perdería entonces su esencia. Lo que sucede es que adapta el programa inicial al ambiente específico en el que está operando. En este sentido el programa inicial debe gozar de cierta flexibilidad o plasticidad, que permita su adaptación. Esto se puede apreciar en la programación de algoritmos para atención al cliente. Estos algoritmos son alimentados por transcripciones de interacciones modélicas entre usuarios y clientes. El acopio sucesivo de más y más datos afina las soluciones del algoritmo, en el sentido de que son más variadas y adaptadas a la complejidad real de las demandas. El programa se afina, pero no es desplazado o anulado por el acopio de datos.

Pues bien, esto es lo que *no* sucede en el aprendizaje humano. «El intelecto humano tiene, desde el primer momento en que aparece sobre el planeta, una capacidad reflexiva tal, una capacidad de aprendizaje tal, que anula o desplaza desde ese momento la programación filogenética, es decir, el instinto, la determinación en concreto de los fines y de los medios básicos para alcanzarlos».¹² Y esto implica que entre los seres humanos, el aprendizaje individual es el factor esencial para la resolución de problemas, para la determinación de fines, y de los medios que conducen a ellos. Por eso llamamos pasmado o estupefacto al ser humano que tiene la actividad intelectual como en suspenso. Y por ello mismo podemos llamar estúpidos

¹² Choza, 1988: 38-39.

a los sistemas de Inteligencia artificial, pues ninguna de las actividades que realizan tiene para ellos significado alguno. Por esa razón Robin Wigglesworth cita a la agencia de inversión en bolsa Gavekal cuando muestra que «en ocasiones lo que hace cualquier ordenador es reemplazar la estupidez humana por la estupidez máquina. Y, gracias a la velocidad y a la convicción preprogramada, la estupidez máquina puede devorar los mercados más rápido de lo que podría hacerlo ningún pánico humano».¹³ De lo que resulta que en ausencia de un yo que sepa el acto intelectual que ejerce, la estupidez máquina es algo que resulta imposible de despejar. El algoritmo nunca logra despejar su estupidez y su pasmo.

Respuesta a las objeciones:

1. La conexión entre comportamiento y datos detectados es una cuestión de cálculo (estadística computacional) y por lo tanto de actividad computante. No se trata de un tipo de actividad en el que la actividad representativa juegue un papel esencial. El director de aprendizaje máquina de Amazon, Ralf Herbrich, indica la dificultad de programar algoritmos para predecir la demanda de ropa y su distribución en diferentes centros logísticos por tallas y colores. Pero el análisis necesario para realizar estas tareas no excede del cálculo estadístico de las compras realizadas por millones de usuarios, de sus preferencias de colores y de sus tallas, y la topología necesaria para su distribución. Esto puede facilitar la tarea logística, pero sería estúpido —y ruinoso— dar completa autonomía a ese programa para realizar los pedidos de ropa y gestionar la logística. Que hayan encontrado un patrón no significa que la moda no sea un fenómeno histórico siempre abierto, y por lo tanto generador de novedades.

El aprendizaje máquina está seriamente limitado por la programación del algoritmo. Se trata de algo obvio, pero sobre lo que hay que reflexionar. Son conocidos los escándalos recientes, cuando *Google Images* clasificó a dos amigos de raza negra como gorilas. Cuando Tito Livio enseñó que todo error

¹³ Wigglesworth, 2016.

humano merece indulgencia —*Venia dignus est error humanus*—,¹⁴ lo hizo porque ese ser humano que erraba podía saber de su error, y con ello de su finitud y culpabilidad. Perdonamos a ese ser humano por su error. Pero no podemos hacer lo mismo con la máquina. Detrás del acto intelectual humano encontramos un yo al que se le imputa el acto. No es así en la máquina.

2. La teoría de Fodor es representativa de lo que se ha denominado el Dogma Central de la Inteligencia Artificial: que toda actividad cerebral es sólo actividad computante. Searle ha calificado la teoría de Fodor como «la versión más extrema [a la] que yo llamo «inteligencia artificial fuerte» o sólo «IA fuerte», el cerebro es un computador digital y la mente es sólo un programa de computador».¹⁵ Autores como Rodney Brooks hablan de *Inteligencia sin representación*.¹⁶ Pero como Minsky y Papert expusieron, una máquina no puede aprender la función XOR o disyunción exclusiva, si no es mediante una representación adecuada.¹⁷ Ya en este nivel se aprecia que la actividad representativa se diferencia de la computacional. Además, la actividad intelectual permite comprender, y es que usar y manejar símbolos de acuerdo con un programa, no significa comprender los símbolos manipulados, como explicó Searle en su celeberrimo experimento de la habitación china.

En el aprendizaje humano de una técnica no se disocia el «hacer» del «saber». Por eso una técnica se puede describir como un «saber-hacer». Estos dos elementos se disocian en el aprendizaje máquina. Así parece que el escaneo masivo de libros que está realizando el proyecto *Google Books* no está dirigido en última instancia a la lectura y comprensión humana de los mismos, sino a la lectura de estos por algoritmos de inteligencia artificial. Estos

¹⁴ Tito Livio, *Historias*, VIII, 35

¹⁵ Searle, 1985: 50.

¹⁶ Brooks.

¹⁷ Minsky, Papert, 1972.

pueden aprender no sólo a reconocerlos ópticamente, sino las técnicas paleográficas de lectura de textos antiguos, y a extraer, descontextualizada, información de aquellos documentos. Aquí falta el elemento de «ensimismamiento» del que hablaba Ortega. El programa «ha aprendido», lee, pero no sabe lo que lee, ni interioriza esa información. Así se pierde «el atributo más esencial» del aprendizaje:

...la posibilidad de meditar, de recogerse dentro de sí mismo para ponerse consigo mismo de acuerdo y precisarse qué es lo que cree; lo que de verdad estima y lo que de verdad detesta. La alteración le obnubila, le ciega, le obliga a actuar mecánicamente en un frenético sonambulismo.¹⁸

3. Reconocer una estructura lógica no es lo mismo que interpretar algo en el mundo como significativo. Como señaló Apel, la semiótica no es la hermenéutica. La semiótica es ahistórica, en cambio la hermenéutica da razón de un intelecto como el humano que está continuamente creando nuevos sentidos, y que sólo interpreta cualquier fenómeno real, o incluso irreal, contando con los procesos históricos abiertos.

Se podría pensar que el *tiempo* de aprendizaje o de entrenamiento es similar al necesitan los seres humanos. El aprendizaje humano es posible en función del desarrollo y la maduración de una serie de factores biológicos y psicológicos: el de su cerebro, la estabilidad emocional en la infancia, y todas las fases de desarrollo cognitivo determinadas por psicólogos como Piaget o Kohlberg. Es decir, el desarrollo cognitivo se realiza siempre en un ámbito de relaciones intersubjetivas abierto y orientado al ámbito de máxima significatividad. Nunca es un aprendizaje en soledad. Sin el concurso de otras subjetividades ningún programa ni ninguna fase del desarrollo cognitivo podría ejecutarse.

¹⁸ Ortega y Gasset, 1983: 83.

Juan J. Padiál

Bibliografía empleada

- R.-A. Brooks, «Intelligence without representation» en <https://people.csail.mit.edu/brooks/papers/representation.pdf>
- J. Choza, *Manual de antropología filosófica*, Madrid, Rialp, 1988.
- D. Dennett, «Artificial Intelligence as Philosophy and as Psychology», en *Philosophical Perspectives in Artificial Intelligence*, 1979.
- D. Hofstadter «Analogy as the core of cognition» en Dedre Gentner *et alia*, *The analogical mind: Perspectives from cognitive science*, 2001.
- M. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry*, The MIT Press, Cambridge MA, 1972 (2ª ed.; 1ª: 1969).
- M.G. Netea, *et alia* «Trained immunity: A program of innate immune memory in health and disease», *Science* vol. 352, nº 6284, Abril 2016.
- J. Ortega y Gasset, «El hombre y la gente» en *Obras completas*, Madrid, Alianza, 1983, Vol. VII.
- R. Penrose, *La nueva mente del emperador*, Madrid, Mondadori, 1991.
- J. Searle, *Mentes, cerebros y ciencia*, Madrid, Cátedra, 1985.
- A. Turing, «Maquinaria computacional e inteligencia», traducido por Cristóbal Fuentes, 2010, versión digital en <http://xamanek.izt.uam.mx/map/cursos/Turing-Pensar.pdf> p. 19.
- R. Wigglesworth, «Trading algorithms bring benefits but fears of accidents grow» en *Financial Times*, 1 de junio de 2016. Versión online: <https://www.ft.com/content/87abe568-00c3-11e6-99cb-83242733f755>

Juan J. Padiál
jjpadiál@uma.es