

# Inteligencia humana y máquinas inteligentes

Manuel Béjar

Cátedra Ciencia, Tecnología y Religión  
Universidad Pontificia Comillas (Madrid)  
Email: mbejar@recuerdo.net

ciencia

Cuando se lee y se oye acerca de la denominada inteligencia artificial (IA) dos son las posiciones más enconadas. La más conservadora rápidamente concluye que no existen las máquinas inteligentes. Esta corriente tiene una manera tan exclusiva de entender la inteligencia humana como una sustancia espiritual, que las ideas acerca de la IA son consideradas casi como de ciencia ficción. Desde esta perspectiva ya puede la IA multiplicar por diez su potencial en las próximas décadas; pues con fina ironía se matiza que diez por cero seguirá ofreciéndonos un resultado nulo pasados los años. En el extremo opuesto se hallan quienes irremediamente ven el devenir del orbe dominado por una nueva inteligencia no biológica de mayor potencia que la humana y con una vertiginosa capacidad para evolucionar, mejorarse y adaptarse a un entorno fuertemente dinámico que ella misma va modelando. Este es el escenario de una IA dominadora, donde lo humano

pronto queda muy desfasado y se prefiere proyectar el triunfo del posthumanismo.

La mayoría se posicionaría entre estas dos visiones tan diferentes del futuro de la IA y de la naturaleza de la inteligencia humana. Ninguna de las dos corrientes tan radicales puede considerarse un posicionamiento propiamente científico. Ni la más conservadora, por integrar en lo humano la idea de un principio no material de la inteligencia; ni la más futurista –que no futurible–, por atreverse a predecir sin conocer qué es la inteligencia humana y aventurarse a hablar sobre el futuro de una realidad sin leyes que avalen sus predicciones. ¿Qué nos queda entonces? Sin duda una ímproba tarea que pasa necesariamente por conocer los avances de la IA y, muy especialmente, por investigar científica y filosóficamente la naturaleza de la inteligencia humana.

### 1. Inteligencia artificial

Mucho se ha escrito sobre el procesamiento computarizado de la información desde la construcción de los primeros ordenadores hace más de 70 años y muy especialmente después de la célebre conferencia de Dartmouth cuando se empezó a hablar de IA. En estos orígenes de la información automatizada se pretendía construir máquinas de cómputo, capaces de simular el comportamiento humano. Nos referimos a la IA clásica, cuya filosofía era hallar semejanzas y puntos de encuentro entre máquinas y humanos. El mismo Marvin Minsky entendía la IA como el arte de conseguir que las máquinas resuelvan problemas al estilo humano.

Ya desde finales del siglo pasado estas ideas clásicas dejaron espacio para el inicio de una frontera en la investigación sobre IA. Surgió con fuerza la idea de que las máquinas bien podrían tener una inteligencia con estilo propio. Así como las abejas, los perros, los delfines tienen una inteligencia propia, las nuevas investigaciones buscan desarrollar una inteligencia de suyo artificial. Más concretamente se pretende dar con los sistemas de procesamiento de la información que permitan la completa autonomía de una máquina para pensar por sí misma y sobrevivir inteligentemente en un entorno.

Pero, ¿qué entendemos por una máquina inteligente?

Los partidarios de la IA consideran que es imprescindible una nueva concepción de inteligencia, mucho menos antropocéntrica que de la IA clásica. Por ello, generalizan el concepto clásico de IA y atienden también otras propiedades más características de las máquinas: el procesamiento distribuido de la información, la autonomía, el acoplamiento entre los sistemas sensoriales y motrices, la capacidad de adaptación al entorno, así como a otras formas de interacción y evolución social. Se trata de una nueva aproximación *bottom-up*. Esto es, partiendo de la reunión de diferentes tecnologías va conformándose una interacción de bloques de inteligencia con capacidad de evolucionar, mejorarse y aprender.

Hoy domina la pretensión de desligarse definitivamente de la idea clásica de que las máquinas necesitan continuamente alimentarse de las bases de datos introducidas por los humanos. Ahora la IA funciona más autónomamente como agentes propios de la información, a través de su capacidad para recabarla del entorno con su sistema sensorial, organizarla en su memoria y analizarla con toda su potencia de procesamiento. Es decir, los actuales sistemas computacionales disponen de una

precisa y actualizada imagen de la información del entorno que la máquina considera relevante. De este modo la moderna IA se entiende mejor a sí misma como una variedad de sistemas de aprovisionamiento y procesamiento de la información, que conjuntamente permiten la emergencia de un comportamiento complejo y la supervivencia autónoma de la máquina en su entorno<sup>1</sup>.

¿Qué hay de todo esto en nuestra sociedad? La IA es hoy una realidad dominante en los proyectos de minería de datos y análisis de los movimientos financieros. La IA puede extraer patrones de conducta entre una enorme cantidad de datos de pagos con tarjeta en un conjunto de comercios de una compañía distribuida por todo el mundo. Los datos están ahí, pero sin IA la empresa no podría adquirir mayor conocimiento para mejorar sus servicios al consumidor, conseguir un aumento de las ventas y hacer más eficiente la distribución de sus productos.

Igualmente podríamos referirnos también a la transformación del modo de ejecución de operaciones bursátiles. Cada vez es más frecuente que sean las máquinas quienes tomen la decisión de ejecutar acciones de compraventa en

el mercado de valores. Un *softbot* (compuesto de *software* y *robot*) compra aquí lo que otro *softbot* vende en otro confín. Sin duda estas implementaciones de la IA son consecuencia de que las máquinas pueden hacer mejor que los humanos muchas cosas. Ahora bien, ya sabíamos que a diferencia de los humanos las máquinas podían volar. ¿Son más inteligentes por volar, analizar más rápidamente las oportunidades del mercado o ver correlaciones entre montañas de datos, que sencillamente son transparentes al entendimiento humano?

## 2. Inteligencia humana

El comportamiento humano no puede explicarse solo en términos puramente racionales. Están las emociones, diríamos. Aunque quizás entendiéramos mejor la determinación de la inteligencia humana a partir de sus propias limitaciones. Los límites de la razón condicionan nuestro comportamiento: no podemos generar antigravedad, y nos caemos; manejamos pobremente la estadística, y nos equivocamos; no alcanzamos a ver más allá de la incertidumbre, y nos fiamos de nuestras emociones totalmente inconscientes de la información escondida entre tanto *big data*. Solo tenemos claro que nuestra naturaleza no hace de

---

<sup>1</sup> Cfr. K. WARWICK, *Artificial Intelligence. The Basics*, Routledge, Oxon 2012.

nosotros los mejores agentes para optimizar nuestras acciones según los patrones de información que se siguen tras un minucioso análisis de datos. No somos así; porque la inteligencia humana, forjada de fracaso en fracaso durante cientos de miles de años de evolución natural, está preparada para elegir en la vida. Elegir bien; pero no lo mejor, ni lo óptimo.

La inteligencia humana ha sido evolutivamente diseñada para tomar buenas decisiones en el quehacer de la vida ordinaria. Sabemos qué comer, cómo vestarnos, con quién juntarnos, a quienes debemos amar, enseñar y proteger; por qué merece la pena el esfuerzo y por quién doblegamos nuestro interés más allá del propio querer. Inventamos tecnología, producimos conocimientos, participamos de la ciencia del mundo. Y nos reconocemos trascendentes al comprobar en carne nuestras limitaciones. Todo esto está cristalizado en nuestra inteligencia humana. Sabemos cómo hacerlo. Es verdad que no somos iguales y lo que más nos diferencia es ese modo personal de enfrentarnos a lo cambiante: a sus amenazas y sus oportunidades. Cuando existen múltiples alternativas inmersas en nubes de incertidumbre, esta inteligencia cristalizada por la experiencia ordinaria ya no es tan competente y cada persona depende crucialmente de su inte-

ligencia líquida para conseguir la mayor ventaja<sup>2</sup>.

Mientras maduramos y envejecemos vamos acumulando conocimiento práctico y teórico del mundo. Es la inteligencia cristalizada por la experiencia de una vida. Nuestros lóbulos temporales almacenan información práctica para operar ante buena parte de las circunstancias que se presentan en la vida. Es notable cómo las personas con daños cerebrales en estas regiones del cerebro pierden la competencia para actuar con normalidad. Siguen contando con competencias vitales básicas, pero son incapaces de concatenarlas para ejecutar con éxito algunas operaciones. En esta tarea de coordinación los lóbulos prefrontales son imprescindibles. Son los directores de orquesta y sin su buen funcionamiento se truncan las operaciones y se pierde la perspectiva para finalizar una empresa con éxito. Las personas con daños cerebrales en los prefrontales pueden saber todo lo que hay que hacer para conseguir algo, pero nunca consiguen que ocurra por falta de estrategia. Es como si carecieran de la visión de conjunto y se pierden por las ramas del bosque de nuestra vida cotidiana.

---

<sup>2</sup> Cfr. J. DUNCAN, *How Intelligence Happens*, Yale University Press, New Haven 2010.

Sabemos que morfológicamente no son los lóbulos frontales humanos sino los parietales los más diferentes de los simios. Ahora bien, los prefrontales humanos tienen mayor cantidad de sustancia blanca. Este hecho favorece la rápida formación de redes neurales y mejora notablemente la adaptabilidad humana para coordinar acciones ante las necesidades exigidas por el entorno. En el córtex prefrontal y parietal se activan dinámicamente conexiones que funcionan como soporte neural para la ejecución de una acción, mientras se coordina de un modo más silente cuándo y cómo se continuará la operación. En el momento preciso se forma un nuevo patrón de elevada intensidad que ejecuta una nueva acción. Al mismo tiempo van apareciendo dinámicas sucesiones de patrones neurales que trabajan para encontrar un comportamiento coherente sostenido. La baja correlación entre estos tipos de conexiones neuronales permite su coexistencia en el cerebro sin interferencias.

Este elevado dinamismo cerebral es la inteligencia líquida. La inteligencia competente ante lo novedoso o lo incierto va perdiéndose y emergen disposiciones neurales de conducta. Cuando envejecemos contamos con una inteligencia cristalizada mayor, que nos capacita para decidir bien en la

vida. Pero, incluso de jóvenes también tenemos experiencia de cómo se debilita la inteligencia líquida cuando nos sentimos cansados, distraídos o confusos ante un problema desconocido. Decimos entonces que no razonamos bien o que sabemos poder hacer algo mejor. Reconocemos la debilidad de nuestra razón y postergamos una decisión a la espera de una restauración de la inteligencia líquida. Curiosamente no basta la razón para decidir. Es necesaria la fe en la razón. Y al decidir tenemos fe en poder llevar algo a cabo. Esta es la esencia de la inteligencia humana: decidirse por algo bueno y comprometerse en que salga adelante, especialmente en situaciones de gran incertidumbre, cuando tiembla la fe en la inteligencia cristalizada. En su lugar, la máquina inteligente simplemente se decantaría por lo óptimo.

### 3. Aprendizaje en máquinas inteligentes

Algo de todo esto debe de conocer el equipo de Google que ha conseguido enseñar a una máquina a comunicarse con otra haciendo todo lo posible por encriptar el mensaje y mantenerlo a salvo de una tercera que trata de descifrarlo. Lo más relevante es que lo han hecho

sin decirle cómo<sup>3</sup>. Es decir, al más puro estilo de la evolución natural. Se han dado cuenta de que los adversarios desempeñan un papel que puede ser clave en el diseño y adaptación de grandes redes de procesamiento artificial de la información. Las redes aprenden automáticamente a encriptar optimizando la transmisión de información; esto es, minimizando el ruido y maximizando la señal sin que sea descifrada por un sistema ajeno.

Este experimento es uno más entre los retos ya conseguidos por múltiples equipos dedicados al *deep learning* de las máquinas inteligentes. Las grandes multinacionales del sector tiene su propio proyecto para que las máquinas aprendan: Watson (IBM), Azure (Microsoft), TensorFlow (Google), Amazon Machine Learning... Entonces, ¿de veras aprenden, como nosotros? ¿Son inteligentes? ¿Nos superarán?

No es cuestión de si somos más o menos inteligentes que las máquinas. Es que nuestra inteligencia cristaliza para ofrecernos cuanto está preparada para hacer: vivir humanamente. ¿Podrá una IA

llevar una vida más humana que la nuestra? Definitivamente estamos intentando contrastar dos incomparables. Donde el cerebro es bueno, eligiendo y organizando sus competencias para la vida ordinaria, la máquina es muy pobre y hasta previsible. El fuerte de la IA es hacer patente la existencia de patrones de información entre montañas de datos. Y, además, conseguir encriptar esta información rescatada y comunicarla a otra máquina, manteniéndola a salvo de intrusos. Parece que su inteligencia les ofrece la posibilidad de vivir como máquinas: optimizando, rastreando, encriptando..., a una velocidad inalcanzable.

Entonces, si la naturaleza de la IA es tan distinta de la inteligencia humana, ¿cabe la posibilidad de que se alce en pos de la dominación mundial y la aniquilación de la humanidad? Nadie nos aseguró que llegáramos para quedarnos. Y si finalmente es así, las máquinas inteligentes siempre serán hijas de la inteligencia humana. ¿Un fatal error final? Es que estamos hechos para elegir bien entre fracaso y fracaso; no lo mejor, ni lo óptimo. ■

---

<sup>3</sup> Cfr. M. ABADI – D. G. ANDERSEN, *Learning to Protect Communications with Adversarial Neural Cryptography* (publicado el 21 octubre de 2016). Acceso en arXiv:1610.06918.