

Reseña del libro: La era de la inteligencia artificial

(The age of artificial intelligence)

Dra. Ailyn Febles Estrada

Presidenta de la Unión de Informáticos de Cuba (UIC)

ailyn.febles@uic.cu

El aprendizaje automático, en particular, está teniendo un fuerte impacto en nuestras vidas y también afectará en gran medida nuestro futuro. Ese es el mensaje de este nuevo y fascinante libro del ex secretario de Estado de los Estados Unidos Henry A. Kissinger, el ex director ejecutivo de Google [Eric Schmidt](#) y el decano del MIT Daniel Huttenlocher. Su hipótesis: la Inteligencia Artificial (IA) pondrá en duda la primacía de la razón humana, que ha existido desde los albores de la ilustración.

El primer ejemplo que presentan es AlphaZero (un programa de IA desarrollado por DeepMind, una subsidiaria de Google) que derrotó a Stockfish, el mejor programa de ajedrez del mundo. El hecho destacable no es que un programa le haya ganado a otro, sino que se enseñó a sí mismo cómo hacerlo. Sus creadores le dieron las reglas del ajedrez e instrucciones para que buscara una estrategia ganadora. Tras solo cuatro horas de aprender jugando contra sí mismo, AlphaZero se convirtió en campeón mundial de ajedrez al derrotar a Stockfish veintiocho veces sin perder un solo partido (hubo 72 empates). Como señalan los autores, con esto la IA trasciende la prueba de Turing: ya no se trata solo de un desempeño que no se pueda distinguir de la inteligencia humana, sino de un desempeño que supera a los humanos.

También describen redes neurales generativas, que tienen la capacidad de crear imágenes o textos nuevos. Los autores mencionan a GPT-3, de OpenAI, como una de las IA generativas más notables de la actualidad. En el 2019, la empresa desarrolló un modelo lingüístico que se entrena a sí mismo consumiendo textos disponibles en Internet, y que, a partir de unas pocas palabras y la detección de patrones en elementos secuenciales, puede extrapolar oraciones y párrafos. El sistema es capaz de componer textos nuevos y originales que pasan la prueba de Turing: mostrar una conducta inteligente que no se pueda distinguir de la de un ser humano.

Los autores consideran que estos temas fundamentales comienzan a adquirir relevancia cuando plataformas de redes globales como Google, Twitter y Facebook están usando la IA para reunir y filtrar volúmenes de información que para sus usuarios serían imposibles de procesar. Pero estos métodos de filtrado conducen a una segregación de los usuarios, al crear cámaras de resonancia sociales que fomentan desacuerdos entre diversos grupos. Lo que una persona considera una descripción exacta de la realidad se vuelve muy diferente de la realidad que ven otras personas o grupos, y eso refuerza y profundiza la polarización. Cada vez más, la IA decide lo que es importante y lo que es verdadero.

Los autores citan el caso de la [halicina](#), un nuevo antibiótico descubierto en 2020 gracias a una investigación del MIT en la que usando IA fue posible crear en unos pocos días modelos de millones de compuestos (una tarea fuera del alcance de los seres humanos) para explorar métodos bactericidas todavía no descubiertos ni explicados.

El libro no habla mucho sobre el fantasma de la IAG (la inteligencia artificial general), es decir, software capaz de realizar cualquier actividad intelectual, incluso relacionar tareas y conceptos en forma interdisciplinaria. Cualquiera sea el futuro a largo plazo de la IAG, ya bastantes problemas tenemos con la IA generativa basada en aprendizaje automático que ya existe. Esta puede extraer conclusiones, ofrecer pronósticos y tomar decisiones, pero no tiene conciencia de sí, ni puede pensar acerca de su lugar en el mundo.