

En el futuro:

Si no se regula la IA, los resultados pueden ser “catastróficos”

Yoshua Bengio, referente mundial en el desarrollo de esta tecnología, hoy busca “mitigar sus riesgos”.

El informático canadiense Yoshua Bengio fue pionero en algoritmos de aprendizaje profundo, y ganador, junto a otros dos científicos, del Premio Turing 2018, el “Nobel” de la informática. De hecho se los suele llamar los “padrinos” de la inteligencia artificial (IA).

Pero en entrevista con Live Science, Bengio habló de por qué pide tanto regular como una “moratoria” a la IA. “Si piensas racionalmente, no hay forma de negar la posibilidad de resultados catastróficos cuando alcancemos un cierto nivel de IA. La razón por la que cambié de actitud es porque antes entendía que hay escenarios malos, pero pensé que lo resolveríamos. Pero el pensar en mis hijos y su futuro me hizo decidir que tenía que actuar diferente y hacer todo lo posible por mitigar los riesgos”.

Sobre los riesgos de la IA, dice que no son de “ciencia fic-



Bengio también recibió en 2022 el premio Princesa de Asturias de Investigación Científica y Técnica.

ción”. “Un estudio reciente demostró que ChatGPT-4 es mucho mejor que los humanos en la persuasión, y eso es solo ChatGPT-4; la nueva versión va a ser peor (...). Si miramos más adelante, cuando alcancemos el nivel de superinteligencia, existen dos riesgos importantes. El primero es la pérdida del control humano; si las máquinas superinteligentes tienen un objetivo de autoconservación, su objetivo podría ser destruir a la humanidad para que no podamos apagarlas. El otro peligro (...) es que los humanos usen el poder de la IA para tomar el control de la humanidad en una dictadura mundial. Puede haber versiones más suaves (...), pero la tecnología dará un poder enorme a quien la controle”.

Bengio dice sentirse en parte responsable por mitigar los

malos pasos de la IA debido a su reconocimiento científico. Pero también porque hay soluciones técnicas. “Las empresas estarían encantadas de incluir estas soluciones técnicas, pero por ahora no sabemos cómo hacerlo. Todavía quieren obtener los billones de ganancias pronosticados de la IA que alcance un nivel humano, así que estamos en una mala posición y necesitamos encontrar respuestas científicas”.

La regulación, dice, es clave: “Es como si toda la humanidad estuviera conduciendo por una carretera que no conocemos muy bien y hay niebla frente a nosotros. (...) Y puede haber un paso muy peligroso”.

Y agrega: “Entonces, ¿qué hacemos? ¿Seguimos adelante con la esperanza de que todo salga bien o intentamos encontrar soluciones tecnológicas? La solución política dice que hay que aplicar el principio de precaución: hay que ir más despacio si no se está seguro. La solución técnica dice que debemos encontrar formas de mirar a través de la niebla y tal vez equipar el vehículo con salvaguardas”.