La inteligencia artificial es la gran revolución del siglo XXI. Las personas siempre hemos tenido un papel predominante en el desarrollo de esta disciplina, desde su nacimiento a mediados del siglo pasado hasta los éxitos cosechados en los últimos años.

La inteligencia artificial es una de las tecnologías más disruptivas del siglo XXI. Ha sido definida como la nueva electricidad e identificada como una de las tres áreas más prometedoras por el mismísimo Bill Gates. Es el ingrediente clave de la cuarta revolución industrial y cada día nos despertamos con noticias sobre sus impresionantes avances, que en muchos casos superan a la pericia humana. Según la Real Academia Española, la inteligencia artificial se define como la disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico. Pero ¿puede la inteligencia artificial superar a la humana?

El nacimiento de la inteligencia artificial

El objetivo de la inteligencia artificial siempre ha sido emular a la inteligencia humana. Aunque es un campo que se ha popularizado entre la sociedad en los últimos años, su nacimiento data de mediados del siglo pasado. En 1950, el matemático británico Alan Turing publicaba su famoso artículo *Computing Machinery and Intelligence*, en el que describía cómo crear máquinas inteligentes. De este artículo nace también el test de Turing que es la prueba de referencia todavía hoy en día para determinar si un sistema artificial exhibe inteligencia. Si una persona mantiene una conversación con una máquina y otra persona y no es capaz de discernir cuál de los dos es realmente una máquina, se dice que la máquina es inteligente.

El otro evento que está ligado al nacimiento de la inteligencia artificial como disciplina, y donde se acuñó oficialmente el término, fue el Proyecto de Investigación de Verano de Dartmouth en 1956. Esta escuela de verano, organizada por Marvin Minksy y John McCarthy, reunió a los investigadores que más tarde se considerarían los fundadores de la inteligencia artificial. El objetivo de este proyecto se basaba en la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, ser descrito con tanta precisión que puede fabricarse una máquina para simularlo.

Después de la escuela de verano de Dartmouth, la historia de la inteligencia artificial sufrió una serie de altibajos. Alternó periodos de éxitos (llamados primaveras) con épocas en las que se cortaba la financiación para seguir investigando en la materia (llamadas inviernos), a menudo debido a las altas expectativas que se creaban y no se cumplían. Como ejemplo, en 1970 Marvin Minksy afirmó que, en un plazo de tres a ocho años, sería posible desarrollar una máquina con la inteligencia general de un ser humano medio. Han pasado más de 50 años y esta afirmación todavía no se ha cumplido.

La nueva primavera

Hoy en día estamos siendo testigos del período de mayores éxitos de la inteligencia artificial. Desde principios del siglo XXI se han conjugado varios factores que forman el caldo perfecto para el progreso de la inteligencia artificial.

El primero de ellos es la aparición del fenómeno Big Data, que alimenta de grandes cantidades de datos a los algoritmos de inteligencia artificial. El segundo factor desencadenante es el aumento en la potencia computacional de las máquinas que permite hacer experimentos más complejos. Gracias a la aparición de estos factores entra en escena el aprendizaje profundo (o *Deep learning*, en inglés). El aprendizaje profundo basa su potencia en las llamadas redes de neuronas, compuestas de capas y capas de unidades de

procesamiento.

Las personas tienen un papel irremplazable en el desarrollo de la inteligencia artificial

Gracias a este subcampo de la inteligencia artificial han aparecido los mayores avances en la disciplina, a menudo ligados al reto de superar la inteligencia o habilidad humana. Así, uno de los grandes hitos en la historia de la inteligencia artificial ocurrió en 2015 cuando AlphaGo, un programa desarrollado por Google DeepMind, fue capaz de vencer al campeón del mundo en el juego de mesa Go. El Go es un juego chino que, según los expertos, es sustancialmente más complejo que el ajedrez. Dos años después, su sucesor AlphaZero logró en solo 24 horas un nivel sobrehumano en varios juegos de mesa (ajedrez, shogi y Go), derrotando a otros programas anteriores y también a campeones humanos.

Otro campo en el que el aprendizaje profundo ha supuesto un salto cualitativo es el procesado del lenguaje natural. Motores de traducción como Google Translate o DeepL han mejorado notablemente en los últimos años. La inteligencia artificial cuenta con éxitos recientes en áreas específicas como la medicina, los vehículos autónomos o los asistentes virtuales.

Si los ordenadores son tan inteligentes ¿cómo es que no pueden leer un libro?

Aunque pueda parecer que la inteligencia artificial está desplazando a las personas y superando su inteligencia, esto no es así. Actualmente, estamos en la era de la inteligencia artificial “estrecha”. Esto significa que tenemos sistemas que son capaces de realizar tareas específicas muy complejas, pero están muy lejos de la inteligencia generalista que tenemos las personas. Este hecho está directamente relacionado con la paradoja de Moravec que en 1988 afirmó: “Es relativamente fácil conseguir que los ordenadores muestren capacidades similares a las de una persona adulta en un test de inteligencia o a la hora de jugar a las damas; y muy difícil lograr que adquieran las habilidades perceptivas y motoras de un bebé de un año”. Más tarde, en 1994, Steven Pinker afirmaba que “la principal lección después de treinta y cinco años de investigación en inteligencia artificial es que los problemas difíciles son fáciles y los problemas fáciles son difíciles”.

Los científicos Gary Marcus y Ernest David explican cómo, a pesar de los recientes e impresionantes avances en procesado del lenguaje natural, la inteligencia artificial todavía no puede leer un libro, entendiendo por «leer» como entender su contenido. Aunque existen herramientas potentes como GTB (Google Talk to Books), éstas se basan en dar respuestas imitando frases que diría una persona, pero no son capaces de entender lo que leen. Cuando las personas leemos una historia, necesitamos seguir una serie de inferencias que están implícitas en ella y que demandan el uso de conocimiento de ámbito general. La aproximación actual de la inteligencia artificial se basa en la representación de probabilidades, de intentar adivinar qué palabras tienden a co-ocurrir en una frase o contexto. De esta forma se pueden producir textos o frases que parezcan dichas por personas, pero no implica conocimiento del dominio ni que se haya entendido un texto.

Confiar o no confiar, esa es la cuestión

Aún en los casos en que los algoritmos de inteligencia artificial consiguen resultados sobrehumanos, no es oro

todo lo que reluce. Las técnicas de aprendizaje profundo tienen la particularidad de que es muy difícil interpretar cómo el sistema ha llegado a una determinada decisión o predicción sobre los datos (lo que se conoce como modelo de caja negra). Este hecho es incompatible con la necesidad de que la inteligencia artificial sea ética y confiable, como demandan las *Directrices éticas para una inteligencia artificial fiable* de la Unión Europea y la *Estrategia Nacional de Inteligencia Artificial* española.

Recientemente han aparecido numerosos casos en los que sistemas de inteligencia artificial se comportaban de forma racista o sexista. Dado que los sistemas inteligentes son creados por personas, parece inevitable que los sesgos humanos se transmitan a los propios algoritmos o a través de los datos. Amazon tuvo que retirar su algoritmo para la contratación de personal al descubrir que era sexista y sólo consideraba candidatos masculinos. Google se disculpó después de descubrir que su algoritmo de reconocimiento de imágenes etiquetaba como “pistola” a un termómetro cuando la mano que lo sujetaba era de una persona de color. Una de las soluciones para mitigar el sesgo que aparece en sistemas de inteligencia artificial e incrementar su confianza es contar con equipos de desarrollo diversos, que supervisen exhaustivamente los resultados obtenidos.

No habrá inteligencia artificial sin inteligencia humana

Las personas tienen un papel irremplazable en el desarrollo de la inteligencia artificial. No solo en su creación, sino también en su supervisión y uso. Como postulan las *Directrices éticas para una inteligencia artificial fiable*, los sistemas inteligentes deben estar centrados en las personas, con un uso para el servicio de la humanidad y del bien común, con el objetivo de incrementar el bienestar humano. Es necesario asegurar que las personas sean capaces de tomar decisiones informadas a partir de los resultados obtenidos por un sistema inteligente. Además, siempre debe haber personas que supervisen el funcionamiento de programas de inteligencia artificial para evitar que se violen derechos fundamentales.

Inteligencia artificial y humana de la mano

No cabe duda de que la inteligencia artificial es una gran revolución con el potencial de transformar nuestra sociedad de formas nunca antes vistas. Pero no podemos olvidar que la inteligencia artificial no entiende, sino que aprende. Debemos ver a esta tecnología como una herramienta que nos puede ayudar en multitud de tareas, pero que no es enemiga de las personas.

El sueño de crear inteligencia se ha convertido en realidad. Solo el tiempo dirá si llegaremos a tener una inteligencia artificial generalista o con consciencia. Pero una cosa está clara: no habrá inteligencia artificial sin inteligencia humana.