

TECNOLOGÍA La coexistencia entre la inteligencia artificial y los seres humanos **P6-7**



La Inteligencia Artificial se encuentra en la base de los nuevos coches autoconducidos que están desarrollando varios fabricantes y compañías tecnológicas.

La Inteligencia Artificial se adapta a las personas

La Inteligencia Artificial es ya una realidad, pero aún queda camino por recorrer a la hora de conseguir una coexistencia eficaz que repercuta positivamente en nuestras vidas.

Richard Waters, Financial Times
 Parece que la especie humana no se puede reemplazar. Al menos de momento. En los últimos años no han faltado las predicciones sobre cómo la Inteligencia Artificial (IA) y la robótica sustituirían a los seres humanos ocupando sus puestos de trabajo. Ahora parece que la mayoría de los expertos en IA no ven un desenlace tan inmediato ni tan drástico.

En su versión del futuro, las personas seguirán desempeñando un papel en colaboración con los sistemas inteligentes: cabe la posibilidad de que la tecnología sea tan avanzada que acabe asumiendo el control, o que sus decisiones tengan consecuencias que sean demasiado importantes como para delegar todo en un proceso automatizado. En la toma de decisiones, una combinación hombre-máquina debería ofrecer mejores re-

sultados que un trabajo individual, opina David Mindell, catedrático del MIT y autor de *Our Robots, Ourselves*. Sólo hay un problema: cuando las personas y los sistemas semi-inteligentes intentan trabajar juntos, las cosas no siempre salen bien.

Este año presenciamos un ejemplo catastrófico en las calles de Tempe, Arizona. Un coche de pruebas de Uber equipado con la tecnología de conducción autónoma de la compañía atropelló a una persona que cruzaba la calle acabando con su vida.

Como en casi todos los coches autónomos, había un conductor de apoyo para intervenir si el software fallaba. Sin embargo, el estudio de la policía local concluyó que el conductor estaba distraído en ese momento viendo un programa de televisión en su *smartphone*.

Es preciso tener una visión más realista de los nuevos sistemas autónomos de conducción

En la identificación de imágenes, por ejemplo, la IA ha demostrado su superioridad

El vehículo de Uber tenía un grado de autonomía que se extenderá a otros modelos a partir del próximo año. El conocido como sistema Level 3 está diseñado para conducir de forma autónoma en la mayoría de situaciones, aunque entregará el control a una persona cuando se enfrente a casos que no pueda gestionar.

“Un sistema que pretenda ser totalmente autónomo pero que de repente se desvía de su cometido puede despistar a una persona”, aseguran los críticos. “Si solo se requiere la intervención humana durante un minuto, no va a funcionar”, lamenta Stefan Heck, consejero delegado de Nauto, una *start up* estadounidense cuya tecnología se utiliza para evitar que los conductores profesionales se distraigan.

Sin un diseño válido, los sistemas inteligentes podrían provocar una clara reacción

en contra de estas tecnologías.

Cuando la gente entienda las limitaciones de los actuales sistemas de aprendizaje automático, las esperanzas depositadas en ellos no tardarán en evaporarse, advierte Roger Schank, experto en IA y en psicología de aprendizaje. “El resultado será una nueva crisis para la IA”, añade, en referencia al periodo de finales de los ochenta, cuando la decepción con la tecnología provocó una retirada masiva de inversiones.

Visión realista

Para evitar que esto suceda será necesario tener una visión más realista de los nuevos sistemas autónomos. No obstante, la propia tecnología presenta una barrera. “Desconocemos tanto la forma en la que funciona la IA como la forma en la que falla”, explica Illah Nourbakhsh, profesor

de robótica de la Carnegie Mellon University.

El coche semiautónomo es un ejemplo de un sistema prácticamente autónomo que sin embargo depende de la intervención de las personas.

A medida que avanza la IA, se introducen sistemas híbridos como estos en distintas áreas. El aprendizaje automático –el tipo de IA responsable de los últimos grandes progresos– es una forma avanzada de reconocimiento de patrones. Ya ha demostrado su superioridad frente a las personas en tareas como la identificación de imágenes o el reconocimiento del habla.

Sin embargo, es menos efectivo cuando tiene que analizar datos concretos sobre los que ha sido entrenado. En el mundo real, la gente suele tomar decisiones sobre situaciones que nunca ha vivido. El problema reside en los

La Inteligencia Artificial ignora cómo tratar con las personas

Durante mucho tiempo ha habido una manera segura de lograr que las personas y las máquinas trabajen juntas de manera más eficaz: hacer que los seres humanos actúen más como las máquinas. Desde los primeros días de la fabricación en masa ha sido más fácil hacer que las personas se adapten al mundo minuciosamente estructurado de los procesos automatizados que conseguir que los sistemas funcionen en el desordenado mundo humano. Esto también es aplicable en la era de la Inteligencia Artificial, (IA) pero con limitaciones. La gente a menudo acepta (y actúa en consecuencia) los resultados que generan dichos sistemas sin cuestionarlos, dice Roger Schank, experto en psicología del

aprendizaje. Como ejemplo menciona el riesgo de que los médicos sigan ciegamente las recomendaciones de los sistemas de diagnóstico inteligentes, incluso cuando se equivoquen. Según algunos expertos en informática, para romper este ciclo y conseguir que haya una relación más creativa entre el ser humano y la máquina se necesitarán nuevas formas de IA. Los sistemas tendrán que tener una comprensión más amplia del mundo para que sus consejos encajen en un contexto más humano, señala Vishal Sikka, ex consejero delegado de Infosys. Schank opina que los expertos en informática deberían dejar de estudiar la inteligencia de las máquinas y centrar su atención en las personas. Señala que algunos



Las cadenas de montaje hicieron posible que las personas se adaptaran al mundo de los procesos automatizados.

fundadores del campo de la IA también eran psicólogos. Los avances necesarios sólo se conseguirán conociendo cómo los humanos aprenden de sus errores

y aplican sus conocimientos a nuevas situaciones. David Mindell, un profesor del MIT que ha escrito sobre el reto de lograr que los humanos y los robots interactúen

de manera efectiva, lo expresa de manera más sucinta: "El mundo de la informática todavía tiene un largo camino por recorrer para saber cómo tratar con las personas".

sistemas capaces de encajar datos pero que no procesan su importancia.

Las nuevas formas de cooperación entre personas y máquinas están echando raíces. En primer lugar, con situaciones en las que las personas actúan de soporte a los robots cuando éstos alcanzan el límite de sus capacidades. Un ejemplo serían los centros de atención al cliente, donde hay sistemas capaces de entender idiomas que resuelven las dudas de los clientes.

El accidente de Uber es un ejemplo extremo de lo que puede fallar, aunque, según un estudio de la Universidad de Stanford, un conductor tarda seis segundos en reaccionar. Pero, incluso en el caso de que haya tiempo suficiente para recuperar la atención, la persona que intervenga podría interpretar las cosas de forma distinta a la máquina. Esto implica que podría haber un problema de coordinación y un posible fallo.

El uso del lenguaje dificulta aún más las cosas. Dos personas pueden comunicar algo en muy pocas palabras: el contexto ayuda. Los informáticos todavía no han conseguido crear esa complicidad entre las máquinas.

Un segundo tipo de cooperación entre personas y máquinas está diseñado para asegurarse de que una tarea complicada siempre dependa de una persona, incluso en situaciones en las que un sistema automatizado podría ser capaz de completar esa tarea. En los drones militares, por ejemplo, siempre se recurre a los pilotos para que tomen la decisión de disparar a un objetivo.

Los sistemas de reconoci-

miento facial –utilizados por los agentes de aduanas para identificar a los viajeros sospechosos– son otro ejemplo. En los dos queda demostrado cómo la IA puede aumentar la eficiencia de las personas.

Una de las críticas de las armas semiautónomas como los drones es que no hay barreras técnicas para que se conviertan en sistemas plenamente autónomos.

Según Stuart Russell, profesor de IA de la Universidad de California, "no se puede decir que la tecnología sólo se puede usar en situaciones defensivas y bajo supervisión humana. No es verdad". Otro ejemplo de sistema con IA pero que requiere la intervención de una persona son los

IBM experimenta con los diagnósticos médicos, aunque la última palabra la tiene el especialista

Es probable que las personas que no son expertas sean reacias a cuestionar a una máquina

algoritmos que agrupan datos para después ofrecer recomendaciones. Estos algoritmos son incapaces de gestionar nuevas situaciones.

Un error de la app que utilizan los conductores de Uber en San Francisco envió a unos pasajeros a la terminal de carga del aeropuerto en lugar de a la de pasajeros. "A veces nos da por seguir las instrucciones de una máquina a ciegas, otras en cambio, pensamos: 'un momento, aquí está pasando algo'. Lleva tiempo adaptarse a algunas tecnologías", explica Tim O'Reilly, experto en Internet y editor.

Estos son casos que por lo general tienen poca importancia, pero ¿qué pasa cuando

hay mucho más en juego? IBM ha experimentado con los diagnósticos médicos, aunque el sistema está diseñado para que la última palabra la tenga un experto. No obstante, los últimos avances de la IA podrían ser más difíciles de cuestionar. Gracias al éxito de técnicas como el reconocimiento de imagen, las expectativas de estos sistemas han ido mejorando. Evaluar la calidad de las recomendaciones de un sistema basado en la IA plantea otros retos.

Es probable que las personas que no son expertas en la materia se sientan reacias a cuestionar a una máquina cuyo funcionamiento no entienden. No es algo nuevo. Hace

más de treinta años, un fallo técnico en el software de una máquina de radioterapia, Therac-25, se cobró la vida de seis personas.

Los técnicos fueron incapaces de detectar el error y la máquina siguió funcionando mucho más tiempo del que hubiera sido deseable.

La tecnología utilizada en la mayor parte de los sistemas de aprendizaje automático presenta nuevos desafíos. En general está basada en una teoría sobre el funcionamiento del cerebro humano, que transmite la información a través de capas de neuronas artificiales hasta que surge un patrón que pueda identificarse.

A diferencia de los circuitos lógicos que se utilizan en un programa de software tradicional, no hay forma de hacer un seguimiento de este proceso para detectar con exactitud por qué un ordenador ofrece una determinada respuesta en concreto. Este es un gran obstáculo en la adopción de redes neuronales. Es lo paradójico de la IA –los mejores sistemas suelen ser los que hoy en día resultan difíciles de entender–. Algunos expertos creen que se están haciendo progresos y que no pasará mucho tiempo antes de que los sistemas de aprendizaje automático sean capaces de reconocer los factores que les han llevado a tomar una determinada decisión. Muchos expertos en la materia están convencidos de que cuando las personas y las máquinas colaboren, se obtendrán resultados mucho mejores de los que se hubieran registrado por separado. Sin embargo, quedan algunos desafíos por resolver antes de que lleguemos a conocer ese futuro prometedor.



La salud será uno de los campos de desarrollo de la IA.



En los drones militares se recurre al ser humano para disparar.