

Los límites éticos de la inteligencia artificial en la sanidad



A pesar de las ventajas que presenta la inteligencia artificial para mejorar la atención sanitaria, es preciso utilizarla con cautela y someterla a una normativa estricta para que no genere efectos no deseados que puedan poner en peligro el bienestar de las personas.

La pandemia puso en evidencia algo que venía siendo una tendencia: la inteligencia artificial se está convirtiendo en una herramienta de gran relevancia para el cuidado de la salud en el siglo XXI. Desde el algoritmo de la empresa canadiense BlueDot que fue capaz de detectar que algo extraño estaba pasando en las inmediaciones del mercado de Wuhan días antes de que la Organización Mundial de la Salud alertase sobre el coronavirus, hasta los sistemas capaces de identificar en segundos la presencia de la COVID-19 en una placa de tórax con un alto porcentaje de acierto, el papel de la tecnología en la medicina no ha hecho más que aumentar a causa de la crisis.

La inteligencia artificial, y en concreto su modalidad más utilizada actualmente, que es el aprendizaje

automático o *machine learning*, ofrece un abanico realmente amplio de aplicaciones en el ámbito de la sanidad. Por ejemplo, la empresa norteamericana GNS Healthcare desarrolla modelos basados en el *machine learning* para determinar los tratamientos más efectivos a las características de cada paciente, y, en España, Sanitas ha puesto en marcha un servicio de seguimiento remoto para pacientes de COVID-19 basado en inteligencia artificial, de forma que, si se detecta cualquier anomalía en el estado del paciente, un médico se pondrá en contacto con él.

Otro campo donde los sistemas inteligentes se perfilan como muy útiles es en el apoyo al diagnóstico médico mediante el reconocimiento de imágenes, especialmente si se trata de patologías muy comunes – como el cáncer de pecho, el de piel o una retinopatía diabética-, pues existen grandes bases de datos para que el algoritmo pueda aprender a identificar los patrones de la enfermedad. Los hospitales Ramón y Cajal, Doce de Octubre y Sant Pau van a implantar un proyecto pionero de Aprendizaje Federado de imágenes radiológicas de tórax, de forma que compartirán la utilización de un modelo de inteligencia artificial para el diagnóstico de imagen médica. Por otro lado, dentro de la medicina preventiva, la inteligencia artificial puede llegar a evaluar el riesgo de padecer enfermedades antes de que se manifiesten, mediante el análisis de factores determinantes, como los antecedentes familiares, el entorno, la profesión desempeñada o el estilo de vida del paciente.

Las tecnologías digitales también apoyan la investigación en el campo médico y farmacológico. En este sentido, el proyecto AIDD (*Artificial Intelligence for Drug Discovery*) de NuMedii utiliza la inteligencia artificial para identificar correlaciones y efectos entre los compuestos químicos de las medicinas y los síntomas de distintas enfermedades. En diciembre de 2020, la compañía DeepMind anunció que su sistema AlphaFold de aprendizaje profundo había resuelto el “problema del plegamiento de las proteínas”, un avance básico para el desarrollo de nuevos medicamentos.

Todo parece indicar el salto adelante para el cuidado de la salud que supone el aplicar estas tecnologías de vanguardia, y, sin embargo, no son pocos los organismos especializados -empezando por la Organización Mundial de la Salud- que han expresado su preocupación por los posibles perjuicios que ello puede acarrear para las personas

Todo parece indicar el salto adelante para el cuidado de la salud que supone el aplicar estas tecnologías de vanguardia, y, sin embargo, no son pocos los organismos especializados -empezando por la Organización Mundial de la Salud- que han expresado su preocupación por los posibles perjuicios que ello puede acarrear para las personas. Gran parte de los peligros asociados a la inteligencia artificial en la medicina están

asociados con el uso de los datos, y la posibilidad de que este sea inadecuado, tenga consecuencias no previstas ni deseadas, o dé lugar a algún tipo de discriminación. Este tema es especialmente relevante partiendo de la evidencia de que los sistemas más utilizados -el *machine learning* y el *deep learning*- son intensivos en datos, pues requieren ingentes cantidades de ellos para “aprender”.

Problemas éticos con los datos

El dato se ha convertido en el combustible de la economía digital. El cuidado de la salud no es una excepción, y la recolección, el análisis y el uso de datos sanitarios -desde los análisis clínicos, hasta los historiales médicos- es la base de la práctica de la medicina y de la investigación en este campo. La progresiva digitalización del sector ha hecho que durante las últimas dos décadas haya aumentado exponencialmente el tipo de datos utilizados por los servicios de salud, incluyendo ahora información personal sobre los individuos procedente de diversas fuentes a veces no directamente relacionadas con la salud, como, por ejemplo, datos sobre el entorno del paciente, sobre su estilo de vida, socioeconómicos o comportamentales.

Qué duda cabe que el potencial de los macrodatos y la inteligencia artificial aplicados al cuidado sanitario es inmenso, tanto en términos de la velocidad y precisión a la hora de realizar diagnósticos, como en la mejora de la calidad del servicio prestado o el apoyo a la toma de decisiones. Sin embargo, este uso de información personal puede acarrear problemas éticos, que es necesario poder identificar y controlar mediante una legislación adecuada, que proteja los derechos de la ciudadanía.

Uno de los principales riesgos que conlleva la utilización de grandes colecciones de datos es la posibilidad de vulnerar el derecho a la privacidad de los individuos. El compartir datos o transferir datos personales cedidos por los pacientes, o generados dentro de un proceso sanitario, puede llevar a que la información recabada pueda ser utilizada para discriminarle o para tratarle injustamente, incluso, si caen en malas manos, pueden utilizarse para cometer actos delictivos. En este sentido, resulta fundamental establecer protocolos y normas estrictas sobre quién puede acceder a los datos personales y bajo qué circunstancias a lo largo de todo el ciclo de vida de la inteligencia artificial.

Uno de los principales riesgos que conlleva la utilización de grandes colecciones de datos es la posibilidad de vulnerar el derecho a la privacidad de los individuos

La Organización Mundial de la Salud (OMS) alerta del peligro adicional de que las instituciones encargadas de prestar los servicios de salud recojan más datos de carácter personal que los que requiere el algoritmo para funcionar, y que ese “superávit de datos”, sea compartido con terceros, por ejemplo, agencias gubernamentales, para llevar a cabo acciones que puedan vulnerar los derechos humanos, como pueden ser la vigilancia y el control de la población, o la aplicación de castigos individuales. La OMS pone de ejemplo de esto la aplicación para rastrear contactos de COVID-19 que puso en marcha el Gobierno de Singapur, cuyos datos podían ser accedidos en el marco de una investigación criminal.

Otro de los grandes problemas que presenta el uso de *big data* en la inteligencia artificial es la posible

aparición de sesgos no deseados en los resultados que ofrecen los algoritmos, algo que puede originarse por la mala calidad de la información utilizada. Un estudio llevado a cabo por Marzyeh Ghassemi en el MIT ha descubierto que los datos utilizados en medicina siempre llevan algún tipo de sesgo de sexo, género o raza, tanto los que proceden de dispositivos clínicos, los asociados a las intervenciones, como los que tienen su origen en las interacciones entre pacientes y personal sanitario. Al alimentar con estos datos el *machine learning*, las conclusiones a las que llegan los algoritmos reproducen esos sesgos. En la práctica el estudio ha desvelado que los modelos de inteligencia artificial sanitarios analizados funcionan de forma diferente en función del tipo de paciente, y ofrecen resultados sesgados en función del género y la raza.

La necesidad de supervisión humana

La lucha contra la aparición de sesgos en los resultados de la inteligencia artificial exige conocer cómo funciona y por qué toma el algoritmo utilizado las decisiones que toma. Es por ello por lo que la inteligencia artificial sanitaria debe ser transparente, es decir, debe permitir recrear los resultados de cualquier decisión que toma, y se deben poder conocer tanto la lógica que rige el modelo como los datos con los que ha sido entrenado. Por lo tanto, es preciso disponer de herramientas de trazabilidad, que muestren cómo se produce el proceso completo, desde el aprendizaje del sistema hasta su forma de analizar un problema. El Hospital Monte Sinaí de Nueva York utiliza un programa de *deep learning* denominado Deep Patient cuyo fin es ayudar a predecir enfermedades. El sistema se nutre de una base de datos de 700 000 pacientes —toda suerte de *big data*: consultas, dolencias, hábitos...— y ha demostrado una gran precisión para anticipar posibles enfermedades que puede sufrir una persona. En concreto, es sumamente eficaz a la hora de diagnosticar antes de que se manifiesten determinados desórdenes mentales, como la esquizofrenia, algo muy difícil de realizar para los psiquiatras. El problema es que nadie entiende cómo Deep Patient puede hacerlo, cómo elabora sus conclusiones, de forma que podamos aprender de él. El director del equipo en el Monte Sinaí, Joel Dudley, llegó a afirmar: “Podemos construir estos modelos, pero no sabemos cómo funcionan”.

La inteligencia artificial sanitaria debe ser transparente, es decir, debe permitir recrear los resultados de cualquier decisión que toma, y se deben poder conocer tanto la lógica que rige el modelo como los datos con los que ha sido entrenado

Resulta primordial que un algoritmo relacionado con el cuidado de la salud resulte lo suficientemente comprensible, transparente y explicable para el personal sanitario que debe trabajar con él, que son aquellos que deben estar al tanto de los datos y variables que ha utilizado el sistema para elaborar su resultado. El problema es que a menudo un sistema de *machine learning* puede ser difícil de comprender por sus usuarios y difícil de explicar por los técnicos que lo han creado. En este sentido, según apunta un informe de DigitalES, se ha comprobado que el personal sanitario prefiere sacrificar un poco de precisión con tal de que el sistema sea más entendible.

Por otro lado, la inteligencia artificial debe ser robusta, es decir, tiene que tener una solidez técnica y haber

sido desarrollada con un enfoque preventivo de riesgos, de forma que se comporte siempre como se espera que lo haga, minimizando los daños involuntarios e imprevistos.

En última instancia, en la aplicación de algoritmos inteligentes en la prestación sanitaria siempre debe regir la supervisión humana. Como indica Javier Mendoza, consultor de Savana Medical, en el citado informe:

“El médico debe entender que la última decisión siempre será suya; la supervisión humana debe prevalecer. Los algoritmos deben tener la suficiente autonomía como para generar resultados y el profesional sanitario la suficiente jerarquía como para controlar la toma de decisiones. Los estudios demuestran que una combinación de las partes siempre resulta en mejores resultados en comparación al trabajo individual”.

Foto de [MART PRODUCTION](#) en [Pexels](#)