

La ética de la inteligencia artificial

Los sistemas de inteligencia artificial son tecnologías de procesamiento de la información constituidos por bases de datos, modelos estadísticos y algoritmos que permiten realizar procedimientos de clasificación y predicción. Valiéndose de estos modelos y con en la información contenida en las bases de datos, se analizarán datos nuevos, los cuales podrán ser clasificados, y sobre los que se podrán hacer predicciones. Los datos nuevos pueden pasar a formar parte de las bases de datos subyacentes y enriquecerlas. Así, los sistemas de inteligencia artificial utilizan el denominado *aprendizaje automático (machine learning)*, que confiere a las computadoras la capacidad de aprender de los datos y mejorar su desempeño sin estar explícitamente programadas.

Los métodos estadísticos empleados por los sistemas de inteligencia artificial son varios: regresiones lineales y logísticas, *support vector machines*, árboles de clasificación y regresión, entre otros, así como también sus combinaciones. Dentro de los métodos empleados por la inteligencia artificial, merece especial atención el denominado *deep learning*, que emplea redes neuronales artificiales con múltiples capas. Las neuronas artificiales se comportan como una red formada por nodos interconectados que, de acuerdo con la información de entrada, debe llegar a la solución de un problema de la manera más rápida y eficiente. Este método se considera una especie de “caja negra” que da un resultado operativamente útil, pero, a diferencia de otros, nadie sabe a ciencia cierta cómo funciona.

En el ámbito de la salud, los sistemas de inteligencia artificial permiten mejorar los diagnósticos, las decisiones terapéuticas y los pronósticos. Las aplicaciones existentes son numerosas y van en constante aumento. En la bibliografía, se encuentran desarrollos para el diagnóstico por imágenes, la optimización de flujos de trabajo y la evaluación de síntomas. De esta manera, la inteligencia artificial se presenta como un recurso valioso para ser usado en el ámbito de la salud con múltiples ventajas. Actualmente, dado el estado de desarrollo de los sistemas, representa un buen respaldo para la toma de decisiones médicas, sin embargo, no puede dejarse en sus manos la toma de decisiones críticas como las que

involucran la vida y la muerte. Las máquinas no pueden pensar, aunque podrían llegar a hacer lo que nosotros, como entidades pensantes, hacemos. Por eso, deben abordarse problemas éticos fundamentales asociados a su uso e implementación.

En primer lugar, el consentimiento informado será uno de los desafíos más inmediatos al integrar la inteligencia artificial a la práctica clínica. Algunas preguntas que deben ser respondidas son: ¿el paciente aceptará que un sistema de inteligencia artificial sea quien realice las recomendaciones sobre su situación de salud o insistirá en ser atendido por un médico, de la manera tradicional?; ¿bajo qué circunstancias el clínico debe notificar al paciente que la inteligencia artificial está siendo utilizada?; ¿hasta qué punto los clínicos tienen la responsabilidad de hablarles a los pacientes sobre las complejidades de la inteligencia artificial?; ¿hasta qué punto un médico clínico debe declarar que no puede interpretar completamente los diagnósticos y tratamientos hechos por sistemas de inteligencia artificial?; ¿qué pasaría con los casos de pacientes reacios a permitir el uso de ciertas categorías de datos como los genéticos o de historia familiar?. A pesar de las múltiples incógnitas, es necesario entender que el desarrollo y el estado de implementación son procesos dinámicos y quizás, en un futuro no muy lejano, las cosas se inviertan y se considere una negligencia no utilizar sistemas de inteligencia artificial para llegar a un diagnóstico o dar una recomendación de tratamiento.

Otro de los desafíos éticos es la cuestión de la seguridad y la transparencia. En este sentido, es importante tener la certeza de que los sistemas de inteligencia artificial harán recomendaciones de tratamiento seguras y correctas. Esto resalta dos puntos clave: la confiabilidad y validez de las bases de datos utilizadas para el adiestramiento del sistema y la transparencia de los algoritmos. Cuanto mejor sea la base de datos de entrenamiento, mejor funcionará el sistema. Además, los algoritmos a menudo necesitan un mayor refinamiento. Idealmente, todos los datos y los algoritmos deberían estar abiertos al público para poder ser examinados. Sin embargo, existen cuestiones legítimas relacionadas con la protección de la inversión realizada para obtener esos datos,

con la propiedad intelectual y con los riesgos relacionados con la ciberseguridad. Es necesario que los desarrolladores de inteligencia artificial sean transparentes en cuanto a la clase de datos utilizados y respecto de cualquier deficiencia del *software*. La transparencia genera confianza entre las partes interesadas y esta es clave para la implementación exitosa de la inteligencia artificial en la práctica clínica.

En tercer lugar, surgen las cuestiones de equidad y sesgos de los algoritmos. La inteligencia artificial tiene la capacidad de mejorar el cuidado de la salud en todos los ámbitos y así, democratizar la experticia, globalizarla y llevarla a áreas remotas. Sin embargo, los datos con los que se entrena el sistema serán los que determinen la confiabilidad, efectividad y equidad del mismo. En este sentido, se corre el riesgo de sesgos y, consecuentemente, de discriminación. Por esto se deben minimizar sesgos potenciales en cada etapa del proceso de desarrollo del producto. En particular, los desarrolladores deben considerar el riesgo de sesgos al elegir la tecnología / procedimiento de *machine learning* a utilizar para entrenar el algoritmo y las bases de datos que van a emplear. Varios ejemplos del mundo real han demostrado que los algoritmos pueden mostrar sesgos que resultan en injusticias con respecto al origen étnico, género, edad o discapacidades. Deben tenerse en cuenta la representatividad de los datos, cómo los mismos son elegidos y analizados y también, el contexto. Cuando está involucrado el fenotipo / genotipo de los pacientes, los sesgos pueden llevar a falsos diagnósticos y a tratamientos inefectivos para algunas subpoblaciones, poniendo en peligro su seguridad. Si tomamos como ejemplo el cáncer de piel, utilizar un algoritmo entrenado en caucásicos podría ser inadecuado para afroamericanos.

Otro problema se asocia al contexto en el que la inteligencia artificial será utilizada. La inteligencia artificial desarrollada para expertos de primera categoría en contextos en donde se tienen todos los recursos podría no recomendar tratamientos adecuados, seguros y que resulten justos en contextos de bajos recursos. Más aún, esta solución podría exacerbar inequidades preexistentes. Por eso, si bien el uso del *big data* es una alternativa tentadora, es también importante considerar la posibilidad de generación de datos locales de calidad y cantidad suficientes para entrenar los algoritmos. En este sentido, es menester que los pacientes acepten brindar sus datos de salud y lo hagan mediante un consentimiento informado, a sabiendas de que serán utili-

zados con ese fin.

Finalmente, se plantea el tema de la privacidad de los datos. Es de fundamental importancia informar adecuadamente a los pacientes sobre el procesamiento de sus datos y propiciar un diálogo abierto para promover confianza. Con respecto a la propiedad de los mismos, es necesario resaltar que el valor de los datos de salud puede ser enorme, sin embargo hay evidencias que sugieren que a la gente no le gusta que las compañías o los gobiernos brinden esa información a cambio de rédito económico. Una opción podría ser que aquellas compañías que buscan utilizar datos de pacientes demuestren que están generando un valor agregado a la salud de los mismos pacientes de quienes se originaron (reciprocidad). Es imperativo también proteger a los pacientes contra los usos que se les pueda dar fuera de la relación médico-paciente, ya que podrían afectar no solo a la salud, sino también las oportunidades laborales, la contratación de seguros y hasta las relaciones personales.

Como conclusión, puede aseverarse que, siempre que se tomen los recaudos necesarios, la inteligencia artificial tiene un potencial enorme para ser usada dentro del ámbito de la salud. En este contexto, los laboratorios de análisis clínicos, vistos como organizaciones productoras de una gran cantidad de datos estructurados, tienen una nueva misión. Los datos generados serán utilizados para la atención de los pacientes individuales y también serán analizados en contexto, sumándolos a la información preexistente y contribuyendo al “aprendizaje” del sistema automático. La información generada bajo estrictas normas de control de calidad adquiere una perspectiva diferente y el profesional bioquímico se desempeñará también en una dimensión diferente, a través de la contribución que realiza desde su disciplina. Al mismo tiempo, es importante su participación en todos los asuntos que conciernen a la ética. Cada uno de sus aspectos deberá ser analizado críticamente y abordado también desde la perspectiva bioquímica.

En mayor o menor medida, la inteligencia artificial ya está presente en el ámbito de la salud y es nuestro deber profesional encontrar la mejor manera de utilizarla para el bien de nuestras comunidades.

Julián Verona

Doctor en Bioquímica, UBA - Bioquímica Clínica Área Lípidos.
Jefe de Servicio, Laboratorio del Hosp. de Balcarce, Buenos Aires.