

## Aplicaciones médico-sanitarias de la inteligencia artificial: una visión desde la práctica clínica\*

### Medical and health applications of artificial intelligence: a view from clinical practice

Juan J. Díez

Académico de Número de la Sección de Medicina de la Real Academia de Doctores de España  
[juanjose.diez@salud.madrid.org](mailto:juanjose.diez@salud.madrid.org)

#### RESUMEN

La inteligencia artificial se ha convertido en uno de los fenómenos sociales más importantes de los últimos años con potenciales influencias en todos los campos de la actividad humana. La sanidad, la medicina y la investigación biomédica no son ajenos a los cambios que inducirá el uso de la tecnología en el reconocimiento de patrones, interpretación de imágenes, toma de decisiones clínicas, tratamientos personalizados o gestión de hospitales. Diferentes modelos de *machine learning* y de *deep learning* son ya una realidad en distintas aplicaciones clínicas. El procesamiento del lenguaje natural se ha introducido en hospitales para facilitar el tratamiento de grandes masas de información clínica. En el ámbito sanitario las tecnologías de inteligencia artificial han mostrado su utilidad en la gestión de los servicios sanitarios, en la medicina predictiva, en la toma de decisiones y en el procesamiento de datos de los pacientes. En el ámbito de la actividad clínica existen diversos enfoques en los que la inteligencia artificial puede facilitar el trabajo de los profesionales, tales como el alivio de determinadas cargas de trabajo rutinario y administrativo, la división y el reemplazo de tareas y la mejora de la práctica clínica en diferentes aspectos. Las aplicaciones de IA se pueden utilizar en un gran número de actividades médicas, como las tareas relacionadas con el diagnóstico, la rehabilitación, los procedimientos quirúrgicos, las tareas de predicción y pronóstico, la medicina personalizada y la investigación clínica y traslacional. Estas nuevas tecnologías no están exentas de inconvenientes y riesgos, por lo que se hace necesaria una regulación que proteja y garantice la privacidad de los datos y la seguridad de los pacientes. Los aspectos éticos más importantes de la inteligencia artificial incluyen el respeto a la dignidad humana, la protección de los derechos humanos y libertades fundamentales y la salvaguarda de los intereses de las generaciones presentes y futuras.

**PALABRAS CLAVE:** Inteligencia artificial, atención sanitaria, práctica clínica, aprendizaje automático.

#### ABSTRACT

In recent years, artificial intelligence has become one of the most important social phenomena with potential influences on all fields of human activity. Healthcare, medicine and biomedical research are not immune to the changes that the use of technology will induce in pattern recognition, image

---

\* Sesión académica de la RADE celebrada el 09-10-2023 con el título *Visiones y aplicaciones de la Inteligencia artificial*.

interpretation, clinical decision making, personalized treatments or hospital management. Various machine learning and deep learning models are already a reality in different clinical applications. Natural language processing has been introduced in hospitals to facilitate the processing of large masses of clinical information. In the healthcare field, artificial intelligence technologies have shown their usefulness in the management of healthcare services, predictive medicine, decision making, and the processing of patient data. In the field of clinical activity, there are various approaches in which artificial intelligence can facilitate the work of professionals, such as alleviating certain routine and administrative workloads, dividing tasks, replacing tasks and improving clinical practice in different aspects. Artificial intelligence applications can be used in a large number of medical activities, such as tasks related to diagnosis, rehabilitation, surgical procedures, prediction and prognosis tasks, personalized medicine, and clinical and translational research. These new technologies are not free of drawbacks and risks. Regulatory standards are necessary to protect and ensure data privacy and patient safety. The most relevant ethical aspects of artificial intelligence include respect for human dignity, the protection of human rights and fundamental freedoms, and the safeguarding of the interests of present and future generations.

**KEY WORDS:** Artificial intelligence, healthcare, clinical practice, machine learning

## INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA SALUD

---

Podemos definir la inteligencia artificial (IA) como la teoría y desarrollo de sistemas informáticos capaces de llevar a cabo tareas que normalmente requieren el concurso de la inteligencia humana, tales como la percepción visual, el reconocimiento de la voz, la toma de decisiones o la traducción entre idiomas (1,2). En ciencias de la computación una máquina inteligente es un agente que percibe su entorno y lleva a cabo acciones que maximicen sus posibilidades de éxito en una tarea. En las ciencias de la salud existen cuatro importantes áreas susceptibles de aplicación de la IA: relación, decisión, descubrimiento y operaciones.

Las aplicaciones relativas a la relación son las que ayudan a que el ser humano se relacione con la máquina. Pueden incluirse aquí varios ejemplos como la tecnología de la diabetes (bombas de infusión de insulina, monitorización de la glucemia en tiempo real), terapias digitales (combinación de terapia farmacológica con apoyo digital), agentes inteligentes (*contact centers*, portales de pacientes), chequeo de síntomas, selección y clasificación automática de pacientes o el cuidado de personas mayores en domicilio. En las aplicaciones de descubrimiento la IA nos ayuda a entender las enfermedades. Un ejemplo puede ser las aplicaciones con datos de vida real (*real world evidence*) para el descubrimiento y reposicionamiento de fármacos, la predicción de complicaciones de enfermedades o la búsqueda de información relevante para descubrir relaciones que no son evidentes a primera vista. Las aplicaciones de decisión ayudan a los profesionales en el proceso de toma de decisiones. Algunos ejemplos incluyen la medicina de precisión en el cáncer (tratamientos personalizados a partir del perfil molecular del paciente), la predicción de reingresos tras el alta o el procesamiento de imágenes dermatológicas, radiológicas o de anatomía patológica. Finalmente, las operaciones permiten mejorar los procesos internos del sistema sanitario y mejorar la eficiencia, como por ejemplo, los sistemas de codificación automática o los de hospitales inteligentes.

## MACHINE LEARNING Y DEEP LEARNING

---

El aprendizaje automático (*machine learning*, ML) es una aplicación de la IA que proporciona a los sistemas la capacidad de aprender y mejorar automáticamente a partir de la experiencia sin ser programados de forma explícita (3). El sistema es, por tanto, capaz de desarrollar hipótesis de soluciones y de aprender a partir de datos de una forma automática, y de establecer comparaciones a partir de esos datos.

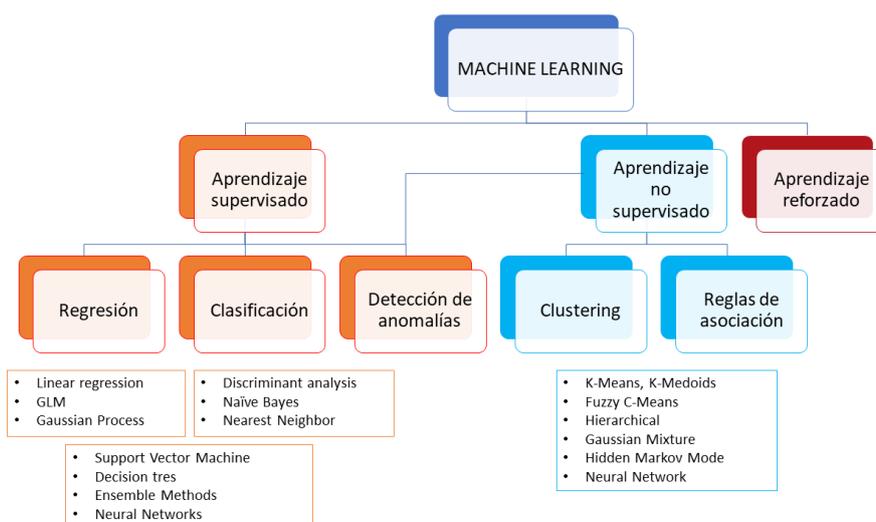
Para entrenar a las máquinas se precisan unos datos de entrenamiento (datos iniciales). Este entrenamiento es el resultado de un proceso de aprendizaje similar al funcionamiento de la inteligencia humana mediante el cual la máquina va incorporando los resultados

obtenidos al modelo estadístico para mejorarlo. De esta forma, la máquina genera modelos optimizados que permitirán explicar el comportamiento o tendencia de los datos, tomar decisiones y limitar el efecto de posibles sesgos.

El aprendizaje profundo (*deep learning*, DL) es un subconjunto de ML que se define formalmente como modelos computacionales que se componen de múltiples capas de procesamiento con múltiples niveles de abstracción y que está inspirado en el funcionamiento de las neuronas del sistema nervioso (4). Es un área de la IA que trabaja con redes neuronales de varias capas. En estas redes, la fase de aprendizaje tiene por objetivo hacer converger los parámetros de los datos hacia una clasificación óptima. Pero requieren de un proceso de entrenamiento intenso y no son válidas para todos los tipos de problemas. El DL es un método muy potente para el procesamiento de imágenes, sonido y vídeo. Utiliza técnicas de optimización y objetivos de capacitación modernos. Por ejemplo, las redes neuronales recurrentes se emplean con frecuencia para modelar datos secuenciales como el lenguaje. Las redes neuronales convolucionales se utilizan a menudo para modelar señales como imágenes (5).

## ALGORITMOS Y MÉTRICAS EN APLICACIONES MÉDICAS

Los algoritmos de ML más habituales en el ámbito biomédico, que se recogen de forma esquemática en la figura 1, se pueden clasificar a partir de las tres formas básicas por las que las máquinas pueden aprender: aprendizaje supervisado, no supervisado y reforzado (6-9).



**Figura 1.** Representación esquemática de los diferentes métodos de aprendizaje en *machine learning* y de los algoritmos más frecuentemente utilizados en estudios del área biomédica.

El *aprendizaje supervisado* se basa en ejemplos etiquetados por un supervisor externo, está dirigido por la tarea y enfocado a la clasificación o a la predicción. Su objetivo es generalizar la predicción de ejemplos no vistos. El sistema trata de responder a preguntas y precisa entrenamiento con datos históricos. De esta forma puede entrenarse con una serie de variables de entrada (por ejemplo, datos clínicos y analíticos de pacientes con Covid-19) a las que se asocian variables de salida reales (ingreso en UCI). Tras el aprendizaje el sistema llegará a una predicción, es decir, obtendrá una variable de resultado (ingreso en UCI) para un caso nuevo.

Las aplicaciones más destacables del aprendizaje supervisado son la clasificación y la regresión. En la primera la predicción sólo tiene un pequeño número de valores (tumor o no tumor, ingreso o no ingreso, etc.). En la regresión el resultado es un número que puede tomar gran cantidad de valores. En los límites con el aprendizaje no supervisado se encuentra la detección de anomalías, que utiliza datos para la búsqueda de alteraciones. Un ejemplo es la tecnología ponible o vestible (*wearable*) en los que se mide el ritmo cardíaco. Con un algoritmo se puede detectar una taquicardia o arritmia.

En el *aprendizaje no supervisado* está basado en ejemplos no etiquetados y su objetivo es encontrar patrones ocultos. Está, por tanto, dirigido por los datos y enfocado al agrupamiento. En este caso no trata de responder a una pregunta, sino entender los datos de entrada. Su labor es determinar características comunes entre los objetos y hacer grupos homogéneos. Se utiliza en dos campos fundamentales: las técnicas de agrupamiento o *clustering* y las reglas de asociación. En el primer caso se trata de encontrar, por ejemplo, grupos de pacientes con características similares o grupos de fármacos con efectos adversos similares. Las reglas de asociación se utilizan en el comercio para la oferta de productos y en medicina para la detección de comorbilidades.

Finalmente, el *aprendizaje por refuerzo* trata de resolver un problema concreto a partir de unas reglas iniciales. El agente recibe una recompensa o un castigo según la decisión tomada. Está basado, por tanto, en la exploración y el aprovechamiento de la interacción y su objetivo es encontrar la mejor acción que maximice la recompensa. Está dirigido por el error y enfocado a enseñar a un agente a aprender. Un buen ejemplo de ello lo constituyen los sistemas de IA para jugar al ajedrez.

Diversos modelos de ML y DL han mostrado su utilidad en distintas aplicaciones médicas y sanitarias (10,11). Según la revisión de Jiang et al. (12), el aprendizaje supervisado es el método recomendado para la atención sanitaria, ya que proporciona resultados clínicamente más relevantes. Según estos autores los algoritmos de máquinas de vectores de soporte (*support vector machine, SVM*) y las redes neuronales (*neural network, NN*) son

los más utilizados para aplicaciones médicas, La tabla 1 recoge las características en que se basan estos y otros algoritmos de uso frecuente en medicina y sanidad.

Por otra parte, el procesamiento del lenguaje natural (*natural language processing*, NLP) es otra área de la IA utilizada con frecuencia en el ámbito sanitario y en la investigación biomédica (13,14). El NLP permite que las computadoras comprendan, analicen, manipulen y potencialmente generen el lenguaje humano. La combinación de algoritmos de NLP y ML puede permitir la realización de tareas complejas, como las que llevan a cabo los asistentes virtuales que se usan en el domicilio o la codificación automatizada de documentos clínicos (15).

En la evaluación del rendimiento de los algoritmos se comparan sus resultados con los obtenidos en la realidad o con los logrados por un patrón oro. Las métricas más habituales son la exactitud (*accuracy*), precisión (P), cobertura (sensibilidad, exhaustividad, *recall*, R) y la puntuación F1. Para calcularlas es preciso identificar los casos verdaderos positivos (tp) como la suma de registros identificados correctamente, los falsos negativos (fn) como la suma de registros no identificados y los falsos positivos (fp) como la suma de registros recuperados incorrectamente. La tabla 2 resume las definiciones de algunas de estas métricas.

La exactitud (*accuracy*) del algoritmo es el cociente entre resultados verdaderos (positivos y negativos) y el total de sujetos analizados. Representa el porcentaje de casos en que el modelo ha acertado y se calcula como  $(tp + tn)/(tp + tn + fp + fn)$ . La precisión indica la fiabilidad de la información recuperada por el sistema, es decir, el valor predictivo positivo o porcentaje de clasificaciones correctas del clasificador. Se calcula como  $P = tp/(tp + fp)$ . Por su parte, la cobertura o sensibilidad (*recall*) es un indicador de la cantidad de información que el sistema recupera. Representa el porcentaje de los positivos reales de nuestro conjunto de datos, detectados o no por el clasificador. Su cálculo es el cociente entre los verdaderos positivos y la suma de verdaderos positivos y falsos negativos, es decir,  $R = tp/(tp + fn)$ .

El valor F1 se utiliza para combinar las medidas de precisión y *recall* en un sólo valor. Se calcula como la media armónica de las dos anteriores, es decir,  $F = 2 \times P \times R / (P + R)$ . Este parámetro ofrece un indicador del rendimiento global de la recuperación de la información (13,14).

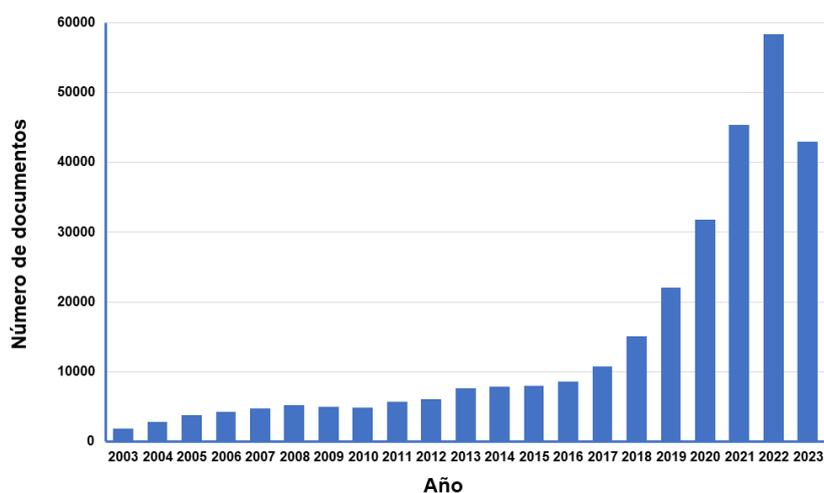
Con los valores anteriormente descritos pueden trazarse las curvas PR (*precision-recall*) y las curvas ROC (*receiver operating characteristic*). Las curvas PR relacionan precisión con la cobertura y nos permite ver gráficamente a partir de qué *recall* tenemos una degradación de la precisión y viceversa. La curva ideal en esta representación es la que se acerca lo máximo posible a la esquina superior derecha (alta precisión y alto *recall*). La curva ROC es una representación gráfica de la sensibilidad frente a la inespecificidad (1-especificidad)

para un sistema clasificador binario según se varía el umbral de discriminación. En otras palabras, la curva ROC relaciona la tasa de verdaderos positivos (*recall*) con la tasa de falsos positivos, es decir, relaciona la sensibilidad de nuestro modelo con los fallos optimistas (clasificar los negativos como positivos). Generalmente, a medida que se aumenta el *recall* el modelo tenderá a ser más optimista e introducirá más falsos positivos en la clasificación. Por ello, la curva ROC ideal es la que se aproxima lo máximo posible a la esquina superior izquierda (alto *recall* y baja tasa de falsos positivos).

## IMPORTANCIA DE LA INTELIGENCIA ARTIFICIAL EN MEDICINA

El empleo de programas y computadoras como ayuda al diagnóstico médico no es nuevo en medicina y ya se empleó en la segunda mitad del siglo XX (16,17). Sin embargo, ha sido en los últimos años cuando el interés y los avances en las aplicaciones médicas de IA han aumentado en paralelo al incremento en la potencia de los modernos ordenadores y a la gran cantidad de datos disponibles para su recopilación y utilización (18).

La importancia que va adquiriendo en los últimos años la tecnología de la IA y ML en medicina puede apreciarse a través de una visión de las publicaciones científicas de los últimos años en esta materia. Una búsqueda en la base de datos PubMed realizada el 9 de septiembre de 2023 con los criterios “artificial intelligence” OR “machine learning” OR “deep learning” obtuvo como resultado 285.207 documentos. La evolución en los últimos 20 años puede verse en la figura 2 que muestra claramente una tendencia ascendente en la producción científica en esta área de conocimiento.



**Figura 2.** Evolución de la producción científica anual recogida en la base de datos de ciencias biomédicas PubMed (búsqueda realizada el 9 de septiembre de 2023) con las palabras clave “artificial intelligence” OR “machine learning” OR “deep learning”.

## INTELIGENCIA ARTIFICIAL EN EL ÁMBITO SANITARIO

---

La IA está cambiando gradualmente la práctica de la atención sanitaria. Existen varias aplicaciones de IA que se pueden utilizar en diversos campos médicos, como las prácticas clínicas, de diagnóstico, de rehabilitación, quirúrgicas y predictivas. Otra área crítica de la medicina donde la IA está teniendo un impacto es la toma de decisiones clínicas ante pacientes concretos. Las tecnologías de IA pueden ingerir, analizar e informar grandes volúmenes de datos en diferentes modalidades para detectar enfermedades y guiar las decisiones clínicas (19). Por otro lado, las aplicaciones de IA pueden manejar la gran cantidad de datos producidos en medicina y encontrar nueva información que de otro modo permanecería oculta en la enorme masa de *big data* médicos (20-22).

De una forma general, en el ámbito sanitario y gestor, las principales aplicaciones de la IA se pueden agrupar en cuatro apartados: gestión, medicina predictiva, toma de decisiones y datos del paciente. La *gestión de servicios sanitarios* incluye el apoyo a la gestión integral de los centros sanitarios, las actualizaciones de información en tiempo real con procedencia de diversas fuentes (p. ej., Covid-19), así como la coordinación de herramientas de información para pacientes. También se consideran incluidas dentro de este grupo las aplicaciones destinadas a mejorar la eficiencia de los hospitales, incluyendo las que se dirigen específicamente al personal sanitario (médicos, enfermeras) y las dirigidas a los propios pacientes. Existen también sistemas destinados a la mejora de la logística de los hospitales (distribución de medicamentos, manejo de equipamientos) y a la reducción de la heterogeneidad entre hospitales (23-26).

Algunas de las aplicaciones en el ámbito de la *medicina predictiva* están destinadas a identificar relaciones significativas en datos masivos sin procesar. Otras, sin embargo, tienen por objetivo respaldar los resultados de diagnóstico, tratamiento y predicción que se producen en muchas situaciones médicas. En esta categoría también pueden incluirse las aplicaciones que permiten a los profesionales médicos adoptar el manejo proactivo del inicio de la enfermedad o bien identificar los factores de riesgo de cada paciente para ayudar a orientar las intervenciones de atención médica y obtener así los mejores resultados. Otros empleos en el ámbito de la prevención implican el diseño y desarrollo de nuevos medicamentos, la monitorización de pacientes y la personalización de los planes terapéuticos. Finalmente, algunos autores han desarrollado modelos de predicción para medicamentos y exámenes que permiten la vigilancia y control de los pacientes afectados por padecimientos crónicos (27-29).

Las aplicaciones del ámbito de la *toma de decisiones* persiguen ayudar a médicos y gestores en la toma de decisiones clínicas en una gran variedad de áreas funcionales específicas. De este modo se pretende incrementar el rendimiento de las decisiones clínicas al acelerar los

procesos y la cantidad de prestaciones brindadas, con un impacto positivo en el coste de los servicios de salud. Estos procesos están en línea con una asistencia virtual para la resolución de problemas de gestión clínica (12,30,31).

El área de *datos del paciente y diagnóstico* incluye de manera primordial el manejo de los *big data* médicos, es decir, la gran masa de datos digitales recolectados de diferentes fuentes. Las técnicas de ML y *data mining* puede ayudar en el procesamiento de esta información y facilitar un conocimiento útil para médicos y pacientes. Los sistemas de IA pueden gestionar datos generados a partir de actividades clínicas, como la detección, el diagnóstico y la asignación de tratamientos. Son útiles también en la gestión y protección de datos personales de los pacientes, así como en la automatización de procedimientos diagnósticos para evitar el error humano manteniendo un alto nivel de exactitud y precisión. Finalmente, existen aplicaciones destinadas a facilitar el diagnóstico a distancia mediante telemedicina con observación remota del paciente (32-35).

## IA EN EL ÁMBITO CLÍNICO

---

Conocimiento y experiencia son probablemente dos de las cualidades principales que le exigimos a un clínico para la práctica de su profesión. Sin embargo, las capacidades humanas tienen sus límites. Se calcula que durante 40 años de actividad profesional, un radiólogo podría estudiar aproximadamente 200.000 imágenes de resonancia magnética o tomografía axial (36). Sin embargo, la IA puede comenzar con este número y, en un corto período de tiempo, llegar a millones de exploraciones, mejorando así aún más su precisión. Este ejemplo deja claro que la IA puede emplearse para aumentar notablemente las capacidades de los médicos. La tecnología del ML puede realizar tareas rutinarias y sistemáticas con alta coherencia, liberando tiempo a los médicos para que puedan dedicarlo a problemas clínicos que son más complejos o que requieren una interacción humana sustancial. Se han desarrollado algunas plataformas de uso clínico que funcionan como un sistema de producción que alberga herramientas modulares de ML. Estas plataformas utilizan datos de las historias clínicas electrónicas para hacer predicciones sobre los pacientes. Al mismo tiempo, también pueden actuar como herramientas de investigación al vincular y agregar múltiples bases de datos clínicos y de otra índole para facilitar el desarrollo conjunto de herramientas de aprendizaje automático (37).

Siguiendo los criterios de Eggers et al. (38) existen cuatro enfoques principales para la automatización y la ayuda de la IA a los clínicos: aliviar, separar, reemplazar y aumentar. Las tecnologías de la información permiten simplificar algunas rutinas y, por tanto, *aliviar la carga de trabajo* de los profesionales. Muchas aplicaciones facilitan las tareas administrativas necesarias en todos los ámbitos sanitarios. Otras permiten realizar síntesis

de historias clínicas y resumen de problemas. La IA puede también aliviar la carga de trabajo en las disciplinas con especial dedicación al diagnóstico por la imagen. Es conocido que el DL, por ejemplo, ha mostrado un rendimiento sólido en la detección de diversas afecciones médicas, incluida la retinopatía diabética o el cáncer de mama. Existen aplicaciones que funcionan como herramientas de detección inicial en la interpretación de exploraciones, dando prioridad a los casos que presentan mayor gravedad para que la atención del médico pueda dirigirse a situaciones críticas más rápidamente. En algunas especialidades, como cardiología y radiología, la IA se puede utilizar como herramienta preliminar para examinar imágenes e incluso completar parcialmente informes para que los médicos los revisen y validen. Por otra parte, las nuevas tecnologías pueden facilitar a los clínicos recomendaciones de no hacer pruebas irrelevantes en la práctica clínica, lo cual permitiría ahorrar tiempo, recursos y probables incomodidades a muchos pacientes (39-44).

La IA puede facilitar el trabajo de los clínicos *dividiendo tareas*. Los algoritmos pueden emplearse para reducir los ingresos hospitalarios evitando los innecesarios antes de que ocurran. Algunas aplicaciones de ML incluyen *chatbots* que pueden realizar aproximaciones a un diagnóstico médico a través de una serie de preguntas, dirigiendo a los pacientes hacia la ruta de atención médica adecuada. También existen aplicaciones para teléfonos inteligentes que ofrecen a los pacientes consejos de salud personalizados sin necesidad de su traslado al hospital. Esto puede ser especialmente útil en el manejo de enfermedades crónicas frecuentes como la diabetes o la hipertensión. Aunque existe un debate sobre la seguridad de tales aplicaciones, no debemos dejar de considerar la IA como una oportunidad con gran potencial para pacientes y profesionales (45-47).

Algunas de las tareas que normalmente desempeñan los profesionales sanitarios podrían ser desempeñadas por la IA. El *reemplazo de tareas* de médicos y enfermeras puede ser especialmente relevante en algunas labores repetitivas como el cribado de la retinopatía diabética que es fácil de automatizar. Ello permitiría dirigir la actividad de los profesionales a tareas complejas que la IA no puede replicar, especialmente aquellas que implican una interacción directa con el paciente. Además, nuevamente este tipo de aplicaciones tendrían como consecuencia una reducción de costes en determinadas prestaciones (48,49).

La IA ha adquirido un enorme potencial para *mejorar la práctica clínica* y la atención al paciente. La tecnología puede llegar a presentar habilidades cuantitativas más allá de lo que los humanos son capaces de hacer y con un mayor nivel de precisión y detalle. Algunos ejemplos son las aplicaciones de diagnóstico por imagen o los métodos predictivos de comorbilidades o complicaciones. Las mejoras en las precisiones cuantitativas permitirían una reducción de los errores humanos a los que los médicos están expuestos en sus tareas diarias, especialmente en situación de sobrecarga de trabajo y estrés. Aparte de colaborar en la reducción del número de errores accidentales, la IA también puede permitir una mejor

toma de decisiones clínicas al proporcionar orientación actualizada sobre directrices o guías clínicas. Ello facilitaría un aumento del conocimiento médico objetivo y una reducción de sesgos inconscientes. Las técnicas de NLP también pueden incrementar el conocimiento clínico mediante la caracterización y clasificación de enfermedades al descubrir palabras clave relevantes que aparecen comúnmente en el diagnóstico o que se vinculan con resultados adversos. Es decir, estos procedimientos pueden identificar patrones que pueden influir en los resultados de los pacientes y generar hipótesis para explicarlos. Finalmente, la IA es extremadamente útil para la inferencia a partir de datos complejos, por lo que potencialmente es una herramienta apropiada para el análisis de conjuntos de datos biológicos. Un ejemplo puede ser el análisis de datos genómicos y clínicos, la clasificación de variantes y las correspondencias fenotipo-genotipo. La integración de diferentes conjuntos de datos como la genómica, transcriptómica, proteómica y otras *ómicas* para la comprensión de la complejidad biológica de la enfermedad es un objetivo de investigación en muchas áreas de la medicina actual en las que la IA tendrá mucho que aportar en los próximos años (12,48,50).

## INCONVENIENTES Y RIESGOS DE LA IA

---

Como cualquier otra tecnología la IA no está desprovista de inconvenientes. Algunos de ellos están directamente relacionados con los *datos* que se manejan. Las cuestiones que plantean problemas y que son objeto de debate en este apartado incluyen la disponibilidad de los datos, la propiedad de los datos de cada hospital para manejar sus pacientes individuales y la disponibilidad de los datos después de la introducción de algoritmos. Existen también preocupaciones razonables sobre la privacidad y seguridad de los datos, especialmente en lo que se refiere a datos sensibles como son los relativos a la salud y la enfermedad de las personas. Además, los sistemas informáticos deben asegurar de algún modo la adecuada calidad de los datos para que puedan ser no sólo útiles sino fiables en sus aplicaciones clínicas, y también su longevidad o tiempo de vida útil (12,51-53).

Los inconvenientes relacionados con el *desarrollo tecnológico* incluyen los posibles sesgos en la introducción de datos. Si los datos de entrenamiento de los algoritmos están sesgados por alguna razón, los resultados obtenidos por los sistemas pueden resultar contraproducentes. El *overfitting* o sobreajuste se produce cuando el algoritmo utiliza demasiadas variables en el conjunto de datos de entrenamiento y predice relaciones inadecuadas y resultados inseguros. El *data leakage* o fuga de datos es otro de los problemas relacionados con el desarrollo tecnológico. Los algoritmos de DL con frecuencia no pueden proporcionar explicaciones detalladas de sus predicciones. Esto es lo que constituye la *black-box* o caja negra. En el caso de que un algoritmo genere recomendaciones erróneas se

puede generar un problema desde el punto de vista jurídico, ya que el sistema no puede justificarse por sí mismo. La caja negra constituye una barrera para la comprensión de la conexión entre los datos y las predicciones y, más aún, puede socavar la confianza de los pacientes en el sistema (54-56).

Se deben considerar también las desventajas relacionadas con la *implementación clínica* de la tecnología. La falta de evidencia de las intervenciones basadas en IA en ensayos clínicos prospectivos es un inconveniente serio relacionado con su aplicación directa en la práctica clínica. Pueden existir dudas sobre la calidad de las investigaciones y sobre la forma de integrar los métodos de IA en el flujo de trabajo del médico. La aplicación de las nuevas herramientas debe facilitar, pero nunca entorpecer, el trabajo de los profesionales para que consigan una valoración positiva y una reducción de cargas. Todo ello implica tiempo y recursos dedicados a capacitar a los proveedores de atención sanitaria para utilizar la tecnología. Otra posible barrera para una integración de la IA en la práctica clínica es la participación de los diferentes actores con distintos intereses en el proceso de desarrollo. Las opiniones de múltiples *stakeholders* son vitales para garantizar que el producto resultante pueda integrarse en el flujo de trabajo de los médicos. En ocasiones se han desarrollado muchas innovaciones destinadas a mejorar las predicciones o diagnósticos que no han resultado exentas de dificultades y problemas en su implementación e integración en el trabajo clínico (47,53,57-59).

Existen también impedimentos que tienen una raíz *social*. Es comprensible la preocupación de que la IA en la atención sanitaria reemplace puestos de trabajo, desplazando así a algunos profesionales. Ello se traduce en desconfianza y temor hacia las intervenciones basadas en IA. Se han señalado también en este grupo de desventajas sociales los relacionados con una posible inequidad en las prestaciones sanitarias. La falta de una adecuada comprensión de estas tecnologías, así como la sobreestimación de sus capacidades actuales constituyen también algunos inconvenientes relacionados con factores sociales (52,59).

La adopción generalizada de métodos de IA en la práctica médica podría generar el riesgo de reducción de las habilidades clínicas de los médicos y, en un caso más extremo, el de sustitución de los profesionales por herramientas informáticas. Es cierto que los médicos están cada vez más expuestos a la tecnología con diferentes propósitos y que esta tecnología incita, cada vez con más consistencia, a una toma de decisiones basada en herramientas innovadoras (23). Ello ha llevado a algunos autores a pensar que el uso excesivo de la tecnología podría dificultar la expansión de las habilidades de los médicos y de los procedimientos clínicos tal como los conocemos ahora. La consecuencia que se seguiría de ello sería una posible descualificación del personal sanitario debido a la reducción de la autonomía en la toma de decisiones sobre los pacientes (60). Pero esto no necesariamente tiene que ser así. En el momento actual el riesgo de una posible sustitución de la figura del

médico aún queda lejos. Aunque, desde un punto de vista teórico, la IA puede reemplazar a los médicos en lo que se refiere a la realización de ciertas tareas diagnósticas o terapéuticas, es casi seguro que estas tecnologías nunca podrán brindar el apoyo emocional, social y ético que un médico puede ofrecer al paciente, y nunca reemplazará el vínculo único de una relación médico-paciente (61,62).

## ASPECTOS ÉTICOS

---

La *privacidad* de los datos y la *seguridad* de su custodia son preocupaciones de tipo ético que han acompañado a la IA desde sus inicios y cuya importancia ya se ha comentado anteriormente. Pero quizá el problema de la *responsabilidad* es el aspecto ético que mayores dudas genera en el momento actual. Las decisiones que se toman en el ámbito de la salud, basadas o no en herramientas informáticas, pueden tener consecuencias muy graves si no son adecuados o se demuestra que ha habido algún error o negligencia. En el caso del uso de aplicaciones de IA para la toma de decisiones se plantea la duda de a quién exigir responsabilidad en caso de daño o perjuicio.

Quién es el responsable de las predicciones o recomendaciones generadas es una cuestión que puede ser difícil de contestar con ecuanimidad. No está claro quién debería asumir la responsabilidad en caso de que una herramienta no funcione correctamente. Responsabilizar al clínico puede parecer injusto, ya que los médicos no desarrollan ni controlan los algoritmos. Responsabilizar al desarrollador parece demasiado alejado del contexto clínico. Es necesario que cualquier decisión que se tome con trascendencia para la salud de las personas sea supervisada y tenga aportaciones de clínicos con experiencia (41,51,63).

Según las directrices de la Comisión Europea (64), la IA debe ser legal (debe respetar todas las leyes y regulaciones aplicables), robusta (tanto desde una perspectiva técnica como teniendo en cuenta su entorno social) y ética (con respeto a los principios y valores éticos). Se hacen necesarias, pues, pautas, códigos o guías sobre cómo manejar en la práctica las tecnologías de IA.

Aunque no disponemos de un estándar ético universal que proteja y beneficie a los pacientes, existe un abundante corpus legislativo y regulador que puede ser aplicado a las herramientas de IA y ML. El artículo 2 del Convenio relativo a los derechos humanos y a la biomedicina del Consejo de Europa (Convenio de Oviedo) especifica que el interés y bienestar del ser humano deberán prevalecer sobre el interés exclusivo de la sociedad y de la ciencia. Su objetivo fue impedir el abuso del desarrollo tecnológico en lo que concierne a la biomedicina y proteger la dignidad y los derechos humanos (65). Posteriormente, el

Reglamento del Parlamento Europeo y del Consejo de 27 de abril de 2016, relativo a la protección de datos de las personas físicas (RGPD) (66) considera que la protección de las personas físicas en relación con el tratamiento de datos personales es un derecho fundamental y que los principios y normas relativos a la protección de las personas físicas en lo que respecta al tratamiento de sus datos de carácter personal deben, cualquiera que sea su nacionalidad o residencia, respetar sus libertades y derechos fundamentales, en particular el derecho a la protección de los datos de carácter personal. Por otro lado, la Ley Orgánica 3/2018, de Protección de datos personales y garantía de los derechos digitales (67) considera que la protección de datos personales es un derecho fundamental de las personas físicas y se plantea como objetivos la adaptación del ordenamiento jurídico español al Reglamento Europeo y la garantía de los derechos digitales de la ciudadanía en conformidad con el mandato establecido en el artículo 18.4 de la Constitución Española.

Tiene interés destacar un informe del comité internacional de ética de la UNESCO (68) que considera que la ética en la IA tiene un componente dinámico por la evolución continua de las tecnologías. Este informe reconoce una serie de valores y principios (tabla 3) que deberían ser respetados por todos los actores durante el ciclo de vida de los sistemas de IA y, cuando resulte necesario y conveniente, ser promovidos mediante modificaciones de las leyes, los reglamentos y las directrices empresariales existentes. Todo ello debe ajustarse al derecho internacional, en particular la Carta de las Naciones Unidas y las obligaciones de los estados miembros en materia de derechos humanos, y estar en consonancia con los objetivos de sostenibilidad social, política, ambiental, educativa, científica y económica acordados internacionalmente. En resumen, la ética en IA ha de promover el respeto a la dignidad humana y la igualdad, salvaguardar los intereses de generaciones presentes y futuras, proteger los derechos humanos, y las libertades fundamentales.

## REFERENCIAS BIBLIOGRÁFICAS

---

1. TRAN BX, VU GT, HA GH, VUONG QH, HO MT, VUONG TT, LA VP, HO MT, NGHIEM KP, NGUYEN HLT, LATKIN CA, TAM WWS, CHEUNG NM, NGUYEN HT, HO CSH, HO RCM. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. *J Clin Med*. 2019 Mar 14;8(3):360. doi: 10.3390/jcm8030360.
2. SECINARO S, CALANDRA D, SECINARO A, MUTHURANGU V, BIANCONE P. The role of artificial intelligence in healthcare: a structured literature review. *BMC Med Inform Decis Mak*. 2021 Apr 10;21(1):125. doi: 10.1186/s12911-021-01488-9.
3. BISHOP CM. *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.

4. BEAULIEU-JONES B, FINLAYSON SG, CHIVERS C, CHEN I, MCDERMOTT M, KANDOLA J, DALCA AV, BEAM A, FITERAU M, NAUMANN T. Trends and Focus of Machine Learning Applications for Health Research. *JAMA Netw Open*. 2019 Oct 2;2(10):e1914051. doi: 10.1001/jamanetworkopen.2019.14051.
5. WU S, ROBERTS K, DATTA S, DU J, JI Z, SI Y, SONI S, WANG Q, WEI Q, XIANG Y, ZHAO B, XU H. Deep learning in clinical natural language processing: a methodical review. *J Am Med Inform Assoc*. 2020 Mar 1;27(3):457-470. doi: 10.1093/jamia/ocz200.
6. ARTZI NS, SHILO S, HADAR E, ROSSMAN H, BARBASH-HAZAN S, BEN-HAROUSH A, BALICER RD, FELDMAN B, WIZNITZER A, SEGAL E. Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med*. 2020 Jan;26(1):71-76. doi: 10.1038/s41591-019-0724-8.
7. OROOJENI MOHAMMAD JAVAD M, AGBOOLA SO, JETHWANI K, ZEID A, KAMARTHI S. A Reinforcement Learning-Based Method for Management of Type 1 Diabetes: Exploratory Study. *JMIR Diabetes*. 2019 Aug 28;4(3):e12905. doi: 10.2196/12905.
8. SHOMORONY I, CIRULLI ET, HUANG L, NAPIER LA, HEISTER RR, HICKS M, COHEN IV, YU HC, SWISHER CL, SCHENKER-AHMED NM, LI W, NELSON KE, BRAR P, KAHN AM, SPECTOR TD, CASKEY CT, VENTER JC, KAROW DS, KIRKNESS EF, SHAH N. An unsupervised learning approach to identify novel signatures of health and disease from multimodal data. *Genome Med*. 2020 Jan 10;12(1):7. doi: 10.1186/s13073-019-0705-z.
9. BALLINGER B, HSIEH J, SINGH A, SOHONI N, WANG J, TISON GH, MARCUS GM, SANCHEZ JM, MAGUIRE C, OLGIN JE, PLETCHER MJ. DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction [Internet]. arXiv:1802.025111. <https://doi.org/10.48550/arXiv.1802.02511>.
10. MANICKAM P, MARIAPPAN SA, MURUGESAN SM, HANSDA S, KAUSHIK A, SHINDE R, THIPPERUDRASWAMY SP. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors (Basel)*. 2022 Jul 25;12(8):562. doi: 10.3390/bios12080562.
11. RECH MM, DE MACEDO FILHO L, WHITE AJ, PEREZ-VEGA C, SAMSON SL, CHAICHANA KL, OLOMU OU, QUINONES-HINOJOSA A, ALMEIDA JP. Machine Learning Models to Forecast Outcomes of Pituitary Surgery: A Systematic Review in Quality of Reporting and Current Evidence. *Brain Sci*. 2023 Mar 15;13(3):495. doi: 10.3390/brainsci13030495.
12. JIANG F, JIANG Y, ZHI H, DONG Y, LI H, MA S, WANG Y, DONG Q, SHEN H, WANG Y. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017 Jun 21;2(4):230-243. doi: 10.1136/svn-2017-000101.

13. DÍEZ JJ, CABRERA L, IGLESIAS P, BENAVENT M, ARGÜELLO G, LÓPEZ G, PARRALEJO A, LEAL J. Prevalence of cancer in patients with hypothyroidism: Analysis using big data tools. *Endocrinol Diabetes Nutr (Engl Ed)*. 2023 Sep;70 Suppl 3:50-58. doi: 10.1016/j.endien.2023.08.004. Epub 2023 Aug 17. PMID: 37598005.
14. DÍEZ JJ, CABRERA L, IGLESIAS P, BENAVENT M, LÓPEZ G, ARGÜELLO G, PARRALEJO A, LÓPEZ-VELÁZQUEZ, A. Thyroid carcinoma in elderly people: Characterization using big data tools. *Endocrinol Diabetes Nutr (Engl Ed)*. 2023 Mar;70(3):179-188. doi: 10.1016/j.endien.2023.03.011.
15. FRIEDMAN C, SHAGINA L, LUSSIER Y, HRIPCSAK G. Automated encoding of clinical documents based on natural language processing. *J Am Med Inform Assoc*. 2004 Sep-Oct;11(5):392-402. doi: 10.1197/jamia.M1552.
16. YANG X, WANG Y, BYRNE R, SCHNEIDER G, YANG S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem Rev*. 2019 Sep 25;119(18):10520-10594. doi: 10.1021/acs.chemrev.8b00728.
17. BURTON RJ, ALBUR M, EBERL M, CUFF SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. *BMC Med Inform Decis Mak*. 2019 Aug 23;19(1):171. doi: 10.1186/s12911-019-0878-9.
18. MESKÓ B, DROBNI Z, BÉNYEI É, GERGELY B, GYÓRFFY Z. Digital health is a cultural transformation of traditional healthcare. *Mhealth*. 2017 Sep 14;3:38. doi: 10.21037/mhealth.2017.08.07.
19. CHO BJ, CHOI YJ, LEE MJ, KIM JH, SON GH, PARK SH, KIM HB, JOO YJ, CHO HY, KYUNG MS, PARK YH, KANG BS, HUR SY, LEE S, PARK ST. Classification of cervical neoplasms on colposcopic photography using deep learning. *Sci Rep*. 2020 Aug 12;10(1):13652. doi: 10.1038/s41598-020-70490-4.
20. DOYLE OM, LEAVITT N, RIGG JA. Finding undiagnosed patients with hepatitis C infection: an application of artificial intelligence to patient claims data. *Sci Rep*. 2020 Jun 29;10(1):10521. doi: 10.1038/s41598-020-67013-6.
21. SHORTLIFFE EH, SEPÚLVEDA MJ. Clinical Decision Support in the Era of Artificial Intelligence. *JAMA*. 2018 Dec 4;320(21):2199-2200. doi: 10.1001/jama.2018.17163.
22. MASSARO M, DUMAY J, GUTHRIE J. On the shoulders of giants: undertaking a structured literature review in accounting. *Account Auditing Account J*. 2016;29(5):767-801. <https://doi.org/10.1108/AAAJ-01-2015-1939>.
23. HUSSAIN AA, BOUACHIR O, AL-TURJMAN F, ALOQAILY M. AI Techniques for COVID-19. *IEEE Access*. 2020 Jul 8;8:128776-128795. doi: 10.1109/ACCESS.2020.3007939.

24. CHAKRADHAR S. Predictable response: Finding optimal drugs and doses using artificial intelligence. *Nat Med.* 2017 Nov 7;23(11):1244-1247. doi: 10.1038/nm1117-1244. Erratum in: *Nat Med.* 2017 Dec 7;23 (12 ):1391.
25. FLEMING N. How artificial intelligence is changing drug discovery. *Nature.* 2018 May;557(7707):S55-S57. doi: 10.1038/d41586-018-05267-x.
26. GUO J, LI B. The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries. *Health Equity.* 2018 Aug 1;2(1):174-181. doi: 10.1089/heq.2018.0037.
27. AGRAWAL A, GANS JS, GOLDFARB A. Exploring the impact of artificial Intelligence: Prediction versus judgment. *Inf Econ Policy* 2019;1:1. <https://doi.org/10.1016/j.infoecopol.2019.05.001>.
28. DAVENPORT T, KALAKOTA R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019 Jun;6(2):94-98. doi: 10.7861/futurehosp.6-2-94.
29. COLLINS GS, MOONS KGM. Reporting of artificial intelligence prediction models. *Lancet.* 2019 Apr 20;393(10181):1577-1579. doi: 10.1016/S0140-6736(19)30037-6.
30. BENNETT CC, HAUSER K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. *Artif Intell Med.* 2013 Jan;57(1):9-19. doi: 10.1016/j.artmed.2012.12.003.
31. REDONDO T, SANDOVAL AM. Text Analytics: the convergence of big data and artificial intelligence. *Int J Interact Multimed Artif Intell.* 2016;3:2540. <https://www.ijimai.org/journal/bibcite/reference/2540>.
32. WINTER JS, DAVIDSON E. Big data governance of personal health information and challenges to contextual integrity. *Inf Soc.* 2019;35(1):36-51. doi:10.1080/01972243.2018.1542648
33. NOVAK D, RIENER R. Control strategies and artificial intelligence in rehabilitation robotics. *AI Magazine*, 36(4), 23-33. <https://doi.org/10.1609/aimag.v36i4.2614>.
34. TARASSOLI SP. Artificial intelligence, regenerative surgery, robotics? What is realistic for the future of surgery? *Ann Med Surg (Lond).* 2019 Apr 17;41:53-55. doi: 10.1016/j.amsu.2019.04.001.
35. GU D, LI T, WANG X, YANG X, YU Z. Visualizing the intellectual structure and evolution of electronic health and telemedicine research. *Int J Med Inform.* 2019 Oct;130:103947. doi: 10.1016/j.ijmedinf.2019.08.007.
36. YOKOTA H, GOTO M, BAMBA C, KIBA M, YAMADA K. Reading efficiency can be improved by minor modification of assigned duties; a pilot study on a small team of general radiologists. *Jpn J Radiol.* 2017 May;35(5):262-268. doi: 10.1007/s11604-017-0629-8.

37. NGIAM KY, KHOR IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019 May;20(5):e262-e273. doi: 10.1016/S1470-2045(19)30149-4.
38. EGGERS WD, SCHATSKY D, VIECHNICKI P. AI-augmented government: Using cognitive technologies to redesign public sector work A report from the Deloitte Center for Government Insights [Internet]. 2017. <https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/artificial-intelligence-government.html>.
39. VERGHESE A, SHAH NH, HARRINGTON RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. *JAMA.* 2018 Jan 2;319(1):19-20. doi: 10.1001/jama.2017.19198.
40. SPENCER M. Brittleness and bureaucracy: software as a material for science. *Perspect Sci* 2015;23:466–84. [http://www.mitpressjournals.org/doi/10.1162/POSC\\_a\\_00184](http://www.mitpressjournals.org/doi/10.1162/POSC_a_00184).
41. TING DSW, CHEUNG CY, LIM G, TAN GSW, QUANG ND, GAN A, HAMZAH H, GARCIA-FRANCO R, SAN YEO IY, LEE SY, WONG EYM, SABANAYAGAM C, BASKARAN M, IBRAHIM F, TAN NC, FINKELSTEIN EA, LAMOUREUX EL, WONG IY, BRESSLER NM, SIVAPRASAD S, VARMA R, JONAS JB, HE MG, CHENG CY, CHEUNG GCM, AUNG T, HSU W, LEE ML, WONG TY. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *JAMA.* 2017 Dec 12;318(22):2211-2223. doi: 10.1001/jama.2017.18152.
42. LAKHANI P, SUNDARAM B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology.* 2017 Aug;284(2):574-582. doi: 10.1148/radiol.2017162326.
43. MCKINNEY SM, SIENIEK M, GODBOLE V, GODWIN J, ANTROPOVA N, ASHRAFIAN H, BACK T, CHESUS M, CORRADO GS, DARZI A, ETEMADI M, GARCIA-VICENTE F, GILBERT FJ, HALLING-BROWN M, HASSABIS D, JANSEN S, KARTHIKESALINGAM A, KELLY CJ, KING D, LEDSAM JR, MELNICK D, MOSTOFI H, PENG L, REICHER JJ, ROMERA-PAREDES B, SIDEBOTTOM R, SULEYMAN M, TSE D, YOUNG KC, DE FAUW J, SHETTY S. International evaluation of an AI system for breast cancer screening. *Nature.* 2020 Jan;577(7788):89-94. doi: 10.1038/s41586-019-1799-6.
44. ATTIA ZI, NOSEWORTHY PA, LOPEZ-JIMENEZ F, ASIRVATHAM SJ, DESHMUKH AJ, GERSH BJ, CARTER RE, YAO X, RABINSTEIN AA, ERICKSON BJ, KAPA S, FRIEDMAN PA. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet.* 2019 Sep 7;394(10201):861-867. doi: 10.1016/S0140-6736(19)31721-0.

45. ARMSTRONG S. The apps attempting to transfer NHS 111 online. *BMJ*. 2018 Jan 15;360:k156. doi: 10.1136/bmj.k156.
46. LUPTON D, JUTEL A. 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. *Soc Sci Med*. 2015 May;133:128-35. doi: 10.1016/j.socscimed.2015.04.004.
47. STEWART J, SPRIVULIS P, DWIVEDI G. Artificial intelligence and machine learning in emergency medicine. *Emerg Med Australas* [Internet] 2018;30:870-4. <http://doi.wiley.com/10.1111/1742-6723.13145>.
48. JONES LD, GOLAN D, HANNA SA, RAMACHANDRAN M. Artificial intelligence, machine learning and the evolution of healthcare: A bright future or cause for concern? *Bone Joint Res*. 2018 May 5;7(3):223-225. doi: 10.1302/2046-3758.73.BJR-2017-0147.R1.
49. TUFAIL A, RUDISILL C, EGAN C, KAPETANAKIS VV, SALAS-VEGA S, OWEN CG, LEE A, LOUW V, ANDERSON J, LIEW G, BOLTER L, SRINIVAS S, NITTALA M, SADDAS, TAYLOR P, RUDNICKA AR. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. *Ophthalmology*. 2017 Mar;124(3):343-351. doi: 10.1016/j.ophtha.2016.11.014.
50. CHENG JZ, NI D, CHOU YH, QIN J, TIU CM, CHANG YC, HUANG CS, SHEN D, CHEN CM. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Sci Rep*. 2016 Apr 15;6:24454. doi: 10.1038/srep24454.
51. LOPEZ K, FODEH SJ, ALLAM A, BRANDT CA, KRAUTHAMMER M. Reducing Annotation Burden Through Multimodal Learning. *Front Big Data*. 2020 Jun 2;3:19. doi: 10.3389/fdata.2020.00019.
52. TOPOL EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019 Jan;25(1):44-56. doi: 10.1038/s41591-018-0300-7.
53. DILSIZIAN SE, SIEGEL EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep*. 2014 Jan;16(1):441. doi: 10.1007/s11886-013-0441-8.
54. TING DSW, CHEUNG CY, LIM G, TAN GSW, QUANG ND, GAN A, HAMZAH H, GARCIA-FRANCO R, SAN YEO IY, LEE SY, WONG EYM, SABANAYAGAM C, BASKARAN M, IBRAHIM F, TAN NC, FINKELSTEIN EA, LAMOUREUX EL, WONG IY, BRESSLER NM, SIVAPRASAD S, VARMA R, JONAS JB, HE MG, CHENG CY, CHEUNG GCM, AUNG T, HSU W, LEE ML, WONG TY. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic

- Populations With Diabetes. *JAMA*. 2017 Dec 12;318(22):2211-2223. doi: 10.1001/jama.2017.18152.
55. KLEINBERG J, LUDWIG J, MULLAINATHAN S ET AL. Discrimination in the age of algorithms. *J Leg Anal* 2018;10:113-74.
56. WIENS J, SHENOY ES. Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. *Clin Infect Dis*. 2018 Jan 6;66(1):149-153. doi: 10.1093/cid/cix731.
57. RAMESH AN, KAMBHAMPATI C, MONSON JR, DREW PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004 Sep;86(5):334-8. doi: 10.1308/147870804290.
58. LIU X, FAES L, KALE AU, WAGNER SK, FU DJ, BRUYNSEELS A, MAHENDIRAN T, MORAES G, SHAMDAS M, KERN C, LEDSAM JR, SCHMID MK, BALASKAS K, TOPOL EJ, BACHMANN LM, KEANE PA, DENNISTON AK. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health*. 2019 Oct;1(6):e271-e297. doi: 10.1016/S2589-7500(19)30123-2. Epub 2019 Sep 25. Erratum in: *Lancet Digit Health*. 2019 Nov;1(7):e334.
59. MESKÓ B, HETÉNYI G, GYÓRFFY Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res*. 2018 Jul 13;18(1):545. doi: 10.1186/s12913-018-3359-4.
60. HOFF T. Deskillling and adaptation among primary care physicians using two work innovations. *Health Care Manage Rev*. 2011 Oct-Dec;36(4):338-48. doi: 10.1097/HMR.0b013e31821826a1.
61. GOLDHAHN J, RAMPTON V, SPINAS GA. Could artificial intelligence make doctors obsolete? *BMJ*. 2018 Nov 7;363:k4563. doi: 10.1136/bmj.k4563.
62. RIGBY MJ. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics*. (2019) 21:121-4. doi: 10.1001/amajethics.2019.121
63. WANG F, KAUSHAL R, KHULLAR D. Should Health Care Demand Interpretable Artificial Intelligence or Accept "Black Box" Medicine? *Ann Intern Med*. 2020 Jan 7;172(1):59-60. doi: 10.7326/M19-2548. Epub 2019 Dec 17.
64. Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission. Ethics guidelines for trustworthy AI. European Commission, 2019. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>. Visitado el 8 de septiembre de 2023.
65. Instrumento de Ratificación del Convenio para la protección de los derechos humanos y la dignidad del ser humano con respecto a las aplicaciones de la Biología y la Medicina

- (Convenio relativo a los derechos humanos y la biomedicina), hecho en Oviedo el 4 de abril de 1997. «BOE» núm. 251, de 20 de octubre de 1999, páginas 36825-36830. Disponible en: [https://www.boe.es/eli/es/ai/1997/04/04/\(1\)](https://www.boe.es/eli/es/ai/1997/04/04/(1)). Visitado el 7 septiembre 2023.
66. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). «DOUE» núm. 119, de 4 de mayo de 2016, páginas 1 a 88 (88 págs.). Disponible en: <https://www.boe.es/buscar/doc.php?id=DOUE-L-2016-80807>. Visitado el 7 septiembre 2023.
67. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. «BOE» núm. 294, de 06/12/2018. Disponible en: <https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673>. Visitado el 7 septiembre 2023.
68. UNESCO. Recomendación sobre la ética de la inteligencia artificial. Código de documento: SHS/BIO/PI/2021/1. París, UNESCO, 2022. Disponible en: [https://unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa). Visitado el 8 de septiembre de 2023.

TABLAS

Tabla 1. Algoritmos de inteligencia artificial más empleados en aplicaciones biomédicas\*

Algoritmo	Descripción
<b>Support Vector Machine (SVM)</b>	El algoritmo de la máquina de vectores de soporte (SVN) se utiliza para encontrar las líneas divisorias óptimas entre clases de datos de entrenamiento dentro de un espacio vectorial. Busca un límite de decisión, llamado hiperplano, que separe mejor las características de diferentes dominios. Su objetivo es encontrar el hiperplano que tenga el máximo grado de separación entre dos puntos más cercanos de cada clase, es decir, los vectores de soporte. Estos hiperplanos pueden luego usarse para proporcionar clasificaciones definitivas para nuevos datos de entrada. El algoritmo puede asignar ejemplos a otras dimensiones y operar en relaciones no lineales transformando datos de entrada de baja dimensión en un espacio de alta dimensión.
<b>Neural Network (NN)</b>	Las redes neuronales artificiales son algoritmos no lineales inspirados en las sinapsis del cerebro humano. Las redes neuronales convolucionales, las más comúnmente aplicadas, comprenden nodos de entrada, nodos de salida y capas de nodos intermedias u ocultas, que pueden llegar a 100. Cada nodo dentro de una capa involucra dos o más entradas y aplica una función de activación y ponderación para producir una salida que sirve como datos de entrada para la siguiente capa de nodos.
<b>Naïve Bayes (NB)</b>	Es un algoritmo basado en un modelo bayesiano, es decir, probabilístico. Las predicciones se basan en el cálculo de probabilidades según el teorema de Bayes. Se emplea para calcular la probabilidad de que un elemento pertenezca a cada una de dos o más clases en función de los valores de su vector de entrada. Normalmente, las probabilidades se obtienen a partir de datos de entrenamiento observando con qué frecuencia un elemento con cada valor de entrada pertenece a cada clase. El clasificador Naïve Bayes asume que la presencia o ausencia de una característica particular no está relacionada con la presencia o ausencia de cualquier otra característica, dada la clase variable.
<b>Discriminant análisis (DA)</b>	El análisis discriminante funciona derivando y combinando las funciones que calculan la probabilidad de que los valores de cada variable predictiva estén en cada clase. Tanto el análisis discriminante lineal como el análisis discriminante cuadrático suponen que cada variable predictiva tiene una distribución normal (gaussiana).
<b>k-Nearest neighbors (k-NN)</b>	El algoritmo del vecino más cercano se utiliza para clasificar elementos dentro de un espacio multidimensional para que correspondan a las clasificaciones de elementos preexistentes dentro de ese espacio. Funciona examinando los k vecinos más cercanos del nuevo elemento y asignándole la clasificación que ocurre con mayor frecuencia entre sus vecinos.
<b>Decision tree (DT)</b>	Los árboles de decisión se utilizan para la clasificación y predicción de valores como una alternativa a los métodos de mínimos cuadrados y regresión logística. En lugar de generar una función matemática compleja para predecir clasificaciones y valores, se genera un árbol donde se toma una decisión en cada bifurcación sobre qué camino tomar en función de los valores de una o más de las variables predictoras.
<b>Random forest (RF)</b>	Es un conjunto de DT, aunque se diferencia de los DT habituales al utilizar entradas o combinaciones de entradas seleccionadas aleatoriamente en cada nodo para hacer crecer cada árbol en lugar de un conjunto consistente. Se trata de ceder para evitar el sobreajuste que suele presentarse en los DT profundos. La distribución aleatoria de las entradas proporciona, cuando se promedia, menores tasas de error en el resultado final y una varianza reducida.

\*Adaptado de referencias 10 y 11

**Tabla 2.** Algunos conceptos y métricas utilizadas en los algoritmos de IA empleados en aplicaciones biomédicas\*

<b>Algoritmo</b>	<b>Descripción</b>
<b>Exactitud (accuracy)</b>	Mide el porcentaje de casos en que el modelo ha acertado. Se calcula como la relación entre el número total de predicciones que son correctas (verdaderos positivos y verdaderos negativos) y el total de predicciones.
<b>Cobertura (Sensibilidad / Recall)</b>	Proporción de predicciones verdaderamente positivas entre los casos que realmente son positivos, es decir, proporción de positivos detectado por nuestro clasificador en el conjunto de datos verdaderamente positivos.
<b>Especificidad</b>	Proporción de verdaderos negativos pronosticados correctamente que se identifican correctamente.
<b>Precisión (valor predictivo positivo, PPV)</b>	Proporción de verdaderos positivos pronosticados correctamente por el clasificador en el conjunto de pronosticados como positivos, es decir, porcentaje de verdaderos positivos entre los que el clasificador define como positivos correcta o incorrectamente.
<b>Valor predictivo negative (NPV)</b>	Proporción de negativos pronosticados correctamente entre todas las predicciones negativas.
<b>Puntuación F1</b>	Métrica compuesta definida como la media armónica entre precisión (PPV) y la cobertura ( <i>recall</i> ).
<b>Validación interna</b>	Evaluación del rendimiento de un modelo con los mismos datos o población, si es prospectiva, utilizada en el proceso de desarrollo.
<b>Validación externa</b>	Evaluación del rendimiento de un modelo en un conjunto de datos que difiere del utilizado en su desarrollo desde el punto de vista geográfico o temporal
<b>Validación cruzada</b>	Técnica de validación interna en la que el conjunto de datos se divide aleatoriamente en k-1 grupos de tamaño similar. El desempeño se evalúa en el grupo restante y todo el proceso se repite n veces. El rendimiento del modelo se toma como el promedio de n iteraciones.
<b>Arranque (bootstrapping)</b>	Abordaje de validación interna como validación cruzada pero basado en muestreo aleatorio con reemplazo; cada muestra tiene el mismo tamaño que el conjunto de datos de desarrollo del modelo.
<b>Muestra dividida</b>	Abordaje de validación interna en el que el conjunto de datos de desarrollo disponible se divide en dos conjuntos de datos: uno para desarrollar el modelo y el otro para validar el modelo; la división puede ser aleatoria o no aleatoria.

\*Adaptado de referencias 10, 11, 13 y 14.

**Tabla 3.** Valores y principios éticos recomendados por la UNESCO para todos los actores implicados en la inteligencia artificial\*

	<b>Descripción</b>
<b>Valores</b>	Respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana
	Prosperidad del medio ambiente y los ecosistemas
	Garantizar la diversidad y la inclusión
	Vivir en sociedades pacíficas, justas e interconectadas
<b>Principios</b>	Proporcionalidad e inocuidad
	Seguridad y protección
	Equidad y no discriminación
	Sostenibilidad
	Derecho a la intimidad y protección de datos
	Supervisión y decisión humanas
	Transparencia y explicabilidad
	Responsabilidad y rendición de cuentas
	Sensibilización y educación
	Gobernanza y colaboración adaptativas y de múltiples partes interesadas

\*Adaptado de referencia 68