UNIVERSIDAD ESPÍRITU SANTO

ESCUELA DE POSTGRADO EN DERECHO

TÍTULO:

Inteligencia artificial aplicada en diagnósticos médicos: modelos de caja negra y la protección al derecho a formar un consentimiento informado

TRABAJO DE TITULACIÓN QUE SE PRESENTA COMO REQUISITO PREVIO A OPTAR EL GRADO
DE MAGÍSTER EN DERECHO, CON MENCIÓN EN DERECHO DE EMPRESA

AUTOR:

María Belén Prieto Correa

SAMBORONDÓN, 02 MES FEBRERO, AÑO 2026

Resumen

El presente trabajo revisa el uso de los sistemas de inteligencia artificial Large Language Models -LLM- de tipo caja negra que son utilizados por los profesionales de la salud cuando introducen los datos clínicos de sus pacientes y emplean dichos sistemas como fundamento para la formulación del diagnóstico médico, así como también, la consecuencia que esto produce al ejercicio del derecho del paciente a tener un consentimiento informado. Para el análisis, se utilizó, principalmente, el método inductivo, aplicando la técnica cualitativa de revisión documental sobre las disposiciones de la Constitución de la República del Ecuador, la Ley Orgánica de Salud, la Ley Orgánica de Protección de Datos Personales y demás normativa pertinente, incluyendo instrumentos internacionales. Principalmente, se concluye que estos sistemas, por su opacidad, al ser utilizados como base de un diagnóstico médico, obstaculizan el ejercicio del derecho del paciente a formar un consentimiento informado, y también limitan la capacidad explicativa del profesional de la salud. **Palabras clave:**
Inteligencia artificial en la medicina; modelos de caja negra; Large Language Models -LLM-; consentimiento informado; autonomía del paciente.

AbstractThe present paper reviews the use of black-box Large Language Model (LLM) artificial intelligence systems that are used by healthcare professionals when they enter their patients' clinical data and use such systems as a basis for the formulation of a medical diagnosis, as well as the consequences that this produces for the exercise of the patient's right to informed consent. For the analysis, the inductive method was mainly used, applying the qualitative technique of documentary review of the provisions of the Constitution of the Republic of Ecuador, the Organic Health Law, the Organic Law on Personal Data Protection, and other relevant regulations, including international instruments. It is mainly concluded that these systems, due to their opacity, when used as a basis for a medical diagnosis, hinder the exercise of the patient's right to form informed consent and also limit the explanatory capacity of the healthcare professional.

Keywords:
Artificial intelligence in medicine; black-box models; Large Language Models (LLM); informed consent; patient autonomy.**Introducción**

Con el paso del tiempo, la inclusión de los sistemas de inteligencia artificial dentro del

campo de la medicina ha generado un gran impacto en la relación tradicional entre médico y paciente. En la actualidad, muchos de los profesionales de la salud utilizan esta herramienta como apoyo para formular diagnósticos médicos.

Varios autores reconocen que los sistemas de inteligencia artificial han beneficiado al sistema sanitario al permitir diagnósticos más rápidos y precisos, sin embargo, entre estos sistemas hay algunos que operan bajo el modelo de caja negra. Un modelo opaco que dificulta que se conozca el procedimiento interno utilizado para llegar a un resultado en particular, lo cual, en el ámbito de la medicina genera elevados riesgos éticos y jurídicos. De tal forma, los profesionales de la salud, al carecer de conocimiento sobre el procedimiento interno que utilizó el sistema para dar un diagnóstico, no puede garantizarle al paciente el derecho a tener un consentimiento informado puesto que no cuentan con la información suficiente y adecuada para estos puedan tomar decisiones libres y autónomas sobre sus tratamientos clínicos. En este sentido, el profesional de la salud que utiliza un sistema opaco no solo afecta el consentimiento informado y la autonomía del paciente, sino que también limita su capacidad médica con respecto a su deber objetivo de cuidado.

En concordancia con lo anterior, el problema jurídico del presente trabajo consiste en analizar cómo el uso de modelos de inteligencia artificial LLM de tipo caja negra, como herramientas de fundamento para al diagnóstico médico por parte de los profesionales de la salud en seres humanos, afecta al ejercicio efectivo del derecho del paciente a formar su consentimiento informado. Este problema es relevante puesto que el consentimiento informado, es un derecho reconocido en el ordenamiento jurídico ecuatoriano, sin embargo, no existen regulaciones específicas que establezcan estándares claros de transparencia y explicabilidad en el uso de sistemas de inteligencia artificial dentro de la medicina.

El presente artículo, refiriere a utilización de sistemas basados en inteligencia artificial, particularmente Large Language Models -LLM-, utilizados como herramientas para fundamentar al diagnóstico clínico, en los casos en que el profesional de la salud introduce previamente los datos clínicos y toma la decisión final. La investigación fue desarrollada principalmente a través del uso del método inductivo, utilizando la técnica cualitativa de revisión documental, principalmente sobre las disposiciones normativas de la Constitución de la República del Ecuador, la Ley Orgánica de Salud, la Ley Orgánica de Protección de Datos Personales, entre otras normas pertinentes, incluyendo instrumentos internacionales.

Marco teórico

Inteligencia artificial en la medicina

El presente marco teórico revisará los modelos de inteligencia artificial basados en Large Language Models -LLM- que participan exclusivamente en la formulación de diagnósticos clínicos. Esta delimitación excluye cualquier otro uso de la IA en el ámbito médico. Asimismo, el desarrollo se centrará en aquellos casos en los que los profesionales de la salud ingresan previamente los datos clínicos ya procesados al sistema, y no en aquellos en los que la inteligencia artificial procesa directamente las muestras o datos obtenidos de los pacientes, con el fin de examinar las consecuencias jurídicas derivadas de la forma en que la inteligencia artificial recibe la información -input- y del uso que el médico hace de la tecnología.

Vera-Rubio et al. (2023) señalan que la inteligencia artificial es parte de un campo de investigación orientado a configurar programas capaces de realizar tareas propias de los seres humanos. Esta busca apoyar a los humanos en la ejecución eficiente de las tareas cotidianas, en la solución de problemas e incluso en procesos creativos. Como indica la Corporación Andina de Fomento (2021), la inteligencia artificial, a través del uso de tecnologías digitales, permite el desarrollo de sistemas capaces de ejecutar tareas que normalmente requieren inteligencia humana.

Desde otra perspectiva, el International Science Council (2025) sostiene que la inteligencia artificial no es una tecnología única, sino que comprende diversos paradigmas, en donde cada uno cuenta con enfoques propios de aprendizaje, inferencia y generación de conocimiento, aplicables al ámbito científico. Esta diversidad conceptual justifica la necesidad de precisar el tipo de IA que será analizado en la presente.

Dentro de la inteligencia artificial hay modelos esenciales que le permiten integrarse dentro del campo de la medicina, entre ellos el aprendizaje automático o mejor conocido como machine learning. Esta técnica posibilita que los sistemas de inteligencia artificial analicen grandes cantidades de datos y aprendan a partir de la experiencia, para que así mejoren su desempeño sin necesidad de ser programados de forma específica para cada tarea (Díaz Cantón et al., 2025).Asimismo, el denominado aprendizaje profundo o deep learning es un conjunto de modelos avanzados de inteligencia artificial diseñados para reproducir el funcionamiento del cerebro humano, lo que permite el análisis y el procesamiento de información de forma eficaz (Díaz Cantón et al., 2025).

En la actualidad, los modelos de aprendizaje automático, particularmente aquellos basados en deep learning, tienen una gran capacidad para poder analizar datos complejos y heterogéneos generados en la práctica de la medicina moderna, tales como las notas médicas, imágenes clínicas y otros registros sanitarios, lo que contribuye al desarrollo de predicciones medicas relevantes (Rajkomar et al., 2019).

En la actualidad, una gran parte de los sistemas de inteligencia artificial utilizados en el ámbito de la medicina funciona como modelos de caja negra. No obstante, Prince y Lim (2025) señalan que, a pesar de que estos modelos no pueden proporcionar explicaciones claras sobre las decisiones que adoptan, la inteligencia artificial aplicada a la medicina ha demostrado aportar beneficios clínicos relevantes, al constituirse como una tecnología precisa y confiable que mejora la calidad de la atención al paciente y apoya los procesos diagnósticos.

La literatura ha señalado que ambos modelos, el de aprendizaje automático y aprendizaje profundo se complementan para asimilar grandes volúmenes de información de los pacientes y analizar sus resultados, lo que permite formular predicciones médicas que, en ciertos contextos pueden alcanzar niveles de precisión igual o incluso superiores a los de los profesionales de la salud humanos. De tal forma que, Ultsch et al. (2024) afirman que los sistemas algorítmicos basados en aprendizaje automático, a partir de grandes volúmenes de datos médicos, pueden contribuir de forma eficiente a los procesos de diagnósticos en la medicina actual.

Aunque existen distintos modelos de inteligencia artificial aplicados a la medicina, el objeto de estudio de la presente revisión son los LLM o mejor conocidos como los Large Language Models, que son modelos de aprendizaje profundo basados en arquitecturas transformer, entrenados mediante grandes volúmenes de datos textuales. Estas características les permiten comprender el contexto del lenguaje, producir textos con rasgos similares al lenguaje humano y ejecutar tareas lingüísticas de alta complejidad (Alkalbani et al., 2025).

Rodríguez Fernández et al. (2024) explican que los Large Language Models se entrenan a partir de grandes cantidades de datos textuales, lo que les permite desarrollar capacidades avanzadas de comprensión y generación de lenguaje natural. En consecuencia, estos modelos pueden reproducir textos coherentes y contextualizados, lo que los convierte en herramientas útiles para los diferentes campos de aplicación.

De igual forma, Chen et al. (2025) señalan que en los últimos años se han desarrollado sistemas de agentes de inteligencia artificial basados en Large Language Models, en los cuales

estos modelos funcionan como el núcleo del sistema y se integran diferentes herramientas de IA especializadas. Esto permite que los agentes, de forma autónoma, puedan comprender las instrucciones del usuario, tomar decisiones y seleccionar las herramientas más convenientes para la realización de tareas medicas complejas.

Esto ha sido demostrado en estudios recientes. En particular, Díaz Cantón et al. (2025) describen los resultados del estudio MASAI, publicado en The Lancet en 2023, el cual analizó aproximadamente 80.000 mamografías y comparó el método tradicional de doble lectura por especialistas con un modelo en el que un médico especialista realizaba la lectura asistido por un sistema de IA. Los resultados evidenciaron una reducción significativa del tiempo de lectura, cercana al 44 %, así como un incremento en la detección de casos de cáncer de mama en el grupo ayudado por inteligencia artificial. Estos hallazgos muestran que la IA no sustituye al profesional médico, sino que actúa como una herramienta de apoyo que mejora la eficiencia y la precisión diagnóstica. En la actualidad, el uso de la inteligencia artificial en el campo de la medicina ha transformado profundamente la relación tradicional entre el médico y el paciente. Daniyal et al. (2024) señalan que las tecnologías de inteligencia artificial permiten procesar, analizar y generar información a partir de grandes cantidades de datos provenientes de diversas fuentes, lo que contribuye a la detección de enfermedades y al apoyo en la toma de decisiones clínicas.

De tal forma, la inteligencia artificial se ha consolidado como una herramienta valiosa para los profesionales de la salud; sin embargo, su utilización plantea desafíos éticos y jurídicos, especialmente en lo relativo a garantizar que los pacientes otorguen un consentimiento verdaderamente libre e informado. Se considera que la inteligencia artificial siempre debe ir ligada al uso de los principios éticos que garantizan una práctica médica responsable y transparente. Como sostiene García-López et al. (2023), en el ámbito de la salud es fundamental que los profesionales comprendan el funcionamiento de la inteligencia artificial, así como la importancia de la transparencia y explicabilidad, para que estas tecnologías puedan apoyar adecuadamente la toma de decisiones médicas.

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO (2021) indica que la explicabilidad en los sistemas de inteligencia artificial refiere a que los procesos y resultados generados por estos sistemas sean comprensibles para los usuarios, mediante información clara sobre sus entradas, salidas y funcionamiento de los procedimientos internos. De esta forma, la explicabilidad se vincula con la transparencia, puesto que los

resultados obtenidos y los procedimientos algorítmicos que conducen a ellos deben ser comprensibles y trazables.

De igual forma, la UNESCO (2021) señala que la finalidad de la transparencia es otorgarles a los usuarios información suficiente que les permita comprender el funcionamiento de los sistemas de inteligencia artificial. De esta forma, la transparencia ayuda a que los destinatarios conozcan las distintas fases de implementación de los sistemas, considerando el contexto y el nivel de sensibilidad que presentan.

En este sentido, el concepto de explicabilidad en la inteligencia artificial aún genera confusión, ya que carece de una definición exacta debido a que las investigaciones sobre este tema son muy recientes (Ortiz de Zárate Alcarazo, 2022)

En particular, respecto de los modelos de LLM, Gómez-Rodríguez (2025) sostiene que la falta de explicabilidad de los LLM limita su uso como herramientas principales en la toma de decisiones relevantes, por lo que en la actualidad no resulta recomendable utilizarlos de manera directa, sino únicamente como sistemas de apoyo que orienten a la persona responsable de decidir. De igual forma, Gómez-Rodríguez (2025) indica que el desconocimiento sobre cómo los LLM llegan a un resultado especifico genera una falta de transparencia en estos sistemas.Se ha identificado que al incorporarse el uso de la inteligencia artificial al diagnóstico en la medicina humana, se genera un riesgo para la transparencia, ya que es posible que se desconozca los procesos internos que utilizaron los algoritmos para tomar decisiones clínicas, a lo que se identifica coloquialmente como caja negra. Mollura et al. (2020) señalan que los sistemas de inteligencia artificial en el ámbito de la salud presentan grandes riesgos relacionados con la falta de transparencia, debido a que muchos de estos modelos operan como cajas negras, lo que dificulta la explicabilidad de sus decisiones.

Derecho al consentimiento informado y su relación con la bioética

El consentimiento informado es un derecho que poseen los pacientes médicos a producir un consentimiento previo, claro, real y sustantivo respecto de sus diagnósticos y procedimientos a los que serán sometidos. En este sentido, la Asociación Médica Mundial (2023) establece que los profesionales de la salud deben garantizar a los pacientes el cumplimiento de su derecho a ser informados de forma clara y adecuada en cada etapa de la atención médica, con la finalidad de que puedan tomar una decisión libre, informada y voluntaria respecto de la atención propuesta. Además, se señala que los pacientes poseen la facultad de poder conservar o no el

consentimiento, y los médicos deben respetar esta decisión ante cualquier circunstancia.El consentimiento informado puede entenderse como un proceso mediante el cual los profesionales de la salud proporcionan al paciente información suficiente respecto de los riesgos, beneficios y alternativas de un procedimiento médico, con la finalidad de que los pacientes estén lo suficientemente informados para que puedan tomar una decisión voluntaria e informada sobre su atención médica (Shah et al., 2024). Desde una perspectiva bioética, el reconocimiento del derecho del paciente a recibir información suficiente para tomar decisiones voluntarias sobre su atención médica supuso el abandono del modelo tradicional de medicina paternalista, en el cual el médico era la única persona que podía tomar decisiones clínicas basándose en su propio criterio profesional (Fraix, 2022).

La bioética es indispensable para resolver dilemas morales y éticos en la práctica médica. Esta se rige por principios que son fundamentales para los profesionales de la salud y la práctica clínica, entre ellos están: La justicia, La No Maleficencia, La Beneficencia y la Autonomía. Este último principio reconoce el derecho de los pacientes a tomar decisiones voluntarias tras haber recibido información adecuada sobre su atención médica, y establece la obligación de los profesionales de la salud de garantizar la comprensión de las opciones de tratamiento y de sus posibles consecuencias para el paciente (Comisión Nacional de Bioética, 2023).

En Ecuador, el consentimiento informado se encuentra garantizado de forma expresa en la Constitución y en diversas legislaciones sanitarias. El artículo 362 de la Constitución, establece que el sistema sanitario deberá garantizar al paciente el acceso a un servicio de salud seguro, cálido y de calidad, donde reciba información clara y suficiente sobre su atención médica, con la finalidad de garantizar su derecho al consentimiento informado y a que se mantenga en reserva su información personal.

De igual forma, la Ley Orgánica de Salud reconoce el derecho a que los pacientes obtengan un consentimiento informado previo al igual que a recibir información, clara y suficiente con respecto de los procedimientos médicos a los que serán sometidos. De tal forma que el artículo 208, dispone que las investigaciones científicas de la salud deberán ser supervisadas por las autoridades y organismos sanitarios competentes, respetando los principios de la bioética y los derechos humanos. En este sentido, el consentimiento informado y por escrito es un requisito previo a la realización de la investigación, así como la obligación de mantener la información confidencial y bajo reserva (Ley Orgánica de Salud, 2006).

Asimismo, los artículos 68 y 212 de dicha ley, hacen prevalecer el consentimiento informado de la salud y la autonomía del paciente ante cualquier tipo de procedimiento médico.

En conjunto, la Ley Orgánica de Donación y Trasplante de Órganos, Tejidos y Células regula expresamente el consentimiento informado en los procedimientos de donación y trasplante. En este sentido, el artículo 5, literales a) y b) reconocen el derecho de los pacientes a recibir información previa y oportuna sobre los efectos que puede generar la donación o trasplante de algún órgano, así como la obligación que tienen de establecer de forma expresa y por escrito su consentimiento para ser parte de estos procesos (Asamblea Nacional del Ecuador, 2011). Mientras que el artículo 33, literal c), establece que la persona que dona o quien recibe la donación debe haber sido informada con anterioridad sobre los efectos de su decisión, así como debe haber expresado de forma escrita y notariada su consentimiento libre y voluntario (Asamblea Nacional del Ecuador, 2011).

El consentimiento informado es una obligación ética y legal que los profesionales de la salud tanto nacionales como internacionales deben cumplir para validar la autonomía del paciente.

El consentimiento informado está constituido como una garantía esencial dentro del ámbito de los derechos fundamentales a la dignidad, libertad, y salud de los seres humanos. Este debe ser cumplido de tal forma que no dé apertura a que se le considere un cumplimiento deficiente, o peor aún, su omisión o desconocimiento. Tales situaciones generan dilemas éticos graves, acompañados de consecuencias legales concretas, pues configuran una vulneración al derecho del paciente con respecto a la autonomía y facultad de tomar decisiones correspondientes a su cuerpo. La falta de información previa a la hora de ejecutar una decisión médica, puede generar responsabilidades legales, no solo por los daños causados, sino también por la vulneración del derecho de las personas a tomar decisiones libres y autónomas sobre su propio cuerpo, una vez que han sido debidamente informadas (Pallocci, 2023).

En síntesis, el consentimiento informado constituye un derecho esencial para una práctica médica responsable. Desde el punto de vista jurídico, se encuentra reconocido en la Constitución del Ecuador, así como en instrumentos internacionales que respaldan este derecho, el cual protege el principio de autonomía del paciente y obliga a los profesionales de la salud al deber legal de informar y respetar la decisión autónoma de cada paciente.

Los límites derivados del fenómeno de caja negra en los sistemas de decisión automatizada

El fenómeno de caja negra se puede presentar en ciertos modelos de inteligencia artificial, en particular en sistemas de aprendizaje profundo -deep learning- en donde se realizan procedimientos internos que resultan opacos o casi imposibles de ser interpretados incluso por sus propios desarrolladores.

Esta opacidad dificulta conocer el procedimiento interno que llevo el sistema a dictar un determinado resultado. En este sentido, Imam et al. (2024) indican que las predicciones y decisiones producidas por los modelos de caja negra se encuentran dentro de capas computacionales internas, lo que dificulta su comprensión. Como consecuencia, la trazabilidad de la predicción se pierde en dichas capas, dificultando que los usuarios puedan comprender el razonamiento que condujo a ese resultado.

Belloto Gómez (2022) explica que los modelos de caja negra, debido a su compleja estructura, dificultan la interpretación de sus resultados por parte de los seres humanos. Esta carencia de explicabilidad hace que su utilización no sea recomendable en la toma de decisiones relevantes y de alto riesgo.

Este fenómeno de caja negra también puede presentarse en modelos de inteligencia artificial utilizados en el ámbito de la medicina. En este sentido, Cortés Rodríguez (2024) señala que muchos de los algoritmos actuales, especialmente aquellos basados en aprendizaje automático profundo, funcionan como sistemas de caja negra, ya que generan resultados a partir de grandes volúmenes de datos sin que resulte evidente el proceso mediante el cual alcanzan una conclusión.

Estado del arte

En los últimos años, la medicina ha evolucionado de forma exponencial, tanto así, que actualmente algunos profesionales de la salud formulan sus diagnósticos bajo su propia responsabilidad apoyándose en herramientas complementarias que utilizan la inteligencia artificial.

Como señala Gallardo-Cochifas (2025), los avances actuales de la inteligencia artificial permiten que estos sistemas analicen imágenes médicas con alta precisión, lo que posibilita la detección temprana de enfermedades como el cáncer, las patologías cardiovasculares y los trastornos neurológicos, incluso antes de que dichos signos sean perceptibles para los profesionales de la salud.

A pesar de sus beneficios, la incorporación de la inteligencia artificial en la medicina ha generado tensiones éticas y jurídicas, en vista que puede dificultar el ejercicio efectivo del derecho al consentimiento informado por parte del paciente.

Las implementaciones de los modelos de Inteligencia Artificial han transformado la dinámica tradicional que existía entre paciente y médico, generado que el consentimiento informado cada vez sea más complejo de obtener.

Baget Bernaldiz et al. (2023) advierten que la incorporación de la inteligencia artificial en la práctica clínica actual, transformará los protocolos de cribado, diagnóstico y tratamiento, alterando la relación tradicional y directa entre el médico y el paciente, lo que hará necesario redefinir los procesos de consentimiento informado.

Asimismo, García-López et al. (2023) señalan que el uso de sistemas de inteligencia artificial en la toma de decisiones clínicas trae consigo responsabilidades éticas, especialmente cuando estos generan resultados erróneos. En dichos casos, resulta indispensable establecer mecanismos claros de rendición de cuentas que permitan garantizar un uso responsable de estos sistemas.

En este contexto, muchos médicos que utilizan sistemas automatizados de inteligencia artificial en la toma de sus decisiones clínicas están expuestos a grandes riesgos, dado que dichos sistemas presentan opacidad algorítmica. Rivas Velasco (2023) sostiene que, si un sistema no cuenta con un proceso transparente que permita comprender su razonamiento, resulta difícil confiar en sus decisiones, especialmente porque a mayor complejidad del sistema, disminuye su explicabilidad, lo que hace necesario buscar un equilibrio entre ambas.

Esto implica que si una decisión medica adolece de falta de explicabilidad, esta no debería ser considerada definitoria, dado que es el fruto del fenómeno de caja negra en donde ni el médico ni el paciente pueden comprender con precisión el porqué de la decisión, es decir no se puede inferir los procesos internos utilizados para llegar a tal resultado. En la medicina, al pasar por alto la explicabilidad y la interpretabilidad de una decisión médica, produciría la vulneración al derecho que tiene el paciente sobre tener un consentimiento informado, dando como resultado la necesidad de analizar como los marcos regulatorios abordan esta tensión.

De igual forma, la ausencia de explicabilidad en los sistemas de inteligencia artificial que son utilizados como apoyo a la toma de decisiones clínicas compromete principios éticos esenciales de la práctica médica y puede generar efectos negativos tanto en la atención de los

pacientes como en la protección de la salud colectiva (Madai et al., 2020). Esta opacidad algorítmica dificulta que el profesional de la salud pueda comprender y justificar adecuadamente el razonamiento adoptado para tomar las decisiones clínicas, lo que incide directamente en el cumplimiento del deber objetivo de cuidado que rige el ejercicio profesional médico. En efecto, el Consejo General de Colegios Oficiales de Médicos señala que, a pesar de que los sistemas inteligentes pueden funcionar como herramientas de apoyo para que los profesionales de la salud participen en la toma de decisiones clínicas, esto no reemplaza la responsabilidad que tiene el medico de hacer uso de todas las técnicas y procedimientos necesarios para garantizar una práctica profesional correcta y adecuada a su paciente (CGCOM, 2022).

En Ecuador, la falta de transparencia en los sistemas de inteligencia artificial coloca al paciente en una situación de desconocimiento dado que el derecho al consentimiento informado no puede ejercerse correctamente ante la ausencia de regulaciones especificas en este ámbito.

En este sentido, Iñiguez Pineda y Maldonado Ruiz (2025) sostienen que la falta de transparencia en el uso de sistemas de inteligencia artificial dificulta que los usuarios comprendan el tratamiento de sus datos personales. Asimismo, cuando estos sistemas operan bajo modelos de “caja negra”, se limita la comprensión del razonamiento utilizado para la toma de las decisiones automatizadas, incluso para quienes los han desarrollado. En el ordenamiento jurídico ecuatoriano, la ausencia de regulaciones que obliguen a las entidades que emplean inteligencia artificial en procesos automatizados a proporcionar explicaciones claras y comprensibles sobre cómo se adoptan dichas decisiones restringe el ejercicio del derecho al consentimiento informado frente a posibles decisiones perjudiciales.

Al referirnos al plano normativo, Pereira Sáez (2025) sostiene que la falta de comprensión sobre el funcionamiento de los sistemas de inteligencia artificial representa un riesgo para las personas, ya que la ausencia de información suficiente puede impedir que estas otorguen un consentimiento libre y válido. Lo que pone en evidencia el vínculo directo que se establece entre el principio de autonomía del paciente y el principio de transparencia.

En la misma línea, tenemos a la Organización Panamericana de la Salud y Organización Mundial de la Salud (2024) quienes señalan que, para resguardar la autonomía humana, los servicios de la salud y las decisiones clínicas deben permanecer bajo control humano. En este sentido, los pacientes deben otorgar un consentimiento informado mediante sistemas legales que garanticen la protección y confidencialidad de sus datos personales.

Por otro lado, cuando nos referimos a la normativa a nivel nacional, el artículo 20 de la Ley Orgánica de Protección de Datos Personales reconoce que el sujeto de derecho no podrá ser sometido a decisiones adoptadas ya sea parcial o totalmente, mediante sistemas automatizados, puesto que esto le podría generar perjuicios o afectaciones a sus derechos fundamentales (Asamblea Nacional del Ecuador, 2021).

En la dimensión del derecho comparado, el Reglamento (UE) 2024/1689 sostiene que los sistemas de inteligencia artificial que puedan generar un nivel elevado de riesgos deberán ser programados de tal forma que permitan la supervisión por parte de los seres humanos durante su funcionamiento (Unión Europea, 2024).Respaldando que los profesionales de la salud son quienes tienen la responsabilidad de evaluar los resultados generados por la inteligencia artificial para así, bajo su propio criterio aceptar o rechazar dichos resultados. Y que las decisiones automatizadas en el campo médico deben realizarse bajo la supervisión y control de los seres humanos, puesto que, de no ser así, los pacientes podrían adoptar un estado de riesgo e indefensión.

Asimismo, la Corporación Andina de Fomento (CAF, 2024) señala que el uso responsable de la tecnología, orientado a la protección de los datos personales y en dirección con los principios éticos propuestos por la Recomendación de la UNESCO en materia de inteligencia artificial, puede contribuir al reconocimiento de la diversidad de género y a la promoción de la igualdad étnico-racial, generando un impacto en el fortalecimiento de la atención preventiva. En este sentido, el desarrollo de la tecnología debe darse bajo principios éticos fundamentales que le permitirán la correcta y responsable aplicación en el campo médico.

Por último, en lo que respecta al ámbito jurisprudencial, la Corte Constitucional de Colombia (2024), en la Sentencia T-323/24, se manifestó que los operadores de justicia, a lo largo del procedimiento judicial, deben garantizar el uso de herramientas que de ninguna manera reemplacen el criterio humano, y que, además, cumplan con los principios de transparencia, privacidad y responsabilidad. Aunque esta sentencia se refiere al ámbito judicial, el criterio de no sustitución de la racionalidad humana, así como las cargas de transparencia, responsabilidad y privacidad resultan también aplicables al ámbito medico en la medida que, fortalecen la exigencia de que la racionalidad humana no sea sustituida por ningún tipo de sistema de inteligencia artificial que pueda vulnerar el consentimiento informado y los derechos humanos de los pacientes.

En síntesis, el derecho comparado demuestra que otros países han desarrollado avances regulatorios que fortalecen el uso de la inteligencia artificial en salud, siempre bajo supervisión humana, reafirmando que ningún sistema puede sustituirla. En contraste, en Ecuador persisten vacíos normativos que no protegen adecuadamente al paciente ni al profesional de la salud, lo que evidencia la necesidad de establecer regulaciones que definan con claridad cómo debe emplearse la inteligencia artificial y que, al mismo tiempo, garanticen los derechos fundamentales de las personas y la seguridad jurídica para los médicos frente a su responsabilidad profesional derivada del deber objetivo de cuidado que deben de cumplir, sobre todo en casos en los que pudieran responder penalmente, como los previstos en el en el artículo 146 del Código Orgánico Integral Penal (Asamblea Nacional del Ecuador, 2025).

El presente estado del arte confirma que los modelos que incluyen el fenómeno de caja negra representan un riesgo significativo para la autonomía del paciente y para la obtención válida de un consentimiento informado, así como obstaculizan el cumplimiento del deber objetivo de cuidado del que el medico es responsable.

Planteamiento del problema

Con base en los antecedentes descritos, se propone el siguiente problema de investigación:

¿La utilización de modelos de inteligencia artificial de tipo “caja negra”, como herramientas de apoyo fundamental al diagnóstico médico por parte de los profesionales de la salud en seres humanos, puede afectar al ejercicio efectivo del derecho del paciente al consentimiento informado?

Análisis

Los sistemas de inteligencia artificial como herramienta de apoyo fundamental al diagnóstico médico han modificado la relación tradicional entre un profesional de la salud y el paciente. Esta transformación no implica la sustitución de la responsabilidad que tiene el médico en la toma de decisiones clínicas, sin embargo, conlleva riesgos que pueden dificultar tanto el ejercicio del derecho del paciente al consentimiento informado, como el cumplimiento de un deber objetivo de cuidado por parte del médico.

Desde el ámbito jurídico, los profesionales de la salud son responsables de otorgar al paciente información suficiente, clara y comprensible sobre su diagnóstico médico, con el objetivo que estos puedan adoptar decisiones médicas voluntarias, libres y autónomas. En

consecuencia, si el paciente no recibió la información suficiente o si el profesional de la salud no fue claro y comprensible al otorgarla, el ejercicio del derecho a un consentimiento informado por parte del paciente podría verse comprometido, al igual que el cumplimiento del deber objetivo de cuidado del profesional.

En este sentido, los modelos de inteligencia artificial basados en el fenómeno de caja negra dificultan el ejercicio efectivo del derecho del paciente al consentimiento informado. La falta de conocimiento de los procesos internos que el sistema tomo para llegar a cierto resultado por parte de los profesionales de la salud, puede dificultar la explicación que pueden brindarle al paciente.

En los sistemas de inteligencia artificial basados en el fenómeno de caja negra existen dos premisas que coexisten. En primer lugar, que la utilización de herramientas tecnológicas ha contribuido para generar diagnósticos médicos más precisos y eficientes. En segundo lugar, los profesionales de la salud son quienes tienen la responsabilidad jurídica de garantizar al paciente un consentimiento informado, claro y suficiente respecto de las decisiones clínicas que afectan su salud. La coexistencia de ambas premisas genera una tensión jurídica. En el sentido que, a pesar de que los sistemas de inteligencia artificial de caja negra han aportado beneficios clínicos, su carácter dificulta que los profesionales de la salud expliquen de manera clara y comprensible los fundamentos que sustentan un determinado diagnóstico. Esta limitación pone en riesgo el ejercicio efectivo del derecho del paciente al consentimiento informado, al restringir la posibilidad de comprender las razones que motivan la decisión clínica.

Por otro lado, aunque la legislación ecuatoriana reconoce el consentimiento informado como un derecho fundamental en el ámbito de la salud y protege la autonomía del paciente, no existen regulaciones específicas que dispongan el uso de sistemas de inteligencia artificial basados en el fenómeno de caja negra en la práctica médica. La ausencia de criterios normativos claros sobre transparencia y explicabilidad en el uso de estas tecnologías genera un escenario de inseguridad jurídica, en el que el cumplimiento efectivo del consentimiento informado depende exclusivamente de la capacidad explicativa del profesional de la salud.

A partir de lo expuesto, puede sostenerse como hipótesis de solución que el uso de modelos de inteligencia artificial que presentan el fenómeno de caja negra en el diagnóstico médico solo resulta compatible con el derecho al consentimiento informado cuando el profesional de la salud tiene la capacidad de explicar, de manera comprensible para el paciente,

los fundamentos médicos del diagnóstico adoptado. En aquellos casos en los que dicha explicación no sea posible, la autonomía del paciente y el ejercicio efectivo del derecho al consentimiento informado pueden verse comprometidos.

Conclusiones

Con base en el análisis precedente, se concluye que el uso de los modelos de inteligencia artificial que presentan el fenómeno de caja negra si bien pueden servir de apoyo fundamental al diagnóstico médico, introducen una tensión jurídica con el derecho al consentimiento informado, pero no por el simple uso de las herramientas tecnológicas, sino por las limitaciones que impone a la capacidad explicativa del profesional de la salud. Esta limitación dificulta el ejercicio efectivo del derecho al consentimiento informado del paciente, en vista que, obstaculiza la comprensión de las razones que sustentan la decisión clínica y el ejercicio de su autonomía.

Igualmente cabe advertir que la actual legislación ecuatoriana, carece de criterios normativos específicos sobre los grados de transparencia y explicabilidad en el uso de inteligencia artificial en el diagnóstico médico generando un vacío regulatorio que debilita la garantía al derecho de obtener un consentimiento informado. En este sentido, se establece que la compatibilidad entre el uso de modelos de caja negra y el consentimiento informado depende de que el profesional de la salud conserve un control efectivo sobre la trazabilidad y razonabilidad de la decisión clínica y pueda explicar de forma clara y comprensible sus fundamentos médicos al paciente. Cuando esta condición no se cumple, el ejercicio efectivo del derecho al consentimiento informado resulta comprometido.

Referencias bibliográficas

Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20**,** 310. https://doi.org/10.1186/s12911-020-01332-6

Alkalbani, A. M., Alrawahi, A. S., Salah, A., Haghighi, V., Zhang, Y., Alkindi, S., & Sheng, Q. Z. (2025). A systematic review of large language models in medical specialties:

Applications, challenges and future directions. Information, 16(489). https://doi.org/10.3390/info16060489

Asociación Médica Mundial. (2023). Código Internacional de Ética Médica. https://www.wma.net/es/policies-post/codigo-internacional-de-etica-medica/

Baget Bernaldiz, M., Téllez Vázquez, J., & Romero Aroca, P. (2023). Aspectos éticos y legales de la inteligencia artificial. Annals d'Oftalmologia, 31(4), 310–315. https://www.annalsoftalmologia.com/articulos/a19117/CDS_2023_CAP_6-2.pdf

Belloto Gómez, J. M. (2022). Explicabilidad e inteligencia artificial responsable. En L. Cotino Hueso (Coord.), Transparencia y explicabilidad de la inteligencia artificial (pp. 211–218). Universitat de València. https://www.uv.es/cotino/publicaciones/libroabiertotp22.pdf

Chen, X., Xiang, J., Lu, S., Liu, Y., He, M., & Shi, D. (2025). Evaluating large language models and agents in healthcare: Key challenges in clinical applications. Intelligent Medicine, 5**,** 151–163. https://doi.org/10.1016/j.imed.2025.03.002

Cheng, Y., Chang, Y., & Wu, Y. (2025). A survey on data contamination for large language models (Version 2) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2502.14425

Comisión Nacional de Bioética. (2023). Gaceta 50: Consideraciones sobre bioética y ciencia. Actualidades y perspectivas. https://www.gob.mx/cms/uploads/attachment/file/914870/Gaceta_50_Consideraciones_sobre_bioetica_y_ciencia_actualidades_y_perspectivas.pdf

Consejo General de Colegios Oficiales de Médicos. (2022). Código de deontología médica: Guía de ética médica. https://www.cgcom.es/sites/main/files/minisite/static/41a75446-a6e8-4df0-af58-7ca9cb625f1e/Codigo_Deontologia_2022/index.html

Corporación Andina de Fomento (CAF). (2024). La IA y los ODS en América Latina y el Caribe. Banco de Desarrollo de América Latina y el Caribe. https://scioteca.caf.com/bitstream/handle/123456789/2326/La%20IA%20y%20los%20ODS%20en%20Am%C3%A9rica%20Latina%20y%20el%20Caribe.pdf?isAllowed=y&sequence=1

Corporación Andina de Fomento. (2021). Experiencia: Datos e inteligencia artificial en el sector público. https://scioteca.caf.com/handle/123456789/1793

Cortés Rodríguez, B. (2024). Inteligencia artificial y práctica de la medicina. Actualidad Médica, 109(819), 115–117. https://doi.org/10.15568/am.2024.819.ami01

Daniyal, M., Qureshi, M., Rillera Marzo, R., Aljuaid, M., & Shahid, D. (2024). Exploring clinical specialists' perspectives on the future role of AI: Evaluating replacement perceptions, benefits, and drawbacks. BMC Health Services Research, 24, Article 587. https://doi.org/10.1186/s12913-024-10928-x

Díaz Cantón, E., Rossi, M., & Fernández Sande, J. (2025). Inteligencia artificial y medicina: Transformando el cuidado de la salud con innovación. Academia Nacional de Medicina de Buenos Aires. https://anm.edu.ar/wp-content/uploads/2025/02/Libro-IA-Avanzada.pdf

Fraix, T. (2022). Reflexión bioética: autonomía y toma de decisiones médicas. Revista Médica de Rosario, 88(1), 27–29. https://revistamedicaderosario.org/index.php/rm/article/download/169/279

Gallardo-Cochifas, J. (2025). Inteligencia artificial y su uso en el diagnóstico: ¿El fin del ars medica? Revista Chilena de Salud Pública, 29, e78286. https://doi.org/10.5354/0719-5281.2025.78286

García-López, A., Girón-Luque, F., & Rosselli, D. (2023). La integración de la inteligencia artificial en la atención médica: Desafíos éticos y de implementación. Universitas Médica, 64(3), e36231. https://doi.org/10.11144/Javeriana.umed64-3.inte

Gómez-Rodríguez, C. (2025). Grandes modelos de lenguaje: ¿de la predicción de palabras a la comprensión? En La inteligencia artificial hoy y sus aplicaciones con Big Data (pp. 73–98). Funcas. https://www.funcas.es/wp-content/uploads/2025/02/04.-Gomez-Rodriguez-Carlos_1.pdf

Imam, N. M., Ibrahim, A., & Tiwari, M. (2024). Explainable artificial intelligence (XAI) techniques to enhance transparency in deep learning models. IOSR Journal of Computer Engineering, 26(6), 29–36. https://doi.org/10.9790/0661-2606012936

International Science Council. (2025). Types of AI and their use in science. Centre for Science Futures. https://doi.org/10.24948/2025.09

Iñiguez Pineda, M. J., & Maldonado Ruiz, L. M. (2025). Protección de datos personales frente al desarrollo de inteligencia artificial en Ecuador: análisis y propuestas legislativas. Visionario Digital, 9(2), 24–48. https://doi.org/10.33262/visionariodigital.v9i2.3366

Mollura, D. J., Culp, M. P., Pollack, E., Battino, G., Scheel, J. R., Mango, V. L., & Dako, F. (2020). Artificial intelligence in low- and middle-income countries: Innovating global health radiology. Radiology, 297(3), 513–520. https://doi.org/10.1148/radiol.2020201434

Organización Panamericana de la Salud (OPS) & Organización Mundial de la Salud (OMS). (2024). Inteligencia artificial en la salud pública: Evaluación y mejora de la preparación para la integración de la IA en la salud pública (v. 2.0). OPS. https://www.paho.org/sites/default/files/2024-08/ai-ra-tool-rev-final-esp-aug-1_0.pdf

Ortiz de Zárate Alcarazo, L. (2022). Explicabilidad (de la inteligencia artificial). Eunomía. Revista en Cultura de la Legalidad, (22), 328–344. https://doi.org/10.20318/eunomia.2022.6819

Pallocci, M., Treglia, M., Passalacqua, P., Tittarelli, R., Zanovello, C., De Luca, L., Caparrelli, V., De Luna, V., Cisterna, A. M., Quintavalle, G., & Marsella, L. T. (2023). Informed consent: Legal obligation or cornerstone of the patient–physician relationship? International Journal of Environmental Research and Public Health, 20(3), 2118. https://doi.org/10.3390/ijerph20032118

Pereira Sáez, C. (2025). Empleo de la inteligencia artificial y consentimiento informado en la asistencia sanitaria. Cuadernos de Bioética, 36(117), 117–127. https://aebioetica.org/revistas/2025/36/117/117.pdf

Prince, S., & Lim, J. E. (2025). Black-box AI and patient autonomy. Minds and Machines, 35, Article 24. https://doi.org/10.1007/s11023-025-09729-w

Rajkomar, A., Dean, J., & Kohane, I. S. (2019). Machine learning in medicine. The New England Journal of Medicine, 380(14), 1347–1357. https://doi.org/10.1056/NEJMra1814259

Rivas Velasco, M. J. (2023). Uso ético de inteligencia artificial en justicia. Centro de Documentación del CGPJ. https://magistratura.es/wp-content/uploads/2024/10/APM-3.9.-MARIA-JOSE-RIVAS-VELASCO.-USO-ETICO-DE-INTELIGENCIA-ARTIFICIAL-EN-JUSTICIA.pdf

Rodríguez Fernández, L., Fernández Mena, A. L., Torres Magaña, M. P., Rodríguez Magaña, M. A., & Rodríguez Fernández, M. A. (2024). Inteligencia artificial en la educación: Modelo de lenguaje de gran tamaño (LLM) como recurso educativo. IPSUMTEC, 7(2), 157–164. https://doi.org/10.61117/ipsumtec.v7i2.321

Shah, P., Thornton, I., Kopitnik, T. A., & Hipskind, J. E. (2024). Informed consent. In StatPearls. StatPearls Publishing. National Center for Biotechnology Information (NCBI) Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK430827/

Ultsch, A., Hoffmann, J., Röhnert, M. A., von Bonin, M., Oelschlägel, U., Brendel, C., & Thrun, M. C. (2024). An explainable AI system for the diagnosis of high-dimensional biomedical data. BioMedInformatics, 4, 197–218. https://doi.org/10.3390/biomedinformatics4010013

UNESCO. (2021). Recomendación sobre la ética de la inteligencia artificial. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. https://unesdoc.unesco.org/ark:/48223/pf0000379920_spa

Vera-Rubio, P. E., Bonilla-González, G. P., Quishpe-Salcán, A. C., & Campos-Yedra, H. M. (2023). La inteligencia artificial en la educación superior: un enfoque transformador. Polo del Conocimiento, 8(11), 67–80. https://doi.org/10.23857/pc.v8i11.6193

Normativa jurídica

Asamblea Nacional del Ecuador. (2025). Código Orgánico Integral Penal. Registro Oficial.

Asamblea Nacional del Ecuador. (2011). Ley Orgánica de Donación y Trasplante de Órganos, Tejidos y Células. Registro Oficial Suplemento No. 398, 4 de marzo de 2011

Asamblea Nacional del Ecuador. (2021). Ley Orgánica de Protección de Datos Personales (R.O. Suplemento 459, 26 de mayo de 2021). Registro Oficial del Ecuador. https://www.finanzaspopulares.gob.ec/wp-content/uploads/2021/07/ley_organica_de_proteccion_de_datos_personales.pdf

Asamblea Nacional del Ecuador. (2006*)*. Ley Orgánica de Salud. Registro Oficial Suplemento No. 423, 22 de diciembre de 2006.

Constitución de la República del Ecuador. (2008). Constitución de la República del Ecuador (Registro Oficial 449, 20 de octubre de 2008). https://www.oas.org/juridico/pdfs/mesicic4_ecu_const.pdf

Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). Diario Oficial de la Unión Europea L, 12.7.2024, 1–144. https://www.boe.es/doue/2024/1689/L00001-00144.pdf

Jurisprudencia

Corte Constitucional de Colombia. (2024). Sentencia T-323/24: Sala Segunda de Revisión (Expediente T-9.301.656). https://www.corteconstitucional.gov.co/relatoria/2024/t-323-24.htm