



**ACREDITACIÓN
INSTITUCIONAL EN
ALTA CALIDAD**
Resolución 008607 de mayo 16 de 2022

**La justicia predictiva en el derecho penal colombiano: expectativas de jueces y juristas
frente al uso de la Inteligencia Artificial -IA- en el proceso penal**

Presentado por:

Diana Marcela Londoño Henao

Jorge Iván García Peláez

Trabajo de grado presentado para optar al título de

Abogado

Asesor:

Robert Anzola León

Universidad Autónoma Latinoamericana (UNAULA)

Facultad de Derecho

Medellín (Antioquia, Colombia)

2025

José Rodrigo Flórez Ruiz

Rector

Universidad Autónoma Latinoamericana

Hernán Darío Aguiar Garcés

Decano

Robert Anzola León

Tutor

Línea de investigación

Nombre de la línea

Martha Isabel Gómez Vélez

Miguel Ángel Cortes García

Evaluadores

El trabajo de grado fue sustentado el día 12 de noviembre de 2025 y obtuvo una aprobación unánime de conformidad con el Acuerdo 195 del Consejo Académico de 2016, lo cual quedó en el consignado en el acta de evaluación de trabajos de grado # 00 de 0000.

Referencia formato APA 7 edición	Londoño H., D., & García P., J. (2025). <i>La justicia predictiva en el derecho penal colombiano: expectativas de jueces y juristas frente al uso de la Inteligencia Artificial -IA- en el proceso penal</i> [Monografía]. Universidad Autónoma Latinoamericana (UNAULA).
El contenido de la presente obra corresponde al derecho de expresión del autor y no compromete el pensamiento institucional de la Universidad Autónoma Latinoamericana (UNAULA). Los autores asumen la responsabilidad por los derechos de autor y conexos	

Dedicatoria

A nuestras familias, pilares inquebrantables de amor, paciencia y fe, quienes nos enseñaron que el conocimiento no solo se busca, sino que se honra.

Agradecimientos

Este trabajo representa no solo el cierre de una etapa académica, sino también el reflejo de un camino recorrido con esfuerzo, convicción y acompañamiento.

Agradecemos a nuestros docentes y asesores, quienes con su guía, exigencia y generosidad intelectual nos impulsaron a pensar críticamente, a investigar con rigor y a escribir con responsabilidad. Su labor trasciende el aula y deja huella en nuestra formación como profesionales y ciudadanos.

Tabla de contenido

	Pág.
Resumen.....	10
Abstract.....	11
Introducción.....	12
1. Alcances de la Inteligencia Artificial -IA- predictiva frente al ejercicio del derecho	16
2. Usos y aplicaciones de la Inteligencia Artificial -IA- en el proceso penal colombiano	24
3. Riesgos, errores y sesgos derivados de los modelos de justicia predictiva frente al ejercicio del derecho procesal penal colombiano.....	38
Conclusiones.....	46
Bibliografía.....	48

Lista de figuras

	Pág.
Figura 1. Fases del procedimiento de aprendizaje autónomo	17

Lista de tablas

	Pág.
Tabla 1. Tareas que puede realizar la IA	18
Tabla 2. Principios fundamentales que deben respetarse en el campo de la IA y la justicia procesal penal.....	25
Tabla 3. Posibles usos de la IA en el ámbito de la justicia	26
Tabla 4. Diferencias entre la justicia predictiva flexible y la justicia predictiva difícil	32
Tabla 5. Herramientas basadas en el uso de algoritmos predictivos en Estados Unidos	34
Tabla 6. Sesgos en la IA que pueden incidir en una decisión penal	42

Abreviaturas

CONPES: Consejo Nacional de Política Económica y Social.

DNP: Departamento Nacional de Planeación.

IA: Inteligencia Artificial.

PretorIA: Sistema Inteligente de la Corte Constitucional colombiana para el apoyo en la selección de tutelas.

RAE: Real Academia Española.

Resumen

La justicia predictiva se entiende como la posibilidad de prever resultados en un juicio a través de cálculos algorítmicos como si se tratara de decisiones robóticas; estos modelos de justicia despiertan una serie de expectativas frente a la forma como se ejercerá el derecho a futuro, donde se presume que la automatización de los procesos decisorios permitirá alcanzar ventajas significativas, pero, a su vez, generar riesgos, debido a la existencia del denominado sesgo algorítmico. Teniendo en cuenta este panorama, la presente monografía tiene como objetivo analizar las expectativas de jueces y juristas frente al uso de la Inteligencia Artificial - IA- en el proceso penal colombiano; para alcanzar este propósito se propone un estudio de enfoque cualitativo bajo un método analítico, basado en fuentes documentales, el cual parte de la descripción de los alcances que tiene la IA predictiva frente al ejercicio del derecho; así mismo, se identifican los diferentes usos y aplicaciones de este tipo de herramientas tecnológicas en el proceso penal del país; y, por último, se reconocen los riesgos, errores y sesgos derivados de los modelos de justicia predictiva frente al ejercicio del derecho procesal penal.

Palabras clave: automatización del proceso decisorio, decisiones robóticas, inteligencia artificial, justicia predicativa, proceso penal, sesgo algorítmico.

Abstract

Predictive justice is understood as the possibility of predicting outcomes in a trial through algorithmic calculations, as if it were a robotic decision-making process. These models of justice raise a number of expectations regarding how the legal system will function in the future, where it is assumed that the automation of decision-making processes will yield significant advantages, but will also generate risks due to the existence of what is known as algorithmic bias. Given this context, this monograph aims to analyze the expectations of judges and legal professionals regarding the use of Artificial Intelligence (AI) in the Colombian criminal justice system. To achieve this objective, a qualitative study using an analytical method based on documentary sources is proposed. This study begins with a description of the scope of predictive AI in the administration of justice; it then identifies the different uses and applications of this type of technological tool in the Colombian criminal justice system; and finally, it examines the risks, errors, and biases inherent in predictive justice models within the context of criminal procedure.

Keywords: decision-making automation, robotic decisions, artificial intelligence, predictive justice, criminal proceedings, algorithmic bias.

Introducción

La Inteligencia Artificial -IA- es un modelo de tecnología que poco a poco se ha abierto espacios en las distintas esferas de la sociedad, incluida la justicia y Colombia no ha escapado a esa realidad, de ahí que desde tiempo atrás se han venido expidiendo diferentes normativas que han preparado al país para la adopción de esta tecnología. En este sentido, se destacan antecedentes normativos como la Ley 527 de 1999, aún vigente, que se ha encargado de regular y reglamentar el acceso y uso de mensajes de datos; igualmente, con la expedición del Código de Procedimiento Administrativo y de lo Contencioso Administrativo, contenido en la Ley 1437 de 2011, se reguló el uso de la tecnología al servicio de procedimientos administrativos; luego se expedirían las leyes 1581 de 2012 y 1712 de 2012, sobre protección de datos personales y transparencia y acceso a la información pública, respectivamente, así como la Ley 2108 de 2021, en la que se dio reconocimiento al internet como servicio público esencial y universal.

Si bien dichas normas no regulaban de manera específica la IA, claramente establecieron las bases para el desarrollo de una política pública sobre la materia que se estructuró en distintos documentos Conpes, dentro de los que se destaca el Conpes 3920 de 2018, que contiene la Política Nacional de Explotación de Datos Big data; el Conpes 3975 de 2019, que establece la Política Nacional para la Transformación Digital e Inteligencia Artificial; el Conpes 4024 de 2021, que formula la Estrategia de Desarrollo de la Inteligencia Artificial en Colombia; y, de manera reciente, el Conpes 4144 de 2025, que contiene la Política Nacional de Inteligencia Artificial. Todas estas políticas públicas han venido preparando a Colombia para el ejercicio de

una adecuada gobernanza de la IA en las distintas esferas del poder público, incluido el poder judicial.

No se puede desconocer que en el marco de la administración de justicia colombiana desde tiempo atrás se ha comenzado a emplear la mediación tecnológica como herramienta al servicio de la actividad judicial; de hecho, en la Corte Constitucional ya se han venido implementando herramientas automáticas basadas en el uso de la IA para la sistematización y selección de acciones de tutela en sede de revisión, como es el caso de la plataforma PretorIA; este tipo de tecnologías, señala Porcelli (2020), se valen de la IA para generar sus propios metadatos, ayudando a resolver complejos problemas de sistematización para quienes están encargados de administrar justicia.

Sin embargo, es de advertir que Colombia aún está lejos de una incorporación profunda de la IA al servicio de la justicia, lo cual se debe a que no existen reglas lo suficientemente claras para determinar los alcances que puede llegar a tener esta tecnología para el ejercicio del derecho en el futuro; tanto es así que la discusión sobre la justicia predictiva apenas se está comenzando a dar en el ámbito académico, es decir, el debate sobre los modelos de justicia anticipativa basados en algoritmos se encuentra en una fase inicial.

Ya algunos autores han comenzado a reconocer la posibilidad de que, a futuro, los modelos de justicia predictiva puedan llegar a reemplazar a jueces, fiscales, abogados y demás actores del derecho; en este sentido, doctrinantes como González & Sanabria (2019), Battelli (2021), Rouvière (2021), Gimeno (2023), Herrera (2024), Pinilla (2024), entre otros, reconocen

que las aplicaciones futuras de justicia predictiva resultan inconmensurables; y aunque el modelo actual de justicia en Colombia no permite inferir una aplicación intensa y generalizada de estas tecnologías al servicio del derecho, no se puede negar que la transformación digital en este ámbito ya ha comenzado y que poco a poco las distintas ramas del derecho irán adoptando este tipo de tecnologías.

No se puede desconocer, por tanto, que los modelos de justicia predictiva prácticamente estarían llamados a reemplazar el componente humano, por lo que en el presente estudio se busca indagar el impacto que tendría poner la IA al servicio del proceso penal colombiano, es decir, si es posible llegar a decisiones robóticas y automatizadas que puedan ser legitimadas y legalizadas por los jueces y si dicha tecnología podría reemplazar el papel que cumplen, por ejemplo, fiscales, abogados defensores, el Ministerio Público y peritos para ciertos casos.

Por lo anterior, el presente estudio da respuesta al siguiente interrogante: ¿cuáles son las expectativas de jueces y juristas frente al uso de la Inteligencia Artificial -IA- en el proceso penal colombiano? Para abordar esta cuestión se plantean los siguientes objetivos específicos: 1) describir los alcances de la Inteligencia Artificial -IA- predictiva frente al ejercicio del derecho; 2) identificar los diferentes usos y aplicaciones de la Inteligencia Artificial -IA- en el proceso penal colombiano; y 3) reconocer los riesgos, errores y sesgos derivados de los modelos de justicia predictiva frente al ejercicio del derecho procesal penal colombiano.

Metodológicamente, el estudio se desarrolla en el marco de una investigación jurídica donde, a partir de herramientas interpretativas y desde un enfoque cualitativo, se hace posible la

valoración conceptual del impacto que puede tener un modelo de justicia interpretativa al servicio del proceso penal colombiano; así mismo, se adopta un método analítico, el cual ayuda a descomponer los distintos elementos del objeto de estudio para tener una mayor comprensión sobre la manera como la IA puede generar decisiones automatizadas en casos concretos abordados por la justicia penal colombiana, recurriendo a la técnica de la revisión y rastreo documental, que permite identificar contenidos, empleando descriptores y palabras clave consultados en bases de datos de revistas indexadas, motores de búsqueda especializados, bibliotecas universitarias y sitios web de entidades estatales.

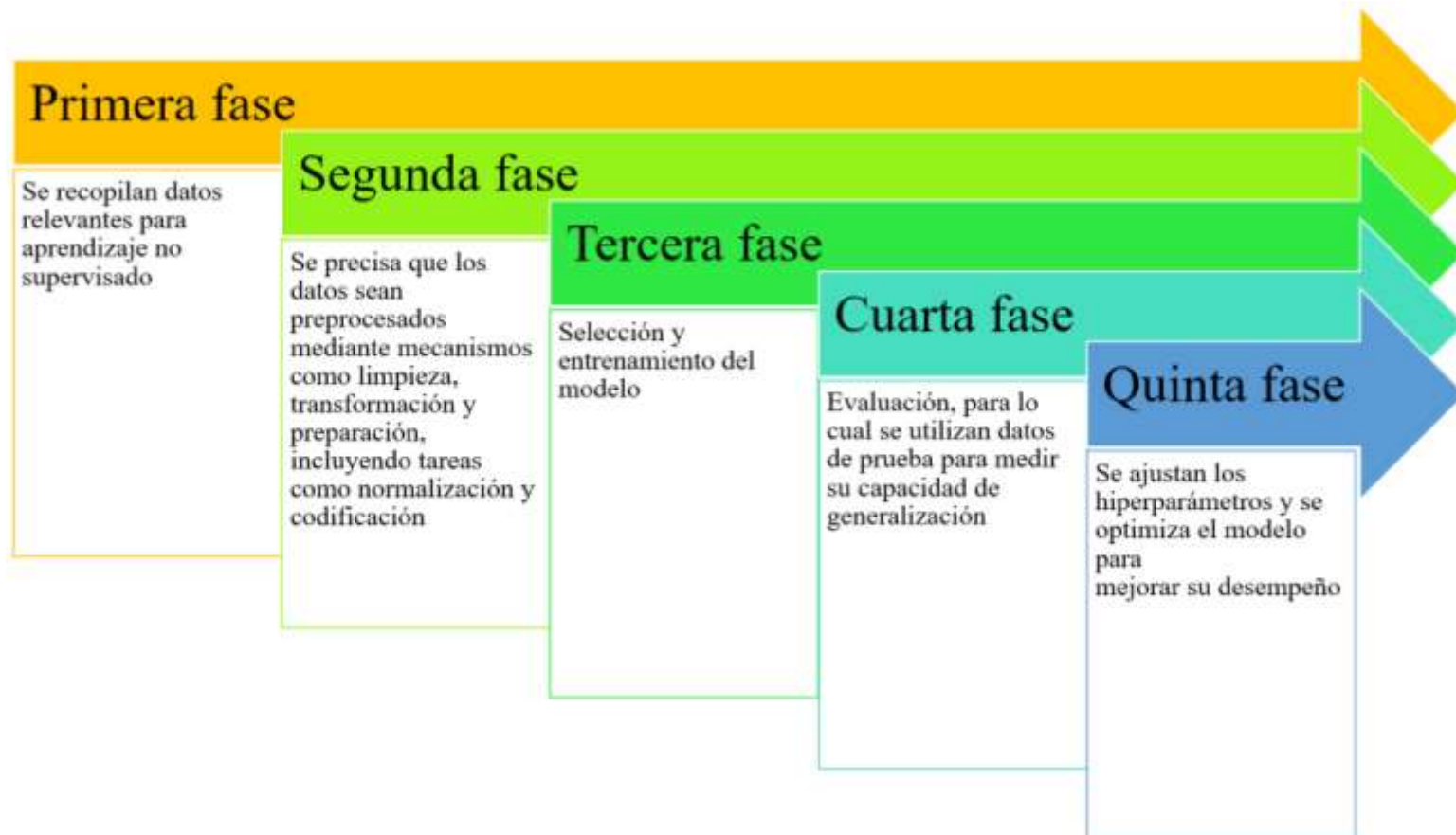
1. Alcances de la Inteligencia Artificial -IA- predictiva frente al ejercicio del derecho

La informática y la tecnología se han convertido en los últimos años en un verdadero reto para los Estados y, por ende, es un ámbito relacional que no escapa a la profundización de la globalización, aspecto marcado por la rapidez de los cambios tecnológicos y el desarrollo de las comunicaciones, lo cual ha dado lugar al surgimiento de nuevas tecnologías, como es el caso de la IA.

La IA es un campo multidisciplinario en que se desarrollan y diseñan sistemas que tiene por objeto emular la inteligencia humana en sus diversas actividades; este ámbito, según García (2024), está comprendido tanto por algoritmos como por avanzados sistemas de aprendizaje profundo, también denominados *deep learning*, los cuales necesitan datos y, por ende, se alimentan del alto tráfico de información que circula y aumenta año a año en la internet.

En el caso de la IA predictiva, es una rama de la IA enfocada en la creación de algoritmos y modelos con los que se busca anticiparse a eventos futuros o determinados resultados a partir de datos históricos y patrones de búsqueda de información; se vale de técnicas de aprendizaje automático y análisis estadístico, los cuales pueden ser aplicados a diversos campos, siendo uno de ellos precisamente el derecho; dicho aprendizaje autónomo, de acuerdo con Alpaydin (2014), conlleva una serie de fases o etapas, las cuales se articulan una con otra a medida que la IA predictiva se ajusta a las necesidades del entorno.

Figura 1. *Fases del procedimiento de aprendizaje autónomo*



Fuente: elaboración propia a partir de Alpaydin (2014).

Así mismo, Bobadilla (2020) señala que a través de la IA es posible realizar múltiples tareas, tales como aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semisupervisado y aprendizaje por refuerzo.

Tabla 1. *Tareas que puede realizar la IA*

Tarea	Descripción
Aprendizaje supervisado	A través de este es posible entrenar un modelo a partir de imágenes con etiquetas, empleando algoritmos de clasificación. Mediante esta tarea se pueden extraer miles de datos que se encuentran etiquetados automáticamente, teniendo como referencia las interacciones entre las personas a través de redes sociales.
Aprendizaje no supervisado	Está basado en información no etiquetada y tiene como objetivo simplificar los datos, de manera que cuando estos se encuentran muy dispersos y no aportan mucha información pueden ser comprimidos y conservados de forma condensada, obteniendo resultados de búsqueda más precisos.
Aprendizaje semisupervisado	Aglutina datos etiquetados y no etiquetados; por tanto, mezcla aprendizaje supervisado y no supervisado, valiéndose de reseñas para aprender a identificar palabras y patrones que se pueden asociar con ciertas posiciones.
Aprendizaje por refuerzo	Se trata de un tipo de aprendizaje automático a través del cual la IA se encarga de tomar decisiones o llevar a cabo acciones,

con el propósito de maximizar una recompensa a lo largo del tiempo.

Fuente: elaboración propia a partir de Bobadila (2020).

Las anteriores tareas pueden estar perfectamente al servicio de la resolución de conflictos jurídicos, lo cual puede suponer que las nuevas tecnologías requieran necesariamente de la imposición de límites y prohibiciones, especialmente en el marco del derecho constitucional para garantizar derechos como la dignidad de las personas y el libre desarrollo de la personalidad; por tanto, tal y como afirman Cajiga, & Vergara (2021), la discusión hoy en día se ha centrado sobre si la IA debe ser regulada y cómo se debe diseñar su regulación, frente a lo cual se ha exigido bastante cautela, máxime si se tiene en cuenta que esta tecnología está guiada por reglas de comportamiento autónomo.

La IA, por tanto, exige una regulación explícita que permita sentar las bases para la implementación de modelos de justicia predictiva, concepto amplio que “se promueve como una herramienta para la eficiencia procesal, combinando una mejora en la calidad de la toma de decisiones y una reducción de la actividad judicial” (Armenta, 2021, p. 262); así mismo, dicha justicia predicativa se encuentra asociada al concepto de jurimetría, el cual obedece a la “aplicación de la inteligencia artificial y *machine learning* a los tradicionales buscadores legales y jurisprudenciales obteniendo nuevas funcionalidades que dan lugar al desarrollo de lo que también se ha denominado justicia predictiva” (Peralta, 2021, p. 37).

Por lo tanto, tal y como señala Pinilla (2024), la justicia predictiva busca la previsión de resultados de un juicio empleando algoritmos, el problema es que se trata de un instrumento que enfrenta desafíos relacionados con su seguridad, racionalidad de la incertidumbre, límites estructurales y problemas relacionados con la recolección del dato, el costo de la digitalización y la evolución de las orientaciones jurisprudenciales.

De acuerdo con Rouvière (2021), la justicia predictiva es un proyecto que consiste en movilizar la IA para tratar de anticipar la resolución de una controversia; presupone el uso de algoritmos, es decir, reglas de la forma “si...”, “entonces...”, que estructuran toda la programación informática. Todavía impensable en el siglo XX, la pregunta hoy en día ya no es si la ley puede reducirse a teorías generales, sino a algoritmos.

N este sentido, la IA procura la sustitución del juez por la máquina, el desarrollo de la mediación anticipando el resultado probable del juicio y la realización del ideal de seguridad jurídica eliminando la incertidumbre judicial. En segundo plano están los ahorros esperados por la reducción del número de jueces o abogados, cuestiones teóricas que son igual de fundamentales porque se relacionan tanto con la definición de la ley como con la teoría del razonamiento jurídico.

La mayoría de las contribuciones que tratan sobre la justicia predictiva tienen un tono angustioso. Constantemente señalan los riesgos: reproducción estática de datos y errores, estandarización de soluciones, reemplazo de humanos por máquinas, etc. Para lograr protección de la amenaza de una sustitución de seres humanos por máquinas, se invocan murallas legales.

Consisten principalmente en argumentar que la máquina ciertamente no es un tribunal en el sentido del Convenio Europeo de Derechos Humanos o que los algoritmos deben hacerse públicos por razones de transparencia.

Las comunicaciones sobre el tema son a menudo fantasía en el verdadero sentido del término: la mayoría de las veces es un debate sobre una apariencia sin sustancia. ¿No se basa el sistema de justicia completamente robótico en la misma caricatura teórica que reduce el razonamiento legal a un silogismo? De la misma manera, ¿cómo se puede analizar y criticar lo que existe solo en el estado de suposición? ¿No sería necesario recopilar experiencias de campo en términos de programación?

En verdad, la realidad es mucho más realista: incluso con los recursos de la IA, no es fácil desarrollar un buscador de jurisprudencia que sea cien por ciento relevante. De hecho, llevar a cabo una investigación jurídica presupone la delimitación posterior de un problema que en sí mismo sólo puede plantearse con la ayuda de conceptos. Por lo tanto, el ser humano está tanto río arriba (para hacer la pregunta) como río abajo (para interpretar las respuestas). Esto está muy lejos de lo totalmente algorítmico.

Uno de los principales malentendidos del debate proviene de la definición del objeto de la predicción. Al transponer el término inglés "*predictive*", se le da el fuerte significado de una predicción, es decir, una predicción o anticipación de la solución misma de la disputa (quién gana, quién pierde y qué convicciones se pronuncian).

Sin embargo, las predicciones meteorológicas no son realmente predicciones meteorológicas a pesar de su nombre. Se basan en extrapolaciones y en la interpretación de diferentes escenarios de desarrollo: por lo tanto, corresponde al hombre determinar cuál le parece más probable, teniendo en cuenta todo el contexto de la información. Básicamente, no es muy diferente para el abogado que trata de entender lo que decidirá un juez según los precedentes, el tribunal al que se le conoce o incluso su personalidad. Por razones obvias relacionadas con la vida privada, no es concebible conocer las opiniones políticas, morales o religiosas de un magistrado. El único dato disponible es el texto de la sentencia, que es el único asunto que puede tratarse de manera sistemática y razonable.

Lo cierto es que la justicia predictiva está diseñada como una apuesta de futuro, pero, tal y como lo reconoce Gimeno (2023), debe ser a corto plazo, de ahí que los legisladores de todo el mundo deben comenzar a aprobar normas que tengan por objeto hacer uso de la IA para la modernización de la administración de justicia, como ya viene ocurriendo en algunas órbitas del derecho comparado, como es el caso de España, donde ya se han propuesto iniciativas sobre la materia, mediante las cuales se pretende utilizar instrumentos de jurimetría, así como técnicas analíticas de datos para los jueces.

No obstante, el panorama es incierto frente al uso de herramientas de justicia predictiva decisorias para órganos judiciales, imponer medidas cautelares o dictar sentencias, ya que el uso real de la IA solo está dirigido para técnicas auxiliares que buscan garantizar eficiencia judicial digital y quizás en el largo plazo podría pensarse en una justicia predictiva decisoria.

Al respecto de lo anterior, Battelli (2021) afirma que aún no es posible imaginar una justicia ejercida exclusivamente por un juez robot, ya que una aplicación de la ley bajo un proceder simple y mecánico significaría desdibujar la representación de la realidad del derecho; lo que sí es cierto es que la actividad interpretativa a futuro irá adquiriendo cada vez mayor relevancia, sobre todo porque la IA puede procurar un modo correcto de concebir la toma de decisiones, pero aquí a la tecnología solo le correspondería ofrecer apoyo y ayuda para que sean los seres humanos los que tomen las decisiones sin que lleguen a ser reemplazados en una actividad tan delicada como es el ejercicio de la función judicial.

2. Usos y aplicaciones de la Inteligencia Artificial -IA- en el proceso penal colombiano

Al analizar los usos y aplicaciones de la IA en el proceso penal es necesario reconocer varios enfoques: el primero implica establecer de qué manera la IA influye en el proceso penal colombiano y cómo está incidiendo en los juristas y el segundo tiene que ver con los problemas éticos y jurídicos que plantea la IA, sobre todo en materia de derechos humanos. Ambos enfoques generan una serie de interrogantes: ¿el marco jurídico colombiano está adecuado para hacer frente a los avances de la robótica? ¿Se pueden regular con suficiencia las nuevas tecnologías robóticas, sobre todo aquellas que se caracterizan por promover un comportamiento autónomo? ¿Es posible una disciplina jurídica autónoma y propia para los robots? ¿En qué situaciones es posible justificar que un sistema autónomo no respete la voluntad humana?

Frente a estos interrogantes Beloso (2019) plantea que la IA, al estar sustentada en algoritmos, resulta incomprensible para la gran mayoría de las personas. Habría que tener en cuenta que las personas deben poder interpretar los algoritmos que determinan las decisiones automatizadas; por lo tanto, preocupa los usos que se le puede dar a esta tecnología y al funcionamiento de los algoritmos, máxime si se tiene en cuenta que una de las áreas susceptibles de aplicar la IA es precisamente los sistemas judiciales. Así, al aplicar la IA a áreas del derecho como el procedimiento penal, hay que tener en cuenta una serie de principios fundamentales que se deben respetar, los cuales son básicos para el procesamiento de las decisiones judiciales y los datos mediante algoritmos.

Tabla 2. *Principios fundamentales que deben respetarse en el campo de la IA y la justicia procesal penal*

Principio	Descripción
Respeto a los derechos humanos	Tiene por objeto brindar garantías y soluciones que desde su diseño hasta su aplicación tengan pleno apego a los derechos establecidos en tratados y convenios internacionales; a su vez, se relaciona con la no discriminación, la cual se expresa en las capacidades de los procesos y operaciones en materia penal para poner en evidencia alguna discriminación al sistematizar datos que tengan que ver con ciertos individuos.
Principio de calidad y seguridad	Implica la posibilidad de procesar datos a través de aprendizajes automáticos a partir de garantías que se ofrecen en distintas partes del procesamiento del dato.
Principio de transparencia	Se aplica a las diferentes metodologías y técnicas utilizadas para procesar decisiones de carácter judicial, haciendo un especial énfasis en las condiciones de acceso y comprensión de las distintas formas del procesamiento del dato ante la posibilidad de que expertos o autoridades en la materia adelanten alguna auditoría externa.
Principio “bajo el control del usuario”	Permite convertir al usuario en un agente activo para

que se sienta a cargo de sus elecciones; en este sentido, el juez debe poder volver en cualquier momento a las decisiones judiciales y acceder a los datos que se han empleado para generar un resultado; así mismo, el usuario debe ser informado del carácter vinculante o no vinculante de las soluciones que proponga la IA.

Fuente: elaboración propia a partir de Beloso (2019).

El cumplimiento de estos principios se debe garantizar a lo largo del procesamiento de las decisiones judiciales y de los datos a través de algoritmos; por lo tanto, el uso de la IA en los sistemas judiciales, especialmente las solicitudes de IA que procesan decisiones y datos judiciales, recogen al menos cuatro tipos de posibles usos en el ámbito de la justicia de acuerdo con los posibles riesgos derivados de su uso.

Tabla 3. *Posibles usos de la IA en el ámbito de la justicia*

Uso	Descripción
Usos que deben ser fomentados	<ul style="list-style-type: none"> - Mejora en la búsqueda de jurisprudencia. - Acceso a la legislación. - Creación de nuevas herramientas estratégicas.
Posibles usos que requieren considerables precauciones metodológicas	<ul style="list-style-type: none"> - Ayuda en la elaboración de baremos para ciertos asuntos litigiosos. - Apoyo a mecanismos alternativos de solución de conflictos en materia civil.

	- Resolución de disputas en línea.
	- Uso de algoritmos en la investigación penal para la identificación de lugares donde se están cometiendo delitos.
Usos que se pueden considerar después de realizar estudios científicos adicionales	- Perfilado de jueces. - Anticipación de decisiones judiciales.
Usos a considerar con la más extrema reserva	- Utilización de algoritmos para elaborar perfiles individuales en materia penal. - Brindar información a los jueces para facilitar el fundamento de sus decisiones, según las adoptadas por otros jueces.

Fuente: elaboración propia a partir de Belloso (2019).

Ahora bien, Vaillant et al. (2004) señalan que decir que una inteligencia es artificial es insistir en el hecho de que no es humana. En otras palabras, no reproduce el proceso real de juzgar a un ser humano, sino solo el resultado que se supone que el ser humano debe lograr. Se ha establecido que el software de inteligencia artificial que identifica animales en imágenes no lo hace razonando como los seres humanos. Por lo tanto, lo que es objeto de una imitación no es el proceso de razonamiento sino su conclusión: es probable que el programa llegue a él por caminos completamente diferentes y que quizás parezcan totalmente exóticos para un usuario humano.

Frente a esto, según señala Rouvière (2021), la inteligencia artificial retoma los estudios de psicología cognitiva de la década de 1980 relacionados con los sistemas expertos. Se suponía que estos últimos debían reproducir el razonamiento de un experto (por ejemplo, un diagnóstico médico), pero este razonamiento primero tenía que traducirse en reglas. La insuficiencia de la potencia de cálculo de los ordenadores de la época condujo a una explosión combinatoria, es decir, a la imposibilidad de que el ordenador gestionara la multiplicidad de reglas y sus articulaciones. Si bien la potencia de cómputo ha aumentado desde entonces, este camino metodológico ha sido abandonado: el objetivo ya no es imitar el proceso de pensamiento, sino solo su resultado. En este sentido, podemos ver que la diferencia entre la inteligencia humana y la artificial no es una diferencia de grado (la máquina calcula más rápido) sino de naturaleza (son cualitativamente diferentes).

Por lo tanto, cada vez es más relevante el papel de los algoritmos: las instrucciones simples que constituyen un programa de computadora, los algoritmos son el resultado de la formalización de ciertos procedimientos de pensamiento. Su estudio constituye una rama de las matemáticas y la lógica, ya que en sí mismos, son una técnica neutral. Lo que importa es el propósito que persiguen y la lógica subyacente detrás de su diseño. Como señala O'Neil (2017), son solo opiniones insertadas en el código de la computadora.

Por lo tanto, tal y como plantea Rouvière (2021), es importante acordar el propósito de los algoritmos y su propósito. Se pueden utilizar para la organización de citas, corrección ortográfica y justicia predictiva. En sí mismos, los algoritmos sirven para lograr un fin específico, son solo herramientas.

Su principal interés en la ley sería permitir la gestión de datos masivos (big data). Según estadísticas del Ministerio de Justicia para el año 2019, la justicia civil genera un promedio de más de 2,5 millones de decisiones al año; La justicia penal es cercana a 1,2 millones de decisiones por año, la justicia administrativa 250.000. Por último, sólo el Tribunal de Casación produce más de 30.000 resoluciones al año, aproximadamente una cuarta parte de las cuales son en materia penal. Aunque algunas decisiones sean de poco interés (valoraciones resumidas, órdenes de unión, etc.), los algoritmos podrían permitir explotar estos datos a gran escala, que se acumulan año tras año (p. 99).

Desde este punto de vista, de acuerdo con Schauer (2018), toda justicia predictiva se basa necesariamente en datos que permitirán extrapolar los resultados y aplicarlos al mundo real, y cualquier proyecto en esta dirección implica utilizar al menos las constituidas por las propias decisiones judiciales, a falta de poder obtener datos personales sobre los magistrados. Este último punto es de hecho parte del programa teórico del realismo jurídico estadounidense: lograr, a través de las ciencias sociales, predecir las soluciones que tomará el juez.

Por su parte, Rouvière (2021), señala que la Big Data también es un problema desde el punto de vista interno de la máquina. Deberían permitir mejorar el funcionamiento de los algoritmos. Lo que se llama “aprendizaje automático” requiere el uso de datos para entrenar o educar a la máquina. La precisión de los resultados es validada por un ser humano (aprendizaje supervisado) o el proceso es completamente automático (aprendizaje no supervisado). El “aprendizaje profundo” es el que más fantasías provoca ya que lleva a la máquina a desarrollar

su propia lógica (en forma de redes neuronales, reglas que la máquina puede crear por sí misma). Estos se denominan modelos de “caja negra” porque es imposible describir de manera simple e instantánea cómo funciona realmente el programa. Si bien estos modelos ya están implementados en los campos de los juegos o el reconocimiento de imágenes, parecen mucho más difíciles de aplicar al razonamiento legal. Y por una buena razón: la justicia predictiva prevista presupone un modelo de razonamiento al que es necesario adherirse. Toda justicia predictiva depende de dicho modelo teórico. La relación es exactamente la misma que entre el plan de un arquitecto y la construcción de una casa. El modelo teórico del razonamiento jurídico es este plan. Por analogía, la programación y el uso de algoritmos corresponden a trabajos de albañilería.

Agrega Rouvière (2021) que para comprender los proyectos de justicia predictiva, es necesario explicar sus presupuestos teóricos. Hay, por supuesto, toda una serie de matices en las diferentes presentaciones que se le dan a la justicia predictiva, pero se pueden distinguir dos tendencias principales. Como no existe un calificativo indiscutible para designarlos, podríamos hablar convencionalmente de justicia predictiva “blanda” o “dura” (“débil/moderada” o “fuerte” según otros autores).

Al igual que las ciencias duras, la justicia predictiva “dura” se basa en la idea de que es posible encontrar una solución a una disputa determinada con certeza. Al igual que la ley “dura”, a diferencia de la “ley blanda”, la justicia predictiva dura es normativa: promueve la sustitución

del hombre por la máquina y promete un “jus ex machina”*. Esta representación, de acuerdo con Abiteboul & G'ssell (2019), parece ser dominante, no solo en el periodismo destinado al público en general, sino también en la mayoría de las contribuciones francesas sobre el tema [6]. Por el contrario, la justicia predictiva "blanda", que tendría como objetivo predecir no el resultado de la disputa sino los juicios relevantes y los argumentos probables que se invocarán, casi nunca se califica como tal. Sin embargo, los dos proyectos son bastante distintos en su objeto, su finalidad, la herramienta utilizada, el método, el modelo de razonamiento preferido, la visión del derecho que transmiten y su grado de normatividad.

Por tanto, tal y como advierte Rouvière (2021), el propósito de la justicia predictiva severa es determinar la solución de la disputa y reemplazar al hombre con la máquina. El propósito de la justicia predictiva flexible es explorar los datos para, sobre todo, reducir el tiempo que lleva buscar información. La justicia predictiva dura promete una herramienta centrada en estadísticas cuantitativas y una lógica para lidiar con litigios masivos; Por el contrario, la justicia predictiva flexible tiene como objetivo establecer una clasificación cualitativa de los datos y se centra en una lógica de caso.

Es necesario un enfoque cuantitativo para obtener estadísticas significativas (no promediamos los daños con tres o cuatro casos). El enfoque cualitativo, por otro lado, presupone una fina precisión en la búsqueda de información y una lógica de singularidad: si solo hay una decisión relevante, el programa debe ser capaz de identificarla. Se puede ver así que la lógica cualitativa se basa en un punto de vista jurídico interno: el razonamiento jurídico se concibe ante

* Variante de la expresión latina “Deus ex machina”, la cual se refiere a una situación difícil e irresoluble en el derecho que se resuelve de manera inesperada y artificial por un elemento, personaje o fuerza externa que no hace parte del problema.

todo como un argumento. El enfoque estadístico es más bien desde un punto de vista externo que pretende encontrar regularidades de forma científica según una lógica causal: detrás del discurso jurídico, es la recurrencia de un pensamiento y una práctica lo que es decisivo para concluir el resultado futuro de la disputa.

Por lo tanto, está claro que el enfoque cualitativo valora la ley como su propia lógica, mientras que el enfoque cuantitativo tiende a reducir la ley al ejercicio del poder y, por lo tanto, a captarla en su dimensión política. En resumen, es la normatividad del resultado, su necesidad, lo que se busca, mientras que el enfoque flexible reconoce una normatividad en el método de argumentación sin poder pronunciarse directamente sobre el resultado final de la decisión judicial.

Tabla 4. *Diferencias entre la justicia predictiva flexible y la justicia predictiva difícil*

	Justicia predictiva flexible	Justicia predictiva difícil
Objeto	Exploración de datos	Resolución de disputas
Propósito	Ahorro de tiempo	Sustitución de humanos y máquinas
Herramienta	Clasificación cualitativa	Estadísticas cuantitativas
Método	Lógica de casos	Lógica de masas
Modelo de razonamiento	Argumentación	Causalidad
Visión de la ley	La ley como lógica propia	La ley como gestión o política
Visión de la normatividad	Normatividad del método	Normatividad del resultado

Fuente: elaboración propia a partir de Rouvière (2021).

Independientemente del modelo preferido, el objetivo de previsibilidad permanece pero no se relaciona con los mismos objetos y, sobre todo, no tiene las mismas consecuencias. Por lo tanto, no es la misma justicia predictiva. El modelo flexible se mantiene al servicio de la lógica jurídica tradicional: recopilar textos (leyes y sentencias) para utilizarlos desde una perspectiva argumentativa. Por el contrario, el modelo duro promueve una nueva forma de pensar el derecho: la herramienta técnica es la oportunidad (o el pretexto) para hacer prevalecer otra concepción de lo legal.

Agrega Rouvière (2021) que, mientras que el modelo flexible trabaja para preservar la identidad jurídica, el modelo duro promueve nuevos conocimientos jurídicos extraídos de las series observadas en las masas de decisiones. Si bien el proyecto flexible conserva una visión relativamente clásica de las fuentes del derecho, el proyecto difícil, para lograr la plena eficacia, también presupone encontrar y agregar datos fuera de las fuentes del derecho, vinculados a la posición y personalidad de los jueces. La lógica de "a medida" que preside la creación de argumentos específicos para cada caso cedería, en el modelo duro, a una lógica de gestión del estándar en la que la estandarización daría sustancia al dogma de la solución única.

Autores como Pinilla (2024) señalan que en la práctica ya es posible identificar el uso de algoritmos predictivos al servicio de la justicia a nivel global y uno de los países que está a la vanguardia es Estados Unidos, el cual cuenta con numerosas herramientas sobre la materia.

Tabla 5. *Herramientas basadas en el uso de algoritmos predictivos en Estados Unidos*

Herramienta	Funcionalidad
Public Safety Assessment -PSA-	Evalúa si un acusado puede ser dejado en libertad antes de un juicio a partir de la valoración del riesgo de reincidencia o falta de comparecencia.
Level of Service Inventory - Revised -LSI-R-	Es empleado para otorgar permisos de salud y libertad condicional a partir de la ponderación de variables como los antecedentes penales, la reincidencia, la educación, el empleo, antecedentes de drogas o alcohol, entre otros.
Harm Assessment Risk Tool - HART-	Predice los riesgos de que un individuo cometa delitos en un periodo de dos años, utilizando datos relacionados con la edad, el sexo y el historial delictivo y a partir de ello decidir si un individuo es de bajo riesgo y si debe ser enviado a programas de rehabilitación.
Crime Anticipation System - CAS-	Predice delitos de alto impacto en zonas específicas a partir del uso de información demográfica, socioeconómica y antecedentes delictivos.
National Data Analytics Solution -NDAS-	Se encarga de predecir delitos violentos, delitos con armas y posibles víctimas, mediante el registro de detenciones, identificación de sospechosos y oferta de servicios sociales o de salud para la prevención del crimen.
Correctional Offender Management Profiling for	Es empleado para valorar los riesgos de reincidencia e identificar necesidades criminológicas de los sujetos.

Alternative Sanctions -

COMPAS-

Fuente: elaboración propia a partir de Pinilla (2024).

No obstante, en otros países es posible identificar diversos usos y aplicativos basados en modelos de IA predictiva. En China, según Agudelo (2021), se cuenta desde tiempo atrás con la herramienta Xiao Fa, que es un asistente de IA que tiene por objeto el análisis de crímenes para la estandarización de condenas y generación de borradores de sentencia; este asistente permitió que en 2017 se creara el primer tribunal virtual de ese país, seguido por tribunales similares especializados en operaciones en red, comercio electrónico y propiedad intelectual. En Estonia se han venido realizando pruebas piloto con jueces robots para casos menores, que tiene por objeto la simplificación de los servicios, pero siempre las decisiones deben ser revisadas por un juez humano.

En Holanda, tal y como lo reconoce Battelli (2021), se cuenta con una plataforma probabilística de justicia robótica para temas relacionados con juicios de separación denominada *Rechtwijzer uit elkaar* (guía legal sobre la separación)*, la cual le ayuda a los usuarios a llevar a cabo divorcios en línea de manera accesible, asequible y confiable, donde se brinda orientación de un administrador de casos robótico y una verificación de los acuerdos por parte de un abogado especializado en derecho de familia; la plataforma también calcula la manutención derivada de los costos de alimentos y brinda asistencia para que los niños, niñas y adolescentes no sufran las consecuencias del divorcio de sus padres.

* <https://rechtwijzer.nl/uit-elkaar>

Se destaca igualmente el modelo experimental basado en algoritmos del Tribunal Europeo de Derechos Humanos, el cual, según Battelli (2021), tiene un grado de precisión cercano al 80% en la predicción de sentencias, hasta el punto de llegar a conclusiones bastante similares a las adoptadas por tribunales conformados por jueces humanos; en España, en el sector privado se han creado herramientas de análisis jurisprudencial estadístico y predictivo como es el caso de Jurimetría^{*}, que es una plataforma que ofrece respuestas a preguntas complejas de carácter legal, el cual determina las posibilidades de éxito de un caso y la probable duración de un juicio.

En América Latina, de acuerdo con Corvalán (2017), se encuentra el sistema de justicia predictiva Prometea, que es un sistema de recepción de expedientes del Tribunal Superior de Justicia de Buenos Aires (Argentina), el cual se encarga de realizar la lectura de sentencia de primera y segunda instancia, analizar dictámenes y emitir predicciones sobre un modelo determinado para resolver el caso; este sistema fue tomado como modelo en Colombia para la implementación del sistema PretorIA por parte de la Corte Constitucional.

Para un adecuado uso y aplicación de la IA en el proceso penal colombiano es necesario contar con un eficiente sistema de expertos, los cuales, de acuerdo con Cajiga & Vergara (2021), están compuestos de una base de conocimientos de los hechos y de las experiencias de los expertos en un tema determinado; por lo tanto, el diseño de un modelo de justicia penal predictiva debe contar con equipos multidisciplinarios, compuestos no solo por expertos en informática, sino también en derecho penal, derecho probatorio y derecho procesal penal.

* <https://www.aranzadilaley.es/productos/jurimetria.html>

Estos sistemas deben ser cautelosos, sobre todo en la forma como deben estar llamados a salvaguardar los derechos, la democracia y la justicia; por lo tanto, en las etapas de desarrollo de estas tecnologías es preciso que las personas comprendan a ciencia cierta cómo se llevan a cabo, para lo cual es fundamental contar con acciones transparentes, especialmente en materia de acceso a la información y la garantía de la libertad de expresión, lo que exige un enfoque orientado al servicio de la administración de justicia.

3. Riesgos, errores y sesgos derivados de los modelos de justicia predictiva frente al ejercicio del derecho procesal penal colombiano

Con los avances de la IA es claro que los algoritmos se utilizarán cada vez más en los tribunales de justicia, especialmente en materia penal, para fijar fianzas, dictar sentencias e incluso para contribuir a la determinación de la culpabilidad o inocencia de un procesado. Se trata de un cambio que le permitiría a la IA aumentar el juicio humano, lo cual puede ser extremadamente beneficioso para el sistema judicial en general; sin embargo, esto también podría conllevar riesgos, errores y sesgos, sobre todo porque los modelos de justicia predictiva pueden generar decisiones deshumanizantes en procura de un buscar justicia objetiva y plena.

Al respecto de lo anterior, Belloso (2019) señala que los métodos predictivos adolecen de inconvenientes y fallas porque eliminan los componentes subjetivos que aportan los seres humanos para la calificación de una conducta; de hecho, la experiencia ha demostrado que los sistemas de IA que ayudan a neutralizar el riesgo pueden lograr incluso resultados más certeros que los aportados por las decisiones de los jueces; no obstante, algunos sistemas algorítmicos ya han demostrado sesgos que, en la práctica, han sido objeto de críticas. Algunas de estas herramientas, por ejemplo, pueden sugerir decisiones racistas, así como predicciones no del todo fiables, lo cual se evidencia sobre todo en países que han empleado estas herramientas donde existen altos índices de criminalidad en población negra.

Hay que tener en cuenta que en la práctica los sistemas predictivos de IA operan a partir del rastreo de antecedentes judiciales para luego recopilar providencias y procesar los datos, empleando algoritmos que permiten anticipar la posible decisión para una situación específica. Desde esta perspectiva, señala Herrera (2024), estas tecnologías no estarían sustituyendo la tarea del componente humano en el sistema judicial, sino que, a partir de insumos jurisprudenciales, se están prediciendo decisiones que deben ser adoptadas para una determinada controversia o causa jurídica; en tal contexto, el riesgo sería mucho menor, pues no se busca sustituir a los jueces, ni mucho menos la máquina está sustituyendo los procesos mentales de quienes administran justicia, lo que se busca es que esos sistemas proyecten la decisión que judicialmente se adoptaría para un determinado caso.

Para Rouvière (2021), la preocupación y la cultura del miedo que se ha apoderado de los algoritmos está llevando a pedir a los diversos actores, especialmente en el ámbito de la justicia estatal, que publiquen sus algoritmos; de hecho, puede ser tentador para un actor legal proceder de acuerdo con los estándares de los gigantes digitales (el GAFAM*), clasificando los juicios por popularidad (número de clics, frecuencia de citas, etc.) y por actualización (favoreciendo la última búsqueda) en lugar de que sea por relevancia semántica (los conceptos involucrados en el análisis) o incluso para ocultar ciertas decisiones; sin embargo, podría ser económicamente contradictorio exigir a las empresas que obtienen su riqueza de la creación de algoritmos que los hagan públicos y se expongan así a perder lo que constituye el corazón mismo de su investigación e inversiones.

* Acrónimo para las mayores empresas tecnológicas del mundo: Google, Amazon, Facebook (Meta), Apple y Microsoft.

Además, hay que relativizar la exigencia de transparencia, porque solo sería necesaria en el caso de que los asuntos sean juzgados por una máquina. La imparcialidad del juez encontraría su contrapartida en la neutralidad de los algoritmos y, más aún, de los datos; de hecho, en el caso de que los motores de búsqueda agreguen datos legislativos y jurisprudenciales podría dar lugar a una forma de selección natural. La automatización de determinadas búsquedas (como las de jurisprudencia reiterada) no se refiere directamente a la resolución del litigio, sino únicamente a la forma en que se argumenta.

En última instancia, lo que revela la cuestión de los algoritmos es la ausencia de una investigación real sobre una clasificación sistemática de los diferentes modelos de razonamiento jurídico. Algunos son compatibles con algoritmos, otros probablemente no o muy débilmente. En todo caso, las decisiones en la ley obedecen (al menos en parte) a una lógica propia, pues hay diferentes formas de razonar para el mismo conjunto de hechos.

La idea de un modelo lógico y silogístico, incluso uno basado en la lógica deóntica (la lógica de las normas), es en gran medida una forma de caricatura del razonamiento legal. Aunque las sentencias se presentan en forma silogística, todos los juristas saben que se trata de hipocresía piadosa.

Por otro lado, el modelo político, el del realismo jurídico, lleva a escudriñar la relación de causa y efecto entre los motivos (no declarados) de los jueces y la solución que favorecen. Es esta misma influencia la que actualmente empuja el razonamiento jurídico hacia una forma de ponderación de intereses, o incluso a veces de toma de decisiones, en el sentido de que la

actividad judicial se concibe más como una actividad política que como una forma de pensamiento por derecho propio.

Un enfoque moral ético o axiológico del razonamiento jurídico, es decir, como una búsqueda del significado correcto o verdadero de los textos, se enfrenta a la dificultad de poner una evaluación moral adecuada en una forma computable (incluso recurriendo a ponderaciones). Esta dificultad también se debe al hecho de que el análisis moral ciertamente tiene un componente emocional importante.

Finalmente, el modelo de argumentación, el de las razones para decidir y el de la casuística del trabajo, es difícilmente traducible en reglas fijas. Los conceptos legales están evolucionando esencialmente, como lo demuestra toda la historia del derecho. La IA ciertamente puede ayudar a movilizar datos relevantes, pero no puede interpretar cada uno en su singularidad.

Los algoritmos parecen apuntar negativamente a todo lo que sigue siendo propiamente humano en el razonamiento jurídico: evaluar la calidad de un argumento, desarrollar una estrategia, imaginar y construir un problema jurídico, dominar y revisar los conceptos jurídicos; en otras palabras, la única ley que se puede reducir a los algoritmos es la que no es propiamente humana.

Aunque la IA puede parecer una herramienta que permite ahorrar tiempo y, en cierta medida, puede dar lugar a la prevención de errores humanos, su imagen de infalibilidad no es del todo cierta, ya que puede presentar fallos en su funcionamiento, al igual que sesgos, los cuales

claramente podrían afectar la toma de decisiones en un proceso penal. Ya en la práctica, según García (2024), es posible identificar múltiples casos en los cuales la IA ha sido generadora de riesgos, errores y sesgos, los cuales pueden ser replicados en el ámbito del derecho penal.

Por lo tanto, es fundamental considerar quién y cómo se debe programar la IA para una toma de decisiones en un juicio, máxime si se tiene en cuenta que los algoritmos de esta clase de tecnología no son, ni serán desarrollados propiamente por juristas, sino por empresas que ponen su experiencia en tecnología al servicio de las necesidades de la justicia; por ello, para el derecho es necesario conocer cómo la IA debe ser entrenada y cómo ajustar los resultados para evitar los sesgos, esto es, “errores sistemáticos en los que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras” (RAE, 2024).

Los sesgos derivados de los modelos de justicia predictiva frente al ejercicio del derecho procesal penal colombiano pueden ser de diversa naturaleza: por un lado, están los relacionados con la gestión algorítmica del dato y, por el otro, los sesgos cognitivos, los cuales pueden influir de manera directa o indirecta frente a una decisión penal.

Tabla 6. *Sesgos en la IA que pueden incidir en una decisión penal*

Sesgos relacionados con la gestión algorítmica del dato	
Sesgo de datos	Se presenta cuando los datos que se utilizan no son objetivos y se han seleccionado de una manera temeraria para favorecer a determinadas personas, grupos o soluciones. En un proceso penal, esto puede llevar a decisiones incorrectas,

	al inferir relaciones inexistentes, por ejemplo, entre una prueba y unos hechos.
Sesgos del algoritmo	Se produce cuando las decisiones que se toman durante el diseño del algoritmo dan lugar a un resultado sesgado; por ejemplo, si un algoritmo empleado para una decisión penal considera que ciertas características afines con el género o la etnia están relacionadas con la comisión de delitos, esto puede llevar a una decisión injusta contra determinadas personas.
Sesgo de selección de características	Se da cuando se seleccionan características o variables para el entrenamiento del modelo, lo cual genera desviaciones en los resultados, de tal forma que, si un sistema considera, por ejemplo, los antecedentes penales como criterio para sancionar un nuevo delito, esto puede perpetuar desigualdades en el juzgamiento de quienes tienen dichos antecedentes.
Sesgo de evaluación	Se presenta cuando se emplean características o variables para entrenar modelos que generan resultados intencionalmente sesgados; de esta manera, si un sistema decisional de una pena utiliza como criterio principal o único criterio los antecedentes penales, esto podría excluir de manera sistemática a otros presuntos implicados que se han visto involucrados en el delito.

Sesgos cognitivos

Sesgo retrospectivo	Ocurre cuando un individuo, al analizar hechos pasados, no puede deslindarse de sus efectos. Por ejemplo, si a una persona se le investiga por violencia intrafamiliar y en el pasado ya ha sido condenada por este delito, el sistema podría dictar una decisión a priori de condena por el antecedente.
Sesgo de representatividad	Corresponde a errores estadísticos y matemáticos en los cálculos de probabilidad, los cuales derivan en ignorar probabilidades a los resultados. Si en un proceso se toman decisiones con base en datos estadísticos, se pueden presentar fallos en la comprensión de la aleatoriedad.
Sesgo de anclaje	Se presenta cuando se realiza un juicio tomando como punto de referencia un valor que va ajustando progresivamente al incorporar nueva información. En un proceso penal puede haber un prejuzgamiento si la Fiscalía desde un inicio infiere la culpabilidad del acusado.
Sesgo de confirmación	Se manifiesta cuando hay una interpretación de una información que respalde ideas previas o hipótesis iniciales. Por ejemplo, en materia penal, este sesgo puede dar lugar a que no se tengan en cuenta las diferencias de instancia entre un juez de garantías y un juez de conocimiento.
Sesgo de grupo	Ocurre cuando se evalúan informaciones basadas en la

pertenencia de una persona a un grupo específico. Si un juicio está marcado por un prejuicio por la pertenencia a una organización determinada, la decisión podría resultar injusta.

Fuente: elaboración propia a partir de García (2024).

Los riesgos, errores o sesgos de una IA predictiva en un proceso penal revelan las limitaciones de esta tecnología, pero también ponen en evidencia los aspectos que se presentan como oportunidades que se deben mejorar, es decir, los errores se presentan como una debilidad que requiere ser abordada, pues no solo la IA como tecnología está llamada a ser ajustada, sino también los sesgos cognitivos e inconscientes de los seres humanos al dar total credibilidad a decisiones automatizadas o dudar y cuestionar en todas las situaciones de estas.

Conclusiones

Son diversas las expectativas de jueces y juristas respecto al uso de la IA en el derecho penal colombiano, ya que la justicia predictiva supone realizar estimaciones anticipadas frente al criterio humano judicial, sin que ello lleve a que sean decisiones robóticas las que se tomen de manera autónoma sin la intervención del hombre: se trata de un elemento de capital importancia, ya que poner al servicio del derecho penal la justicia predictiva no debe suponer en ningún caso prescindir del componente humano. El real alcance de este modelo es convertirse en una herramienta que, a partir de algoritmos de IA, permita predecir resultados jurídicos a través del análisis de decisiones judiciales anteriores.

La justicia predictiva está aún lejos de reemplazar a jueces y juristas, lo que busca es convertirse en una herramienta que permita establecer patrones de diagnóstico a decisiones judiciales futuras, sirviendo de apoyo y modulación estratégica para las partes, esto es, para la Fiscalía y para los abogados defensores, facilitándoles su labor judicial, lo cual puede servir para extinguir la alta litigiosidad que opera en Colombia, permitiendo que la justicia predictiva se convierta en complemento y sirva de herramienta a la justicia cognitiva de los seres humanos.

La aplicación de la IA al ámbito del derecho penal colombiano implica desarrollar un proceso de cambio en los distintos niveles que conforman esta rama de la justicia y para ello es fundamental contar con un marco normativo sólido de gobernanza de esta tecnología, de tal forma que sea asequible para todos los usuarios de la justicia, garantizando derechos humanos y

acceso libre a la información, lo cual exige la articulación del derecho con expertos en tecnología que procuren modelos fiables al servicio de la justicia.

Bibliografía

Abiteboul, S., & G'ssell, F. (2019). *¿Podrían los algoritmos reemplazar a los jueces?* Dalloz.

Agudelo, S. (2021). Inteligencia artificial en el proceso penal: análisis a la luz del Fiscal Watson. *Pensamiento Jurídico*, (54), 147-164.

Alpaydin, E. (2014). *Introduction to Machine Learning*. The MIT Press.

Armenta D., T. (2021). *Derivas de la justicia*. Marcial Pons.

Battelli, E. (2021). La decisión robótica: algoritmos, interpretación y justicia predictiva. *Revista de Derecho Privado*, (40), 45-86.

Belloso M., N. (2019). Algoritmos predictivos al servicio de la justicia: ¿una nueva forma de minimizar el riesgo y la incertidumbre? *Revista da Faculdade Mineira de Direito*, 22(43), 1-31.

Bobadilla, J. (2020). *Machine Learning y deep learning*. Usando Python, Scikit y Keras. Ediciones de la U.

Cajiga, C., & Vergara, S. (2021). Justicia predictiva y derechos humanos. *Revista de Derechos Humanos*, 3, 81-92.

Congreso de la República. (1999, 18 de agosto). *Por medio de la cual se define y reglamenta el acceso y uso de los mensajes de datos, del comercio electrónico y de las firmas digitales, y se establecen las entidades de certificación y se dictan otras disposiciones [Ley 527 de 1999]*. DO: 43.673.

Congreso de la República. (2011, 18 de enero). *Por la cual se expide el Código de Procedimiento Administrativo y de lo Contencioso Administrativo [Ley 1437 de 2011]*. DO: 47.956.

Congreso de la República. (2012, 17 de octubre). *Por la cual se dictan disposiciones generales para la protección de datos personales [Ley 1581 de 2012]*. DO: 48.587.

Congreso de la República. (2014, 6 de marzo). *Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones [Ley 1712 de 2014]*. DO: 49.084.

Congreso de la República. (2021, 29 de julio). *Ley de Internet como servicio público esencial y universal” o por medio de la cual se modifica la Ley 1341 de 2009 y se dictan otras disposiciones [Ley 2108 de 2021]*. DO: 51.750.

Consejo Nacional de Política Económica y Social. (2018, 17 de abril). *Política Nacional de Explotación de Datos (BIG DATA) [Documento Conpes 3920 de 2018]*. Departamento Nacional de Planeación.

Consejo Nacional de Política Económica y Social. (2019, 8 de noviembre). *Política Nacional para la Transformación Digital e Inteligencia Artificial [Documento Conpes 3975 de 2019]*. Departamento Nacional de Planeación.

Consejo Nacional de Política Económica y Social. (2021). *Concepto favorable a la nación para contratar una operación de crédito público externo con la banca multilateral hasta por USD 100 millones, o su equivalente en otras monedas, destinados a financiar el Programa para la Transformación Digital de la Justicia en Colombia- Fase I [Documento Conpes 4024 de 2021]*. Departamento Nacional de Planeación.

Consejo Nacional de Política Económica y Social. (2025, 14 de febrero). *Política Nacional de Inteligencia Artificial [Documento Conpes 4144 de 2025]*. Departamento Nacional de Planeación.

Corvalán, J. (2017). La primera inteligencia artificial predictiva al servicio de la justicia: Prometea. *Pensamiento Penal*, 6(477), 1-12.

Criado E., J. (2023). Sistemas de justicia predictiva en la toma de decisiones en la instrucción penal: nuevos retos en sede procesal. Una mirada desde Europa. *Revista General de Derecho Europeo*, (60), 1-29.

Galeano M., M. (2011). *Diseño de proyectos en la investigación cualitativa*. Fondo Editorial Universidad EAFIT.

García T., M. L. (2024). La inteligencia artificial predictiva al servicio de la prevención e investigación del delito y del proceso penal. *Ciencia Policial*, 183, 91-132.

Gimeno, J. (2023). Instrumentos actuales de policía y justicia predictiva en el proceso penal español: análisis crítico y reflexiones de *lege ferenda* ante aplicaciones futuras. *Estudios Penales y Criminológicos*, 44(ext), 1-20.

González A., A., & Sanabria C., A. (2019). La justicia predictiva: entre mito y realidad. *Revista de Derecho y Nuevas Tecnologías*, (1), 21-50.

Herrera M., H. (2024). ¿La “justicia predictiva” (la artificial) contra la “justicia cognitiva” (la humana)? <https://www.ambitojuridico.com/noticias/columnista-impreso/constitucional-y-derechos-humanos/la-justicia-predictiva-la-artificial>

O'Neil, C. (2017). *La era de la fe ciega en los datos masivos ha de terminar*.

<https://mujeresconciencia.com/2017/09/03/la-era-de-la-fe-ciega-en-los-datos-masivos-ha-de-terminar/>

Peralta G., A. (2021). *Diálogos para el futuro judicial: jurimetría y justicia predictiva*. Diario La Ley.

Pinilla M., N. (2024). *Aplicación de herramientas de justicia predictiva en Asuntos de Propiedad Intelectual en Colombia*. Universidad Cooperativa de Colombia.

Porcelli, A. M. (2020). La inteligencia artificial y la robótica: sus dilemas sociales, éticos y jurídicos. *Derecho Global. Estudios sobre Derecho y Justicia*, 6(16), 49-105.

Real Academia Española. (2025). *Diccionario de la Lengua Española*. <https://dle.rae.es/>

Rouvière, F. (2021). La justice prédictive peut-on réduire le droit en algorithmes? En N. Belloubet (Ed.), *La justice: regards critiques* (pp. 97-107). SEUIL.

Schauer, F. (2018). *Thinking as a Jurist. Nouvelle introduction au raisonnement juridique*. Dalloz.

Suárez X., P. (2024). Perspectivas entrelazadas: justicia predictiva y técnica procesal. *Revista General de Derecho Procesal*, (64), 1-5.

Vaillant, R., Monrocq, C., & Le Cun, Y. (2004). Original Approach for the Localization of Objects in Images. *Proceedings: Vision, Image and Signal Processing*, 141(4), 245-250.