



FACULTAD DE DERECHO

**DIGITALIZACIÓN DE LA JUSTICIA  
PENAL: EVALUACIÓN DE LA  
IMPLEMENTACIÓN DE  
ALGORITMOS PREDICTIVOS EN LA  
TOMA DE DECISIONES SOBRE  
PRISIÓN PROVISIONAL**

Autor: Laura Villadangos Lainz

5º E3 C

Área de Derecho Penal

Tutor: Gabriel Castro Salillas

Madrid

Mayo 2024

## **ABSTRACT:**

Los algoritmos han demostrado superar las capacidades humanas en la tarea predictiva y se han convertido ya en actores clave en diversas jurisdicciones. Este Trabajo de Fin de Grado tiene por objeto evaluar la viabilidad de implementar un algoritmo predictivo para asistir en la toma de decisiones relativas a la adopción de la medida de prisión provisional en el sistema procesal penal español. Del análisis del proceso mediante el cual se toman estas decisiones se infiere que la utilización de un algoritmo podría mejorar la objetividad y consistencia de las decisiones judiciales. El estudio examina cómo esta tecnología puede integrarse eficazmente en el sistema judicial, detallando el proceso necesario para desarrollar dicha herramienta. Además, se identifican potenciales desafíos de la aplicación de estos algoritmos, y se proponen directrices para una implementación ética y efectiva, que complemente, pero no sustituya, la decisión judicial.

A pesar de los problemas que plantea, el algoritmo ofrece una oportunidad única para mejorar la toma de decisiones en el ámbito de la justicia penal, promoviendo una mayor objetividad y consistencia en el proceso judicial.

**Palabras clave:** *Prisión preventiva, algoritmo predictivo, evaluación del riesgo, variables, transparencia, sesgo, COMPAS.*

---

Algorithms have demonstrated the ability to surpass human capabilities in predictive tasks and have already become key players in various jurisdictions. This Final Degree Project aims to evaluate the feasibility of implementing a predictive algorithm to assist in the decision-making process regarding the adoption of pretrial detention measures in the Spanish criminal procedural system. From the analysis of the decision-making process for pretrial detention, it is inferred that the use of an algorithm could improve the objectivity and consistency of these decisions. The study examines how this technology can be effectively integrated into the judicial system, detailing the necessary process for developing such a tool. Additionally, the potential challenges of applying these algorithms are identified, and guidelines are proposed for an ethical and effective implementation that complements but does not replace judicial decision-making.

Despite the challenges it presents, the algorithm offers a unique opportunity to improve decision-making in the field of criminal justice, promoting greater objectivity and consistency in the judicial process.

**Keywords:** *Pretrial detention, predictive algorithm, risk assessment, variables, transparency, bias, COMPAS.*

## **ÍNDICE DE CONTENIDOS**

1.	INTRODUCCIÓN.....	5
1.1	Contextualización del tema .....	5
1.2	Objetivos de la investigación.....	6
1.3	Metodología utilizada.....	7
2.	REGIMEN DE LA PRISIÓN PROVISIONAL .....	7
2.1	Concepto y finalidad.....	7
2.2	Principios.....	8
2.2.1	Principio de excepcionalidad.....	8
2.2.2	Principio de subsidiariedad.....	8
2.2.3	Principio de provisionalidad.....	8
2.2.4	Principio de proporcionalidad .....	9
2.2.5	Principio de exclusividad jurisdiccional.....	9
2.3	Requisitos .....	9
2.4	Criterios para la toma de la decisión cautelar.....	10
2.4.1	Riesgo de Fuga .....	11
2.4.2	Riesgo de ocultación, alteración o destrucción de pruebas .....	12
2.4.3	Riesgo de atentado contra bienes jurídicos de la víctima.....	12
2.4.4	Riesgo de reincidencia.....	12
2.5	Plazos máximos .....	13
2.6	Recurso .....	14
3.	EVALUACIÓN DEL RIESGO POR PARTE DE LOS JUECES .....	14
3.1	Fase procesal .....	14
3.2	Proceso mental.....	15
3.3	Aplicación de la prisión preventiva en la actualidad .....	16
4.	EL ALGORITMO.....	18
4.1	Elección del tipo de algoritmo.....	18

4.2 Construcción del algoritmo .....	21
4.2.1 Recopilación y procesamiento de datos.....	21
4.2.2 Selección de variables .....	23
4.2.3 Desarrollo y validación del modelo predictivo.....	27
5. IMPLEMENTACIÓN EN LAS DECISIONES JUDICIALES.....	28
5.1 Medición del riesgo .....	29
5.2 Interpretación del riesgo .....	31
5.3 Alcance del algoritmo .....	31
5.4 Plazos máximos .....	34
5.5 Recurso .....	34
6. PROBLEMAS Y CONSIDERACIONES ÉTICAS .....	36
6.1 Desconfianza en el algoritmo .....	36
6.1.1 Percepción social .....	36
6.1.2 Percepción de los jueces .....	37
6.2 Transparencia, interpretabilidad y explicabilidad.....	38
6.3 Sesgos y discriminación .....	39
7. CONCLUSIONES Y RECOMENDACIONES .....	40
7.1 Recapitulación de hallazgos .....	40
7.2 Implicaciones prácticas.....	41
7.3 Limitaciones del estudio y futuras líneas de investigación .....	42
8. REFERENCIAS .....	44
8.1 Legislación .....	44
8.2 Jurisprudencia.....	44
8.3 Bibliografía.....	44

# 1. INTRODUCCIÓN

## 1.1 Contextualización del tema

La evolución de las tecnologías ha marcado profundamente el curso de la historia humana. En las últimas décadas, este ritmo de innovación se ha acelerado exponencialmente (Ordóñez, 2007), obligando a la sociedad a adaptarse rápidamente a los nuevos avances tecnológicos. Uno de los desarrollos más significativos ha sido en el campo de la inteligencia artificial. Esto se evidencia con la reciente aprobación del nuevo Reglamento Europeo sobre Inteligencia Artificial, que aún no ha entrado en vigor pero que ya busca regular la utilización de esta tecnología, anticipándose a los desafíos y oportunidades que presenta.

Esta transformación tecnológica también plantea nuevas oportunidades y desafíos dentro del ámbito del derecho. Según Barona Vilar (2023), estamos presenciando una "hibridación" creciente de la justicia, donde máquinas y humanos no solo coexisten, sino que colaboran. Los algoritmos no solo asisten a operadores legales, sino que, en algunos casos, están caminando hacia la sustitución progresiva del ser humano en el sistema judicial por máquinas o robots.

Un área de especial interés ha sido el uso de algoritmos predictivos para asistir en decisiones judiciales clave, como la prisión provisional. La decisión de decretar esta medida cautelar se basa en la evaluación subjetiva del juez, quien debe considerar una variedad de factores, incluido el riesgo de fuga, la posibilidad de obstrucción a la justicia, y el riesgo de reincidencia. Sin embargo, el potencial de sesgo y error humano en tales evaluaciones es significativo (Álvarez Buján, 2023). En este contexto, los algoritmos de toma de decisiones emergen como herramientas que pueden ofrecer análisis basados en datos, potencialmente aumentando la objetividad, precisión y consistencia de las decisiones judiciales.

Los algoritmos predictores de riesgo, conocidos como Risk Assessment Tools, han ganado amplia aceptación, especialmente en Estados Unidos. Un ejemplo prominente es el modelo COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), que estima la probabilidad de que un acusado reincida dentro de los próximos dos años. Este modelo se utiliza frecuentemente para informar decisiones judiciales sobre fianzas y sentencias. En Europa, el algoritmo HART (*Harm Assessment Risk Tool*) se destaca en el Reino Unido por su capacidad para evaluar el riesgo de reincidencia,

apoyando así decisiones similares a las de COMPAS en el sistema judicial británico (Notaro, 2023). En España, particularmente en el ámbito de la Administración penitenciaria catalana, la herramienta RisCanvi lleva más de 15 años en uso. Este sistema está diseñado para valorar el riesgo de reincidencia de los convictos, siendo una herramienta clave en la toma de decisiones relativas a la concesión de permisos de salida, el tercer grado y la libertad condicional (Alemán Aróstegui, 2023).

Sin embargo, la implementación de estas tecnologías no está exenta de desafíos. Las preocupaciones sobre la transparencia de los algoritmos, su capacidad para perpetuar sesgos existentes, y las implicaciones éticas de su uso en el contexto penal plantean preguntas significativas que necesitan ser exploradas.

## **1.2 Objetivos de la investigación**

El objetivo general de este trabajo es evaluar la viabilidad de implementar un algoritmo predictor de riesgo para informar sobre la toma de decisiones en la adopción de la prisión provisional en el sistema procesal penal español, y proponer directrices específicas para su implantación efectiva y ética.

Los objetivos específicos son los siguientes:

- Analizar el sistema actual de toma de decisiones sobre prisión preventiva en España: Investigar y documentar cómo se toman actualmente estas decisiones, identificando los actores clave, los criterios y los requisitos utilizados, y cómo estos elementos se integran en el proceso judicial.
- Describir el proceso de desarrollo de un algoritmo predictor de riesgo: Detallar las etapas de creación de un algoritmo, desde la recopilación y análisis de datos hasta el diseño y pruebas del modelo, con un enfoque en las especificaciones técnicas y los desafíos metodológicos.
- Simular y proyectar la integración del algoritmo en el sistema judicial español: Explorar cómo se vería el sistema tras la incorporación del algoritmo y mencionar los cambios necesarios para su funcionamiento.
- Identificar y analizar los desafíos potenciales en la aplicación del algoritmo: Considerar los problemas éticos, técnicos y prácticos que podrían surgir de su uso en el contexto judicial.

### **1.3 Metodología utilizada**

El estudio se llevará a cabo mediante una metodología deductiva, centrada en un análisis cualitativo profundo y una revisión de la literatura existente sobre el tema. Nótese que la mayor parte de esta proviene de Estados Unidos debido a la reciente introducción de la cuestión en el contexto español. Además, se realizará un examen de las fuentes legales españolas relevantes para entender cómo se interpreta y se aplica la normativa en el ámbito nacional. Este enfoque permitirá explorar adecuadamente el tema dentro del marco jurídico español, asegurando la relevancia y aplicabilidad de los resultados obtenidos.

## **2. REGIMEN DE LA PRISIÓN PROVISIONAL**

### **2.1 Concepto y finalidad**

Dentro de las medidas cautelares personales previstas en la Ley de Enjuiciamiento Criminal, en el Capítulo III encontramos la prisión provisional o prisión preventiva, concretamente desarrollada en los artículos 502 a 519 de la LECrim. Consiste en la total privación de la libertad del derecho fundamental a la libertad deambulatoria del imputado, ingresándole en un centro penitenciario, antes de que se lleve a cabo el juicio y se emita sentencia definitiva.

Esta medida se establece con una finalidad cautelar y con una finalidad preventiva. Por un lado, tiene por objetivo garantizar la presencia del imputado en el proceso, facilitar la celebración de este con todas las pruebas necesarias y asegurar la ejecución de la sentencia (finalidad cautelar). Por otro lado, la prisión provisional puede aplicarse con el fin de prevenir que el imputado cause daño a bienes jurídicos de la víctima o cometa nuevos delitos (finalidad preventiva). Es crucial destacar que su aplicación es excepcional y debe ajustarse a los principios de proporcionalidad y necesidad, respetando siempre el derecho a la presunción de inocencia.

Es una medida que preocupa a la Doctrina, pues es la intromisión más severa que puede ejercer el estado al limitar la libertad del individuo sin que medie aun sentencia penal firme que condene al mismo. Sin embargo, se contempla como un mal necesario para asegurar los fines previamente mencionados (Hernández, 2011).

## 2.2 Principios

### 2.2.1 Principio de excepcionalidad

El principio de excepcionalidad, plasmado en el artículo 502.2 LECrim, exige que la aplicación de la prisión provisional sea siempre la excepción y nunca la norma, solo pudiendo darse cuando sea estrictamente necesario para asegurar la integridad del proceso judicial. Por lo tanto, esta medida no debe tener fines punitivos ni debe emplearse para promover la investigación del delito (Turturro-Pérez de los Cobos, 2021). Como ilustra la sentencia del Tribunal Constitucional, núm. 210/2013, de 16 de diciembre de 2013, se aplica el principio *in dubio pro libertate*, el cual establece que la prisión preventiva debe aplicarse de manera restrictiva, priorizando el derecho fundamental a la libertad.

Para asegurar la idoneidad de la decisión sobre la prisión provisional, es imprescindible realizar una ponderación meticulosa de las circunstancias específicas del caso. Sin embargo, la magistrada Gallego Sánchez (2006) destaca que un problema frecuente en la práctica judicial es la falta de una evaluación exhaustiva de estas circunstancias y la adecuada motivación en las resoluciones que imponen esta medida cautelar.

### 2.2.2 Principio de subsidiariedad

El principio de subsidiariedad, también incluido en el artículo 502.2 LECrim, implica que la prisión provisional debe considerarse como una medida subsidiaria, aplicándose únicamente cuando no haya disponibles opciones menos perjudiciales para la libertad que permitan lograr los objetivos procesales. La LECrim recoge distintos tipos de medidas cautelares menos gravosas para los derechos fundamentales del investigado. Las más utilizadas en la práctica son la retirada del pasaporte, la fianza, el embargo, la presentación periódica ante la autoridad judicial y el arresto domiciliario (Morillas Cueva, 2016).

### 2.2.3 Principio de provisionalidad

Reconocido en el artículo 504. LECrim, el principio de provisionalidad establece que la prisión preventiva es una medida temporal y sujeta a revisión. Este principio se fundamenta en la necesidad de ajustar la prisión provisional a cambios en las circunstancias que justificaron su implementación inicial. Es decir, está sometido a la cláusula *rebus sic stantibus*.

El artículo mencionado también estipula límites temporales para la prisión provisional: para delitos con una pena privativa de libertad de tres años o menos, la duración no puede

exceder un año; para delitos con penas superiores a tres años, el límite es de dos años. En casos excepcionales, se permite una prórroga de seis meses en el primer caso y de hasta dos años en el segundo.

#### *2.2.4 Principio de proporcionalidad*

Estrechamente ligado al principio de excepcionalidad, el principio de proporcionalidad establece que dado que la medida en cuestión limita derechos fundamentales como son la libertad y la presunción de inocencia, debe ser proporcional a los fines que persigue. El principio de proporcionalidad exige un equilibrio entre el derecho a la libertad y cualquier restricción a esta libertad. Este concepto fue refinado en el fundamento jurídico cuarto de la Sentencia del Tribunal Constitucional 341/1993, que especifica que la ley no debe permitir supuestos de privación de libertad que no se alineen con la finalidad de proteger derechos, bienes o valores constitucionalmente reconocidos.

#### *2.2.5 Principio de exclusividad jurisdiccional*

Este principio implica que la decisión de decretar la prisión provisional corresponde exclusivamente a los órganos judiciales. Concretamente, el artículo 502.1 LECrim establece que será el juez o magistrado de instrucción, el juez que inicia las primeras diligencias o el juez o tribunal de lo penal que conozca de la causa.

### **2.3 Requisitos**

El artículo 503 de la Ley de Enjuiciamiento Criminal establece tres requisitos fundamentales para decretar la prisión provisional. En primer lugar, es necesario que en la causa consten hechos que presenten características de delito sancionado con pena cuyo máximo sea igual o superior a dos años de prisión, o con pena privativa de libertad de duración inferior, siempre y cuando el investigado tenga antecedentes penales no cancelados por delito doloso (503.1 LECrim).

El segundo requisito implica que deben aparecer en la causa motivos suficientes para creer responsable criminalmente del delito a la persona contra la que se pretende dictar el auto de prisión. Es decir, se requiere una base sólida que respalde la creencia de la culpabilidad del investigado (503.2 LECrim). Esta base sólida se obtendrá de la información y las pruebas aportadas por la autoridad investigadora. Para salvaguardar el principio de presunción de inocencia es fundamental que las mismas sean suficientemente firmes para asociar al investigado a los hechos delictivos, y no pudiendo ser decretada sin

que exista un mínimo elemento que conecte directamente al imputado con el delito (Noriega Estrada, 2023).

En tercer lugar, la prisión provisional solo puede decretarse cuando persiga uno de los siguientes fines. El primero de ellos es asegurar la presencia del investigado en el proceso, especialmente cuando exista un riesgo razonable de fuga (503.1 3º a) LECrim). El segundo posible fin es evitar la ocultación, alteración o destrucción de las fuentes de prueba relevantes, siempre que exista un peligro fundado y concreto (503.1 3º b) LECrim). En tercer lugar, se contempla la posibilidad de decretar la prisión provisional para evitar que el investigado pueda actuar contra bienes jurídicos de la víctima, especialmente cuando esta sea alguna de las personas referidas en el apartado segundo del artículo 173 del Código Penal (503.1 3º c) LECrim). Además, se establece un supuesto especial en el que la medida puede adoptarse para evitar el riesgo de que el investigado cometa otros hechos delictivos, siempre que el delito imputado sea doloso (503.2 LECrim).

#### **2.4 Criterios para la toma de la decisión cautelar**

El apartado tercero del artículo 502.1 LECrim establece que, al tomar la de decisión de decretar la prisión provisional, el órgano judicial debe tener en cuenta dos aspectos fundamentales. En primer lugar, debe evaluar cómo afectará la medida al individuo, considerando sus circunstancias personales y las particularidades del hecho que está siendo investigado. Esto es, se debe tener en cuenta si la imposición de la medida cautelar va a causar un grave perjuicio al individuo dadas sus circunstancias. En segundo lugar, debe ponderar la gravedad de la pena que podría imponerse en el caso, y cuanto mayor sea la pena a la que se enfrente, mayor será la probabilidad de ser encarcelado preventivamente.

Además, como se ha mencionado en el apartado anterior, únicamente se podrá decretar la prisión provisional cuando persiga asegurar la presencia del investigado en el proceso, evitar la ocultación, alteración o destrucción de pruebas, evitar que actúe contra bienes jurídicos de la víctima y evitar que cometa otros delitos. Para valorar si existen tales riesgos, la Ley establece criterios específicos en función de la finalidad que la prisión preventiva pretenda cumplir en cada caso concreto. Este análisis integral, considerando la interrelación de los diferentes factores en su conjunto, busca garantizar que la decisión de prisión provisional sea justa y proporcional a la situación específica que se está

evaluando. Entender qué factores se ponderan en el sistema actual es relevante para determinar qué factores debería tener en cuenta el algoritmo.

Aunque la Ley de Enjuiciamiento Criminal identifica cuatro factores de riesgo, es crucial subrayar que, en la práctica, para garantizar que no se vulnere el principio que estipula que nadie puede ser condenado sin una sentencia firme, y así preservar el carácter excepcional de esta medida cautelar, Javier Muñoz Cuesta (2019), fiscal del Tribunal Supremo, señala que únicamente dos causas justifican esta excepcionalidad: el riesgo de fuga y el riesgo de reincidencia. La presencia de cualquiera de estas causas es suficiente para decretar prisión provisional sin que sea necesaria la concurrencia de ambas. Sin embargo, cuando no concurra ninguna de las mencionadas causas, con independencia de la gravedad de un delito o el nivel de indignación social que pueda generar, el fiscal recalca que la situación de libertad provisional del investigado debe prevalecer. Únicamente podría exceptuarse en situaciones donde la severidad del crimen, por sí sola, demuestre una clara imposibilidad de mantener al acusado en libertad antes del juicio, como es el caso en crímenes de terrorismo o intentos de desestabilización estatal. Dicho esto, a continuación se detallan los cuatro factores de riesgo según la LECrim.

#### *2.4.1 Riesgo de Fuga*

La Ley exige que pueda inferirse racionalmente un riesgo de fuga. La determinación del riesgo de fuga por parte del juez se realiza mediante una evaluación de factores que permitan inferir la probabilidad de que el investigado o encausado eluda la acción de la justicia evadiéndose del proceso. Algunos de los elementos que el juez debe considerar para evaluar este riesgo incluyen:

- La naturaleza del hecho: Se analiza la gravedad y las circunstancias del delito imputado. Delitos más graves pueden sugerir un mayor riesgo de fuga.
- La gravedad de la pena: Se evalúa la posible pena que pudiera imponerse al investigado. A mayor gravedad de la pena, se podría inferir un mayor incentivo para eludir la acción de la justicia.
- La situación personal y familiar del investigado: Se consideran aspectos como arraigo familiar, laboral y económico. La estabilidad en estos aspectos puede disminuir el riesgo de fuga.

- **Antecedentes del investigado:** Si existen al menos dos requisitorias previas para el llamamiento y búsqueda del investigado en los dos años anteriores, dictadas por cualquier órgano judicial, se puede inferir que hay riesgo de fuga.
- **Inminencia del Juicio Oral:** Si el juicio oral está próximo significa que el proceso judicial está en una fase avanzada y que pronto se tomará una decisión definitiva sobre la culpabilidad o inocencia del investigado. La proximidad temporal al juicio y, por ende, a la adopción de una decisión definitiva sobre la culpabilidad o inocencia del investigado puede aumentar su ansiedad o el temor respecto a las posibles consecuencias legales e incentivar su ánimo evasivo.

#### *2.4.2 Riesgo de ocultación, alteración o destrucción de pruebas*

El artículo 503.1 3º c) LECrim exige que el peligro de ocultación, alteración o destrucción de pruebas sea fundado y concreto. Se prohíbe expresamente inferir dicho riesgo del ejercicio del derecho de defensa o de la falta de colaboración del investigado o encausado con la investigación. Para determinar el riesgo, se tienen en cuenta los siguientes factores:

- **Capacidad de acceso a las fuentes de prueba:** El juez analizará si el investigado o encausado tiene la capacidad directa o a través de terceros para acceder a las pruebas que podrían ser relevantes para el caso. Esto implica evaluar si la persona tiene la posibilidad real y efectiva de manipular, ocultar o destruir evidencia.
- **Capacidad de influencia sobre otros:** Se considerará si el investigado o encausado tiene la capacidad de influir en otras personas involucradas en el caso: otros investigados, peritos, testigos o cualquier otra persona con información relevante.

#### *2.4.3 Riesgo de atentado contra bienes jurídicos de la víctima*

La Ley de Enjuiciamiento Criminal no especifica los factores a tener en cuenta al valorar el riesgo del investigado o encausado de atentar contra bienes jurídicos de la víctima.

#### *2.4.4 Riesgo de reincidencia*

En primer lugar, la ley exige que, para decretar la prisión provisional por riesgo de cometer otros delitos, el hecho delictivo imputado sea doloso. Esto es, los imputados por negligencia no podrán ser encarcelados provisionalmente por riesgo de reincidencia. Dentro de los hechos delictivos dolosos, el juez valora los siguientes factores:

- Circunstancias del hecho: El juez analiza las circunstancias específicas del delito imputado. Factores como la naturaleza del delito, la forma en que se llevó a cabo y la gravedad de este son tomados en cuenta.
- Gravedad de los delitos que pudiera cometer en caso de permitir su libertad: Esta evaluación implica considerar el impacto y las consecuencias de los posibles delitos, así como la amenaza que representan para la sociedad.
- Colaboración organizada para la comisión de actos delictivos: El legislador dota de importancia al hecho de participar en una organización delictiva, pues el artículo 503.2 LECrim permite la posibilidad de aplicar la medida de prisión provisional por riesgo de reincidencia cuando hay evidencia de que se trata de un caso de colaboración organizada para la comisión de actos delictivos, aunque la pena a la que expone sea inferior a dos años.
- Habitualidad en la comisión de delitos: al igual que con la colaboración organizada, la ley prescinde del mínimo de dos años de pena cuando se pueda racionalmente inferir que el investigado o encausado comete delitos con habitualidad. Para ello, se tendrán en cuenta los antecedentes delictivos del individuo.

Además de los mencionados en la ley, Muñoz-Cuesta (2019) añade las tendencias psicológicas diagnosticadas por profesionales de la psicología o psiquiatría, el comportamiento social, las conductas en ciertos entornos y la actividad delictiva manifestada en las redes sociales, un fenómeno cada vez más recurrente.

## **2.5 Plazos máximos**

El Tribunal Constitucional, en su sentencia 41/1982 estableció que *“toda persona detenida preventivamente tiene derecho a ser juzgada en un plazo «razonable o a ser puesta en libertad durante el procedimiento»*”. El mismo tribunal exige que se establezca un plazo máximo para la prisión provisional, e imponiendo el deber de poner en libertad a un acusado una vez transcurridos los plazos legales establecidos para la detención (STC 40/1987).

Por consiguiente, el artículo 504 de la LECrim subraya que la prisión provisional debe durar solamente el tiempo imprescindible para alcanzar los fines para los que fue decretada y solo mientras persistan las circunstancias que justificaron su adopción. Además, impone los siguientes plazos máximos de duración a la prisión provisional:

- Para delitos con penas hasta de tres años, la duración máxima que puede estar el imputado en prisión provisional de un año, ampliable a seis meses adicionales.
- Para delitos con penas superiores a tres años, la duración máxima es de dos años, ampliable a dos años adicionales.

## **2.6 Recurso**

El artículo 507 de la LECrim, que remite al artículo 766 de la misma ley, simplifica el procedimiento de apelación de los autos de prisión provisional, alineándolo con el proceso de apelación del Procedimiento Abreviado. Según este artículo, contra los autos del Juez de Instrucción y del Juez de lo Penal que decreten la prisión provisional y que no estén exceptuados de recurso, se pueden presentar tanto el recurso de reforma como el de apelación.

# **3. EVALUACIÓN DEL RIESGO POR PARTE DE LOS JUECES**

## **3.1 Fase procesal**

Para evaluar adecuadamente la viabilidad de aplicar la Inteligencia Artificial en los procesos de toma de decisiones judiciales, resulta crucial comprender cuál es el momento procesal en que los jueces deciden sobre la imposición de medidas cautelares. Es esencial comprender la naturaleza de la información que el juez posee y las pruebas que han sido practicadas en el momento en que deciden sobre si se decreta la prisión preventiva, para así evaluar la base sobre la cual se fundamenta su decisión.

El artículo 505 LECrim establece que, tras la puesta a disposición del juez o tribunal competente, se debe convocar una audiencia a menos que se decrete la libertad provisional sin fianza. Esta audiencia, que debe celebrarse en un plazo máximo de 72 horas, contará con la presencia del investigado asistido por un letrado, el Ministerio Fiscal y las demás partes personadas. Durante la audiencia, las partes pueden presentar alegaciones y proponer pruebas, las cuales deben realizarse en el acto o dentro del plazo entre la puesta a disposición judicial y la audiencia. El abogado del investigado debe tener acceso a los elementos esenciales del caso para impugnar eficazmente cualquier propuesta de privación de libertad. Finalmente, el juez o tribunal decidirá si procede la prisión provisional o la imposición de fianza, y si ninguna parte solicita estas medidas, se ordenará la liberación inmediata del detenido.

### **3.2 Proceso mental**

Las decisiones judiciales son a menudo vistas como demasiado inconsistentes, lo que ha alimentado un debate sobre la necesidad de equilibrar la justicia individualizada con la demanda de mayor previsibilidad y equidad en el proceso judicial (McKay, 2019). Para constatar la utilidad de la Inteligencia Artificial en la toma de decisiones sobre prisión preventiva, es importante analizar el proceso mental que llevan a cabo los jueces a la hora de tomar tales decisiones y ver en qué medida se ven afectados por factores externos.

Estas decisiones se toman en un contexto particularmente desafiante, puesto que, al momento de evaluar el riesgo del individuo, a pesar de haber tenido la oportunidad de practicar ciertas diligencias de prueba, las decisiones deben tomarse sin el beneficio de toda la información que típicamente estaría disponible en fases posteriores. En este escenario, los jueces deben evaluar si es necesario aplicar la prisión preventiva, basándose en indicios que pueden no ser totalmente sólidos o completos. A pesar de que se espera que los jueces, altamente capacitados y con amplia experiencia, basen sus decisiones estrictamente en los hechos y en las leyes pertinentes a cada caso, investigaciones en el campo de la psicología y la economía conductual señalan que, al igual que cualquier ser humano, están expuestos a sesgos cognitivos (DeMichele et al., 2021).

Si bien es cierto que el proceso de la toma de decisión es casi automático, diferentes heurísticos y afecciones juegan un papel importante en la misma. Un heurístico se define como un conjunto de reglas o estrategias diseñadas para solucionar problemas en situaciones de incertidumbre (Rondón Galvis, 2022). Como señala Neira Pena (2021), en la toma de estas decisiones, los jueces son influenciados por casos anteriores con los que encuentran similitudes, o por procedimientos a los que se han visto expuestos recientemente. Asimismo, las primeras impresiones y las emociones de los jueces respecto al caso también afectan al subconsciente. Estos heurísticos, si bien son útiles para afrontar la incertidumbre y la sobrecarga de trabajo a la que se exponen los jueces en España, pueden llevar a decisiones potencialmente injustas (Calaza-López & Llorente-Sánchez-Arjona, 2022).

Además de los sesgos subconscientes, los jueces pueden tener una tendencia consciente a favorecer los falsos positivos, optando ante la duda por la prisión preventiva. Esto significa que, por temor a perjudicar su reputación o por evitar ser responsables de un daño potencial, es probable que dicten prisión provisional en situaciones donde el riesgo

asociado al acusado no sea particularmente alto. Y lo que es más importante, todos los factores que intervienen en el proceso de toma de decisiones de los jueces son incuantificables y permanecerán indescifrables (Elyounes, 2020). La propuesta de incorporación de la inteligencia artificial en el proceso de toma de decisiones busca mejorar el proceso al ofrecer análisis basados en datos más objetivos.

### 3.3 Aplicación de la prisión preventiva en la actualidad

A pesar de la insistencia del Tribunal Constitucional en que la prisión provisional es una medida aplicable únicamente en casos excepcionales (STC 128/1995), la última reforma realizada en 2003 empeoró la situación previa al reducir el límite penológico de tres a dos años para la adopción de esta medida. Según el estudio llevado a cabo por la Asociación Pro Derechos Humanos de España (2015), en el que se entrevistó a cinco magistrados, no existe una opinión unificada por parte de los jueces respecto a la aplicación de la medida. Dos de ellos consideraron que la prisión preventiva se utiliza de manera excesiva e injustificada, otros dos indicaron que no se puede generalizar, y hubo uno que negó rotundamente que se esté haciendo un uso excesivo de la misma.

En relación con delitos de índole económica, García Martínez (2013) sostiene que la prisión provisional se ha convertido en una práctica excesivamente común, citando como ejemplos los casos de los Eres, Gürtel y Miguel Blesa. Frecuentemente, esta medida se justifica por la necesidad de mitigar la alarma social y satisfacer demandas de seguridad pública, en lugar de basarse en las razones legalmente prescritas. Esta tendencia persiste a pesar de que el Tribunal Constitucional ha descartado la "alarma social" como fundamento válido para la aplicación de la prisión provisional.

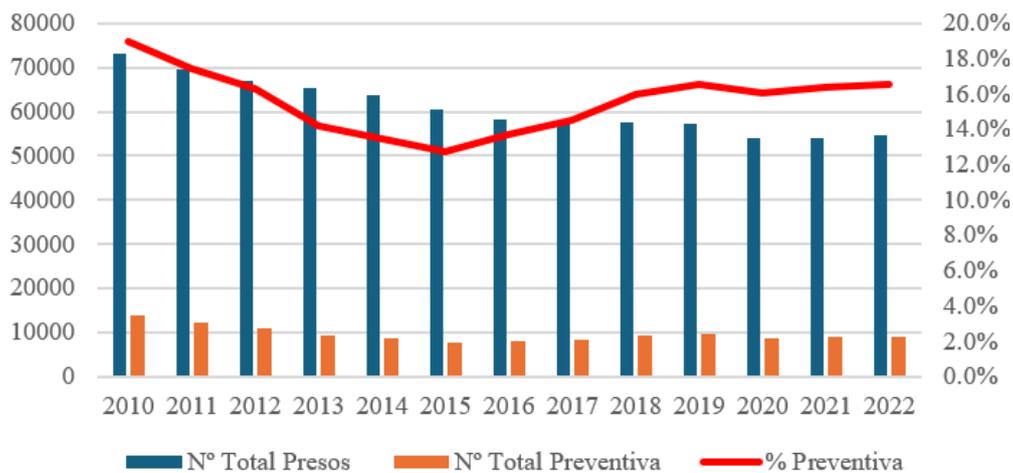
**Figura 1.**

*Datos estadísticos sobre la población reclusa*

Año	Nº Total Presos	Nº Total Preventiva	% Preventiva
2010	73088	13837	18.9%
2011	69588	12148	17.5%
2012	67047	10938	16.3%
2013	65395	9292	14.2%
2014	63709	8595	13.5%
2015	60395	7684	12.7%
2016	58360	7996	13.7%
2017	57488	8362	14.5%
2018	57553	9205	16.0%
2019	57213	9452	16.5%
2020	54053	8672	16.0%
2021	53944	8849	16.4%
2022	54587	9025	16.5%

*Fuente: elaboración propia*

**Figura 2.**  
*Proporción de reclusos en prisión preventiva*



*Fuente: elaboración propia*

Según los datos estadísticos sobre la población reclusa proporcionados por el Poder Judicial, en 2022, 9,025 personas se encontraban en prisión preventiva, lo que representa el 16.5% de los reclusos sin condena firme. Al analizar las cifras presentadas en las figuras 1 y 2, se observa un descenso significativo en el número total de presos en los últimos 12 años. Sin embargo, resulta preocupante que el número de reclusos en prisión preventiva no haya disminuido en la misma proporción. Aunque es cierto que experimentó una reducción considerable entre 2013 y 2017, el número de personas en prisión provisional ha permanecido constante a lo largo de los años, a pesar de la continua disminución en el número total de presos.

En adición a la información proporcionada, sería valioso examinar las estadísticas relacionadas con la proporción de presos en prisión preventiva que, al final del proceso judicial, resultan absueltos. Lamentablemente, no se han hallado estadísticas oficiales que aborden este aspecto. No obstante, la investigación de Sánchez et al. (2017), que analizó casos de individuos declarados culpables e inocentes sometidos a prisión preventiva entre 2005 y 2015, revela que la gran mayoría de quienes han sido incorrectamente confinados bajo esta medida cautelar son varones mayores de 26 años. Asimismo, el estudio destaca una notable sobrerrepresentación de personas extranjeras en la prisión preventiva, lo que sugiere un escrutinio más intenso hacia este grupo poblacional.

Del análisis realizado, se puede inferir que existen diversos motivos para proponer la aplicación de un algoritmo que asista a los jueces en la toma de decisiones respecto a la

prisión provisional. En resumen, en primer lugar, la sociedad se encuentra en un momento de auge irreversible de la Inteligencia Artificial, para el cual es conveniente adaptar los sistemas procesales-penales. En segundo lugar, las opiniones divergentes entre los magistrados con respecto al uso de la prisión provisional y los factores que influyen en el proceso de toma de la decisión sugieren que la aplicación de esta medida no es uniforme, lo que implica la existencia de subjetividad en el proceso judicial. Por último, se observa una discrepancia entre la disminución general de la población carcelaria y la persistencia de una proporción significativa de reclusos en prisión preventiva. La propuesta de implementar un algoritmo se presenta como una solución potencial para reducir o eliminar la subjetividad en las decisiones judiciales y obtener resultados más precisos.

#### **4. EL ALGORITMO**

Asquerino Lamparero (2022) simplifica la definición de un algoritmo al equiparlo con una receta, donde los datos son el ingrediente esencial. A diferencia de los seres humanos, los algoritmos no tienen límites en su capacidad de aprendizaje, ya que pueden ser entrenados con una cantidad ilimitada de datos. Esta capacidad permite que el algoritmo no solo proporcione respuestas, sino que también evolucione, mejore y tenga la capacidad de predecir comportamientos futuros.

Como se ha explicado anteriormente, para tomar la decisión de decretar prisión provisional, es necesario que el Juez valore el riesgo del acusado en diferentes aspectos. Es importante destacar que esta evaluación del riesgo implica hacer una predicción. La investigación a lo largo de muchos años ha reiterado que incluso los modelos estadísticos más simples superan a los humanos en la capacidad de predecir. Según Miller (2018), los humanos son notoriamente deficientes en la toma de decisiones, y los algoritmos han demostrado producir resultados más precisos y menos sesgados cuando los reemplazan. Por todo lo anterior, se plantea la introducción de algoritmos en la toma de decisiones sobre prisión preventiva. En los siguientes apartados se introducirán diferentes aspectos relevantes del potencial algoritmo.

##### **4.1 Elección del tipo de algoritmo**

Si bien sabemos que la Inteligencia Artificial es capaz de hacer predicciones de forma más precisa que los humanos, existen diferentes tipos de algoritmos y no hay uno que sea dominante en todos los contextos. Dependiendo de los datos de los que se disponga, un

tipo u otro funcionará mejor (James et al., 2013). Además, para valorar la implementación de estos, es esencial tener en mente que no todos funcionan de la misma forma, y presentan diferencias en precisión, transparencia y otros aspectos. La elección del tipo más adecuado escapa del alcance de este trabajo, puesto que dependerá de la cantidad de datos de los que se disponga y de los resultados que se obtengan al ser probados. No obstante, se ofrecerá una breve explicación sobre algunas de las posibilidades.

En primer lugar, existen los algoritmos de regresión lineal, que pesar de su simplicidad y antigüedad, siguen siendo ampliamente utilizados debido a su utilidad y facilidad de interpretación. Sirven para predecir un resultado basándose en variables continuas. Esto es, en su forma más básica, un modelo de regresión lineal simple intenta establecer una relación lineal entre algunas variables independientes (X) y una variable dependiente (Y). Matemáticamente, puede escribirse como  $Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \epsilon$ , siendo las  $\beta$  unos coeficientes que representan la relación entre las variables y que se hayan en función de los datos sobre los que se crea el modelo (James et al., 2021).

En segundo lugar, existen algoritmos de árboles de decisión. Funcionan de manera intuitiva, dividiendo el conjunto de datos en subconjuntos más pequeños basados en características específicas, con el objetivo de hacer predicciones. Se comienza seleccionando la mejor característica para dividir el conjunto de datos de forma que los subconjuntos sean lo más diferentes posibles entre sí, pero que los casos dentro de la partición se asemejen lo máximo posible. El proceso de división se repite recursivamente en cada subconjunto generado hasta que se cumple algún criterio de parada, como, por ejemplo, que la profundidad del árbol alcance un límite predefinido. Una vez construido el árbol, se utiliza para hacer predicciones recorriendo el árbol desde la raíz hasta las hojas, evaluando las condiciones en cada nodo interno y asignando la etiqueta de clasificación o el valor de regresión correspondiente en las hojas (Sangüesa, 2005).

En tercer lugar, los más complejos son los algoritmos de Deep Learning y las Redes Neuronales Artificiales, pues imitan la estructura de las redes neuronales biológicas. Estas redes neuronales son extremadamente grandes y utilizan pequeñas cantidades de datos etiquetados y cantidades mucho mayores de datos no etiquetados. Tienen muchos inputs que pasan por varias capas ocultas antes de producir uno o más outputs. Estas conexiones forman un ciclo específico que refleja la forma en que el cerebro humano procesa la información y establece conexiones lógicas (Shrestha & Mahmood, 2019). Aunque estos

algoritmos son capaces de aprender y modelar relaciones complejas entre los datos, también presentan desafíos, como la dificultad de interpretación y la necesidad de grandes volúmenes de datos para un entrenamiento preciso.

Al decidir qué tipo de algoritmo utilizar, es crucial tener en cuenta sus diferencias en varios aspectos, incluyendo la precisión y la transparencia. Aunque se han mencionado solamente tres tipos, es importante destacar que existen muchos más y que la elección del tipo adecuado dependerá en gran medida de la disponibilidad y naturaleza de los datos. Para optimizar la comprensión del estudio, se detallará el proceso de desarrollo del algoritmo partiendo de la premisa de que se utiliza un modelo de regresión lineal. La elección de este tipo de algoritmo se debe a su capacidad para proporcionar una mayor transparencia, un aspecto fundamental en decisiones que implican alto riesgo, como es decretar prisión provisional.

La mayor transparencia de los modelos de regresión lineal se debe a que este tipo de modelos son intrínsecamente interpretables, lo que probablemente también los hace explicables. En los mismos, la relación entre las variables independientes (edad, antecedentes, etc.) y la variable dependiente (riesgo) se expresa a través de una ecuación lineal. Cada variable independiente recibe un peso o coeficiente. La predicción de si una persona investigada representa un alto riesgo se calcula multiplicando cada dato concreto por su coeficiente correspondiente y sumando estos valores ponderados. Por lo tanto, comprender qué valores concretos se tienen en cuenta y cuáles fueron los más importantes no representa un desafío en el algoritmo propuesto. Optar por un modelo que es inherentemente interpretable, en lugar de depender de técnicas de *explainable Machine Learning* permite que el modelo ofrezca explicaciones que son fieles a los cálculos realizados.

Una crítica frecuente a los modelos de regresión lineal es que, pese a su alta interpretabilidad, tienden a ofrecer un rendimiento inferior en comparación con los modelos de *Deep Learning* (Selbst y Barocas, 2018). Sin embargo, esta afirmación no es absoluta y depende de características relacionadas con los datos. Primero, el tamaño del conjunto de datos es determinante; en casos donde este es reducido, modelos más sencillos como la regresión lineal podrían ser más efectivos, dado que los modelos más complejos necesitan de una cantidad de datos posiblemente muy superior a la que se tendría acceso (Wilstrup & Kasak, 2021). Segundo, la comparación de precisiones también depende cómo sea la relación entre variables independientes y dependiente. Si

los datos reflejan patrones simples y lineales, la regresión lineal podría incluso superar a los modelos de *Deep Learning*. Además, si los datos están bien estructurados y las variables han sido seleccionadas cuidadosamente, es probable que la diferencia entre ambos modelos sea mínima (Rudin, 2019).

## 4.2 Construcción del algoritmo

Antes de sumergirnos en la construcción del algoritmo predictivo, es necesario explicar brevemente el funcionamiento de estos. El entendimiento profundo de los mismos no es necesario para comprender el presente estudio, pero sí es conveniente tener una base mínima.

La fórmula general es  $Y=f(X)+\epsilon$  y establece que la variable dependiente  $Y$  puede ser modelada como una función  $f$  de varios predictores  $X_1, X_2, \dots, X_p$ , más un término de error  $\epsilon$ , donde:

- $f$ : Es una función desconocida que describe cómo los predictores  $X$  afectan a la respuesta  $Y$ . Esta función captura la relación sistemática entre las variables.
- $\epsilon$ : Representa el error aleatorio en el modelo, que se asume independiente de  $X$ . Este término refleja la variabilidad en  $Y$  que no puede ser explicada por  $X$ .

Debido a que la forma exacta de  $f$  es desconocida, los investigadores deben estimarla a partir de datos observados y utilizando técnicas de aprendizaje estadístico. Cuando se obtiene tal estimación, se utiliza para hacer predicciones de  $Y$  (James et al., 2013).

En el caso concreto del algoritmo de evaluación de riesgo,  $Y$  es el riesgo de que un individuo cometa un nuevo delito o no cumpla con los requisitos del proceso judicial si no se le retiene en prisión preventiva. Los predictores  $X_1, X_2, \dots, X_p$  incluyen variables como la edad del individuo, su historial criminal, la gravedad de la ofensa actual, entre otras. De esta forma, en base a datos históricos de tales predictores se podrá predecir el riesgo de un nuevo individuo.

A continuación, se explicará el proceso de construcción del algoritmo:

### 4.2.1 Recopilación y procesamiento de datos

El éxito de los algoritmos depende en gran parte de la habilidad que se tenga para acceder a datos (Velupillai et al., 2019). De hecho, cuanto más grandes sean las bases de datos utilizadas para entrenar el algoritmo, mejor será su capacidad de predicción (Turner,

2018). Por lo tanto, es crucial recopilar tantos datos como sea posible para asegurar la calidad del algoritmo. El método de recolección debería ser a través de un estudio retrospectivo, que consiste en analizar datos históricos de individuos que fueron objeto de decisiones de prisión provisional en España en el pasado. En caso de que se inicie el proceso de creación de este algoritmo, se realizaría a través de la autorización o concesión de un organismo estatal. Por lo tanto, se debería tener acceso a los datos de la Administración de Justicia y el Consejo General del Poder Judicial. De los archivos recopilados, se deberán extraer, por un lado, características del individuo, y por otro, información sobre si se decretó prisión preventiva. En caso de que no se haya aplicado la prisión preventiva, también es importante determinar si el resultado fue satisfactorio, es decir, si el individuo no reincidió, si se presentó al juicio, etc.

Uno de los principales problemas que presenta el método de recolección de datos retrospectivo es la facilidad para que ciertos datos relevantes no estén disponibles. En este caso, es posible que las bases de datos no contengan determinada información de los individuos que pueda resultar interesante, como puede ser su situación laboral. Además, los informes judiciales a veces no están suficientemente detallados, lo que dificulta la comprensión completa de cada caso. Otro problema es que estos datos históricos no fueron recopilados con la intención de ser utilizados en un algoritmo predictivo, lo que puede generar discrepancias entre la intención original de la información y su uso actual. A pesar de estas posibles limitaciones, el método de estudio retrospectivo sigue siendo el más adecuado para este contexto, ya que la seriedad de las decisiones sobre prisión provisional y gran impacto que tienen en las personas descartan la posibilidad de realizar un estudio experimental (Montgomery et al., 2021).

Una vez tiene acceso a los datos históricos, se debe efectuar el procesamiento de datos, que tiene como objetivo preparar los datos para el análisis y corregir las inconsistencias que se encuentren en los mismos. Normalmente los problemas que se corrigen son la existencia de datos incompletos, de datos con ruido y de posibles inconsistencias (Hernández y Rodríguez Rodríguez, 2008). Para procesar datos históricos de imputados que han sido sujetos de la decisión de si decretar la prisión preventiva, se deben seguir varios pasos. A título de ejemplo, se exponen los siguientes:

En primer lugar, se debe realizar una limpieza de datos, que incluye corregir o eliminar registros erróneos o inconsistentes, así como rellenar los valores faltantes. Estos objetivos se pueden alcanzar mediante técnicas de imputación, como el uso de la media o mediana

para estimar los valores que faltan. Esto es, durante este proceso, se revisan todos los archivos para corregir errores, tales como nombres mal escritos o fechas incorrectas, eliminar datos duplicados y completar datos faltantes. Por ejemplo, si falta la edad de un sujeto, se podría rellenar este campo utilizando la media de edad de todos los individuos con historiales similares.

En segundo lugar, se ha de llevar a cabo una integración de datos, en la que se unifican datos de múltiples fuentes, homogeneizando formatos y terminologías para asegurar coherencia. Es decir, dado que la información puede venir de diferentes fuentes y no estar organizada de la misma manera, será necesario ajustar los datos para obtener un formato unificado. Por ejemplo, si algunas bases de datos utilizan "Fecha de nacimiento" y otras "F. Nac." para referirse a la misma información, se escogería un término estándar. Igualmente, las fechas podrían estar en formatos diferentes como "DD/MM/AAAA" o "AAAA-MM-DD", y en este paso se homogeneizarían.

Además, un paso esencial es la transformación de datos. Se trata de normalizar los atributos numéricos para asegurar que estén en escalas comparables y generalizar las variables categóricas a niveles más abstractos, lo que facilita su análisis. Por ejemplo, respecto a la situación laboral del imputado, las diversas opciones pueden simplificarse en categorías generales como empleado, desempleado, estudiante y jubilado.

La recopilación y procesamiento adecuados de los datos es fundamental para el éxito del algoritmo de evaluación de riesgo en el contexto de la prisión provisional. A pesar de los retos inherentes al uso de datos retrospectivos, como la falta de información específica y la inconsistencia en los detalles de los registros judiciales, la meticulosa limpieza, integración y transformación de los datos permiten superar estas barreras, asegurando que el algoritmo pueda operar con la mayor precisión posible. Este enfoque no solo mejora la calidad del conjunto de datos, sino que también refuerza la objetividad y la equidad del proceso judicial, contribuyendo a decisiones más justas y fundamentadas.

#### *4.2.2 Selección de variables*

La construcción de algoritmos de evaluación de riesgo se realiza mediante el proceso de desarrollo de modelos predictivos. Independientemente del tipo de algoritmo que se escoja, este desarrollo requiere "entrenar" a los ordenadores utilizando extensas bases de datos. Se analizan registros de antiguos casos criminales, considerando el fallo de las sentencias dictadas y el posterior comportamiento de cada perfil, para determinar qué

variables son las recurrentes. Posteriormente, este procedimiento se invierte: los estadísticos buscan estas variables específicas en nuevos casos con el objetivo de efectuar predicciones. Cuando estas predicciones logran una precisión adecuada en entornos de prueba controlados, se considera que el algoritmo tiene capacidad predictiva y, por tanto, puede ser aplicado a casos actuales (Christin et al., 2015).

Así, la selección de las variables que integrarán el algoritmo se realizará a través del análisis de datos históricos. El objetivo de esta selección de variables es que el algoritmo utilice el menor número posible de variables de las que se obtenga la máxima información, sin comprometer la precisión del modelo. Es crucial prescindir de aquellas variables que resulten irrelevantes o redundantes, ya que estas únicamente añaden complejidad y ruido innecesario al modelo (Galelli et al., 2014). Para ello, se han de seleccionar variables con alta correlación con la variable dependiente, es decir, con el riesgo del individuo, y baja correlación con las variables independientes (Bodur & Atsa'am, 2019).

En el estudio de Elyounes (2020) se analizan las variables consideradas en siete diferentes algoritmos de evaluación del riesgo para la toma de decisiones sobre prisión provisional en Estados Unidos. Cada uno de ellos tiene en cuenta entre tres y doce variables, aunque el que tan solo utiliza tres no se utiliza en la práctica en ningún Estado. Se concluye que el factor más popular es el historial delictivo del investigado, en todas sus diferentes variantes (antecedentes penales, número arrestos previos, número de condenas previas de más de 30 días, casos pendientes de sentencia, etc.). Entre los factores populares también

**Figura 3.**

*Factores utilizados en siete herramientas de evaluación de riesgo preprocesal estadounidenses*

Herramienta	Historial delictivo	Delito actual	Factores socioeconómicos	Edad	Abuso de sustancias	Otros
Pretrial Risk Assessment (PTRA)	Sí	Sí	Sí	No	Sí	Vínculos con el extranjero
Public Safety Assessment (PSA)	Sí	Sí	No	Sí	No	
Virginia Pretrial Risk Assessment Instrument (VPRAI)	Sí	Sí	Sí	No	Sí	Supervisión activa comunitaria
Colorado Pretrial Risk Assessment Tool (CPAT)	Sí	Sí	Sí	Sí	Sí	Problemas de salud mental
Ohio Pretrial Assessment Tool (PAT)	Sí	No	Sí	Sí	Sí	
COMPAS	Sí	Sí	No	No	Sí	Tiempo en la comunidad actual
Kleinberg et al. Tool	Sí	Sí	No	Sí	No	

Fuente: elaboración propia con datos de Elyounes (2020)

destacan el delito del que se le acusa, la edad, la estabilidad laboral, la estabilidad económica y el historial de abuso de sustancias.

A diferencia de los enfoques tradicionales, donde la consideración de cuáles son los factores relevantes a la hora de tomar la decisión puede estar limitada por preconcepciones o la experiencia previa, los algoritmos analizan los datos de manera exhaustiva y pueden identificar patrones y correlaciones que no son evidentes a primera vista, y así decidir de qué forma es más efectivo tener en cuenta un factor para el modelo concreto. Este enfoque se destaca claramente al observar las diferencias en la selección y tratamiento de variables entre distintas herramientas de evaluación de riesgo. Por ejemplo, respecto a la variable de estabilidad laboral, la herramienta PTRRA utilizada en el sistema federal de Estados Unidos distingue dicotómicamente entre acusados empleados y desempleados, mientras que el VPRAI de Virginia amplía las categorías de empleo al considerar estados como ser estudiante, cuidador principal y jubilado, categorías estas que no acarrearán puntos negativos para los acusados. Este enfoque se extiende también a la categorización de la residencia: el PTRRA solo considera a aquellos que son propietarios de una vivienda o que están en proceso de compra, mientras que otras herramientas, como la utilizada en Virginia, también toman en cuenta a los que alquilan su residencia. Este tipo de diferenciación refleja la habilidad del aprendizaje automático para adaptar y afinar los algoritmos con el fin de capturar una imagen más completa y precisa de la situación de cada individuo (Elyounes, 2020).

Es relevante también mencionar el estudio realizado por DeMichele et al. (2018), en el cual se analiza la percepción de distintos actores del sistema de justicia penal sobre la importancia de ciertos factores al tomar decisiones respecto a la libertad bajo fianza antes del juicio. Para ello, se encuesta a jueces, fiscales, abogados defensores, y otros profesionales involucrados en el proceso, preguntando cuán relevantes consideraban distintos factores a la hora de tomar la decisión. La Figura 3 muestra el porcentaje de encuestados que consideran que una determinada variable es "Extremadamente" o "Muy" importante. Se concluye que, según la perspectiva de los profesionales involucrados en el proceso, los factores considerados más relevantes incluyen el historial criminal y sus variantes, el número de ocasiones en que el imputado no se ha presentado previamente en el proceso, la existencia de lesiones causadas a la víctima, y la implicación de armas en el hecho.

**Figura 4.**

*Porcentaje de Profesionales de la Justicia Penal que Perciben que los Elementos son "Extremadamente" o "Muy" Importantes en la Decisión de Liberar/Retener Antes del Juicio*

Variable	Todos	Jueces	Fiscales	Defensores	Personal de libertad previa al juicio
Cargos actuales	76%	85%	100%	42%	68%
Cargos pendientes	90%	100%	94%	42%	91%
Historial criminal	91%	88%	100%	58%	98%
Fallas previas a comparecer	81%	83%	71%	33%	93%
Lesiones a la víctima	73%	75%	100%	27%	73%
Involucramiento de armas	80%	88%	100%	36%	77%
Edad del acusado	44%	40%	18%	33%	58%
Condición mental del acusado	58%	63%	65%	55%	54%
Historial de uso de sustancias	33%	40%	29%	29%	30%
Capacidad de la cárcel	6%	3%	0%	0%	7%
Otros	10%	15%	17%	5%	20%

*Fuente: elaboración propia con datos de DeMichele et al. (2018)*

Por otro lado, además de seleccionar las variables más relevantes mediante el análisis de datos históricos, la Ley de Enjuiciamiento Criminal establece una serie de criterios para la evaluación del riesgo, tal y como se explicó anteriormente. Muchos de estos criterios son cuantificables de manera directa y su importancia es claramente perceptible. Por ejemplo, la ley sugiere considerar la participación del investigado en una colaboración organizada para perpetrar el delito. Esta consideración se traduce en una variable dicotómica, asignándosele un valor de 0 o 1 dependiendo de la ausencia o presencia de dicha colaboración. En consecuencia, se deberá establecer que la presencia de colaboración organizada indica un mayor riesgo por parte del imputado.

Ahora bien, algunos de estos factores presentarán problemas a la hora de ser introducidos en el algoritmo. Por ejemplo, uno de los factores que la ley indica que se debe tener en cuenta a la hora de evaluar el riesgo de fuga del individuo es la naturaleza del hecho delictivo. Es posible que a la hora de entrenar al algoritmo se detecte un problema de multicolinealidad con el factor “gravedad de la pena”. La multicolinealidad significa que hay una alta correlación entre dos variables, es decir, las variables “naturaleza del hecho” y “gravedad de la pena” proveen información muy parecida (Bayman & Dexter, 2021). Este problema se solucionaría eliminando una de las variables. Lo mismo ocurre con la variable “historial criminal” y todas sus variantes, por lo que para evitar el problema de multicolinealidad solo se puede incluir en el modelo una de ellas.

Asimismo, pueden surgir problemas de cuantificación, puesto que no todos los factores que la ley exige que se tengan en consideración son fácilmente cuantificables. Un ejemplo

de esto sería la capacidad de acceso a las fuentes de prueba o la capacidad de influencia sobre otros.

En resumen, la identificación de variables relevantes constituye una base crucial para la construcción de algoritmos predictivos eficaces. No obstante, la selección de variables concretas excede el alcance de este estudio. Este análisis pretende ofrecer una visión general de los principios y consideraciones involucrados en el proceso, dejando la determinación final de variables específicas expertos en el campo.

#### *4.2.3 Desarrollo y validación del modelo predictivo*

Con las variables seleccionadas, se construye un modelo predictivo. Aquí se pueden emplear diversos algoritmos y, como se ha explicado anteriormente, no hay un algoritmo que sea necesariamente mejor que los demás; la elección debe basarse en los datos específicos y el contexto de uso.

En el caso de los algoritmos predictivos, el principal objetivo es buscar la máxima precisión. Esto es, que cuando prediga un riesgo bajo y por tanto se deje en libertad al imputado, el mismo no reincida, se presente al juicio, etc., y, al contrario, que cuando prediga un riesgo elevado, se corresponda con la realidad. Sin embargo, hay otros factores relevantes como su simplicidad y su interpretabilidad, que es importante valorar a la hora de escoger el tipo (Cano de Amo et al., 2005).

Una vez creado el algoritmo, se procederá a su validación, que tiene por objeto asegurar que el mismo predice correctamente los resultados o comportamientos futuros basándose en datos que no fueron utilizados durante la fase de entrenamiento del modelo. Esta etapa es crucial para evaluar cómo el modelo se comportará en situaciones reales, fuera de las condiciones controladas de entrenamiento. Para validar un algoritmo, es necesario reservar una parte de los datos históricos disponibles, los cuales no se utilizan en la etapa de entrenamiento. La idea es que estos datos sirvan como nuevos para el modelo durante la fase de validación, permitiendo simular una situación real en la que el modelo se enfrenta a información que no ha visto antes (Steyerberg et al., 2001).

Introduciendo las variables de los expedientes pasados en nuestro modelo y observando si las predicciones coinciden con los resultados reales, se puede medir la precisión del modelo. Si el modelo puede predecir correctamente el riesgo del imputado en la mayoría de los casos, se considera que tiene una buena capacidad predictiva. Por el contrario, si el

modelo falla en predecir los desenlaces reales de manera consistente, podría necesitar ajustes o mejoras.

En resumen, este apartado ha proporcionado una visión general limitada, aunque ilustrativa, del proceso de elaboración de un algoritmo predictor de riesgo. A continuación, exploraremos cómo se vería esto aplicado a nuestro sistema procesal penal.

## **5. IMPLEMENTACIÓN EN LAS DECISIONES JUDICIALES**

Aunque se trate de un tema relativamente novedoso en la justicia española, los algoritmos de evaluación del riesgo se han integrado profundamente en el sistema de justicia penal de Estados Unidos (Viljoen et al., 2019). Para valorar la posible implementación de un algoritmo de este estilo a nuestro sistema penal, conviene analizar cómo resultaría su aplicación en la práctica. A continuación, se plasmará una imagen representativa de cómo sería el proceso, detallando los pasos a seguir por parte de los jueces, cómo deben medir e interpretar el riesgo y cuál sería el peso del algoritmo a la hora de tomar la decisión.

De acuerdo con lo dispuesto en la Ley de Enjuiciamiento Criminal, el proceso para decretar la prisión provisional comenzaría con la detención del individuo por parte de la policía, la cual puede llevarse a cabo cuando existen indicios razonables de la comisión de un delito. El detenido deberá ser puesto a disposición del Juez de Instrucción más cercano al lugar de la detención dentro de un plazo máximo de 72 horas. Durante este periodo, el detenido tendrá la opción de solicitar un *habeas corpus* si considera que su detención ha sido ilegal, un procedimiento que se tramita de manera urgente para verificar la legalidad de la detención. Una vez ante el juez, el detenido declarará, asistido por su abogado, en una audiencia donde se podrán presentar pruebas y testimonios. En caso de que el fiscal identifique uno de los cuatro riesgos previamente explicados, solicitará al juez la prisión provisional como medida cautelar para asegurar la presencia del acusado en el juicio y evitar la obstrucción del proceso judicial (Díaz Torrejón, 2013).

Es en este momento en el que el Juez deberá tomar la decisión. Para ello, debe tener en cuenta los tres requisitos que establece la Ley de Enjuiciamiento Criminal previamente explicados. Primero, verificar el cumplimiento del límite penológico requerido por el artículo 503.1. 1º LECrim. Segundo, valorar la existencia de indicios racionales de criminalidad por parte del investigado. Por último, justificar la prisión preventiva por

riesgo de fuga, reiteración delictiva, obstrucción de la instrucción penal o atentado contra bienes jurídicos de la víctima.

El algoritmo propuesto está específicamente diseñado para asistir en la evaluación del riesgo, que constituye el tercer requisito en el proceso de toma de decisiones judiciales. Sin embargo, es importante destacar que la utilización de este algoritmo no compromete la autonomía del juez o magistrado a la hora de verificar el cumplimiento del límite penológico ni a la hora de evaluar si existen motivos suficientes para considerar al imputado como responsable del delito. Dicho esto, a continuación, se explicará en detalle cómo se lleva a cabo la evaluación del riesgo del investigado mediante el uso del algoritmo.

### **5.1 Medición del riesgo**

El algoritmo de predicción debe determinar si una persona debe ser sometida a prisión preventiva basándose en su nivel de riesgo. Este proceso se realiza mediante la asignación de un porcentaje por parte del algoritmo, que actúa como un indicador de la probabilidad de riesgo basado en los datos introducidos. Este porcentaje refleja la probabilidad de que se materialice el riesgo evaluado (Moore et al., 2023).

Conforme a lo establecido por la legislación, el juez debe analizar el riesgo asociado al imputado considerando cuatro aspectos fundamentales: el riesgo de fuga, el riesgo de alterar pruebas, el riesgo de daño a los bienes jurídicos de la víctima, y el riesgo de reincidencia. Para alinear el funcionamiento del algoritmo con estos requisitos legales, sería necesario desglosar los resultados en estas cuatro categorías específicas. Esto implicaría el desarrollo de cuatro algoritmos especializados, cada uno enfocado en evaluar una de las categorías de riesgo mencionadas, asignando un porcentaje que refleje la probabilidad de que el individuo incurra en cada tipo de riesgo. Los factores considerados por cada algoritmo variarán en importancia dependiendo del tipo de riesgo evaluado. Un posible resultado de esta evaluación se presentaría de la siguiente manera:

## Figura 5.

*Modelo de evaluación del riesgo dividida en riesgos específicos*

### **EVALUACIÓN DEL RIESGO ESPECÍFICA**

Nombre: ABC

Apellidos: UVW XYZ

DNI: 00000000X

Riesgo de fuga	<b>65%</b>
Riesgo de ocultación, alteración o destrucción de pruebas	<b>22%</b>
Riesgo de atentado contra bienes jurídicos de la víctima	<b>15%</b>
Riesgo de reincidencia	<b>80%</b>

*Fuente: elaboración propia*

Aunque ese sería el escenario ideal para adaptar este nuevo sistema a lo exigido por la Ley de Enjuiciamiento Criminal, es posible que no dispongamos de suficientes datos para llevarlo a cabo. Esto es, para hacer una clara diferenciación entre cada tipo de riesgo debemos de tener recopilada información sobre el tipo de riesgo que suponía cada persona que fue encarcelada provisionalmente. En caso de no disponer de esta información, la evaluación del riesgo se tendría que hacer de forma genérica, considerando todos los factores en un mismo algoritmo, y se mostraría como resultado un solo porcentaje. Véase el ejemplo de cómo se mostraría.

## Figura 6.

*Modelo de evaluación del riesgo “general”*

### **EVALUACIÓN DEL RIESGO GENERAL**

Nombre: ABC

Apellidos: UVW XYZ

DNI: 00000000X

Riesgo	<b>56%</b>
--------	------------

*Fuente: elaboración propia*

Por tanto, se recomienda que la evaluación del riesgo se lleve a cabo de manera integral. Es crucial reconocer que, aunque la legislación no prioriza un tipo de riesgo sobre otro, en la práctica, el riesgo de fuga y el riesgo de reincidencia suelen tener una relevancia considerablemente mayor (Muñoz Cuesta, 2019). Esta realidad debe reflejarse adecuadamente en la evaluación global del riesgo, asegurando que ambos factores sean

debidamente ponderados frente a los otros para garantizar una toma de decisiones justa y equilibrada.

## **5.2 Interpretación del riesgo**

Aunque el algoritmo proporcione un porcentaje, la tecnología no tiene la capacidad de determinar en qué punto el riesgo justifica el encarcelamiento de una persona. La opinión sobre el umbral de riesgo aceptable varía significativamente: para algunos, una probabilidad de reincidencia del 51% podría justificar el encarcelamiento, mientras que otros podrían considerar un umbral del 80% como más apropiado. A diferencia de la evaluación de riesgo explicada en el apartado anterior, que es un ejercicio de ciencia de datos enfocado en predecir probabilidades, la interpretación del riesgo no es en absoluto científica. Por el contrario, las consideraciones jurídicas y éticas son primordiales (Martínez et al., 2019).

Por ende, será necesario definir legalmente la cantidad y el umbral de las categorías de riesgo, así como determinar cuáles implican la necesidad de prisión preventiva. Se podría optar por un enfoque dicotómico, diferenciando entre "riesgo bajo" y "riesgo alto", donde un "alto riesgo" desaconseje la liberación del detenido por la alta probabilidad de reincidencia, fuga, etc. Alternativamente, se podrían establecer tres categorías, añadiendo una de "riesgo moderado" para una graduación más precisa del riesgo (Hamilton, 2022). De nuevo, esta elección dependerá de la política judicial adoptada.

## **5.3 Alcance del algoritmo**

Supongamos que el algoritmo predice un riesgo del 95% para un caso específico, categorizándolo como "alto riesgo". A pesar de esta evaluación, el juez puede determinar que no hay suficiente riesgo para justificar una medida cautelar. ¿En qué medida debe considerarse el resultado del algoritmo en la decisión judicial? Esta sección explora el alcance o peso que el resultado del algoritmo debe tener en la toma de decisiones judiciales, abordando la cuestión de hasta qué punto los jueces están obligados o pueden seguir las recomendaciones de este.

Una posibilidad es que el algoritmo tenga un carácter vinculante. Esto implica que la calificación proporcionada por el algoritmo automáticamente determine si se cumple o no el tercer requisito para ordenar la prisión provisional. Si bien es cierto que los algoritmos ofrecen una objetividad útil a la hora de tomar decisiones, su rigidez puede resultar contraproducente debido a la falta de adaptación individualizada a cada caso

concreto. En la toma de decisiones legales de tal calibre, la capacidad de individualización es crucial para asegurar que prevalezca la justicia y la equidad. En este sentido, el juicio humano tiene la habilidad de considerar un espectro más amplio circunstancias únicas a cada caso, permitiendo decisiones más ajustadas a la realidad específica de cada individuo (Chiao, 2019). Por tanto, es fundamental asegurar que el factor humano continúe desempeñando cierto papel en la toma de decisiones, para evitar la aplicación mecánica de la justicia y fomentar un enfoque más equitativo y sensible al contexto. Por tanto, se recomienda descartar tal nivel de influencia.

En el otro extremo está la posibilidad de que sea meramente informativo. Esto es, el algoritmo actúa meramente como una herramienta informativa, proporcionando a los jueces una evaluación cuantificada del riesgo, pero sin imponer restricciones sobre sus decisiones. En este caso, los jueces tienen plena discreción para considerar o desestimar los resultados del algoritmo según su criterio y el contexto del caso. Al aplicar este tipo de modelo, existe preocupación sobre cómo las personas interactúan con las recomendaciones provenientes de un algoritmo. Por un lado, existe el temor de que los jueces ignoren su propio criterio y sigan ciegamente las sugerencias del algoritmo, lo que podría llevar a errores. Por otro lado, preocupa que, por aversión a la automatización, los jueces descarten las recomendaciones del algoritmo sin considerar su utilidad. Sin embargo, un estudio llevado a cabo por De-Arteaga et al. (2020) concluyó que estos problemas no se manifestaron. Los trabajadores, en este caso, consideraron las puntuaciones proporcionadas por la herramienta, pero no las siguieron de manera automática. En su lugar, demostraron ser capaces de integrar las recomendaciones del algoritmo con su propio criterio y otras fuentes de información, adoptando un enfoque equilibrado en la toma de decisiones.

Se plantea una opción intermedia: el modelo semiobligatorio. En este enfoque, las predicciones del algoritmo tienen un peso considerable y solo pueden ser contrarrestadas si se justifica adecuadamente la decisión. Se considera que este modelo encuentra el mejor equilibrio entre los fines perseguidos por la implantación del algoritmo y los riesgos que el mismo presenta. Así, se aspira a mejorar la objetividad y precisión de las decisiones, al tiempo que se asegura la posibilidad de corregir aquellas que resulten injustas mediante una justificación adecuada.

Es crucial considerar la importancia que nuestro sistema legal concede al derecho a la libertad personal. Como se ha explicado, el uso de la prisión provisional debe regirse por

los principios de excepcionalidad, subsidiariedad y provisionalidad, siempre dando prioridad a la libertad del acusado. Por lo tanto, al aplicar el algoritmo, es necesario reflejar esta clara preferencia por la libertad.

Sin perjuicio de que todo auto que decreta la aplicación de la prisión provisional debe estar motivado, en el modelo semiobligatorio propuesto se requerirán distintos niveles de motivación. Cuando el órgano judicial esté de acuerdo con la sugerencia del algoritmo, la motivación será la que se viene haciendo en la actualidad, pero podrá ayudar a justificar su decisión mediante la señalización de las variables con más peso en la decisión, extraídas del propio algoritmo. Para los casos en los que el juez o magistrado opte por contradecir lo establecido por el algoritmo en favor de la libertad, también se requerirá de esa motivación ordinaria.

En el caso de que el juez contravenga al algoritmo en detrimento de la libertad del individuo, será exigido un plus de motivación. Esto significa que el juez deberá proporcionar una explicación detallada que justifique su decisión de imponer la prisión provisional, a pesar de la recomendación del algoritmo de conceder la libertad. Esta motivación deberá abordar no solo las razones jurídicas y fácticas tradicionales, sino también explicar por qué las variables y análisis proporcionados por el algoritmo no son aplicables o resultan insuficientes en el caso concreto.

La propuesta concreta consiste en el siguiente esquema, para la que se asume que se cumplen los otros dos requisitos para decretar prisión provisional:

### Figura 7.

*Propuesta del nivel de influencia del algoritmo*

	<b>JUEZ ESTIMA LIBERTAD</b>	<b>JUEZ ESTIMA PRISIÓN PROVISIONAL</b>
<b>ALGORITMO ESTIMA LIBERTAD</b>	Libertad (Motivación)	Prisión provisional (Exige plus de motivación)
<b>ALGORITMO ESTIMA PRISIÓN PROVISIONAL</b>	Libertad (Motivación)	Prisión provisional (Motivación)

*Fuente: Elaboración propia*

#### **5.4 Plazos máximos**

Como se mencionó anteriormente, la Ley de Enjuiciamiento Criminal impone límites máximos al período de prisión provisional, basados en la gravedad del delito imputado y en el tipo específico de riesgo que se busca mitigar. Específicamente, el artículo 504 establece plazos más breves para los casos en los que se pretende prevenir la alteración, ocultación o destrucción de pruebas.

Esta distinción es especialmente relevante al evaluar la implementación de herramientas predictivas en el proceso judicial. El algoritmo propuesto, al operar sobre la base de un riesgo genérico, no distingue entre los diferentes tipos de riesgos contemplados por la ley actual. En consecuencia, para integrar eficazmente esta tecnología en el sistema legal, sería necesario modificar el artículo 504 de la LECrim. Esto implicaría eliminar la diferenciación actual por tipo de riesgo y establecer plazos máximos generales, o quizás desarrollar nuevos criterios que se basen en otros factores relevantes.

Un cambio legislativo de esta magnitud requeriría un análisis detallado de sus impactos potenciales, además de una amplia consulta con profesionales legales para asegurar el cumplimiento de la exigencia constitucional de definir un límite máximo para la duración de la prisión provisional, según establece el artículo 17.4 de la Constitución Española. Sin embargo, se ha de destacar que tal y como señala Enrique López López (2003), vocal del Consejo General del Poder Judicial, para garantizar el cumplimiento de la norma constitucional es necesario y suficiente que el legislador establezca plazos específicos para la duración de la prisión provisional y que estos plazos sean respetados. Por lo tanto, modificar el plazo establecido no implicaría, en principio, una violación de la Constitución.

#### **5.5 Recurso**

La decisión de un juez de ordenar prisión preventiva, incluso con el uso de un algoritmo predictivo, está sujeta a recurso. La introducción de este algoritmo no debe alterar esta garantía procesal esencial. Sin embargo, los fundamentos para recurrir dicha decisión podrían modificarse ligeramente debido a la influencia del algoritmo en el proceso decisonal. A continuación, se explora cómo podría variar el panorama de los recursos interpuestos por las partes afectadas contra los autos de prisión preventiva cuando se ha utilizado un algoritmo en la decisión.

Recordemos que para acordar prisión preventiva deben concurrir tres requisitos: el delito del que se le acusa debe tener una pena de prisión igual o superior a dos años, debe existir una base sólida que respalde la creencia de culpabilidad del encausado, y debe constatar la existencia de uno de los mencionados riesgos. De estos tres requisitos, el algoritmo influye únicamente en la evaluación de los riesgos, por lo que su implementación no afectará los argumentos basados en los dos primeros requisitos, que continuarán tratándose como hasta ahora.

El escenario se altera al recurrir la decisión basada en el tercer requisito—la evaluación de riesgos. En ese caso, se pueden presentar dos escenarios distintos:

- A. Que se recurra una decisión en la que el algoritmo identifique un alto riesgo y el juez concuerde con esta evaluación. En este supuesto, dado que el juez posiblemente se apoye en el algoritmo para fundamentar su motivación, los motivos en los que se basa un recurso cambian. En este contexto, no cabría recurrir la existencia del algoritmo *per se*. El recurso podría basarse únicamente en la falta de uno o varios inputs. Esto es, la parte afectada podría argumentar que el algoritmo no ha incluido ciertas variables o circunstancias pertinentes al caso que, de haber sido consideradas, podrían haber llevado a una conclusión diferente respecto a la imposición de la prisión provisional.
- B. Que se recurra una decisión en la que el algoritmo sugiera un bajo riesgo, pero el juez, no obstante, opte por imponer la prisión preventiva. Dado que en este supuesto es imperativo que el juez proporcione una justificación exhaustiva para su decisión, los motivos para recurrir este tipo de decisiones serían los mismos que se emplean en la actualidad, principalmente la insuficiencia en la fundamentación del auto. Si la motivación del juez no es lo suficientemente detallada para contradecir la evaluación de riesgo propuesta por el algoritmo, se podría argumentar que se ha producido una violación del derecho a la tutela judicial efectiva.

En resumen, aunque el uso del algoritmo se limita a evaluar los riesgos asociados al imputado y no altera los fundamentos legales para decretar la prisión provisional, modifica en algunos casos los argumentos que las partes pueden utilizar al interponer recursos. Este nuevo escenario demanda una revisión y posiblemente nuevas directrices sobre cómo deben estructurarse y fundamentarse los recursos para garantizar que

continúen protegiendo efectivamente los derechos del imputado en un sistema procesal penal que comienza a abrazar la tecnología avanzada.

## **6. PROBLEMAS Y CONSIDERACIONES ÉTICAS**

La integración de algoritmos en la toma de decisiones judiciales, particularmente en lo que respecta a la prisión preventiva, plantea una serie de desafíos éticos y problemas prácticos que no pueden ser ignorados. Mientras que los algoritmos ofrecen la promesa de mejorar la objetividad y eficiencia de los procesos judiciales, también introducen preocupaciones relacionadas con la desconfianza, la transparencia y la potencial perpetuación de sesgos preexistentes. Este capítulo se dedica a explorar estas problemáticas, evaluando cómo las percepciones tanto de la sociedad como de los profesionales del derecho pueden influir en la aceptación de estas herramientas tecnológicas avanzadas. Además, se discutirán los retos relacionados con la interpretabilidad y la explicabilidad de los algoritmos, fundamentales para su justa implementación dentro del marco legal y ético existente.

### **6.1 Desconfianza en el algoritmo**

#### *6.1.1 Percepción social*

Uno de los principales desafíos que enfrentan los algoritmos de evaluación de riesgo en el contexto de la prisión provisional es la marcada desconfianza social hacia estos sistemas tecnológicos. Los algoritmos, aunque potencialmente más eficientes, a menudo enfrentan escepticismo sobre si pueden tomar decisiones justas, especialmente en comparación con el juicio humano (Lee, 2018). A la hora de implantar un algoritmo de evaluación de riesgo en el contexto de decisiones sobre prisión preventiva, será importante comprender cómo valora la sociedad el uso de este tipo de herramientas, puesto que esta percepción influye significativamente en la aceptación de las mismas.

El estudio llevado a cabo por Mok et al. (2023) demuestra que las personas perciben las evaluaciones de riesgo realizadas por algoritmos como menos justas en comparación con evaluaciones idénticas realizadas por humanos. Esta percepción se sostiene incluso cuando los algoritmos demuestran un rendimiento igual o superior al de las evaluaciones humanas. Además, según Gabbiadini et al. (2024), existe una relación inversamente proporcional entre el riesgo social percibido y el valor social percibido. Esto es, cuanto

mayor es el riesgo que la gente percibe de una herramienta IA, menor es el valor social que se le atribuye.

Para combatir la desconfianza en el algoritmo propuesto, es crucial demostrar su eficacia a la sociedad, puesto la confianza del público tiende a incrementarse al presenciar resultados precisos y predicciones acertadas (Hobson et al., 2021). Generalmente, la aceptación de la tecnología surge cuando su efectividad es evidente. Un ejemplo claro es el uso de aviones; la mayoría de las personas no cuestionan su mecanismo interno, sino que confían en las explicaciones básicas sobre su funcionamiento y en las garantías de los expertos sobre su seguridad. Para alcanzar este nivel de confianza con los algoritmos, es esencial comunicar de manera efectiva tanto su capacidad predictiva como sus características básicas. Esto implica un compromiso con la transparencia, aspecto que será discutido más adelante.

#### *6.1.2 Percepción de los jueces*

Desde el punto de vista del poder judicial, es posible que exista una resistencia por parte de los jueces a tomar decisiones de tal calibre basándose únicamente en probabilidades. En este sentido, un estudio de Wells (1992) analizó cómo diferentes individuos, entre los que se incluían jueces, procesan y aplican las pruebas basadas en estadísticas matemáticas en la toma de decisiones legales. Dicho estudio concluyó que, incluso en los casos en que la probabilidad es alta, hay una notable reticencia a fundamentar decisiones de alto riesgo exclusivamente en datos estadísticos.

No obstante, no podemos olvidar que la implementación del algoritmo se ha de comparar con el *status quo* y no con una supuesta situación ideal. Si bien es cierto que la idea de encarcelar a alguien basándose solo en la recomendación de un ordenador puede resultar incómoda, la realidad es que las decisiones de encarcelamiento basadas en predicciones de riesgo también lo son. Y, de hecho, estas decisiones son una práctica común hoy en día, tomadas en base a juicios subjetivos de los jueces.

Por lo tanto, la pregunta clave no es si el algoritmo es la solución óptima a la problemática de la prisión provisional. Más bien, debemos partir del hecho de que las decisiones de encarcelamiento basadas en predicciones son un mal necesario e inevitable en nuestro sistema. A partir de esta realidad, se debe evaluar si la introducción del algoritmo contribuye a mejorar el panorama actual.

## 6.2 Transparencia, interpretabilidad y explicabilidad

La necesidad de explicabilidad en los algoritmos de evaluación de riesgo utilizados en el contexto de la prisión provisional obedece a múltiples razones. En primer lugar, como se menciona anteriormente, sirve para cultivar y mantener la confianza en el sistema. Cuando las personas pueden entender el razonamiento detrás de las decisiones automatizadas, la confianza en el proceso automatizado aumenta. En segundo lugar, es esencial para identificar y corregir los sesgos sistémicos que pueden estar presentes tanto en los datos como en el propio algoritmo. Entender cómo se ponderan las variables permite ajustar cualquier desigualdad que surja. En tercer lugar, los investigadores utilizan estos algoritmos para aprender, descubrir, establecer causalidades y realizar investigaciones científicas, y para poder aprovechar su potencial en estas áreas es necesario comprender su funcionamiento (Lipton, 2017).

La explicabilidad de un modelo se establece cuando existe un mecanismo capaz de ofrecer, al menos, información parcial sobre el proceso de predicción. Esto habitualmente se traduce en la capacidad de identificar los factores con mayor peso en la decisión, o en poder determinar qué cambios en los factores concretos podrían alterar el resultado de la predicción. Los modelos intrínsecamente interpretables, como los de regresión lineal, suelen ser explicables. Esto se debe a la facilidad con la que se pueden identificar las variables utilizadas y su importancia relativa al examinar los coeficientes del modelo (Kästner, 2021).

En contraste, muchos modelos de aprendizaje automático o *machine learning* funcionan como sistemas opacos, ofreciendo predicciones sin explicaciones comprensibles para los humanos, comúnmente descritos como "cajas negras". La opacidad de este tipo de modelos ha tenido ya consecuencias graves. Por ejemplo, se han tomado decisiones erróneas sobre fianzas que han liberado a individuos potencialmente peligrosos, o se ha categorizado de forma incorrecta el aire altamente contaminado como seguro para respirar. Al no poder comprender de dónde sale la predicción, es imposible evitar errores. En respuesta a esto, se ha intensificado el desarrollo del *explainable Machine Learning*, donde se genera un segundo modelo *post hoc* que sirve para explicar el funcionamiento del modelo inicial de caja negra. No obstante, las explicaciones ofrecidas por esta versión simplificada del algoritmo a menudo no son fiables y pueden llevar a malinterpretaciones (Fernandez-Quilez, 2023).

La selección del algoritmo específico que se utilizará dependerá de cuál se adapte mejor a los datos, lo que no es posible determinar hasta que comience el desarrollo y se prueben los diferentes modelos. No obstante, como se ha explicado ya en el apartado 4.1, de elección del tipo de algoritmo, en términos de transparencia, un modelo de regresión lineal sería preferible, ya que resuelve eficazmente las preocupaciones sobre explicabilidad.

Sin perjuicio de la preferencia por el modelo de regresión lineal, es crucial garantizar dos valores fundamentales en términos de transparencia con independencia del modelo que finalmente se demuestre más adecuado. En primer lugar, es esencial que el público esté informado sobre el uso de un algoritmo en las decisiones de libertad condicional, garantizando que las personas afectadas comprendan que sus casos son evaluados, en parte, mediante esta tecnología. En segundo lugar, debe facilitarse el acceso a la información sobre cuáles fueron las variables más influyentes en el proceso de toma de decisiones, motivando así la decisión. Estas medidas asegurarán que el proceso no solo sea efectivo, sino también transparente y justo.

### **6.3 Sesgos y discriminación**

Una de las grandes preocupaciones con este tipo de algoritmos reside en su potencial tanto para perpetuar los sesgos ya existentes como para introducir nuevas formas de discriminación. Según Goel et al. (2018), el problema surge de los datos históricos que se utilizan para entrenar el algoritmo, que traen consigo dos principales complicaciones: el error de medición y el sesgo en la muestra.

El error de medición se refiere a las inexactitudes que ocurren cuando los datos utilizados para desarrollar un algoritmo no reflejan correctamente la realidad o están inherentemente sesgados. A título de ejemplo, los datos podrían incluir información sobre arrestos previos, lo cual es problemático si esos registros de arrestos reflejan un sesgo contra ciertos grupos. Se ha demostrado que las minorías raciales, especialmente las personas negras, tienen más probabilidades de ser detenidas y arrestadas por la policía en comparación con las personas blancas por el mismo comportamiento (Eckhouse et al., 2018). Por lo tanto, si un algoritmo se entrena con estos datos, probablemente heredará estos sesgos. No es que el algoritmo sea discriminatorio, sino que se ha entrenado con unos datos que reflejan una discriminación institucional. Esto puede suceder incluso si el algoritmo no incluye explícitamente la raza como una variable, puesto que el algoritmo

podría usar otros puntos de datos que actúan como sustitutivos de la raza, como pueden ser códigos postales o historiales delictivos (Kleinberg et al., 2018). Estos factores sustitutivos, aunque no estén vinculados directamente con la variable raza, pueden seguir reflejando los sesgos inherentes a los datos. Por lo tanto, el algoritmo podría aprender y reforzar las desigualdades existentes, perpetuando así la discriminación en sus resultados.

Por otro lado, el sesgo en la muestra ocurre cuando los datos utilizados para entrenar un algoritmo no representan adecuadamente a la población a la que se aplicará el algoritmo (Goel et al., 2018). En España, como en muchos otros países, la población incluye diversas comunidades con diferencias en etnicidad, nivel socioeconómico, educación y otros factores demográficos. Si un algoritmo de evaluación de riesgos, por ejemplo, se desarrolla utilizando datos provenientes, en su mayoría, de una comunidad específica o de ciertas áreas geográficas, podría no ser representativo de la población general y proporcionar resultados sesgados al ser aplicado a ciertas regiones. Los desarrolladores a menudo tratan de crear algoritmos que sean lo más portables y abstractas posibles. Esto es lo que Selbst et al. (2019) consideran como la trampa de portabilidad, pues esta búsqueda de abstracción puede ignorar las sutilezas y particularidades de cada contexto social.

No obstante, no son todo críticas por parte de la literatura. Muchos expertos coinciden en que, en lugar de simplemente señalar los defectos de los algoritmos, es crucial comparar estos defectos con los de los humanos que estos algoritmos están destinados a reemplazar. En este contexto, Miller (2018) sostiene que los algoritmos suelen ser menos sesgados y más precisos que los humanos a los que sustituyen. Además, poseen la capacidad de identificar, mitigar e incluso eliminar los sesgos inherentes, siempre y cuando se adhieran a los principios de transparencia y explicabilidad (Kleinberg et al., 2018). Por lo tanto, a diferencia de los procesos de decisión humanos, que tienden a ser más opacos, el empleo de algoritmos puede facilitar la auditoría y lucha contra la discriminación en nuestros sistemas.

## **7. CONCLUSIONES Y RECOMENDACIONES**

### **7.1 Recapitulación de hallazgos**

La investigación ha demostrado una clara necesidad de aumentar la objetividad y consistencia en las decisiones judiciales relativas a la prisión provisional,

tradicionalmente basadas en evaluaciones subjetivas sujetas a sesgos personales y variabilidad entre jueces. Los algoritmos predictivos, utilizados adecuadamente, tienen el potencial de mejorar significativamente la precisión de las predicciones sobre los riesgos asociados con los acusados, ofreciendo análisis basados en datos objetivos. Sin embargo, aunque se verifica la posibilidad de desarrollar el algoritmo predictor de riesgo dado los datos suficientes, se reconoce que no todas las variables relevantes son cuantificables, lo que podría limitar la precisión del algoritmo. Expertos tendrán que testar y ajustar el algoritmo considerando estas limitaciones.

Se ha analizado cómo la integración de algoritmos predictivos podría transformar el panorama judicial en decisiones sobre prisión provisional, proponiendo directrices críticas para su implementación. Estas directrices abarcan desde la interpretación adecuada de los resultados del algoritmo hasta la limitación de su alcance, asegurando que el algoritmo actúe como una herramienta de apoyo y no como un sustituto de la decisión judicial. Además, se enfatiza la necesidad de cambios en la Ley de Enjuiciamiento Criminal para facilitar una integración efectiva de estas tecnologías, garantizando que las decisiones puedan ser objeto de recurso para proteger los derechos de los acusados.

A pesar de su potencial, la implementación de algoritmos predictivos enfrenta desafíos significativos, especialmente en términos de ética, transparencia y aceptación por parte de la comunidad judicial y la sociedad. Estas preocupaciones son críticas y deben abordarse cuidadosamente para asegurar que la implementación de tecnologías predictivas en el sistema judicial no solo sea efectiva, sino también justa y equitativa.

## **7.2 Implicaciones prácticas**

- La adopción de algoritmos predictivos puede contribuir a estandarizar las decisiones sobre prisión provisional, reduciendo la variabilidad entre jueces y asegurando una mayor consistencia en todo el sistema judicial. Esto llevaría [a](#) un proceso más equitativo donde casos similares reciben tratamientos similares.
- Al basarse en datos y hechos concretos, los algoritmos pueden ayudar a mitigar los sesgos personales inherentes a las decisiones humanas, aunque es crucial asegurar que los propios algoritmos estén libres de sesgos, lo que requeriría una constante revisión y ajuste de los modelos predictivos utilizados.

- El uso de algoritmos puede acelerar el proceso de toma de decisiones, permitiendo que las audiencias se enfoquen en aspectos más sustantivos del caso. Esto podría traducirse en una reducción de los tiempos de detención previa al juicio y, potencialmente, en una disminución de la sobrepoblación carcelaria.
- Implementar algoritmos predictivos requerirá una inversión significativa en capacitación para jueces y otros operadores judiciales, así como la adecuación de infraestructura tecnológica. Será esencial desarrollar programas de formación que preparen adecuadamente a estos profesionales para utilizar estas herramientas de manera efectiva y ética.
- Para integrar plenamente estos algoritmos en el sistema judicial, será necesario modificar la Ley de Enjuiciamiento Criminal y otras regulaciones pertinentes para definir claramente el papel de la tecnología predictiva en las decisiones judiciales. Esto debería incluir normas sobre la transparencia, la posibilidad de apelar las decisiones y medidas para la protección de los datos personales.
- Es crucial garantizar que el desarrollo de tecnologías judiciales no quede en manos de multinacionales privadas. La justicia debe mantener su independencia sin depender de intereses comerciales que podrían comprometer la equidad y la integridad del proceso judicial. Esto implica establecer regulaciones estrictas sobre la propiedad y el control de las tecnologías empleadas.
- Una vez implementados, los sistemas algorítmicos deberán ser objeto de una vigilancia constante para evaluar su efectividad y justicia. Esto incluye establecer mecanismos independientes que puedan realizar auditorías regulares y proporcionar informes sobre el desempeño y los efectos de estos sistemas.

### **7.3 Limitaciones del estudio y futuras líneas de investigación**

A pesar de los importantes hallazgos, este estudio presenta ciertas limitaciones que deben considerarse al interpretar los resultados. En primer lugar, si bien se ha explicado el proceso de creación del algoritmo, no se ha elaborado un prototipo del mismo. En segundo lugar, el estudio asume la disponibilidad y calidad de los datos judiciales históricos. No obstante, en caso de no ser así, no sería posible desarrollar el algoritmo. Por último, cambios futuros en las leyes que rigen la prisión provisional y el uso de herramientas tecnológicas en el sistema judicial podrían alterar la aplicabilidad de los resultados del estudio y la implementación del algoritmo en la práctica judicial.

Para futuras líneas de investigación, se sugiere en primer lugar que expertos técnicos desarrollen un prototipo de modelo predictivo. Este prototipo permitiría probar la efectividad del algoritmo en la toma de decisiones y ayudaría a identificar cuáles son las variables más relevantes para la predicción del riesgo en base a los datos. Otra línea de investigación interesante sería explorar las implicaciones de la implementación de algoritmos en derechos fundamentales, como el derecho a la presunción de inocencia y el derecho a la tutela judicial efectiva, para asegurar que el algoritmo no suponga la vulneración de los mismos. Finalmente, se sugiere estudiar la necesidad de posibles reformas legislativas para regular el uso de algoritmos predictivos en el ámbito judicial, estableciendo garantías y controles que eviten posibles abusos y aseguren la transparencia y rendición de cuentas.

## 8. REFERENCIAS

### 8.1 Legislación

- Constitución Española (1978). Boletín Oficial del Estado núm. 311, de 29 de diciembre de 1978.
- Ley de Enjuiciamiento Criminal (1882). Real Decreto de 14 de septiembre de 1882, por el que se aprueba la Ley de Enjuiciamiento Criminal. Boletín Oficial del Estado núm. 260, de 17 de septiembre de 1882.

### 8.2 Jurisprudencia

- Sentencia del Tribunal Constitucional núm. 41/1982, de 2 de julio
- Sentencia del Tribunal Constitucional núm. 127/1984, de 26 de diciembre
- Sentencia del Tribunal Constitucional núm. 32/1987, de 12 de marzo
- Sentencia del Tribunal Constitucional, núm. 40/1987, de 3 de abril
- Sentencia del Tribunal Constitucional, núm. 128/1995, de 26 de julio
- Sentencia del Tribunal Constitucional núm. 210/2013, de 16 de diciembre

### 8.3 Bibliografía

Alemán Aróstegui, L. (2023). El uso de RISCANVI en la toma de decisiones penitenciarias. *Estudios Penales Y Criminológicos*, 44(Ext.), 1–43.

<https://doi.org/10.15304/epc.44.8884>

Álvarez Buján, M. V. (2023). Inteligencia artificial y medidas cautelares en el proceso penal: Tutela judicial efectiva y autodeterminación informativa en potencial riesgo. *Revista Española de Derecho Constitucional*, 43(127), 177–207.

<https://dialnet.unirioja.es/servlet/articulo?codigo=8906903>

Asociación Pro Derechos Humanos de España. (2015). *La práctica de la prisión provisional en España*. <https://n9.cl/asociacionproderechoshumanos>

Asquerino Lamparero, M. J. (2022). Algoritmo y discriminación. *Noticias CIELO*, 4, 7.

<https://dialnet.unirioja.es/servlet/articulo?codigo=8430567>

- Barona Vilar, S. (2023). Dataización de la justicia (Algoritmos, Inteligencia Artificial y Justicia, ¿el comienzo de una gran amistad?). *Revista Boliviana de Derecho*, 36, 14–45. <https://dialnet.unirioja.es/servlet/articulo?codigo=9043836>
- Bayman, E. O., & Dexter, F. (2021). Multicollinearity in logistic regression models. *Anesthesia & Analgesia*, 133(2), 362–365. <https://doi.org/10.1213/ane.00000000000005593>
- Bodur, E. K., & Atsa'am, D. D. (2019). Filter variable selection algorithm using risk ratios for dimensionality reduction of healthcare data for classification. *Processes*, 7(4), 222. <https://doi.org/10.3390/pr7040222>
- Calaza-López, S., & Llorente-Sánchez-Arjona, M. (2022). Inteligencia artificial legal y administración de justicia. *Thomson Reuters Aranzadi*. <https://dialnet.unirioja.es/servlet/libro?codigo=870011>
- Cano de Amo, J. R., Herrera Triguero, F., & Lozano Márquez, M. (2005). *Extracción de modelos predictivos e interpretables en conjuntos de datos de tamaño grande mediante la selección de conjuntos de entrenamiento*. 145–152. Taller Nacional de minería de datos y Aprendizaje. <https://dialnet.unirioja.es/servlet/articulo?codigo=7975177>
- Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, 15(2), 126–139. <https://doi.org/10.1017/s1744552319000077>
- Christin, A., Rosenblat, A., & Boyd, D. (2015). *Courts and predictive algorithms*. Data & Civil Rights. [https://www.law.nyu.edu/sites/default/files/upload\\_documents/Angele%20Christin.pdf](https://www.law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf)
- De Arteaga, M., Fogliato, R., & Chouldechova, A. (2020). A case for humans-in-the-loop: Decisions in the presence of erroneous algorithmic scores. *Proceedings of*

*the 2020 CHI Conference on Human Factors in Computing Systems.*

<https://doi.org/10.1145/3313831.3376638>

DeMichele, M., Baumgartner, P., Barrick, K., Comfort, M., Scaggs, S., & Misra, S. (2018). What Do Criminal Justice Professionals Think About Risk Assessment at Pretrial? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3168490>

DeMichele, M., Comfort, M., Barrick, K., & Baumgartner, P. (2021). The intuitive-override model: Nudging judges toward pretrial risk assessment instruments. *Federal Probation*, 85, 22. <https://heinonline.org/HOL/LandingPage?handle=hein.journals/fedpro85&div=18&id=&page=>

Díaz Torrejón, P. (2013). *El Fiscal ante el servicio de Guardia. Medidas cautelares personales: detención, prisión provisional y Habeas Corpus*. Ministerio Fiscal. <https://www.fiscal.es/documents/20142/277968/Ponencia+Pedro+Diaz+Torrej%20n.pdf/bb32de8f-a220-ae6e-540d-ae41c7992cd5?t=1562234506959>

Eckhouse, L., Lum, K., Conti-Cook, C., & Ciccolini, J. (2018). Layers of bias: A unified approach for understanding problems with risk assessment. *Criminal Justice and Behavior*, 46(2), 185–209. <https://doi.org/10.1177/0093854818811379>

Elyounes, D. A. (2020). Bail or jail? Judicial versus algorithmic decision-making in the Pretrial System. *Science and Technology Law Review*, 21(2), 376–445. <https://doi.org/10.7916/stlr.v21i2.6838>

Fernandez-Quilez, A. (2023). Deep learning in radiology: Ethics of data and on the value of algorithm transparency, interpretability and explainability. *AI Ethics*, 3, 257–265. <https://doi.org/10.1007/s43681-022-00161-9>

Gabbiadini, A., Durante, F., Baldissarri, C., & Andrighetto, L. (2024). Artificial intelligence in the eyes of society: Assessing social risk and social value perception in

a novel classification. *Human Behavior and Emerging Technologies*, 2024.

<https://doi.org/10.1155/2024/7008056>

Galelli, S., Humphrey, G. B., Maier, H. R., Castelletti, A., Dandy, G. C., & Gibbs, M. S.

(2014). An evaluation framework for input variable selection algorithms for environmental data-driven models. *Environmental Modelling and Software*, 62,

33–51. <https://doi.org/10.1016/j.envsoft.2014.08.015>

Gallego Sánchez, G. (2006). *Acerca de la aplicación de la prisión provisional, a través de sus características*. 4. *Revista de Jurisprudencia El Derecho*.

García Martínez, M. J. (2015, March 20). *La prisión provisional ¿excepción o regla general?* Garrigues. [https://www.garrigues.com/es\\_ES/noticia/la-prision-provisio-](https://www.garrigues.com/es_ES/noticia/la-prision-provisio-nal-excepcion-o-regla-general)

[nal-excepcion-o-regla-general](https://www.garrigues.com/es_ES/noticia/la-prision-provisio-nal-excepcion-o-regla-general)

Hamilton, M. (2022). Algorithmic risk assessment: A progressive policy in pretrial re-

lease. *Idaho Law Review*, 57(3). [https://digitalcommons.law.uidaho.edu/idaho-](https://digitalcommons.law.uidaho.edu/idaho-law-review/vol57/iss3/8/)

[law-review/vol57/iss3/8/](https://digitalcommons.law.uidaho.edu/idaho-law-review/vol57/iss3/8/)

Hernández Gómez, I. (2011). Prisión provisional y garantías. *Revista de Ciencias Jurídicas*,

p. 37-82. <http://hdl.handle.net/10553/12228>

Hernández, C., & Rodríguez Rodríguez, J. E. (2008). Procesamiento de datos estructu-

rados. *Revista Vínculos*, 4(2), 27–48. <https://doi.org/10.14483/2322939X.4123>

Hobson, Z., Yesberg, J. A., Bradford, B., & Jackson, J. (2021). Artificial fairness? Trust

in algorithmic police decision-making. *Journal of Experimental Criminology*,

19(1), 165–189. <https://doi.org/10.1007/s11292-021-09484-9>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical

learning. In *Springer Texts in Statistics* (Vol. 103). [https://doi.org/10.1007/978-](https://doi.org/10.1007/978-1-4614-7138-7)

[1-4614-7138-7](https://doi.org/10.1007/978-1-4614-7138-7)

- Kästner, C. (2021, July 8). *Interpretability and explainability*. Medium.  
<https://ckaestne.medium.com/interpretability-and-explainability-a80131467856>
- Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis*, 10, 113–174.  
<https://doi.org/10.1093/jla/laz001>
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 205395171875668. <https://doi.org/10.1177/2053951718756684>
- Lipton, Z. C. (2017). The mythos of model interpretability. *Arxiv*, 3. Cornell University.  
<https://doi.org/10.48550/arxiv.1606.03490>
- López López, E. (2003). La nueva prisión provisional. *Revista Galega de Administración Pública*, 1(34). <https://egap.xunta.gal/revistas/REGAP/article/view/3100>
- Martinez, B. P., Petersen, N., & Omori, M. (2019). Time, money, and punishment: Institutional racial-ethnic inequalities in pretrial detention and case outcomes. *Cardozo J. Conflict Resol.*, 21. <https://doi.org/10.1177/0011128719881600>
- McKay, C. (2019). Predicting risk in criminal procedure: Actuarial tools, algorithms, AI and judicial decision-making. *Current Issues in Criminal Justice*, 32(1), 1–18.  
<https://doi.org/10.1080/10345329.2019.1658694>
- Miller, A. P. (2018). Want less-biased decisions? Use algorithms. *Harvard Business Review*. <https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms>
- Mok, L., Nanda, S., & Anderson, A. (2023). People perceive algorithmic assessments as less fair and trustworthy than identical human assessments. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), 1–26.  
<https://doi.org/10.1145/3610100>

- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. (6th ed.). John Wiley.
- Moore, C., Ferguson, E., & Guerin, P. (2023). Pretrial risk assessment on the ground: Algorithms, judgments, meaning, and policy. *MIT Case Studies in Social and Ethical Responsibilities of Computing, Summer 2023*.  
<https://doi.org/10.21428/2c646de5.b016a7b3>
- Morillas Cueva, L. (2016). Reflexiones sobre la prisión preventiva. *Digitum*, 34(1).  
<https://digitum.um.es/digitum/handle/10201/48564>
- Muñoz-Cuesta, J. (2019). La prisión provisional como medida excepcional. *Actualidad Jurídica Aranzadi*, 948. <https://dialnet.unirioja.es/servlet/articulo?codigo=6793369>
- Neira Pena, A. M. (2021). Inteligencia artificial y tutela cautelar. Especial referencia a la prisión provisional. *Revista Brasileira de Direito Processual Penal*, 7(3).  
<https://doi.org/10.22197/rbdpp.v7i3.618>
- Noriega Estrada, G. A. (2023). Prisión preventiva. *Revista Ciencia Multidisciplinaria CUNORI*, 7(1), 131–140. <https://doi.org/10.36314/cunori.v7i1.213>
- Notaro, L. (2023). “Algoritmos predictivos” y justicia penal desde una perspectiva italiana y europea. In *Derecho Penal, Inteligencia Artificial y Neurociencias* (pp. 191–211). Roma Tre-Press. <https://dialnet.unirioja.es/servlet/articulo?codigo=8825012>
- Ordóñez, L. (2007). El desarrollo tecnológico en la historia. *Areté*, 19(2), 187–210.  
[http://www.scielo.org.pe/scielo.php?script=sci\\_arttext&pid=S1016-913X2007000200001](http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1016-913X2007000200001)

- Rondón Galvis, S. (2022). Procesos heurísticos y sesgos cognitivos en el razonamiento del Juez. *Tecnológico de Antioquia, Institución Universitaria*.  
<https://dspace.tdea.edu.co/handle/tdea/3185>
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- Sánchez, N., Sobral, J., & Seijo, D. (2017). El error judicial en el uso de la prisión preventiva: Personas en prisión que nunca llegan a ser condenadas. *Revista Iberoamericana de Psicología Y Salud*, 8(1), 36.  
<https://doi.org/10.23923/j.rips.2017.08.004>
- Sangüesa, R. (2005). Clasificación: árboles de decisión. In *Data mining*. Editorial UOC.  
<https://dialnet.unirioja.es/servlet/libro?codigo=730090>
- Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham L. Rev.*, 87(3). <https://doi.org/10.2139/ssrn.3126971>
- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery.  
<https://doi.org/10.1145/3287560.3287598>
- Shrestha, A., & Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access*, 7, 53040–53065. <https://doi.org/10.1109/access.2019.2912200>
- Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. Dik. F. (2001). Internal validation of predictive models. *Journal of Clinical Epidemiology*, 54(8), 774–781.  
[https://doi.org/10.1016/s0895-4356\(01\)00341-9](https://doi.org/10.1016/s0895-4356(01)00341-9)

- Turner, L. (2018, June 27). *Machine learning: A primer*. Medium. <https://medium.com/@lizziedotdev/lets-talk-about-machine-learning-ddca914e9dd1>
- Turturro-Pérez de los Cobos, S. (2021). Libertad personal versus prisión provisional: Un derecho minusvalorado por razones anacrónicas. *Revista de Las Cortes Generales*, 111, 313–344. <https://doi.org/10.33426/rcg/2021/111/1615>
- Velupillai, S., Hadlaczky, G., Baca-Garcia, E., Gorrell, G. M., Werbeloff, N., Nguyen, D., Patel, R., Leightley, D., Downs, J., Hotopf, M., & Dutta, R. (2019). Risk assessment tools and data-driven approaches for predicting and preventing suicidal behavior. *Frontiers in Psychiatry*, 10. <https://doi.org/10.3389/fpsy.2019.00036>
- Viljoen, J. L., Jonnson, M. R., Cochrane, D. M., Vargen, L. M., & Vincent, G. M. (2019). Impact of risk assessment instruments on rates of pretrial detention, postconviction placements, and release: A systematic review and meta-analysis. *Law and Human Behavior*, 43(5), 397–420. <https://doi.org/10.1037/lhb0000344>
- Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology*, 62(5), 739–752. <https://doi.org/10.1037/0022-3514.62.5.739>
- Wilstrup, C., & Kasak, J. (2021). Symbolic regression outperforms other models for small data sets. *ArXiv*. Cornell University. <https://doi.org/10.48550/arxiv.2103.15147>