



Flavia Luxemburgo Poy Barrio
presenta el resumen del trabajo

**Por qué no se puede automatizar la equidad:
cerrando la brecha entre la legislación
de no discriminación de la UE y la IA**
Wachter, S., Mittelstadt, B., & Russell, C
Computer Law Security Review, 41, 105567. (2021).

3 de mayo de 2024

SERIE DE RESÚMENES EN ESPAÑOL

Palabras clave: Algoritmo, Aprendizaje automático, Unión Europea, No discriminación, Justicia, Paridad demográfica, Inteligencia artificial

Introducción

La idea básica parte del Desafío de Automatizar la Equidad. El artículo aborda el desafío fundamental de automatizar la equidad en el contexto de la inteligencia artificial (IA) y el derecho antidiscriminatorio de la Unión Europea (UE). Destaca la complejidad de traducir los principios legales de no discriminación en algoritmos automatizados, dada la naturaleza multifacética y contextual de la equidad en el derecho. Se argumenta que la equidad no puede reducirse a una simple fórmula o regla algorítmica, ya que implica consideraciones éticas, sociales y políticas que son difíciles de codificar en sistemas automatizados.

1. Metodología

La metodología de investigación de carácter descriptivo trata de analizar las condiciones específicas que se dan en un contexto de discriminación automatizada con el uso de la IA, donde parte de una contextualización rigurosa de la conceptualización de la “discriminación” como acto adverso cometido contra un individuo o grupo legalmente protegido y define sus diferencias entre la discriminación directa e indirecta, especialmente advirtiendo dificultad por esta última al requerir una comparación a nivel grupal.

El artículo tiene asimismo un carácter explicativo en el que se incluye una propuesta de un estándar comparativo para evaluar la disparidad y la discriminación, destaca la importancia de refutar la discriminación aparente mediante evidencia sólida y análisis rigurosos y explica la necesidad de establecer procedimientos de evaluación consistentes para abordar la discriminación automatizada que parte de la complejidad de desarrollar tales procedimientos debido a la naturaleza dinámica y evolutiva de la tecnología.

Es relevante el uso de la evidencia estadística para evaluar casos de discriminación pues arroja luz sobre los requerimientos y situación actual de los procedimientos de evaluación consistentes en lo que a la discriminación automatizada requiere.

2. Resultados

Con la metodología mencionada los propios autores reconocen hacer tres contribuciones puntuales al respecto de esta cuestión. Por un lado, revisan los requisitos probatorios para presentar como debido a la naturaleza dispar de lo algoritmo y lo humano, los requisitos legales son muy contextuales e interpretables. Por otro lado, muestra cómo la protección legal ofrecida por la ley de no discriminación es desafiada, cuestión que se evidencia por la ausencia de un marco homogéneo en la ley para realizar pruebas de discriminación en sistema de IA. Lo anterior le deriva en un examen sobre cómo el trabajo existente sobre la equidad en el aprendizaje automático se alinea con los procedimientos y las dificultades que se materializan en la detección y prueba de una discriminación que se define como “poco intuitiva, sutil e intangible”.

Así, en primer lugar analiza la intersección entre el Derecho de la UE y la IA. El artículo examina la intersección entre el derecho antidiscriminatorio de la UE y el desarrollo de sistemas de IA. Se destaca la importancia de garantizar que los algoritmos y sistemas de IA cumplan con los estándares legales de no discriminación establecidos en la legislación de la UE, como el Reglamento General de Protección de Datos (GDPR) y la Directiva de Igualdad de Trato. Se argumenta que la aplicación efectiva de estos principios en el contexto de la IA requiere un enfoque interdisciplinario que integre la expertise legal, ética y tec-

nológica. En segundo lugar, es relevante la serie de desafíos y recomendaciones que se plantean en la Regulación de la IA en la UE. Wachter et al. identifican una serie de desafíos y ofrecen recomendaciones para abordar la regulación de la IA en el marco de la UE. Se señala la necesidad de desarrollar marcos regulatorios que promuevan la equidad y la transparencia en el desarrollo y uso de sistemas de IA, incluida la evaluación y mitigación de sesgos algorítmicos. Además, se aboga por una mayor colaboración entre las partes interesadas, incluidos legisladores, expertos en IA y defensores de los derechos humanos, para desarrollar normativas que protejan los derechos fundamentales en la era de la IA.

Finalmente, se examina cómo el trabajo existente sobre la equidad en el aprendizaje automático se alinea con los procedimientos. Después de presentar la evidencia de una desventaja particular y la composición de determinadas poblaciones desfavorecidas y comparadas, concluye que efectivamente la jurisprudencia es inconsistente para medir la igualdad jurídica entre grupos. El Tribunal de Justicia Europeo ha propuesto un concepto de discriminación *prima facie*, pero aún no se ha traducido en procedimientos de evaluación estándar para la discriminación automatizada. Los autores proponen la “disparidad demográfica condicional” (CDD) como medida estadística de referencia estándar que se alinea con el “estándar de oro” de la Corte. Establecer un conjunto estándar de evidencia estadística para casos de discriminación automatizada puede ayudar a garantizar procedimientos consistentes para la evaluación, pero no la interpretación judicial, de casos que involucran IA y sistemas automatizados. A través de esta propuesta de regularidad procesal en la identificación y evaluación de discriminación, se aclara cómo incorporar consideraciones de equidad en los sistemas automatizados permitiendo el enfoque contextual de las interrelaciones judiciales.

3. Discusión/Conclusiones

Una cuestión fundamentalmente contextual y política tradicionalmente respondida por los tribunales nacionales y el TJUE, en términos de diseño de sistemas de IA imparciales o no discriminatorios, es que efectivamente ningún sistema puede estar perfectamente calibrado y preparado para el futuro. Para sistemas que trabajan con problemas reales, o que utilizan datos del mundo real, se debe esperar, como mínimo, una pequeña cantidad de sesgo en una dirección u otra. El análisis de la jurisprudencia europea y de los tribunales nacionales en la UE reveló que el uso de evidencia estadística para demostrar *prima facie* que la discriminación es rara e inconsistente, carece de establecer umbrales bien definidos y estandarizados para la paridad que se mantendría en diferentes casos.

Asimismo, se discute cómo las víctimas potenciales

enfrentan desafíos similares a los que se presentan en la actualidad. Demostrar la discriminación algorítmica puede ser particularmente difícil porque no necesariamente se experimentará o “sentirá” de una manera comparable a la discriminación humana. La identidad se define por la intuición judicial, no por las estadísticas. Los tribunales no proporcionan un enfoque consistente y coherente para evaluar discriminación *prima facie*. Como resultado, los desarrolladores de sistemas, controladores, reguladores y usuarios carecen de leyes claras y consistentes, lo que demuestra así la necesidad de mecanismos de gobernanza para detectar, remediar y prevenir la discriminación automatizada.

Por otra parte, se concluye que las pruebas actuales utilizadas por el poder judicial podrían incentivar lo que llamamos un enfoque de “divide y vencerás”. Con este enfoque sería posible defender el comportamiento discriminatorio de un sistema dividiendo al grupo desfavorecido en subgrupos y demostrando que cada subgrupo es tan pequeño que ninguno de ellos supera el umbral legalmente requerido de desventaja significativa y desproporcionada. Se observa así mismo que en la jurisprudencia no se han establecido tampoco umbrales de disparidad ilegal, aspecto que va de la mano con la conclusión central de que el uso progresivo de algoritmos está alterando los procedimientos tradicionales y recursos contra la prevención e investigación de discriminaciones.

Todo lo anterior permite argumentar tres cuestiones clave: (1) equidad es contextual y no puede (y posiblemente no debería) ser automatizada de una manera que respete plenamente la igualdad contextual”. (2) La IA como discriminador está alterando formas conocidas de detección, investigación y prevención de sesgos y, por lo tanto, los procedimientos (judiciales) deben ser más coherentes, estáticos y explícitos en sus definiciones. (3) Las estadísticas resumidas que informan sobre la disparidad demográfica condicional son una primera paso para cerrar la brecha de rendición de cuentas entre la IA y los países no pertenecientes a la UE o bajo diferentes jurisdicción y es por eso que la implementación de un sistema como el propuesto, el uso de CDD como medida estadística de referencia para casos de discriminación automatizada es fundamental para abordar estas cuestiones.

4. Valoración del documento original

El estudio plantea una cuestión central que por su misma naturaleza detalla un tema cada vez más relevante en torno a lo que supone la discriminación automatizada cuyo origen radica en el uso de algoritmos de Inteligencia Artificial. La revisión estadística y su propuesta de un instrumento que permita mejor a los tribunales afrontar esta cuestión es relevante por la multiplicidad de afectados y de elementos que intervienen cuando hablamos de discriminación automati-

zada, independientemente de su grado de afectación a las víctimas. Señala así mismo elementos relevantes que llegarán a la brevedad a tribunales nacionales e

internacionales, lo que permite identificar claras variables que tendrán que considerarse en términos de nueva legislación y actualización de la hoy vigente.