

GMV

ÉTICA E INTELIGENCIA ARTIFICIAL



© GMV, 2019 Propiedad de GMV

Todos los derechos reservados

GMV-LIMITED

El presente documento está clasificado como "GMV-CONFIDENTIAL". Esta clasificación habilita a su receptor al uso de la información contenida en el documento para los fines para los que la empresa la ha facilitado o, en su caso, a lo acordado contractualmente en relación al intercambio de información entre las partes, y ello sin perjuicio del cumplimiento de la normativa sobre propiedad intelectual y sobre protección de datos de carácter personal.

gmv[®]
INNOVATING SOLUTIONS

ÉTICA & IA


**AFRONTANDO
UN NUEVO
ESCENARIO**

SIEMPRE NOS HAN CLASIFICADO

RECREANDO SESGOS HISTÓRICOS



UN ALGORITMO TE PUEDE CONDICIONAR TU FUTURO



BRISHA BORDEN

Prior Offenses
4 juvenile
misdemeanors

Subsequent Offenses
None

HIGH RISK

8



VERNON PRATER

Prior Offenses
2 armed robberies, 1
attempted armed
robbery

Subsequent Offenses
1 grand theft

LOW RISK

3



INQUIETUD EN LOS GOBIERNOS



the **WHITE HOUSE**
PRESIDENT BARACK OBAMA

BRIEFING ROOM

ISSUES

THE ADMINISTRATION

1600 PENN



HOME · BLOG

Big Risks, Big Opportunities: the Intersection of Big Data and

GDPR and AI

According to the [Brookings Institution](#), a US think tank:

The GDPR being implemented in Europe place severe restrictions on the use of artificial intelligence and machine learning. According to published guidelines, “Regulations prohibit any automated decision that ‘significantly affects’ EU citizens. This includes techniques that evaluates a person’s ‘performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements.” In addition, these new rules give citizens the right to review how digital services made specific algorithmic choices affecting people.

a
t
on

ÉTICA & IA MACHINE LEARNING & DISCRIMINACIÓN

ML PARA LA TOMA DE DECISIONES

Desde las solicitudes de crédito hasta las citas on-line, los modelos de **machine learning** están **automatizando** nuestra **toma de decisiones** día a día.



CRECIENTE IMPACTO EN LA SOCIEDAD

Cada vez surgen más **inquietudes** sobre la pérdida de **transparencia**, **responsabilidad** y **equidad** de los algoritmos que toman decisiones.



AI & DISCRIMINACIÓN

Necesitamos intensificar y buscar nuevas formas para **mitigar la discriminación** emergente en nuestros modelos.



Photo by Ben Torres (Bloomberg)

Big Bad Data May Be Triggering
Discrimination



Bernard Parker, left, was rated high risk; Dylan Puget was rated low risk. (Josh Ritchie)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

AI & DISCRIMINACIÓN

Tenemos que estar seguros que nuestras predicciones **no perjudiquen injustamente** a la gente con ciertas **características sensibles** (e.j. género, etnia, etc.).

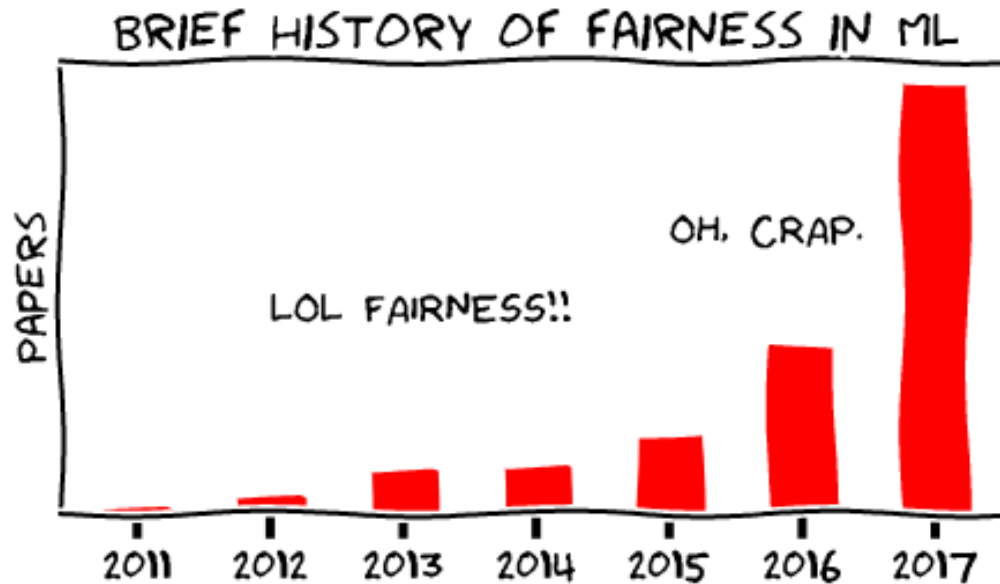


ÉTICA & IA
REACCION
DE LA
COMUNIDAD
CIENTÍFICA

ÁREA ACTIVA DE INVESTIGACIÓN

Cómo realizar modelos predictivos equitativos es un **ámbito de investigación activa**.

Finalmente nos lo **estamos tomando en serio**.



ALGUNAS APROXIMACIONES

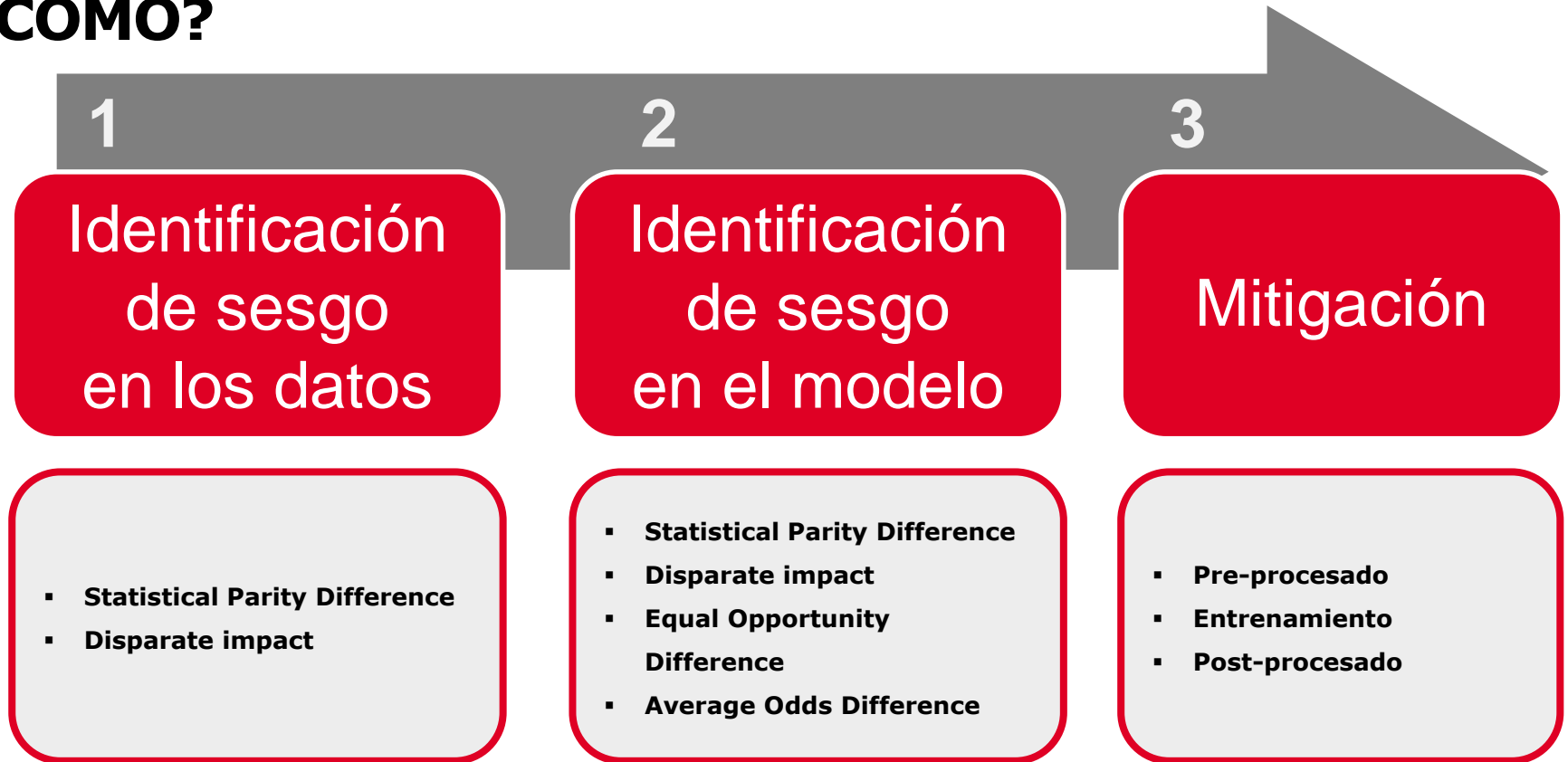
- Foco en la **interpretabilidad** y **transparencia**: Permitir interrogar a los modelos complejos.
- Hacer modelos más **robustos** y **justos** en sus **predicciones**: Modificar la **optimización de las funciones objetivos** y añadir **restricciones**.



ÉTICA & IA ¿CÓMO CONSEGUIR EQUIDAD?

CONSEGUIR ALGORITMOS EQUITATIVOS

¿CÓMO?



ML EQUITATIVO CONCLUSIONES

CONCLUSIONES

- Construir modelos predictivos imparciales **no es tan sencillo como quitar algunos atributos sensibles** de los datos de entrenamiento.



CONCLUSIONES

- Se requieren técnicas ingeniosas para corregir el **profundo sesgo de los datos** y forzar a los modelos a realizar **predicciones más imparciales.**



CONCLUSIONES

- Realizar predicciones imparciales conllevan un coste: **reduce el desempeño** de nuestro modelo.



CONCLUSIONES

- Es un **pequeño precio a pagar** para dejar atrás el mundo sesgado de ayer y **tener un mañana más justo.**





THANKS

@jbaquerot