

INTELIGENCIA ARTIFICIAL

ASPECTOS BÁSICOS

Por Martín Zapiola Guerrico
mzg@zgabogados.com.ar



INTELIGENCIA ARTIFICIAL

Conjunto de técnicas que permiten a los sistemas de computación realizar **tareas que normalmente requieren inteligencia humana.**

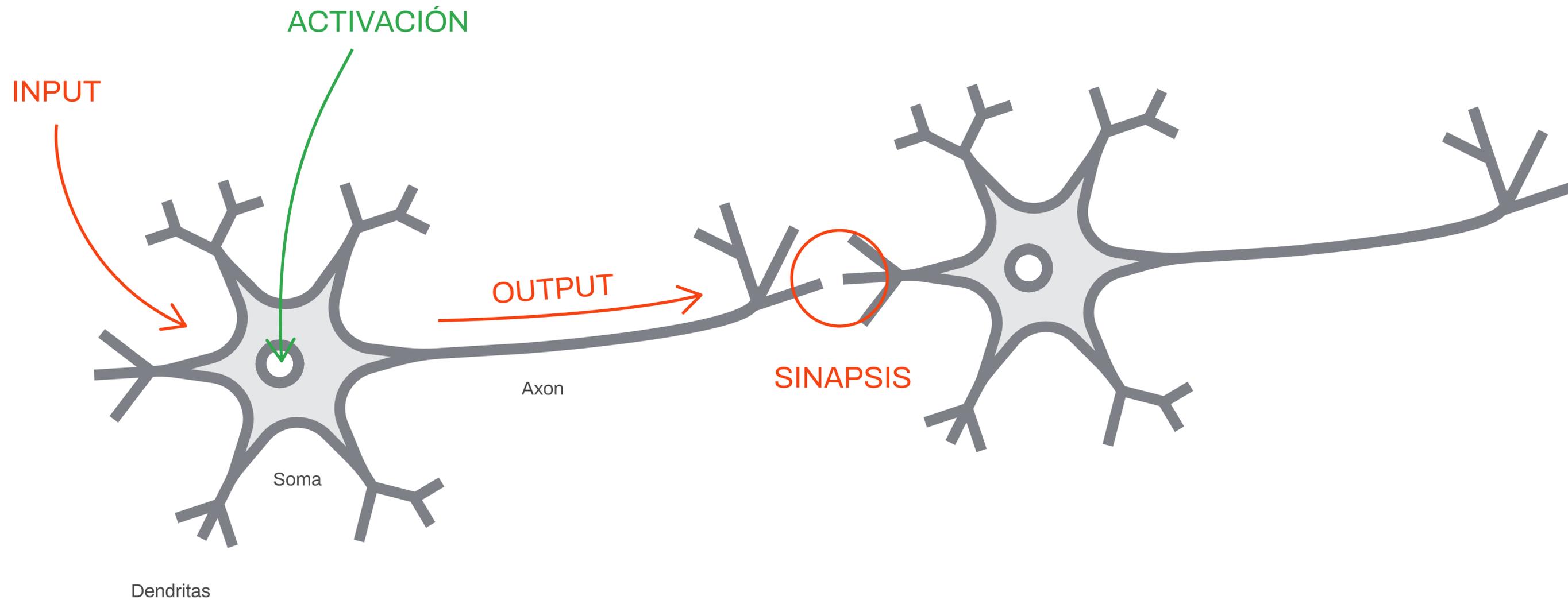
MACHINE LEARNING

Técnica de inteligencia artificial que otorga a las computadoras la habilidad de **aprender una tarea sin ser específicamente programadas** a tal efecto.

DEEP LEARNING

Es la **tecnología más avanzada** de MACHINE LEARNING, con algoritmos que analizan datos con **una estructura lógica similar a la del cerebro humano.** Tiene una estructura de capas de algoritmos llamada RED NEURONAL ARTIFICIAL.

Dinámica de las neuronas humanas



El cerebro humano tiene
100.000 millones de neuronas

Red Neuronal Artificial

BACK PROPAGATION

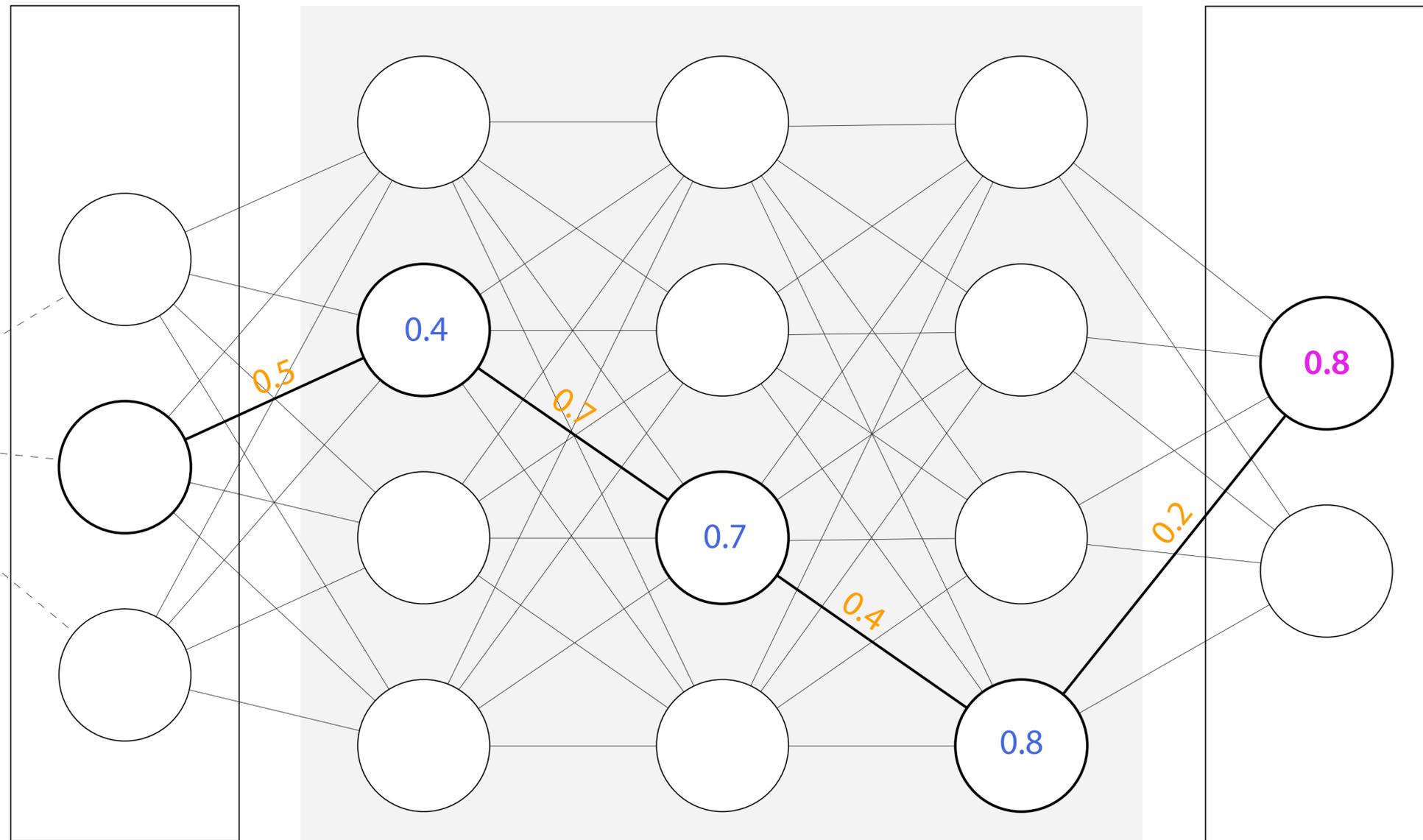
INPUT

CAPAS OCULTAS

OUTPUT



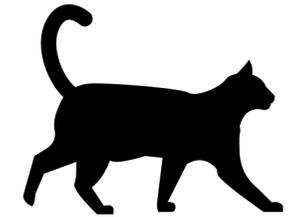
1



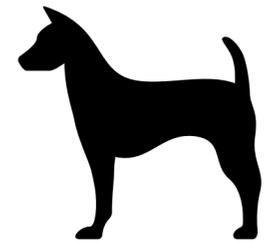
PESOS (P)
FUNCIÓN DE ACTIVACIÓN (FA)

FUNCIÓN DE PERDIDA (FP)
Mide el grado de acierto en la predicción

1



0



FORWARD PROPAGATION

Explicabilidad de los Algoritmos

El problema de la "caja negra"

Los algoritmos más avanzados de Inteligencia Artificial (Redes Neuronales) tienen un alto nivel de exactitud (accuracy), pero **no pueden explicar la forma** en que arriban a una conclusión determinada.

El algoritmo vá **aprendiendo en base a la inmensa cantidad de datos** con que lo alimentan, y no en base a una secuencia de pasos lógicos como los sistemas más tradicionales.

Por eso se lo denomina **"caja negra"**: uno sabe lo que introduce en la caja (INPUT) y lo que saca (OUTPUT), pero **ignora lo que ocurre adentro de la misma**.

Esto resulta problemático cuando el sistema se utiliza para tomar decisiones que afectan derechos fundamentales de terceros, porque las **decisiones pueden tener un sesgo discriminatorio**. Caso COMPAS.



Explicabilidad de los Algoritmos (cont.)

Existe una tensión entre la **explicabilidad** de un algoritmo y su **efectividad** (exactitud).

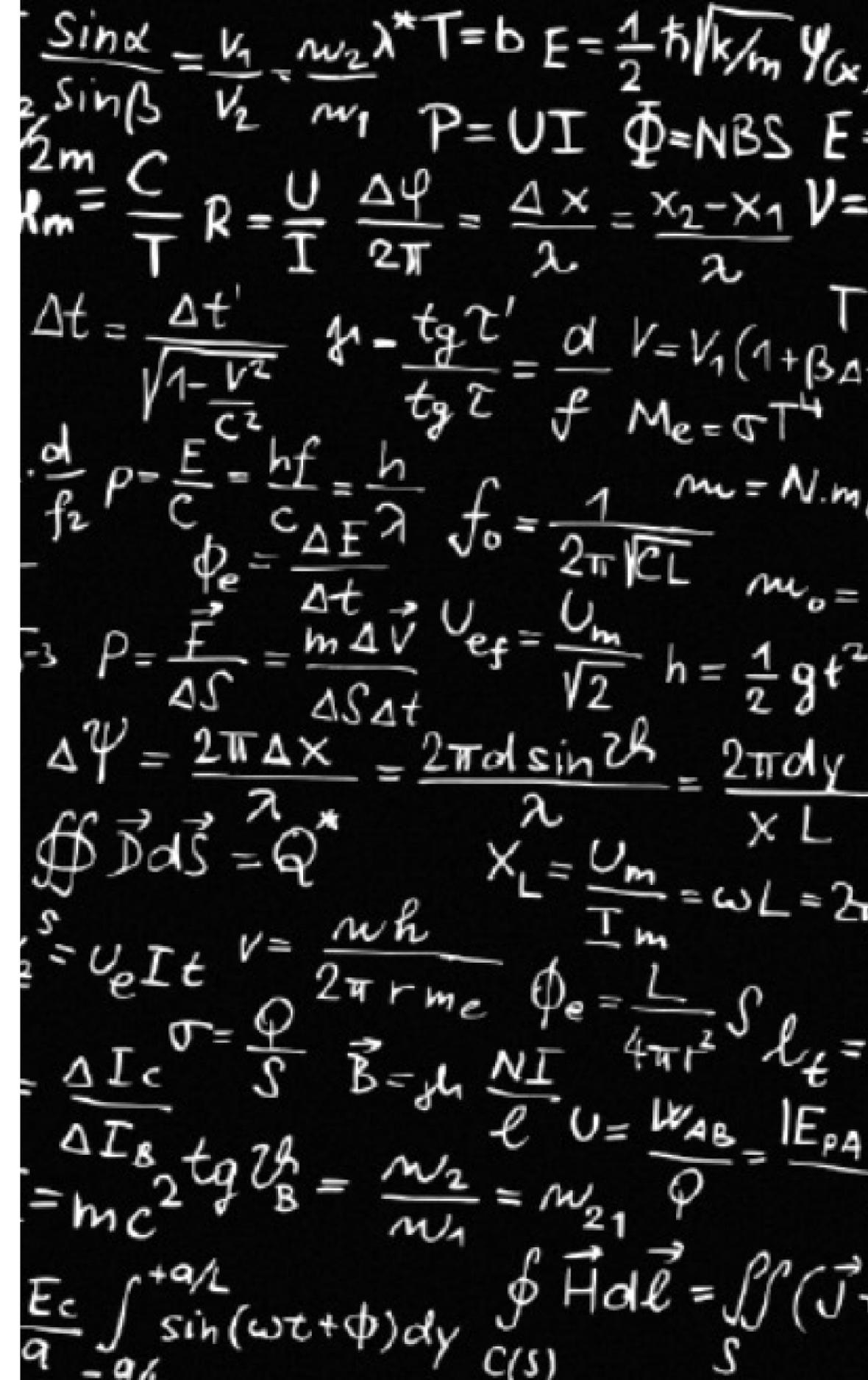
Los **mecanismos** para que un algoritmo pueda explicar su razonamiento necesariamente **afectarán su performance**.

No en todos los casos se requiere que el sistema sea explicable.

Ej: Los sistemas de alta performance que no afectan derechos fundamentales de las personas (y muchas veces son cruciales para el avance científico o tecnológico)

Actualmente existe un movimiento importante para el desarrollo de una **Inteligencia Artificial Explicable** (xAI), que permita a cada sistema aportar explicaciones entendibles para las personas afectadas por sus decisiones.

Ej: Proyectos DARPA y LIME.



¡MUCHAS GRACIAS!

Por Martín Zapiola Guerrico
mzg@zgabogados.com.ar

