

Sesgo de datos

La Inteligencia Artificial está preparada para convertirse en la nueva electricidad del siglo XXI. La definición más simplista de la Inteligencia Artificial es la simulación de procesos de inteligencia humana por parte de las máquinas.

Dicho de otro modo, es la disciplina que trata de crear sistemas capaces de aprender y razonar como un ser humano, aprendan de la experiencia, averigüen cómo resolver problemas ante unas condiciones dadas, contrasten información y lleven a cabo tareas lógicas.

Para conseguir esto es necesario que estos procesos incluyan el aprendizaje (la adquisición de información y reglas para el uso de la información), el razonamiento (usando las reglas para llegar a conclusiones aproximadas o definitivas) y la autocorrección.

El aprendizaje automático (machine learning) es la ciencia de conseguir que una computadora actúe sin programación. El aprendizaje profundo (deep learning) es un subconjunto del aprendizaje automático que, en términos muy simples, puede considerarse como la automatización de la analítica predictiva.

Existen tres tipos de algoritmos de aprendizaje automático:

- el aprendizaje supervisado, en el que los conjuntos de datos están etiquetados para que los patrones puedan ser detectados y utilizados para etiquetar nuevos conjuntos de datos,
- el aprendizaje no supervisado, en el que los conjuntos de datos no están etiquetados y se clasifican de acuerdo con similitudes o diferencias;
- el aprendizaje de refuerzo, en el que los conjuntos de datos no están etiquetados, pero después de realizar una acción o varias acciones, el sistema de AI recibe retroalimentación.

Para que la IA sea realmente un apoyo para el humano, los datos sobre los que se alimentan deben ser muy representativos de la población que representa. Dicho de otra manera, los datos no deben tener sesgo. El sesgo es un error sistemático en el que se puede incurrir cuando, al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras.

Pongamos como caso el ejemplo de la empresa Amazon (<https://www.itpro.co.uk/machine-learning/32083/ai-recruitment-tool-pulled-by-amazon-for-sex-bias>)

Amazon creó un sistema de IA para que seleccionará a las mejores personas en sus procesos de contratación a partir de los curriculums que le enviaban.

Para entrenar al sistema de IA utilizó la base de datos de los trabajadores de la empresa. Debido al porcentaje tan bajo de mujeres que estudian estudios de informática (sobre un 12 %), en la base de datos utilizada por Amazon la mayoría eran hombres. El sistema, debido a este sesgo,

estableció que un parámetro fundamental era el sexo y en el proceso de contratación sólo eligió a hombres.

En la película Elysium, ambientada en 2154, los habitantes de la Tierra más adinerados viven en una plataforma espacial en la que hay una especie de cápsula médica capaz de diagnosticar inmediatamente cualquier enfermedad, curarla y hasta reparar las lesiones del cuerpo sin que intervenga ninguna persona. Sin llegar hasta tan extremo lo cierto es que la medicina es uno de los campos que más y mejor se está beneficiando de la inteligencia artificial (IA).

La mayoría de las bases de datos de tratamiento de enfermedades están soportadas por los países del primer mundo. Los sistemas de IA se deben alimentar de esta información.

Todos entendemos que hay diferencias en la alimentación, en los hábitos de comportamiento, en el ambiente que se respira entre los diferentes países y/o continentes.

Además, la mayoría de los datos están recogidos de personas de las razas, y origen racial, que predominan en ese primer mundo.

CUESTIONES

- Es éticamente correcto utilizar bases de datos, con sesgo, para poder identificar factores que determinan una enfermedad
- Debería analizarse el sesgo que se produce e intentar corregir las desviaciones de los datos.
- La sanidad debe aprovechar la potencia de la IA, a pesar de saber que existen sesgos significativos en la recogida de los datos
- Las empresas públicas/privadas deben hacer un esfuerzo para minimizar los sesgos del análisis de los datos