Comprensión del sesgo algorítmico y las estrategias de mitigación

Qué es ¿Sesgo algorítmico?

Errores sistémicos y repetibles en los sistemas informáticos que conducen a resultados injustos.

Importancia de comprender el sesgo a medida que la IA y el ML se vuelven más frecuentes.

Tipos de sesgo algorítmico

Sesgo basado en datos: ideas injustas recogidas de los datos. Ejemplo: Sistema de contratación de IA sesgado hacia los candidatos masculinos.

Algorítmico o Sesgo basado en modelos: Sesgo introducido accidentalmente en modelos de IA.

Ejemplo: Sistema de reconocimiento facial mejor para reconocer rostros de piel clara.

Sesgo humano o social: Sesgos introducidos por personas que crean o usan IA. Ejemplo: Etiquetas injustas en las descripciones de las imágenes.

Identificación de sesgos

Necesidad de herramientas para medir el sesgo en la IA, similar a probar un pastel para comprobar si la receta es buena.

El sesgo puede colarse en diferentes etapas de la creación de IA (diseño, construcción, uso).

Estrategias de mitigación

Utiliza datos diversos y justos: asegúrate de que los datos incluyen a todo tipo de personas.

Ten cuidado con los resultados sesgados: cuestiona las respuestas injustas de la IA.

Haz preguntas y aprende: Comprende cómo funciona la IA para detectar sesgos.

Conviértete en un defensor de la equidad: Habla en contra de los prejuicios.

Para volver al juego, cierra esta ventana, haciendo clic en la X.

Cofinanciado por la Unión Europea

Financiado por la Unión Europea. Las opiniones y puntos de vista expresados solo comprometen a su(s) autor(es) y no reflejan necesariamente los de la Unión Europea o los de la Agencia Ejecutiva Europea de Educación y Cultura (EACEA). Ni la Unión Europea ni la EACEA pueden ser considerados responsables de ellos.

Universitat de les Illes Balears · ISQe ENGAGING PEOPLE · AARHUS UNIVERSITY · VAMK · helixconnect · Qi Charlie

2022-1-ES01-KA220-HED-000085257