Fundamentos algorítmicos y ética en IA: de la teoría a la práctica

Kit de herramientas para sesiones sincrónicas

UC4 | Equidad y sesgo de los datos en la IA

Soporte de diapositivas de PowerPoint

Número de proyecto | 2022-1-ES01-KA220-HED-000085257

ÍNDICE

- INTRODUCCIÓN - 3
- INTRODUCCIÓN A LA EQUIDAD Y EL SESGO DE LOS DATOS - 7
- TIPOS DE SESGO EN LOS SISTEMAS DE IA - 20
- MÉTODOS PARA IDENTIFICAR Y MEDIR SESGOS EN LOS DATOS - 34
- ESTRATEGIAS PARA ABORDAR Y MITIGAR LOS SESGOS - 41
- ENTENDIENDO LAS MÉTRICAS DE EQUIDAD - 47
- EQUIDAD EN LA RECOPILACIÓN Y EL PREPROCESAMIENTO DE DATOS - 58
- TENDENCIAS FUTURAS Y NUEVAS CUESTIONES - 67
- CONCLUSIÓN - 76

INTRODUCCIÓN

FUENTE DE LA IMAGEN | Generado por DALL-E

Sugerencia

Comience con una encuesta o pregunta interactiva para medir la comprensión inicial o las experiencias personales de los participantes con la equidad y el sesgo de la IA. Esto puede hacer que la introducción sea más atractiva y sentar las bases de por qué el tema es relevante.

- Ejemplo: ¿Cuál es su familiaridad con la equidad y el sesgo en la IA?
  - No estoy familiarizado con él en absoluto.
  - He oído hablar de él, pero no conozco los detalles.
  - Entiendo lo básico, pero me gustaría saber más.
  - Estoy bastante bien informado y he trabajado con sistemas de IA teniendo en cuenta la equidad y el sesgo.

EN ESTA UNIDAD DE COMPETENCIA ENCONTRARÁS: LAS SIGUIENTES MATERIAS:

- Los conceptos básicos de la equidad y el sesgo de los datos en la IA.
- Varios sesgos que pueden influir en los sistemas de IA.
- Cómo detectar y medir sesgos en los datos.
- Técnicas para reducir el sesgo y aumentar la equidad.
- Métricas para evaluar la equidad de la IA.
- Mejores prácticas para la recopilación y el procesamiento justos de datos.
- Tendencias emergentes en la equidad de la IA. equidad y sesgo en la IA.

AL FINAL DE LA UNIDAD DE COMPETENCIA, USTED DEBERÍA SER CAPAZ DE:

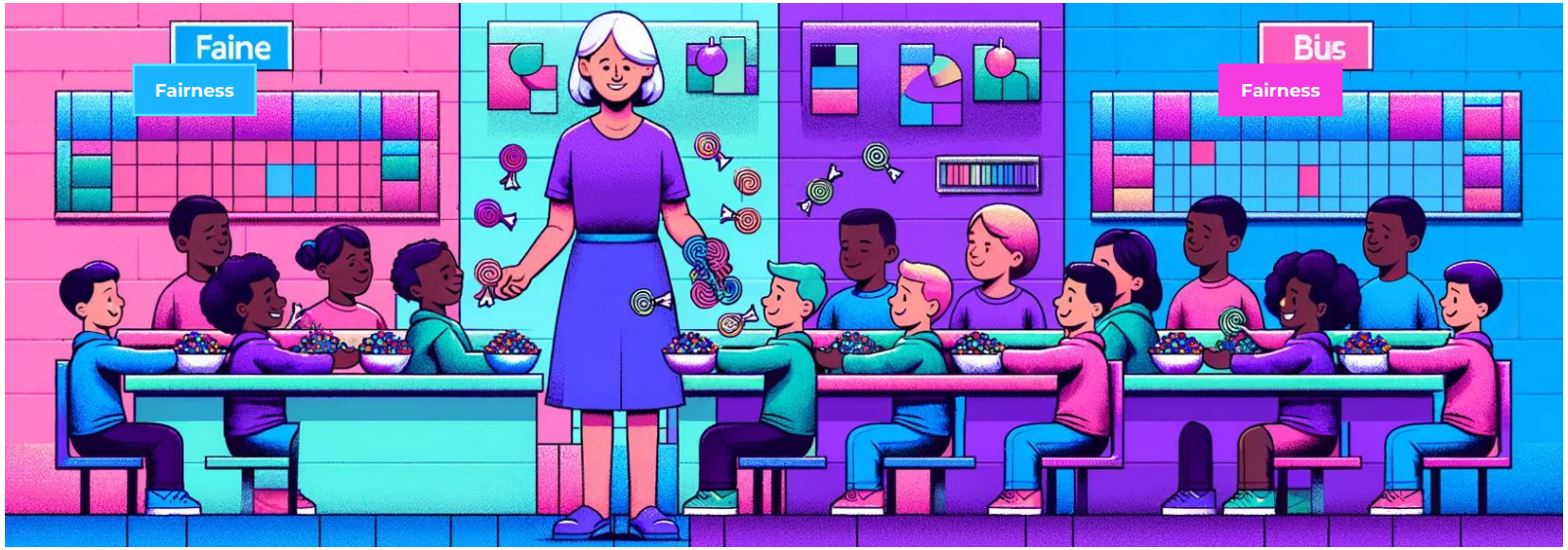
- Explicar los fundamentos de la equidad de los datos y la importancia del sesgo en la IA.
- Identificar las diversas formas de sesgo que pueden afectar a los sistemas de IA.
- Evalúe los sistemas de IA utilizando métricas de equidad.
- Siga las prácticas recomendadas para recopilar y preprocesar datos a fin de promover la equidad.

INTRODUCCIÓN A LA EQUIDAD Y EL SESGO DE LOS DATOS

FUENTE DE LA IMAGEN | Generado por DALL-E

Introducción a la equidad y el sesgo de los datos

Equidad y sesgo: la analogía de los dulces



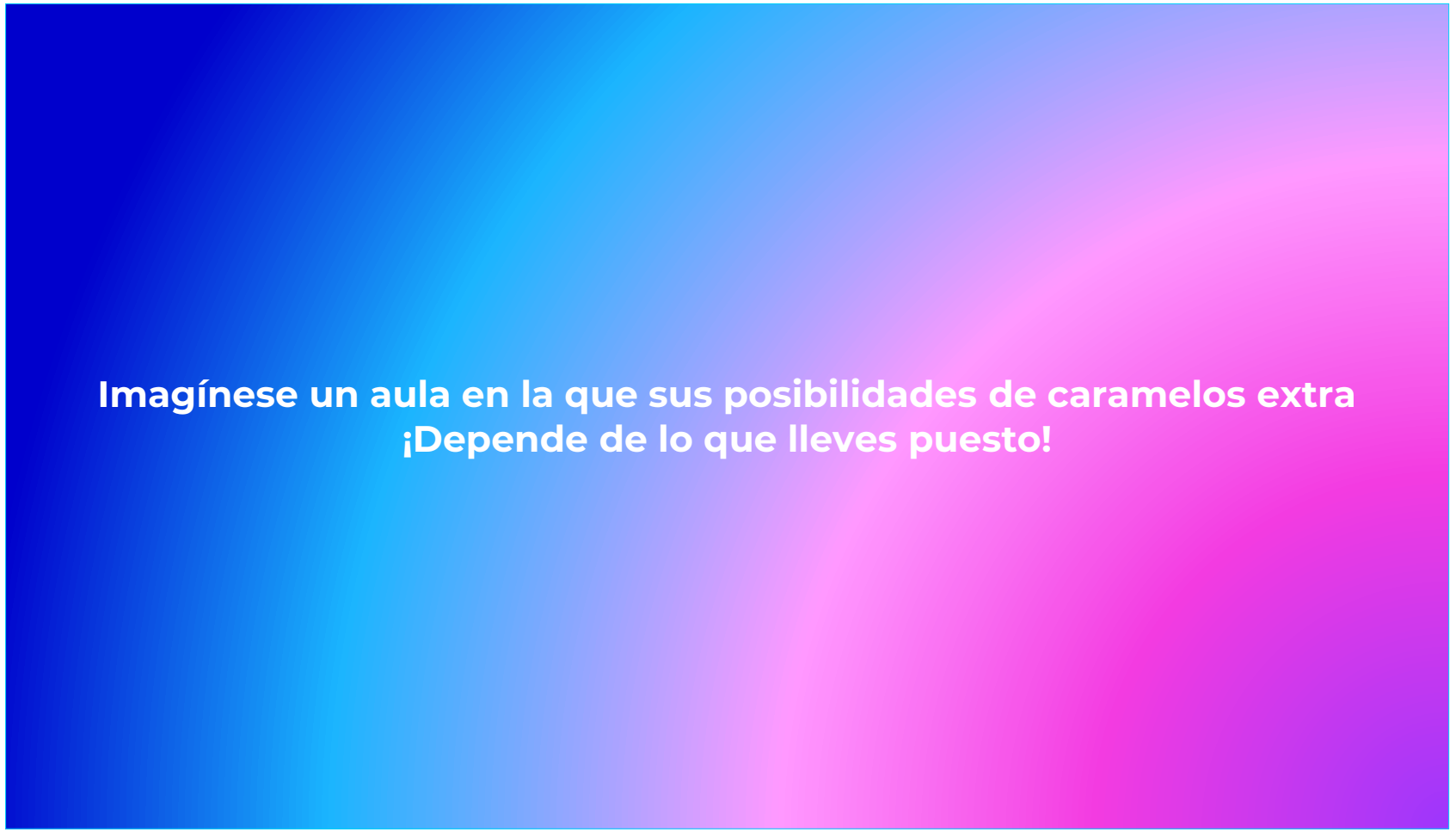
FUENTE DE LA IMAGEN | Generado por DALL-E

Descubra cómo la equidad y el sesgo en los sistemas de IA pueden afectar nuestras decisiones cotidianas, al igual que las reglas de un maestro para repartir dulces pueden dar forma a una experiencia en el aula.

Analicemos el concepto de equidad de la IA: ¡un dulce ejemplo a la vez!

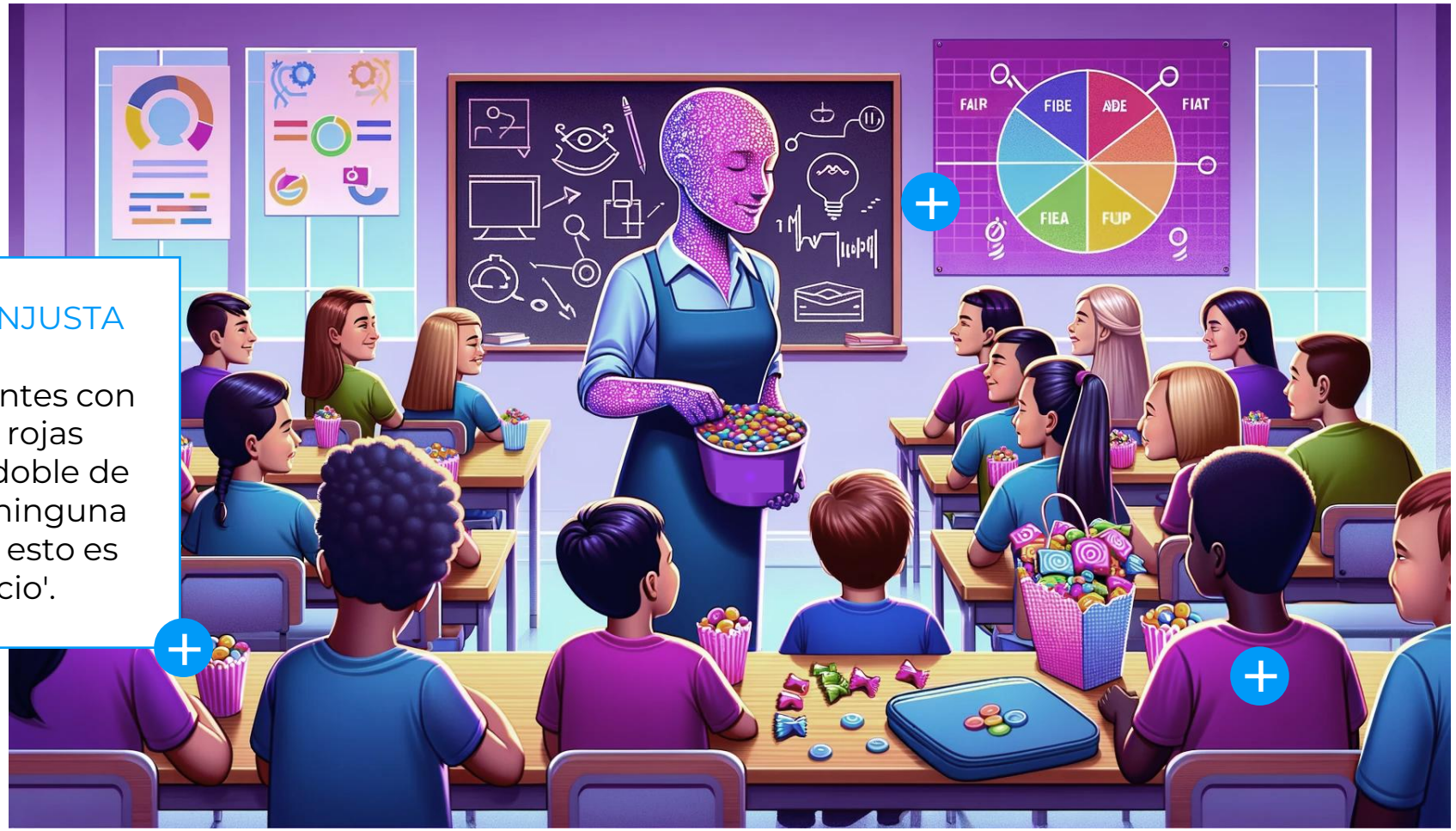
Introducción a la equidad y el sesgo de los datos

Equidad y sesgo: la analogía de los dulces



Introducción a la equidad y el sesgo de los datos

Equidad y sesgo: la analogía de los dulces



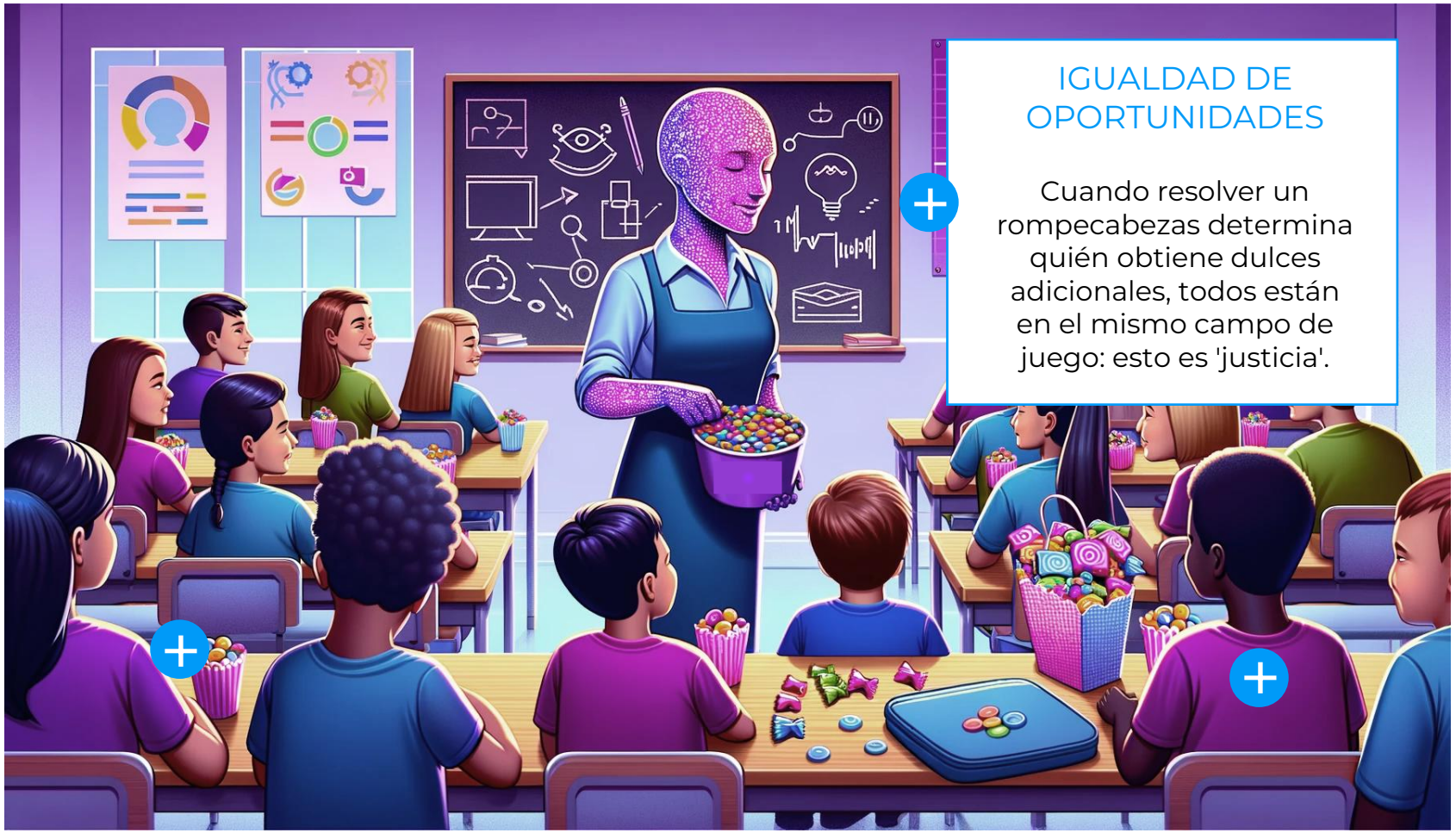
VENTAJA INJUSTA

Los estudiantes con camisas rojas reciben el doble de dulces sin ninguna razón real: esto es 'prejuicio'.

FUENTE DE LA IMAGEN | Generado por DALL-E

Introducción a la equidad y el sesgo de los datos

Equidad y sesgo: la analogía de los dulces



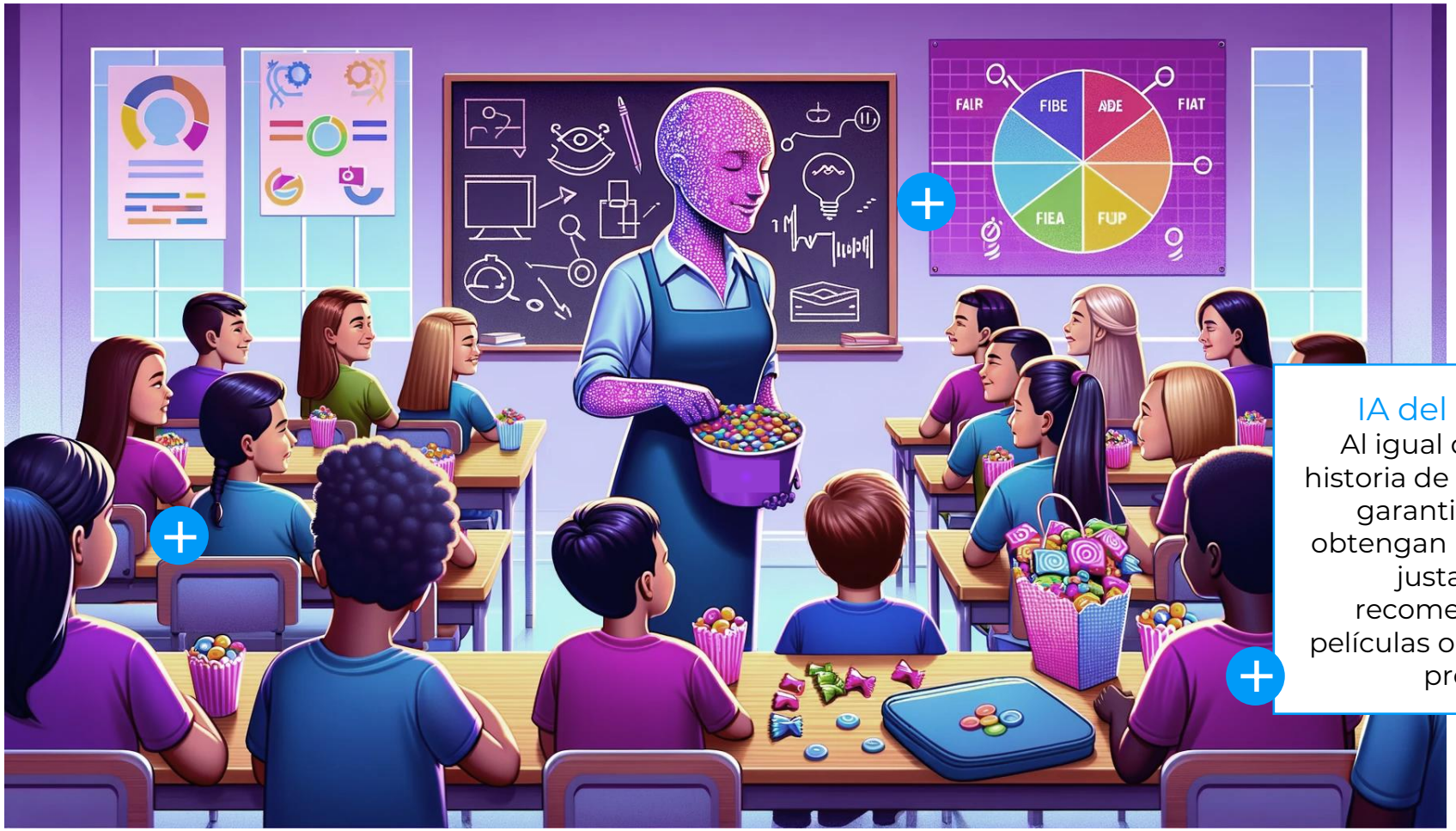
IGUALDAD DE OPORTUNIDADES

Cuando resolver un rompecabezas determina quién obtiene dulces adicionales, todos están en el mismo campo de juego: esto es 'justicia'.

FUENTE DE LA IMAGEN | Generado por DALL-E

Introducción a la equidad y el sesgo de los datos

Equidad y sesgo: la analogía de los dulces



IA del mundo real

Al igual que en nuestra historia de dulces, la IA debe garantizar que todos obtengan una oportunidad justa, ya sea en recomendaciones de películas o aprobaciones de préstamos.

FUENTE DE LA IMAGEN | Generado por DALL-E

Introducción a la equidad y el sesgo de los datos

Equidad de datos

La equidad de los datos consiste en garantizar que todas las personas sean tratadas de manera equitativa por los sistemas de IA, sin que los sesgos afecten a sus resultados. Implica los procesos y prácticas que intervienen en la recopilación, el procesamiento y el uso de datos para evitar efectos discriminatorios.

La equidad de los datos es un aspecto crucial a la hora de diseñar e implementar sistemas de IA que tengan un impacto positivo en la sociedad. Garantiza que las decisiones de la IA sean justas e imparciales, promoviendo la igualdad y evitando la discriminación.

Las prácticas justas de datos conducen a aplicaciones de IA confiables que pueden ser ampliamente aceptadas y beneficiosas en diversas comunidades.

Su objetivo es una IA que sirva a todos de manera equitativa, independientemente de su origen o características demográficas.

Introducción a la equidad y el sesgo de los datos

¿Qué es el sesgo? +

¿Cómo se produce? +

¿Cuáles son las consecuencias del sesgo? +

Introducción a la equidad y el sesgo de los datos

¿Qué es el sesgo?

El sesgo en IA se refiere a errores sistemáticos que favorecen injustamente a un grupo sobre otros. Esto puede surgir de datos sesgados, algoritmos defectuosos o procesos de entrenamiento con prejuicios.

¿Cómo se produce? +

¿Cuáles son las consecuencias del sesgo? +

Introducción a la equidad y el sesgo de los datos

¿Qué es el sesgo? +

¿Cómo se produce? -

Sesgo de datos: cuando los datos introducidos en los sistemas de IA reflejan desigualdades históricas o representaciones incompletas de diversos grupos.
Sesgo algorítmico: cuando las reglas o modelos utilizados para tomar decisiones refuerzan los prejuicios existentes o pasan por alto patrones minoritarios.
Bucles de retroalimentación: cuando las decisiones sesgadas de la IA influyen en los datos futuros, perpetuando el ciclo de sesgo.

¿Cuáles son las consecuencias del sesgo? +

Introducción a la equidad y el sesgo de los datos

¿Qué es el sesgo? +

¿Cómo se produce? +

¿Cuáles son las consecuencias del sesgo? -

Impacto social: puede llevar a perpetuar estereotipos, reforzar las desigualdades sociales y erosionar la confianza en las instituciones esenciales.

Impacto individual: las personas pueden enfrentarse a un trato injusto basado en decisiones defectuosas de la IA en áreas críticas como la contratación, la aplicación de la ley y la aprobación de préstamos.

Introducción a la equidad y el sesgo de los datos

Escenarios del mundo real: sesgo de la IA

Sesgo de la herramienta de reclutamiento

Escenario

En 2018, se informó de que Amazon había desechado una herramienta de contratación de IA que mostraba sesgos contra las mujeres. El sistema de IA se entrenó con datos de currículum recopilados durante 10 años, predominantemente de hombres, lo que refleja el dominio masculino en la industria tecnológica. En consecuencia, la IA se enseñó a sí misma que los candidatos masculinos eran preferibles, penalizando los currículos que incluían palabras como "femenino", como podría aparecer en "capitán del club de ajedrez femenino".

Impacto

Este ejemplo pone de manifiesto cómo la IA puede perpetuar sesgos históricos si los datos de entrenamiento están sesgados. También muestra el potencial de la IA para influir en las oportunidades de empleo y la importancia de garantizar que las herramientas de contratación de la IA promuevan la diversidad y la equidad.

Introducción a la equidad y el sesgo de los datos

Escenarios del mundo real: sesgo de la IA

Sesgo del algoritmo sanitario

Escenario

Un estudio de 2019 publicado por Science reveló que un algoritmo utilizado en muchos sistemas de atención médica de EE. UU. para asignar recursos de atención médica estaba sesgado contra los pacientes negros. El algoritmo predijo las necesidades de atención médica en función de los costos de atención médica, asumiendo inadvertidamente que los costos de atención médica más bajos significaban menores necesidades de atención médica. Sin embargo, históricamente, los pacientes negros incurrieron en menos costos debido a diversas barreras para acceder a la atención, no porque fueran más saludables. Como resultado, el algoritmo favoreció a los pacientes blancos más sanos sobre los pacientes negros más enfermos al identificar a los candidatos para programas de atención adicionales.

Impacto

Este sesgo en el algoritmo dio como resultado que los pacientes negros recibieran un acceso significativamente menor a los programas de atención diseñados para ayudar a manejar afecciones de salud complejas, a pesar de que su necesidad de tales intervenciones era mayor. Este escenario subraya la necesidad crítica de equidad en las aplicaciones de IA sanitaria, donde la toma de decisiones sesgada puede afectar directamente a las vidas y el bienestar.

TIPOS DE SESGO EN LOS SISTEMAS DE IA

FUENTE DE LA IMAGEN | Generado por DALL-E

Tipos de sesgos en los sistemas de IA

Sesgo de selección

Haz clic en las cartas para darles la vuelta

Sesgo de muestreo

Haz clic en las cartas para darles la vuelta

Sesgo de medición

Haz clic en las cartas para darles la vuelta

Sesgo algorítmico

Haz clic en las cartas para darles la vuelta

Sesgo de confirmación

Haz clic en las cartas para darles la vuelta

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vuelta

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

Sesgo de muestreo

Haz clic en las cartas para darles la vuelta

Sesgo de medición

Haz clic en las cartas para darles la vuelta

Sesgo algorítmico

Haz clic en las cartas para darles la vuelta

Sesgo de confirmación

Haz clic en las cartas para darles la vuelta

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vuelta

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

El sesgo de muestreo es un subconjunto del sesgo de selección y surge cuando el método utilizado para recopilar datos favorece a ciertos tipos de individuos o grupos sobre otros.

Sesgo de medición

Haz clic en las cartas para darles la vuelta

Sesgo algorítmico

Haz clic en las cartas para darles la vuelta

Sesgo de confirmación

Haz clic en las cartas para darles la vuelta

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vuelta

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

El sesgo de muestreo es un subconjunto del sesgo de selección y surge cuando el método utilizado para recopilar datos favorece a ciertos tipos de individuos o grupos sobre otros.

El sesgo de medición se produce debido a errores o inconsistencias en el proceso de medición, lo que conduce a resultados inexactos o engañosos. Puede deberse a instrumentos defectuosos, preguntas ambiguas o sesgos del observador.

Sesgo algorítmico

Haz clic en las cartas para darles la vuelta

datos de entrenamiento.

Sesgo de confirmación

Haz clic en las cartas para darles la vuelta

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vuelta

desfavorables

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

El sesgo de muestreo es un subconjunto del sesgo de selección y surge cuando el método utilizado para recopilar datos favorece a ciertos tipos de individuos o grupos sobre otros.

El sesgo de medición se produce debido a errores o inconsistencias en el proceso de medición, lo que conduce a resultados inexactos o engañosos. Puede deberse a instrumentos defectuosos, preguntas ambiguas o sesgos del observador.

El sesgo algorítmico surge en los algoritmos de aprendizaje automático cuando discriminan sistemáticamente a ciertos grupos o individuos en función de la raza, el género u otras características protegidas presentes en los datos de entrenamiento.

Sesgo de confirmación

Haz clic en las cartas para darles la vuelta

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vueta

desfavorables.

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

El sesgo de muestreo es un subconjunto del sesgo de selección y surge cuando el método utilizado para recopilar datos favorece a ciertos tipos de individuos o grupos sobre otros.

El sesgo de medición se produce debido a errores o inconsistencias en el proceso de medición, lo que conduce a resultados inexactos o engañosos. Puede deberse a instrumentos defectuosos, preguntas ambiguas o sesgos del observador.

El sesgo algorítmico surge en los algoritmos de aprendizaje automático cuando discriminan sistemáticamente a ciertos grupos o individuos en función de la raza, el género u otras características protegidas presentes en los datos de entrenamiento.

Sesgo de confirmación: cuando los investigadores o analistas utilizan o interpretan selectivamente datos que confirman sus creencias o hipótesis preconcebidas, mientras ignoran o descartan las pruebas contradictorias.

Sesgo en la presentación de informes

Haz clic en las cartas para darles la vuelta

desfavorables.

Tipos de sesgos en los sistemas de IA

El sesgo de selección se produce cuando la muestra de datos utilizada para el análisis no es representativa de la población a la que se supone que representa. Conduce a conclusiones sesgadas o inexactas.

El sesgo de muestreo es un subconjunto del sesgo de selección y surge cuando el método utilizado para recopilar datos favorece a ciertos tipos de individuos o grupos sobre otros.

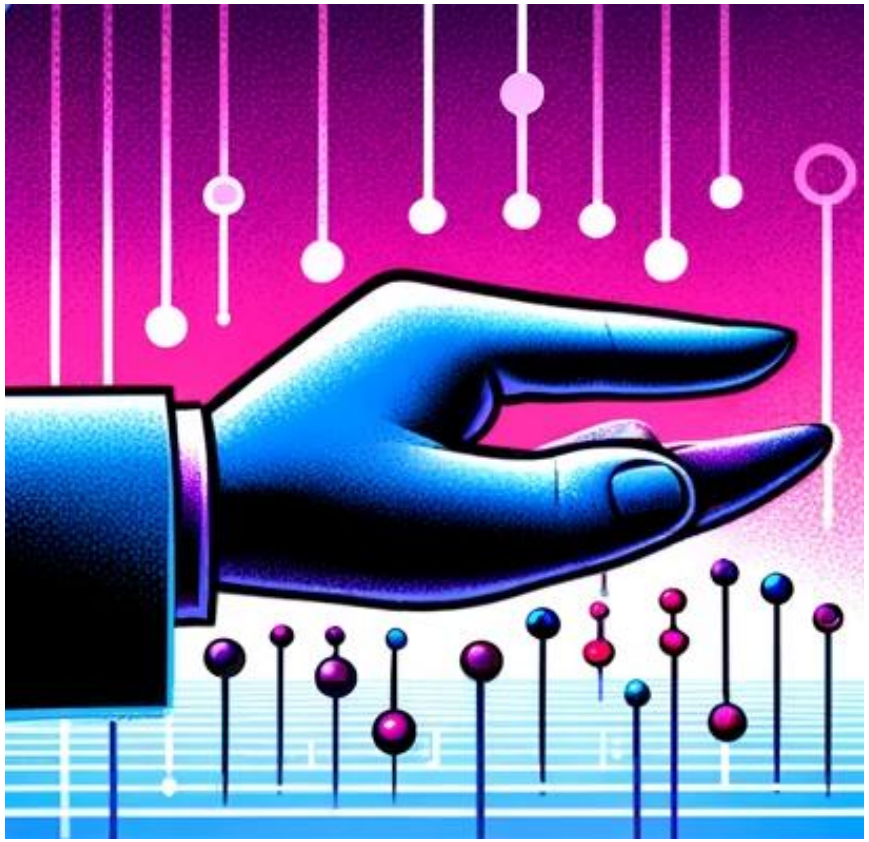
El sesgo de medición se produce debido a errores o inconsistencias en el proceso de medición, lo que conduce a resultados inexactos o engañosos. Puede deberse a instrumentos defectuosos, preguntas ambiguas o sesgos del observador.

El sesgo algorítmico surge en los algoritmos de aprendizaje automático cuando discriminan sistemáticamente a ciertos grupos o individuos en función de la raza, el género u otras características protegidas presentes en los datos de entrenamiento.

Sesgo de confirmación: cuando los investigadores o analistas utilizan o interpretan selectivamente datos que confirman sus creencias o hipótesis preconcebidas, mientras ignoran o descartan las pruebas contradictorias.

El sesgo de informe se refiere a la publicación selectiva o el informe de hallazgos de investigación que son estadísticamente significativos o favorables, mientras que suprime o descuida los resultados no significativos o desfavorables.

Tipos de sesgos en los sistemas de IA
Ejemplos



FUENTE DE LA IMAGEN | Generado por DALL-E

Sesgo de selección

Sesgo de selección | Tecnología de reconocimiento facial

Las tecnologías de reconocimiento facial a menudo se han entrenado predominantemente en conjuntos de datos compuestos principalmente por rostros caucásicos. Esta falta de diversidad en los conjuntos de datos de entrenamiento ha llevado a tasas de error más altas en el reconocimiento de individuos de otros orígenes étnicos. Por ejemplo, un estudio realizado por MIT Media Lab encontró disparidades significativas en la precisión de las tecnologías de reconocimiento de género, particularmente la identificación errónea del género entre las mujeres de piel más oscura.

Tipos de sesgos en los sistemas de IA
Ejemplos

FUENTE DE LA IMAGEN | Generado por DALL-E

Sesgo de muestreo

Sesgo de muestreo | Errores en las encuestas políticas

En las elecciones presidenciales de Estados Unidos de 2016, varias encuestas de opinión subestimaron el apoyo a Donald Trump. Un problema clave fue el sesgo de muestreo, donde la ponderación demográfica en las muestras no representaba con precisión la base de votantes. Los encuestadores habían sobremuestreado poblaciones urbanas y submuestreado áreas rurales, lo que llevó a predicciones inexactas que no dieron en el blanco en el resultado electoral real.

Tipos de sesgos en los sistemas de IA
Ejemplos

FUENTE DE LA IMAGEN | Generado por DALL-E

Sesgo de medición

Sesgo de medición | Oxímetros de pulso y tono de piel

Estudios recientes, incluidos los intensificados por la pandemia de COVID-19, han demostrado que los oxímetros de pulso pueden presentar sesgos de medición al sobreestimar los niveles de oxígeno en sangre en personas con piel más oscura. Este sesgo surge porque los sensores de los dispositivos son menos efectivos para penetrar en la piel más oscura, lo que lleva a diagnósticos erróneos potencialmente peligrosos o retrasos en el tratamiento para los pacientes de color.

Tipos de sesgos en los sistemas de IA
Ejemplos

FUENTE DE LA IMAGEN | Generado por Sesgo algorítmico DALL-E

Sesgo algorítmico | COMPAS en la condena penal

Se descubrió que el software de Perfiles de Gestión de Delincuentes Correccionales para Sanciones Alternativas (COMPAS), utilizado en el sistema de justicia penal de los Estados Unidos para evaluar la probabilidad de que un acusado vuelva a delinquir, exhibe sesgos raciales. El análisis reveló que el software tenía casi el doble de probabilidades de predecir falsamente la criminalidad futura en los acusados negros en comparación con los acusados blancos, lo que influía en las sentencias y las decisiones de fianza basadas en recomendaciones algorítmicas sesgadas.

Tipos de sesgos en los sistemas de IA
Ejemplos

Sesgo de confirmación

FUENTE DE LA IMAGEN | Generado por DALL-E

Sesgo de Confirmación - Modelos de Calificación Crediticia

Las instituciones financieras que utilizan modelos de calificación crediticia impulsados por IA podrían reforzar involuntariamente los sesgos de confirmación si los modelos se entrenan con datos históricos de préstamos que reflejan prejuicios pasados. Por ejemplo, si los prestamistas humanos anteriores tenían sesgos contra ciertos grupos demográficos, el sistema de IA podría replicar estos sesgos favoreciendo a individuos de grupos históricamente favorecidos, suponiendo que sean menos riesgosos en función de los reembolsos de préstamos anteriores.

Tipos de sesgos en los sistemas de IA
Ejemplos

Sesgo en la presentación de informes

FUENTE DE LA IMAGEN | Generado por DALL-E

Sesgos en la denuncia - Efectos secundarios de los medicamentos en foros en línea

Las compañías farmacéuticas y los investigadores de la salud que utilizan la IA para monitorear los foros en línea y las redes sociales en busca de informes de efectos secundarios de medicamentos pueden encontrar sesgos en los informes. Los pacientes que experimentan efectos secundarios graves tienen más probabilidades de informar sus experiencias en línea en comparación con los que experimentan efectos secundarios leves o sin efectos secundarios. Esto puede llevar a los sistemas de IA a sobrestimar la gravedad o la frecuencia de las reacciones adversas, lo que podría influir de forma inexacta en los perfiles de seguridad de los medicamentos.

MÉTODOS PARA IDENTIFICAR Y MEDIR SESGOS EN LOS DATOS

FUENTE DE LA IMAGEN | Generado por DALL-E

Métodos para identificar y medir sesgos en los datos

- Auditoría de datos +
- Análisis de disparidades +
- Pruebas de equidad algorítmica +
- Análisis de sensibilidad +

Métodos para identificar y medir sesgos en los datos

Auditoría de datos -

La auditoría de datos implica examinar sistemáticamente los conjuntos de datos para identificar cualquier disparidad, desequilibrio o anomalía que pueda indicar un sesgo.

Este proceso incluye la comprobación de la igualdad de representación entre los diferentes grupos dentro de los datos. Garantiza que el proceso de recopilación de datos no excluya sistemáticamente a determinadas poblaciones ni refuerce los estereotipos existentes.

Análisis de disparidades +

Pruebas de equidad algorítmica +

Análisis de sensibilidad +

Métodos para identificar y medir sesgos en los datos

Auditoría de datos +

Análisis de disparidades -

El análisis de disparidad se refiere a la comparación de los resultados o predicciones de los modelos de IA en diferentes grupos demográficos o protegidos para identificar cualquier impacto desigual.

Implica el cálculo de medidas estadísticas como la diferencia en las tasas de error, las tasas de aceptación u otras métricas de resultados entre grupos. Esto ayuda a cuantificar y visualizar las disparidades que pueden indicar un sesgo en los procesos de toma de decisiones de la IA.

Pruebas de equidad algorítmica +

Análisis de sensibilidad +

Métodos para identificar y medir sesgos en los datos

Auditoría de datos +

Análisis de disparidades +

Pruebas de equidad algorítmica -

Las pruebas de equidad algorítmica implican la implementación de pruebas o métricas que evalúan la equidad de las decisiones tomadas por los sistemas de IA.

Esto puede incluir el uso de métricas de equidad como la igualdad de oportunidades, la paridad demográfica o los criterios de equidad individual para evaluar el rendimiento de los algoritmos. Estas pruebas ayudan a determinar si los sistemas de IA están tomando decisiones justas e imparciales en diferentes grupos demográficos.

Análisis de sensibilidad +

Métodos para identificar y medir sesgos en los datos

Auditoría de datos +

Análisis de disparidades +

Pruebas de equidad algorítmica +

Análisis de sensibilidad -

El análisis de sensibilidad consiste en probar la solidez de las decisiones de la IA alterando ligeramente las entradas de datos y observando cómo varían los resultados para los diferentes grupos.

Al perturbar las entradas de datos, el análisis de sensibilidad evalúa la sensibilidad de los modelos de IA a los cambios en los datos de entrada. Ayuda a identificar posibles fuentes de sesgo y evalúa la fiabilidad y la coherencia de las predicciones de la IA en diversos conjuntos de datos.

Métodos para identificar y medir sesgos en los datos

Taller interactivo para identificar y medir sesgos en los datos

Sesgo de la herramienta de reclutamiento

Objetivo: este taller permitirá a los participantes aplicar la auditoría de datos, el análisis de disparidades y el análisis de sensibilidad utilizando un conjunto de datos real para identificar y comprender posibles sesgos.

Herramientas y materiales

Conjunto de datos: proporcione un conjunto de datos simple y disponible públicamente con varios atributos demográficos (por ejemplo, el conjunto de datos de ingresos de adultos del Repositorio de aprendizaje automático de UCI, que incluye características como edad, educación, raza, género e ingresos).

Actividad: los participantes cargan el conjunto de datos y realizan una exploración básica de datos (comprobando los valores que faltan y comprendiendo la distribución de los atributos clave).

Objetivo: familiarizar a los participantes con el conjunto de datos y preparar los datos para el análisis.

Auditoría de datos: realice una auditoría de datos analizando la representación de diferentes grupos dentro del conjunto de datos. Calcule la proporción de cada grupo demográfico (por ejemplo, por género, raza). Identifique los grupos subrepresentados.

Análisis de disparidad: los participantes aplican el análisis de disparidad para evaluar qué tan bien un modelo simple (por ejemplo, regresión logística para predecir el nivel de ingresos) predice diferentes grupos. Identifique las disparidades significativas en el rendimiento del modelo entre los grupos y formule hipótesis sobre las posibles causas.

Análisis de sensibilidad: realice un análisis de sensibilidad alterando ligeramente los puntos de datos y observando los cambios en las predicciones del modelo. Modifique los atributos relacionados con variables demográficas sensibles (por ejemplo, cambie ligeramente la edad o el nivel de educación) y observe cómo los cambios afectan las predicciones del modelo.

ESTRATEGIAS PARA ABORDAR Y MITIGAR LOS SESGOS

FUENTE DE LA IMAGEN | Generado por DALL-E

Estrategias para abordar y mitigar los sesgos

Representación diversa

Haz clic en las cartas para darles la vuelta

Algoritmos sensibles al sesgo

Haz clic en las cartas para darles la vuelta

Seguimiento y evaluación periódicos

Haz clic en las cartas para darles la vuelta

Transparencia y explicabilidad

Haz clic en las cartas para darles la vuelta

Estrategias para abordar y mitigar los sesgos

Representación diversa

Los diversos puntos de datos de varios grupos demográficos garantizan que los sistemas de IA se entrenen con un conjunto de datos representativo, lo que reduce la probabilidad de resultados sesgados. La diversidad aporta diferentes perspectivas y perspectivas, lo que ayuda a identificar y mitigar los sesgos durante el proceso de desarrollo.

Algoritmos sensibles al sesgo

Haz clic en las cartas para darles la vuelta

Seguimiento y evaluación periódicos

Haz clic en las cartas para darles la vuelta

Transparencia y explicabilidad

Haz clic en las cartas para darles la vuelta

la equidad de los resultados de IA.

Estrategias para abordar y mitigar los sesgos

Representación diversa

Los diversos puntos de datos de varios grupos demográficos garantizan que los sistemas de IA se entrenen con un conjunto de datos representativo, lo que reduce la probabilidad de resultados sesgados. La diversidad aporta diferentes perspectivas y perspectivas, lo que ayuda a identificar y mitigar los sesgos durante el proceso de desarrollo.

Algoritmos sensibles al sesgo

Diseño de algoritmos con mecanismos incorporados para detectar y mitigar sesgos. Esto puede implicar la incorporación de restricciones de equidad en el proceso de optimización del algoritmo o el uso de técnicas como el entrenamiento de adversarios para minimizar los sesgos en las predicciones del modelo.

Seguimiento y evaluación periódicos

Haz clic en las cartas para darles la vuelta

Transparencia y explicabilidad

Haz clic en las cartas para darles la vuelta

la equidad de los resultados de IA.

Estrategias para abordar y mitigar los sesgos

Representación diversa

Los diversos puntos de datos de varios grupos demográficos garantizan que los sistemas de IA se entrenen con un conjunto de datos representativo, lo que reduce la probabilidad de resultados sesgados. La diversidad aporta diferentes perspectivas y perspectivas, lo que ayuda a identificar y mitigar los sesgos durante el proceso de desarrollo.

Seguimiento y evaluación periódicos

Supervisión continua de los sistemas de IA en entornos del mundo real para identificar los sesgos a medida que surgen. La evaluación periódica implica analizar los resultados de la IA en busca de disparidades entre diferentes grupos demográficos y actualizar algoritmos o conjuntos de datos para abordar con prontitud los sesgos identificados.

Algoritmos sensibles al sesgo

Diseño de algoritmos con mecanismos incorporados para detectar y mitigar sesgos. Esto puede implicar la incorporación de restricciones de equidad en el proceso de optimización del algoritmo o el uso de técnicas como el entrenamiento de adversarios para minimizar los sesgos en las predicciones del modelo.

Transparencia y explicabilidad

Haz clic en las cartas para darles la vuelta

la equidad de los resultados de IA.

Estrategias para abordar y mitigar los sesgos

Representación diversa

Los diversos puntos de datos de varios grupos demográficos garantizan que los sistemas de IA se entrenen con un conjunto de datos representativo, lo que reduce la probabilidad de resultados sesgados. La diversidad aporta diferentes perspectivas y perspectivas, lo que ayuda a identificar y mitigar los sesgos durante el proceso de desarrollo.

Algoritmos sensibles al sesgo

Diseño de algoritmos con mecanismos incorporados para detectar y mitigar sesgos. Esto puede implicar la incorporación de restricciones de equidad en el proceso de optimización del algoritmo o el uso de técnicas como el entrenamiento de adversarios para minimizar los sesgos en las predicciones del modelo.

Seguimiento y evaluación periódicos

Supervisión continua de los sistemas de IA en entornos del mundo real para identificar los sesgos a medida que surgen. La evaluación periódica implica analizar los resultados de la IA en busca de disparidades entre diferentes grupos demográficos y actualizar algoritmos o conjuntos de datos para abordar con prontitud los sesgos identificados.

Transparencia y explicabilidad

Proporcionar transparencia y explicaciones para las decisiones de IA permite a las partes interesadas comprender cómo se toman las decisiones e identificar cualquier sesgo presente en el sistema. Las técnicas de IA explicables, como proporcionar puntuaciones de importancia de características o fundamentos de decisiones, permiten a los usuarios interpretar y evaluar la equidad de los resultados de IA.

DESCRIPCIÓN DE LAS MÉTRICAS DE EQUIDAD

FUENTE DE LA IMAGEN | Generado por DALL-E

Descripción de las métricas de equidad

- Paridad estadística +
- Cuotas igualadas +
- Disparidad demográfica condicional +
- Equidad a través de la concienciación +
- Equidad individual +
- Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística

La paridad estadística también se conoce como paridad demográfica, esta métrica evalúa si la proporción de resultados positivos (por ejemplo, aprobaciones de préstamos u ofertas de trabajo) es la misma en diferentes grupos demográficos.

Cuotas igualadas +

Disparidad demográfica condicional +

Equidad a través de la concienciación +

Equidad individual +

Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística +

Cuotas igualadas

Las probabilidades igualadas examinan si la tasa de verdaderos positivos (sensibilidad) y la verdadera tasa negativa (especificidad) son iguales en diferentes grupos demográficos. Garantiza que el rendimiento predictivo del modelo sea coherente en todos los grupos.

Disparidad demográfica condicional +

Equidad a través de la concienciación +

Equidad individual +

Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística +

Cuotas igualadas +

Disparidad demográfica condicional:

La disparidad demográfica condicional evalúa la disparidad en los resultados predictivos condicionados a un atributo protegido. Mide la diferencia en los resultados de la predicción entre diferentes grupos demográficos, teniendo en cuenta otros factores relevantes.

Equidad a través de la concienciación +

Equidad individual +

Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística +

Cuotas igualadas +

Disparidad demográfica condicional +

Equidad a través de la concienciación

La equidad a través de la conciencia incorpora la conciencia de atributos sensibles (por ejemplo, raza o género) en el proceso de toma de decisiones. Garantiza que las decisiones tomadas por los sistemas de IA sean justas e imparciales, teniendo en cuenta el contexto histórico y los impactos sociales.

Equidad individual +

Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística +

Cuotas igualadas +

Disparidad demográfica condicional +

Equidad a través de la concienciación +

Equidad individual

La equidad individual se centra en garantizar que personas similares reciban un tratamiento o resultados similares del sistema de IA, independientemente de sus características demográficas. Su objetivo es minimizar el trato desigual basado en atributos irrelevantes.

Equidad contrafáctica +

Descripción de las métricas de equidad

Paridad estadística +

Cuotas igualadas +

Disparidad demográfica condicional +

Equidad a través de la conciencia +

Equidad individual +

Equidad contrafáctica

La equidad contrafáctica evalúa si cambiar el atributo sensible de un individuo (por ejemplo, raza o género) mientras se mantienen constantes otras características resultaría en un cambio en el resultado de la decisión. Garantiza que las personas sean tratadas de manera justa, independientemente de sus atributos protegidos.

Descripción de las métricas de equidad

Cuestionario interactivo

Creación de un cuestionario interactivo en Mentimeter

Pregunta de la encuesta

Antes de empezar, ¿cuáles son los factores importantes que hay que tener en cuenta a la hora de determinar si un sistema de IA es justo?

Opciones

a) Equidad de resultados; b) Representación demográfica; c) Igualdad de trato; d) Transparencia.

Descripción de las métricas de equidad

Ejercicios interactivos en clase

Tema: Paridad estadística

Calcular la paridad estadística

- Teniendo en cuenta los datos, calcule la tasa de aprobación de préstamos para cada carrera. ¿Hay paridad?
- Proporcione a los participantes los números (o porcentajes) de aprobaciones y rechazos para cada carrera y pídales que calculen si las tasas son aproximadamente iguales.
- Respuestas de opción múltiple: "Las tarifas son iguales; se logra la paridad" y "Las tarifas no son iguales; Hay una disparidad".

Descripción de las métricas de equidad

Ejercicios interactivos en clase

Tema: Cuotas igualadas

Evalúe las probabilidades igualadas

- Suponiendo que la tasa positiva verdadera (préstamo correctamente aprobado para aquellos que deberían ser aprobados) para los blancos es del 80% y para los negros es del 60%, ¿hay probabilidades igualadas?
- Respuestas de opción múltiple: "Sí, las probabilidades son iguales" o "No, las probabilidades no son iguales".

EQUIDAD EN LA RECOPILACIÓN DE DATOS Y EL PROCESAMIENTO PREVIO

FUENTE DE LA IMAGEN | Generado por DALL-E

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Haz clic en las cartas para darles la vuelta

Muestreo inclusivo

Haz clic en las cartas para darles la vuelta

Limpieza de datos

Haz clic en las cartas para darles la vuelta

Detección de sesgos

Haz clic en las cartas para darles la vuelta

Ingeniería de características justas

Haz clic en las cartas para darles la vuelta

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Haz clic en las cartas para darles la vuelta

minoritarios para abordar los

Limpieza de datos

Haz clic en las cartas para darles la vuelta

los formatos de datos.

Detección de sesgos

Haz clic en las cartas para darles la vuelta

grupos demográficos.

Ingeniería de características justas

Haz clic en las cartas para darles la vuelta

discriminatorias.

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Haz clic en las cartas para darles la vuelta

los formatos de datos.

Detección de sesgos

Haz clic en las cartas para darles la vuelta

grupos demográficos.

Ingeniería de características justas

Haz clic en las cartas para darles la vuelta

discriminatorias.

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Limpie cuidadosamente y preprocese los datos para eliminar cualquier sesgo o inconsistencia. Esto incluye la identificación y corrección de errores, el manejo de los valores faltantes y la estandarización de los formatos de datos.

Detección de sesgos

Haz clic en las cartas para darles la vuelta

grupos demográficos.

Ingeniería de características justas

Haz clic en las cartas para darles la vuelta

discriminatorias.

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Limpie cuidadosamente y preprocese los datos para eliminar cualquier sesgo o inconsistencia. Esto incluye la identificación y corrección de errores, el manejo de los valores faltantes y la estandarización de los formatos de datos.

Detección de sesgos

Implementar técnicas para detectar y mitigar sesgos en los datos. Esto puede implicar la realización de auditorías de sesgo o el uso de herramientas automatizadas para identificar disparidades entre diferentes grupos demográficos.

Ingeniería de características justas

Haz clic en las cartas para darles la vuelta

discriminatorias.

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Limpie cuidadosamente y preprocese los datos para eliminar cualquier sesgo o inconsistencia. Esto incluye la identificación y corrección de errores, el manejo de los valores faltantes y la estandarización de los formatos de datos.

Detección de sesgos

Implementar técnicas para detectar y mitigar sesgos en los datos. Esto puede implicar la realización de auditorías de sesgo o el uso de herramientas automatizadas para identificar disparidades entre diferentes grupos demográficos.

Ingeniería de características justas

Considere las implicaciones de la selección de características y la ingeniería en la equidad. Evite el uso de funciones que puedan perpetuar estereotipos o discriminación y esfuércese por obtener funciones que sean relevantes y no discriminatorias.

Transparencia y documentación

Haz clic en las cartas para darles la vuelta

características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Limpie cuidadosamente y preprocese los datos para eliminar cualquier sesgo o inconsistencia. Esto incluye la identificación y corrección de errores, el manejo de los valores faltantes y la estandarización de los formatos de datos.

Detección de sesgos

Implementar técnicas para detectar y mitigar sesgos en los datos. Esto puede implicar la realización de auditorías de sesgo o el uso de herramientas automatizadas para identificar disparidades entre diferentes grupos demográficos.

Ingeniería de características justas

Considere las implicaciones de la selección de características y la ingeniería en la equidad. Evite el uso de funciones que puedan perpetuar estereotipos o discriminación y esfuércese por obtener funciones que sean relevantes y no discriminatorias.

Transparencia y documentación

Documentar los procedimientos de recopilación y preprocesamiento de datos para proporcionar transparencia y responsabilidad. Esto incluye documentar cualquier decisión tomada durante la limpieza de datos o la ingeniería de características que pueda afectar a la equidad.

Seguimiento periódico

Haz clic en las cartas para darles la vuelta

preprocesamiento para garantizar la equidad a lo largo del tiempo.

Equidad en la recopilación de datos y el preprocesamiento

Diversas fuentes de datos

Recopilar datos de diversas fuentes para garantizar una representación completa de la población. Esto ayuda a prevenir la infrarrepresentación o el sesgo hacia grupos demográficos específicos.

Muestreo inclusivo

Utilice técnicas de muestreo inclusivas para garantizar que todos los segmentos de la población estén adecuadamente representados en el conjunto de datos. Esto puede implicar un muestreo estratificado o un sobremuestreo de grupos minoritarios para abordar los

s.

Limpieza de datos

Limpie cuidadosamente y preprocese los datos para eliminar cualquier sesgo o inconsistencia. Esto incluye la identificación y corrección de errores, el manejo de los valores faltantes y la estandarización de los formatos de datos.

Detección de sesgos

Implementar técnicas para detectar y mitigar sesgos en los datos. Esto puede implicar la realización de auditorías de sesgo o el uso de herramientas automatizadas para identificar disparidades entre diferentes grupos demográficos.

Ingeniería de características justas

Considere las implicaciones de la selección de características y la ingeniería en la equidad. Evite el uso de funciones que puedan perpetuar estereotipos o discriminación y esfuércese por obtener funciones que sean relevantes y no discriminatorias.

Transparencia y documentación

Documentar los procedimientos de recopilación y preprocesamiento de datos para proporcionar transparencia y responsabilidad. Esto incluye documentar cualquier decisión tomada durante la limpieza de datos o la ingeniería de características que pueda afectar a la equidad.

Seguimiento periódico

Supervise continuamente la recopilación de datos y el proceso de preprocesamiento para identificar y abordar cualquier sesgo o disparidad emergente. Actualice periódicamente el conjunto de datos y las técnicas de preprocesamiento para garantizar la equidad a lo largo del tiempo.

TENDENCIAS FUTURAS Y NUEVAS CUESTIONES

FUENTE DE LA IMAGEN | Generado por DALL-E

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Haz clic en las cartas para darles la vuelta

Ataques y defensas adversarias

Haz clic en las cartas para darles la vuelta

Equidad en múltiples dimensiones

Haz clic en las cartas para darles la vuelta

Mitigación de sesgos algorítmicos

Haz clic en las cartas para darles la vuelta

Consideración ética y gobernanza

Haz clic en las cartas para darles la vuelta

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

A medida que los sistemas de IA se vuelven más complejos, existe una creciente demanda de transparencia e interpretabilidad. Las tendencias futuras pueden centrarse en el desarrollo de técnicas y estándares para explicar las decisiones de IA y hacerlas más interpretables para las partes interesadas.

Ataques y defensas adversarias

Haz clic en las cartas para darles la vuelta

Equidad en múltiples dimensiones

Haz clic en las cartas para darles la vuelta

Mitigación de sesgos algorítmicos

Haz clic en las cartas para darles la vuelta

Consideración ética y gobernanza

Haz clic en las cartas para darles la vuelta

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Los ataques adversarios, en los que los actores maliciosos manipulan los sistemas de IA alterando sutilmente los datos de entrada, suponen una amenaza importante para la equidad y la solidez. Las investigaciones futuras pueden explorar defensas sólidas y contramedidas contra los ataques adversarios para mejorar la

Equidad en múltiples dimensiones

Haz clic en las cartas para darles la vuelta

Mitigación de sesgos algorítmicos

Haz clic en las cartas para darles la vuelta

Consideración ética y gobernanza

Haz clic en las cartas para darles la vuelta

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Equidad en múltiples dimensiones

La equidad en la IA va más allá de los atributos demográficos y puede abarcar otras dimensiones como el estatus socioeconómico, la discapacidad o el dominio del idioma. Las tendencias futuras pueden implicar el desarrollo de métricas y algoritmos de equidad multidimensionales para abordar los sesgos interseccionales.

Mitigación de sesgos algorítmicos

Haz clic en las cartas para darles la vuelta

Consideración ética y gobernanza

Haz clic en las cartas para darles la vuelta

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Equidad en múltiples dimensiones

Mitigación de sesgos algorítmicos

Los avances en las técnicas algorítmicas de mitigación de sesgos seguirán siendo un punto focal, y la investigación se centrará en el desarrollo de algoritmos y modelos más justos que reduzcan los sesgos en diversos contextos de toma de decisiones.

Consideración ética y gobernanza

Haz clic en las cartas para darles la vuelta

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Equidad en múltiples dimensiones

Mitigación de sesgos algorítmicos

Consideración ética y gobernanza

Las implicaciones éticas de la equidad y el sesgo de la IA seguirán siendo un área clave de preocupación, con un énfasis cada vez mayor en las prácticas éticas de desarrollo de la IA y los marcos de gobernanza. Las tendencias futuras pueden implicar el establecimiento de directrices y normas reglamentarias para garantizar la equidad y la rendición de cuentas en los sistemas de IA.

Diseño centrado en el ser humano

Haz clic en las cartas para darles la vuelta

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Equidad en múltiples dimensiones

Mitigación de sesgos algorítmicos

Consideración ética y gobernanza

Diseño centrado en el ser humano

Los futuros sistemas de IA pueden priorizar los principios de diseño centrados en el ser humano, incorporando los comentarios y preferencias de los usuarios para mejorar la equidad y la usabilidad. Esto podría implicar el diseño conjunto de sistemas de IA con diversas partes interesadas para garantizar la inclusión y la equidad.

Equidad en las tecnologías emergentes

Haz clic en las cartas para darles la vuelta

Tendencias futuras y nuevas cuestiones

Explicabilidad e interpretabilidad

Ataques y defensas adversarias

Equidad en múltiples dimensiones

Mitigación de sesgos algorítmicos

Consideración ética y gobernanza

Diseño centrado en el ser humano

Equidad en las tecnologías emergentes

A medida que la IA continúa cruzándose con tecnologías emergentes como los vehículos autónomos, la IA sanitaria y los sistemas de reconocimiento facial, abordar la equidad y el sesgo se vuelve cada vez más importante. Las tendencias futuras pueden implicar la adaptación de los principios de equidad a las nuevas aplicaciones y avances tecnológicos.

CONCLUSIÓN

FUENTE DE LA IMAGEN | Generado por DALL-E

CONCLUSIÓN

1. **Comprender la complejidad: La equidad y el sesgo de la IA son desafíos multifacéticos que requieren estrategias integrales y atención continua.**
2. **Estrategias de acción: desde la representación diversa en la recopilación de datos hasta la mitigación de sesgos algorítmicos, se dispone de una serie de estrategias para abordar y mitigar los sesgos en los sistemas de IA.**
3. **Anticiparse a los desafíos emergentes: a medida que la IA continúa evolucionando, será crucial mantenerse a la vanguardia de las tendencias y problemas emergentes, como los ataques adversarios, la equidad multidimensional y las consideraciones éticas.**
4. **Hacia un futuro más equitativo: al priorizar el diseño centrado en el ser humano, la transparencia y la responsabilidad, podemos trabajar para crear sistemas de IA que promuevan la equidad, la inclusión y la confianza.**

GRACIAS

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.

2022-1-ES01-KA220-HED-000085257