

Fundamentos algorítmicos y ética en IA: del curso de teoría a práctica

Kit de herramientas para sesiones sincrónicas

UC4 | Equidad de datos y sesgo
Estudio de caso

Estudio de caso



Estudio de casos

Descripción para Profesores

	Descripción	Comentarios
Tarea descripción	Esta sesión tiene como objetivo ayudar a los estudiantes a comprender los fundamentos de la privacidad y la conveniencia, especialmente en actividades cotidianas como dispositivos inteligentes para el hogar, aplicaciones móviles y compras en línea. Incluirá discusiones interactivas, reflexiones y actividades para reforzar el aprendizaje.	
Descripción de cómo realizar la tarea	Presentar el estudio de caso.	Consulte las instrucciones y directrices en las siguientes páginas. Tiempo estimado para realizar la tarea
30-40	minutos. Sugerencia de fuentes para realizar la tarea	
Búsqueda web sobre incidentes relacionados con violaciones de privacidad o filtraciones de datos.	Descripción detallada de cómo entregar la tarea Discusión en grupo + Formularios de Google. Consulte las preguntas sugeridas para los formularios de Google en el documento: UC4_CaseStudy_Tentative_questions_for_google_forms.docx	
Información sobre la fecha límite para la entrega de la tarea	Durante la sesión sincrónica. Información de contacto o cómo aclarar dudas El maestro debe proporcionar una forma de contacto	(puede ser una dirección de correo electrónico, un número de teléfono...)
UC4 Justicia y sesgo de datos	2	
Estudio de caso	Descripción para capacitadores (it could be an email address, a telephone number...)	

Estudio de caso

Caso ficticio para estudiantes

Estudio de caso: Equidad y sesgo de IA en sistemas educativos

Un gran distrito escolar público implementó una herramienta impulsada por IA para ayudar con las evaluaciones de los estudiantes, que incluía predecir el rendimiento estudiantil, identificar a los estudiantes que podrían necesitar ayuda adicional e incluso contribuir a las decisiones sobre admisiones universitarias. La herramienta analizaba una amplia variedad de datos, incluyendo el rendimiento académico pasado, los registros de asistencia, el contexto socioeconómico y la participación en actividades extracurriculares.

Después de un año de implementación, se descubrió que el sistema de IA etiquetaba desproporcionadamente a estudiantes de ciertos antecedentes demográficos (por ejemplo, estudiantes de menor estatus socioeconómico o de ciertas minorías raciales y étnicas) como en riesgo de bajo rendimiento, incluso cuando su rendimiento real estaba a la par o mejor que el de sus compañeros. Estos estudiantes a menudo eran recomendados para programas de recuperación o se les negaban oportunidades para cursos avanzados y actividades extracurriculares, lo que tenía implicaciones a largo plazo para sus trayectorias académicas y perspectivas de admisión a la universidad.

Puntos de discusión

- ¿Deben los beneficios de la IA en educación superar los riesgos de sesgo? ¿Cómo encontramos el equilibrio adecuado?
- ¿Cómo afecta el sesgo de datos los resultados de la IA para los estudiantes? ¿Qué pasos pueden tomar las escuelas para reducir el sesgo?
- ¿Por qué es crucial que las decisiones de IA en educación sean transparentes? ¿Cómo pueden las escuelas hacer que la IA sea más comprensible?
- ¿Deben los educadores humanos tener siempre la última palabra en decisiones impulsadas por IA? ¿Cuáles son los riesgos de la automatización total?



FUENTE DE IMAGEN
<https://plat.ai/blog/bias-and-fairness-in-ai-algorithms/>

Estudio de caso

Consejos para maestros

La IA ha penetrado cada vez más en el sector educativo, desde plataformas de aprendizaje personalizado hasta sistemas de admisión y evaluaciones de estudiantes. Si bien la IA tiene el potencial de mejorar los resultados educativos, hay crecientes preocupaciones sobre la equidad y el sesgo, especialmente en relación con cómo estas tecnologías impactan a los jóvenes y estudiantes. Este estudio de caso explorará un escenario donde se encontró que un sistema impulsado por IA en un contexto educativo era sesgado, las consecuencias de este sesgo y los pasos tomados para abordarlo.

1

Cuestiones clave

- **Sesgo de datos:** el sistema de IA fue entrenado con datos históricos que reflejaban sesgos existentes en el sistema educativo, como la sobrerrepresentación de ciertos demographics en programas de remediación.
- **Falta de transparencia:** los maestros y administradores no pudieron entender o cuestionar las recomendaciones de la IA porque el proceso de toma de decisiones era opaco.
- **Impacto en los estudiantes:** los estudiantes que fueron evaluados erróneamente como de bajo rendimiento enfrentaron estigmatización y perdieron oportunidades de crecimiento, afectando su confianza y perspectivas.
- **Preocupaciones éticas:** el uso de sistemas de IA sesgados planteó preguntas éticas sobre la equidad, la igualdad de oportunidades y la responsabilidad de las instituciones educativas.

2

Acciones tomadas

Una vez que se identificó el sesgo, el distrito escolar tomó los siguientes pasos:

- **Auditar y volver a entrenar el modelo** : el sistema de IA fue auditado para identificar y eliminar datos sesgados y volver a entrenar el modelo con un conjunto de datos más equilibrado que reflejara mejor la diversa población estudiantil.
- **Supervisión humana** : se reintrodujeron educadores humanos en el ciclo de toma de decisiones, con el sistema de IA proporcionando recomendaciones en lugar de determinaciones. Esto permitió a los maestros usar su juicio y conocimiento de los estudiantes para tomar decisiones finales.
- **Transparencia y explicabilidad** : el distrito escolar trabajó con expertos en IA para mejorar la transparencia del sistema de IA, asegurando que los maestros y administradores pudieran entender cómo se tomaban las decisiones y proporcionando a los estudiantes y padres explicaciones para las evaluaciones y recomendaciones.
- **Monitoreo continuo** : se estableció un sistema para el monitoreo continuo del rendimiento de la IA para asegurar que los sesgos no volvieran a surgir, y que el sistema continuara sirviendo a todos los estudiantes de manera justa.

Conclusión

El sistema de IA reentrenado, combinado con la supervisión humana, redujo significativamente el número de estudiantes que eran categorizados de manera injusta. Con el tiempo, esto llevó a resultados educativos más equitativos, permitiendo que más estudiantes sobresalieran. El caso también provocó una conversación más amplia dentro del distrito sobre el uso responsable de la IA en la educación y la importancia de la equidad y la transparencia en la toma de decisiones algorítmicas.

Estudio de caso

Guía del instructor para enseñar el estudio de caso

Objetivo

Esta sesión tiene como objetivo ayudar a los estudiantes a comprender los fundamentos de la privacidad y la conveniencia, especialmente en actividades diarias como dispositivos domésticos inteligentes, aplicaciones móviles y compras en línea. Incluirá discusiones interactivas, reflexiones y actividades para reforzar el aprendizaje.

Consejos generales para maestros

- **Involucrar a los** estudiantes
- **: mantén la sesión interactiva y fomenta la participación.** Relación con la vida real : u
- **sa ejemplos y escenarios de la vida real con los que los estudiantes puedan identificarse.** Sé adaptable
- **: ajusta la sesión según las respuestas y niveles de participación de los estudiantes.** Fomenta el pensamiento crítico
- **: motiva a los estudiantes a pensar críticamente sobre las compensaciones entre privacidad y conveniencia.** Mantente actualizado

Estudio de caso

Guía del instructor para enseñar el estudio de caso

1

Preparación

- Materiales necesarios:
 - Diapositivas de presentación con el estudio de caso
 - Folletos de escenarios del estudio de caso
 - Enlace de Google Forms con preguntas de actividad/evaluación
- Configuración:
 - Organizar los asientos para facilitar las discusiones en grupo.
 - Compartir el enlace de Google Forms con los estudiantes antes/durante el inicio de la sesión.

2

Esquema de la sesión

- Explicar el objetivo de la sesión: explorar cómo la tecnología impacta nuestra privacidad y cómo podemos equilibrar la conveniencia con la protección de la privacidad.
- Actividad de hielo: pedir a los estudiantes que compartan un dispositivo tecnológico que usan a diario y cómo hace sus vidas más fáciles.
- Presentar el Estudio de Caso y discutir la conveniencia y los problemas de privacidad enfrentados.
- **Explicar la experiencia de John con las compras personalizadas y los anuncios dirigidos, y las preocupaciones sobre la privacidad que surgieron.** Dividir a los estudiantes en grupos pequeños de 3-4 y pedir a cada grupo que identifique las principales preocupaciones de privacidad en el estudio de caso asignado.
- Pedirles que discutan posibles soluciones o acciones para mitigar estos riesgos de privacidad.
 - Animarles a pensar sobre el equilibrio entre la conveniencia y la privacidad.
 - Compartir el enlace de Google Forms con los estudiantes.
- Hacer que los estudiantes completen las preguntas individualmente o en parejas.
 - Las preguntas incluirán completar el espacio en blanco, opción múltiple, múltiples respuestas correctas y reflexiones personales.
 - Discusiones basadas en escenarios:
- Presentar escenarios cortos y pedir a los estudiantes que reflexionen sobre lo que harían.
 - Ejemplos: “¿Qué pasos tomarías si descubres que tu dispositivo inteligente estaba escuchando conversaciones privadas?” o “¿Cómo manejarías una notificación de violación de datos de una aplicación de fitness que usas?”

• 6

Estudio de caso

Guía del instructor para enseñar el estudio de caso

3

Reflexión y conclusión

- **Auto-reflexión**

- Pida a los estudiantes que escriban un párrafo corto sobre cómo pueden equilibrar la privacidad y la conveniencia en su vida diaria.

- **Conclusiones clave**

- Resuma los puntos principales discutidos durante la sesión.
- Enfatice la importancia de proteger la privacidad mientras se disfrutan de las comodidades tecnológicas.

GRACIAS

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



El apoyo de la Comisión Europea a la producción de esta publicación no constituye una aprobación de los contenidos, que reflejan únicamente las opiniones de los autores, y la Comisión no puede ser considerada responsable de cualquier uso que se pueda hacer de la información contenida en la misma.

