

A stylized illustration of a person's head and shoulders. The person has long, dark hair and is wearing a glowing blue headband. Below the headband, there is a grid pattern of blue lines radiating downwards, suggesting a digital or futuristic theme.

Fomentando la Inteligencia Artificial en las Escuelas

Guía de inclusión para educadores (no formales) en el contexto de la educación en Inteligencia Artificial

Tabla de contenidos

Introducción	4
Sobre la guía	5
Objetivos de aprendizaje	5
1. Desarrollar programas de educación no formal inclusivos sobre inteligencia artificial	6
1.1. Las ventajas de aprender sobre IAen un entorno no formal	7
1.2. Promoción del género, la diversidad y la inclusión en el entorno de aprendizaje	8
1.3. Prácticas docentes inclusivas en la educación no formal	10
2. Diversidad e inclusión en inteligencia artificial	14
2.1. Inteligencia artificial: ¿qué es?	15
2.2. Sesgo en inteligencia artificial	17
2.3. ¿Por qué el género, la diversidad y la inclusión son importantes para la IA?	23
2.4. Modelos femeninos a seguir que cambian el campo de la IA	24
3. Directrices para actividades que promueven un enfoque inclusivo de la inteligencia artificial	28
3.1 Actividad 1: El juego de etiquetas	30
3.2 Actividad 2: Dibujar la descripción	33
3.3 Actividad 3: Adivina a la persona	37
3.4 Actividad 4: Sesgo de género usando LearningML con conjunto de datos 1	41
3.5 Actividad 5: Sesgo de género usando LearningML con el conjunto de datos 2	47
3.6. Actividad 6: Identificar el sesgo en la IA	52
Conclusiones	57
Definiciones y glosario	59
Referencias	60
Créditos de imágenes	62

Impresión

2023: CollectiveUP info@collectiveup.be

Editores: CollectiveUP (BE), King Juan Carlos University (SP), Teatro Circo Braga (PT), Vrije Universiteit Brussel (BE).

Autores: Liliana Carrillo, Chrysanthi Katrini, Paul Fenton, Loredana Bucseneanu (CollectiveUP), Marjon Blondeel (Vrije Universiteit Brussel), Meritxell Diaz, Antonio Romero, Gregorio Robles (Universidad Rey Juan Carlos), Sara Borges, Joana Miranda (Teatro Circo de Braga).

Maquetación: Jurate Laugalyte (CollectiveUP), Carlos Nivia Otero (CollectiveUP).

Traducido por: Frederick Ducatelle (holandés), Ana Oliveira Monteiro (portugués), Gregorio Gobles y Meritxell Díaz Coque (español).

Sugerimos citar este informe de la siguiente manera: Carrillo, L., Katrini, C., Fenton, P., ... (2023). *Guía de inclusión para educadores (no formales) en el contexto de la educación en Inteligencia Artificial.* https://www.agileforcollaboration.eu/sdc_download/720/?key=shx1aegquebhqkkehrudj7xn0yc8yo

Esta publicación ha sido escrita dentro del proyecto Erasmus + Fostering AI at Schools, que tiene como objetivo apoyar a los educadores para integrar y utilizar conceptos de Inteligencia Artificial en sus lecciones creando y compartiendo recursos sobre IA que puedan usarse en el aula.

CollectiveUP agradece al Digital Belgium Skills Fund (DBSF) por cofinanciar el proyecto AI4InclusiveEducation (www.ai4inclusiveeducation.be) con un número de subvención DBSF2023-14. DBSF cofinanció y apoyó parcialmente la realización del proyecto Erasmus+ FAIAS creando un impacto en Bélgica.

Esta publicación (en inglés, español, holandés y portugués) y otras publicaciones y herramientas producidas en el proyecto se pueden descargar gratuitamente en: www.fosteringai.net

Este trabajo está bajo una licencia Creative Commons Attribution-ShareAlike 4.0 International License. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/4.0/>



Renuncia de responsabilidad

Fostering AI at Schools es una asociación de proyectos, cofinanciada a través de la Comisión Europea. Esta publicación refleja únicamente las opiniones de los autores, y la Comisión Europea no se hace responsable del uso que pueda hacerse de la información contenida en ella.

Fomento de la IA en las escuelas – Proyecto Erasmus+, Tipo de acción «Asociaciones de cooperación en la educación escolar», Número de subvención: 2020-1-ES01-KA201-083047.



Co-funded by the
Erasmus+ Programme
of the European Union



ARTIFICIAL
INTELLIGENCE
RESEARCH GROUP



Introducción

En una era en la que la inteligencia artificial (IA) está cada vez más presente en nuestra vida cotidiana, el desarrollo ético y responsable de esta tecnología transformadora es de vital importancia (1). Fomentando la Inteligencia Artificial en las escuelas (Fostering Artificial Intelligence at Schools, FAIaS en inglés) es un proyecto innovador y transformador dedicado a fomentar una inteligencia artificial (IA) ética y responsable. Su objetivo último es mejorar el conocimiento de la IA en niños y jóvenes. Con ello, FAIaS aspira a cultivar una generación que no sólo conozca la IA, sino que también participe activamente en el pensamiento crítico y la toma de decisiones responsable. Con la creciente prevalencia de la IA en nuestra sociedad, este proyecto reconoce la urgente necesidad de abordar los complejos retos éticos y las implicaciones sociales, en particular la necesidad de que los jóvenes tengan más conocimientos y habilidades sobre el uso de la IA.

A través de artículos que invitan a la reflexión, estudios de casos y directrices prácticas, el proyecto explora diversas facetas de la ética de la IA y el impacto social de la IA (2) para su uso en la educación no formal. Al examinar ejemplos del mundo real y basarse en las experiencias de expertos, FAIaS ofrece valiosas ideas y marcos para orientar el desarrollo y la implantación responsables de los sistemas de IA.

Este documento sirve de guía y proporciona un marco para navegar por los complejos dilemas éticos que surgen a medida que la IA se integra cada vez más en diversos aspectos de nuestras vidas, como la sanidad, las finanzas y la educación, entre otros. "Guía inclusiva para educadores (no formales)" aborda temas como el sesgo algorítmico, la preocupación por la privacidad, la responsabilidad, la transparencia y el impacto de la IA en el mercado laboral. El libro aboga por un enfoque holístico del desarrollo de la IA, que priorice el bienestar y la capacitación de las personas y las comunidades, al tiempo que garantiza que los sistemas de IA se adhieran a los principios de imparcialidad, equidad y justicia.

Sobre la guía

El objetivo principal de esta guía es proporcionar un recurso para educadores no formales que deseen incorporar conceptos de inteligencia artificial (IA) en su currículo y enseñanza, con un enfoque aplicado sobre la diversidad e inclusión. De esta forma, complementa la “Guía para docentes de educación formal” desarrollada en el marco del proyecto “Fomentando la Inteligencia Artificial en las Escuelas” (FAlaS, en inglés).

Capítulo 1 Introduce el contexto de las oportunidades creadas por la educación no formal para aprender sobre la IA, se presentan una serie de conceptos relacionados con la promoción del género, la diversidad y la inclusión en el entorno de aprendizaje, y brinda ejemplos de prácticas docentes inclusivas en entornos no formales.

Capítulo 2 Profundiza en los desafíos de desarrollar sistemas de IA inclusivos, presentando el concepto de sesgo en la IA, sus causas, impacto y posibles soluciones. También analiza por qué el género, la diversidad y la inclusión son importantes al crear tecnologías de IA, y brinda ejemplos de modelos femeninos a seguir que están cambiando el campo de la IA.

Capítulo 3 Consta de guías prácticas para actividades en el aula que promueven un enfoque aplicado de la IA. Incluye una descripción e instrucciones paso a paso para cada actividad.

Finalmente, esta guía presenta un glosario de términos y otros recursos para ser utilizados en la práctica docente.

Objetivos de aprendizaje

- Conocer prácticas que promuevan la diversidad y la inclusión en el contexto educativo no formal.
- Comprender el papel de los educadores no formales en la promoción de un futuro más inclusivo y equitativo para la IA.
- Comprender los conceptos básicos de la IA y su impacto en la sociedad.
- Reconocer la importancia de la inclusión en la IA y sus aplicaciones.
- Desarrollar habilidades para diseñar y ofrecer lecciones de IA atractivas e inclusivas.
- Adquirir la capacidad de facilitar discusiones y debates sobre IA y sus implicaciones.
- Involucrar a los estudiantes para que se conviertan en aprendices críticos y reflexivos en el campo de la IA.



1

Desarrollar programas
de educación no formal
inclusivos sobre
inteligencia artificial

1.1. Las ventajas de aprender sobre inteligencia artificial en un entorno no formal

La IA tiene el potencial de revolucionar la forma en que vivimos, trabajamos y aprendemos. A medida que la IA continúa creciendo en popularidad e importancia, se vuelve cada vez más importante que las personas entiendan y puedan trabajar con la IA. Los programas de educación no formal pueden desempeñar un papel importante en la enseñanza de la IA a una amplia gama de personas, desde estudiantes y profesionales hasta aquellos que recién comienzan a explorar el mundo de la tecnología.

Una de las principales ventajas de la educación no formal es su flexibilidad y adaptabilidad. Los programas de educación no formal se pueden adaptar para satisfacer las necesidades e intereses específicos de los estudiantes individuales, lo que los convierte en una plataforma ideal para enseñar IA. Por ejemplo, los cursos y talleres en línea pueden diseñarse para brindar experiencias prácticas con la IA, mientras que los programas basados en la comunidad pueden ayudar a las personas a desarrollar sus habilidades y conocimientos de IA en un entorno de apoyo y colaboración.

Otra ventaja de la educación no formal es su capacidad para llegar a comunidades subrepresentadas. La IA es un campo en rápido crecimiento y es importante garantizar que las personas de todos los orígenes y experiencias tengan acceso a la educación y la formación en IA. Los programas de educación no formal pueden ayudar a cerrar la brecha entre la educación formal y la informal, brindando oportunidades para que las personas que pueden no tener acceso a la educación tradicional o los programas de capacitación aprendan sobre la IA.

Para enseñar IA de manera efectiva en programas de educación no formal es importante centrarse tanto en los aspectos técnicos como éticos de la IA. Las habilidades técnicas, como la codificación y la programación, son esenciales para las personas que desean trabajar con IA. Sin embargo, también es importante enseñar sobre las implicaciones éticas y sociales de la IA, incluidas las cuestiones relacionadas con la privacidad, la protección de datos y el sesgo en los sistemas de IA. Al incluir aspectos tanto técnicos como éticos de la IA en los programas de educación no formal, las personas pueden desarrollar una comprensión integral de la IA y estar mejor preparados para tomar decisiones informadas sobre su uso.

Por lo tanto, los programas de educación no formal tienen el potencial de desempeñar un papel importante en la enseñanza de la IA a una amplia gama de personas y complementar los programas existentes en la educación formal. Aprovechando las ventajas de la educación no formal, como la flexibilidad y la adaptabilidad, y centrándose tanto en los aspectos técnicos como éticos de la

AI, las personas pueden desarrollar las habilidades y el conocimiento que necesitan para tener éxito en un panorama tecnológico que cambia rápidamente. Al brindar acceso a la educación y capacitación en IA, los programas de educación no formal pueden ayudar

1. Desarrollar programas de educación no formal inclusivos sobre inteligencia artificial

a las personas de todos los orígenes y experiencias a alcanzar su máximo potencial y contribuir al desarrollo de un mundo más inclusivo y equitativo (3).

Algunas de las formas clave en que la educación no formal puede apoyar el aprendizaje sobre tecnología e IA para los jóvenes incluyen:

- **Proporcionar experiencias prácticas:** los programas de educación no formal pueden brindar a los jóvenes experiencias prácticas con la tecnología, como talleres de codificación, programas de robótica y espacios de creación.
- **Fomento de la creatividad y la innovación:** al proporcionar a los alumnos una variedad de herramientas y recursos para experimentar y explorar con la tecnología, la educación no formal puede fomentar sus competencias de creatividad e innovación.
- **Apoyar a las comunidades subrepresentadas:** la educación no formal puede ayudar a abordar la brecha digital al proporcionar acceso a la tecnología y la educación tecnológica a las comunidades que normalmente están subrepresentadas en los campos tecnológicos.
- **Fomentar el aprendizaje a lo largo de toda la vida:** la educación no formal puede ayudar a los jóvenes a desarrollar un amor de por vida por aprender sobre tecnología y prepararlos para el panorama tecnológico que cambia rápidamente.
- **Desarrollo de habilidades prácticas:** los programas de educación no formal pueden brindar a los jóvenes habilidades prácticas en áreas como programación de computadoras, diseño web y desarrollo de aplicaciones, que pueden ayudarlos a desarrollar habilidades comercializables y prepararse para carreras tecnológicas.

En resumen, el papel de la educación no formal en el aprendizaje de la IA y la tecnología, en términos más generales, es proporcionar un entorno complementario, centrado en el alumno y de apoyo para ayudar a los jóvenes a desarrollar las habilidades y el conocimiento que necesitan para prosperar en la era digital (4).

1.2. Promoción del género, la diversidad y la inclusión en el entorno de aprendizaje

Promover el género, la diversidad y la inclusión en programas de educación no formal sobre IA es crucial para lograr una industria tecnológica más equitativa. La IA es un campo en rápido crecimiento y es esencial que las personas de todos los orígenes y experiencias tengan acceso a la educación y la formación en IA. Los programas de educación no formal pueden desempeñar un papel importante en la promoción del género, la diversidad y la inclusión en la IA al crear oportunidades para que las personas de comunidades subrepresentadas aprendan sobre la IA y desarrollen las habilidades que necesitan para tener éxito en este campo.

1. Desarrollar programas de educación no formal inclusivos sobre inteligencia artificial

Una de las principales ventajas de los programas de educación no formal es su flexibilidad y adaptabilidad. Esto permite que los programas se diseñen para satisfacer las necesidades e intereses específicos de los estudiantes individuales, incluidos los de comunidades subrepresentadas.

Para promover la diversidad y la inclusión en los programas de educación no formal sobre IA, es importante garantizar que los programas sean accesibles e inclusivos. Esto se puede lograr ofreciendo opciones de programas flexibles y asequibles, y brindando servicios de apoyo para garantizar que las personas de comunidades subrepresentadas puedan participar. También es importante asegurarse de que los programas sean impartidos por instructores que reflejen la diversidad de la comunidad y que tengan experiencia trabajando con personas de comunidades subrepresentadas.

Otro aspecto importante de la promoción de la diversidad y la inclusión en los programas de educación no formal sobre IA es centrarse tanto en los aspectos técnicos como éticos de la IA. Las habilidades técnicas, como la codificación y la programación, son esenciales para las personas que desean trabajar con IA. Sin embargo, también es importante enseñar sobre las implicaciones éticas y sociales de la IA, incluidas las cuestiones relacionadas con la privacidad, la protección de datos y el sesgo en los sistemas de IA. Al incluir aspectos tanto técnicos como éticos de la IA en los programas de educación no formal, las personas pueden desarrollar una comprensión integral de la IA y estar mejor preparados para tomar decisiones informadas sobre su uso (5).

Al crear un entorno de aprendizaje inclusivo, podemos mejorar la calidad de la educación para todos los alumnos. **Pero, ¿qué significa esto?**

Género, Diversidad e Inclusión en la educación se refiere al reconocimiento y valoración de las diversas identidades, experiencias y perspectivas de estudiantes, educadores y otros actores del sistema educativo. Implica crear un entorno de aprendizaje inclusivo y equitativo donde todos los individuos se sientan valorados y apoyados, independientemente de su género, raza, etnia, religión, orientación sexual, capacidad, edad o antecedentes socioeconómicos. Esto incluye abordar los problemas de discriminación, prejuicio y marginación, y la implementación de estrategias para apoyar el desarrollo académico y éxito personal de todos los estudiantes. Género, Diversidad e Inclusión en la educación también implica promover activamente la equidad y justicia social, y sistemas y estructuras desafiantes que perpetúan la desigualdad y la exclusión. (6)

1.3. Prácticas docentes inclusivas en la educación no formal

Como educador, su papel es clave para ayudar a los alumnos a construir una base sólida para su futuro. Promover un ambiente inclusivo significa reconocer las diferencias entre los estudiantes, así como sus similitudes. Significa asegurarse de que todos los estudiantes se sientan cómodos con quienes son y darles oportunidades para explorar sus intereses fuera de la escuela. Al hacerlo, ayudamos a crear un entorno más inclusivo para todos los involucrados. Aquí hay algunas maneras de llevarlo a cabo:

1. Asegúrese de ser consciente de sus propios prejuicios y estereotipos

Si se encuentra haciendo suposiciones sobre los estudiantes en función de su apariencia o comportamiento, dé un paso atrás y trate de verlos a través de una lente diferente. Haga un esfuerzo para comprobar sus suposiciones. Es posible que ni siquiera se dé cuenta de que ha hecho ciertas suposiciones sobre los estudiantes, pero esta es una buena manera de comenzar a tomar conciencia de sus propios prejuicios.

Todas las personas albergan creencias y actitudes sobre grupos de personas en función de su raza o etnia, sexo, peso corporal y otras características. Esas creencias y actitudes sobre los grupos sociales se conocen como sesgos. Los sesgos son creencias que no están fundadas en hechos conocidos sobre alguien o sobre un grupo particular de individuos (7).

¿Por qué existen los sesgos? En la mayoría de los casos, los sesgos se forman debido a la tendencia del cerebro humano a categorizar a las nuevas personas y la nueva información. Para aprender rápidamente, el cerebro conecta nuevas personas o ideas con experiencias pasadas. Una vez que la cosa nueva ha sido puesta en una categoría, el cerebro responde de la misma forma que lo haría con otras cosas en esa categoría.

Más precisamente:

- **Demasiada información:** nuestro mundo tiene tanta información que necesitamos atajos y trucos para seleccionar las partes importantes;
- **No hay suficiente significado:** para ello llenamos los vacíos en la información usando lo que creemos que ya sabemos;
- **Necesidad de actuar rápido:** esto se agrava por el hecho de que necesitamos evaluar y aplicar información, tratar de predecir el futuro y actuar sobre nuevos conocimientos en un margen de tiempo muy pequeño;
- **¿Qué debemos recordar?** ¡No podemos recordar todo, así que tratamos de retener lo que será importante en el futuro y de olvidarnos del resto!

1. Desarrollar programas de educación no formal inclusivos sobre inteligencia artificial

Estos son algunos ejemplos de sesgos que pueden albergar inconscientemente los educadores y llevarían a entorpecer el proceso educativo:

- **Sesgo racial:** un educador puede tener inconscientemente estereotipos negativos sobre ciertos grupos raciales y esto puede afectar su estilo de enseñanza y comportamiento hacia los estudiantes de estos grupos (8).
- **Sesgo de género:** un educador puede tener estereotipos de género que afecten a sus interacciones con estudiantes masculinos y femeninos. Por ejemplo, puede suponer que los niños son mejores en matemáticas y ciencias, mientras que las niñas son mejores en lenguaje y artes (9).
- **Sesgo de edad:** un educador puede tener prejuicios contra los estudiantes de diferentes edades. Por ejemplo, puede suponer que los estudiantes más jóvenes son inmaduros y carecen de la capacidad de concentrarse, mientras que los estudiantes mayores tienen mayor dificultad para entender conceptos nuevos (10).
- **Sesgo de capacidad:** un educador puede tener prejuicios contra los estudiantes con discapacidades o necesidades especiales, asumiendo que no pueden aprender o participar de la misma manera que otros estudiantes (11).
- **Sesgo socioeconómico:** un educador puede inconscientemente tener sesgos basados en el estatus socioeconómico de un estudiante, asumiendo que los estudiantes de familias de bajos ingresos son menos inteligentes o estaban menos motivados (12).

2. Tenga en cuenta el lenguaje que utiliza

Ser inclusivo significa usar palabras como "todos" y "no importa el qué" en lugar de frases como "niños y niñas" o "aquellos que se identifican como niños o niñas". Trate de evitar hacer suposiciones sobre la identidad de género preguntando a los estudiantes qué pronombres prefieren (por ejemplo: él/él, ella/ella, ellos/ellos) en lugar de utilizar las formas de masculino genérico.

3. Asegúrese de conocer las necesidades de todos los estudiantes

Esto significa tomarse el tiempo para hacer preguntas sobre cómo se sienten y qué quieren aprender durante la clase. Asegúrese de incluir a todos en las discusiones y actividades, así como brindar oportunidades para aquellos que se sientan excluidos por defecto.

4. Haga un esfuerzo para crear un espacio seguro para ellos

Esto puede significar establecer expectativas claras para el comportamiento y hacerlas cumplir de manera consistente. También significa ser consciente de lo que dice y hace frente a sus alumnos, así como asegurarse de no dejarlos de lado sin querer o hacerlos sentir incómodos.

5. Finalmente, si usted es un educador y tiene dificultades con alguno de estos pasos, es importante buscar apoyo

Muchas organizaciones tienen al menos un miembro del personal capacitado para ayudar a los educadores que luchan con problemas de diversidad. Si su organización no lo hace, intente buscar socios locales que brinden este tipo de apoyo (13).

Los programas de educación no formal tienen el potencial de desempeñar un papel vital en la promoción del género, la diversidad y la inclusión en la IA, tanto cuando se trata de crear un entorno de aprendizaje inclusivo como de enfatizar la importancia de la diversidad en el desarrollo de sistemas que usan IA. Aprovechando las ventajas de la educación no formal, como la flexibilidad y la adaptabilidad, y al centrarse tanto en los aspectos técnicos como éticos de la IA, las personas de comunidades subrepresentadas pueden desarrollar las habilidades y el conocimiento que necesitan para tener éxito en este campo de rápido crecimiento. Al brindar acceso a la educación y capacitación en IA, los programas de educación no formal pueden ayudar a construir una industria tecnológica más inclusiva y equitativa, donde las personas de todos los orígenes y experiencias puedan alcanzar su máximo potencial y contribuir al desarrollo de una industria más inclusiva y equitativa, redundando en un mundo más equitativo.

Las prácticas de enseñanza inclusivas son esenciales para crear un ambiente de aprendizaje de apoyo para todos los estudiantes. La inclusión significa que todos los estudiantes son valorados y respetados, independientemente de sus antecedentes, experiencias y habilidades. Este enfoque de la educación ayuda a promover la participación de los estudiantes, el éxito académico y el bienestar social y emocional.

Uno de los principios clave de la enseñanza inclusiva es reconocer y respetar la diversidad de los estudiantes en el aula. Esto incluye reconocer y valorar los diferentes orígenes, culturas y experiencias que los estudiantes aportan al entorno de aprendizaje. Al aceptar esta diversidad, los maestros pueden ayudar a los estudiantes a sentirse vistos y valorados, y pueden crear un sentido de comunidad en el desarrollo de las clases.

Otro aspecto importante de la enseñanza inclusiva es proporcionar adaptaciones para estudiantes con diversas necesidades. Esto podría incluir el uso de tecnología, como subtítulos o descripciones de audio, para apoyar a los estudiantes con discapacidades auditivas o visuales, o proporcionar métodos de evaluación alternativos para estudiantes con dificultades de aprendizaje. Los profesores también deben ser flexibles en sus enfoques de enseñanza, adaptando sus métodos para satisfacer las necesidades de los estudiantes individuales.

La enseñanza inclusiva también implica crear un ambiente de aprendizaje seguro y de apoyo, donde todos los estudiantes se sientan cómodos expresándose y haciendo preguntas. Esto significa crear una cultura en el aula libre de discriminación y acoso, y promover activamente el respeto y la aceptación de todos los estudiantes. Los maestros pueden lograr esto estableciendo expectativas claras de comportamiento, fomentando relaciones positivas entre los estudiantes y abordando cualquier incidente de intimidación o discriminación de manera rápida y efectiva.

1. Desarrollar programas de educación no formal inclusivos sobre inteligencia artificial

Además, las prácticas de enseñanza inclusiva implican involucrar a los estudiantes en el proceso de aprendizaje y alentarlos a asumir un papel activo en su propia educación. Esto significa brindar oportunidades para que los estudiantes participen en discusiones en clase, hagan preguntas y compartan sus ideas y perspectivas. Los maestros también deben alentar a los estudiantes a trabajar en colaboración, ayudándolos a desarrollar importantes habilidades sociales y de comunicación (14).

Estos son algunos ejemplos de prácticas de enseñanza inclusivas que se pueden utilizar en los programas de educación de IA:

- **Representación diversa en el material:** Incluir ejemplos de personas, culturas y perspectivas diversas en el material del curso y estudios de casos.
- **Accesibilidad:** Usar tecnología y métodos de aprendizaje alternativos para que el curso sea accesible para todos los estudiantes, independientemente de su capacidad.
- **Lenguaje inclusivo:** Evitar el lenguaje que pueda ser discriminatorio, como el lenguaje estereotípico o de género, y utilizar un lenguaje inclusivo en las discusiones y actividades de la clase.
- **Fomentar la diversidad de perspectivas:** Crear oportunidades para que los estudiantes compartan sus perspectivas y experiencias únicas, y fomentar la diversidad de perspectivas en las discusiones en clase.
- **Trabajo en grupo:** Animar a los estudiantes a trabajar en grupos diversos, ayudándolos a construir relaciones positivas y desarrollar importantes habilidades sociales y de comunicación.
- **Enfoques de enseñanza flexibles:** Ser flexible en los enfoques de enseñanza y adaptarse a las necesidades de los estudiantes individuales, en lugar de utilizar un enfoque único para todos.
- **Proporcionar contexto:** Incluir el contexto histórico y cultural en la enseñanza de la IA para ayudar a los estudiantes a comprender su desarrollo e impacto en diferentes comunidades.
- **Fomento de debates éticos:** Fomento de debates sobre las implicaciones éticas y sociales de la IA y cómo puede afectar a diferentes comunidades e individuos.
- **Brindar oportunidades para la reflexión:** Brindar oportunidades para que los estudiantes reflexionen sobre sus propios prejuicios y experiencias, y alentarlos a convertirse en defensores activos de la diversidad, la equidad y la inclusión en el campo de la IA.

Al incorporar estas prácticas de enseñanza inclusivas, los educadores pueden ayudar a los estudiantes a desarrollar una comprensión más diversa e inclusiva de la IA y equiparlos con las habilidades y el conocimiento necesarios para contribuir a un futuro más equitativo (15).



2

Diversidad e inclusión
en inteligencia artificial

2.1. Inteligencia artificial: ¿qué es?

Para el propósito de nuestra guía no profundizaremos en los aspectos técnicos de la IA, sino que nos centraremos en la percepción pública de la IA, cómo da forma al mundo que nos rodea y por qué es importante.

Información más detallada sobre la enseñanza de la IA, y los planes de lecciones se encuentra desarrollado en el documento “Guías para docentes en educación formal” realizado como parte del proyecto “Fomentando la inteligencia artificial en la Escuela”, accesible mediante este enlace: <https://fosteringai.github.io/project/result1/>

A La inteligencia artificial (IA) es una disciplina de la informática que trata de replicar y desarrollar la inteligencia humana y sus procesos implícitos a través de los ordenadores. El término fue acuñado en 1956 en la Conferencia de Dartmouth por el reconocido informático John McCarthy, aunque algunas primeras experiencias se remontan a poco después de la Segunda Guerra Mundial.

Hoy en día, la IA se considera una de las tecnologías clave de la última década, y quizás también de la siguiente. Actualmente es un “tema candente”: la cobertura de los medios y la discusión pública sobre la IA es casi imposible de evitar. Google, Facebook y Apple están desarrollando programas y asistentes inteligentes que prometen cambiar la forma en que trabajamos, jugamos y aprendemos. Mires donde mires, hay evidencia de ello: su smartphone escuche sus comandos y entiende lo que dice; los vehículos autónomos se están convirtiendo en una realidad; la robótica sigue revolucionando en las fábricas de todo el mundo. La computación en la nube ha hecho que la IA sea más accesible que nunca. Como resultado, los desarrolladores están creando una amplia gama de productos con IA en su núcleo: interactuar con los clientes mediante chatbots; ayudar a los negocios minoristas a predecir qué quieren los clientes y cuándo lo comprarán; manejar el diagnóstico médico automáticamente en minutos en lugar de días o semanas (16).

Aunque no hay acuerdo sobre una definición única de IA, sí hay consenso sobre cuatro enfoques principales: dos de ellos se basan en humanos (sistemas que piensan como ellos y sistemas que actúan como ellos) y dos en aspectos racionales (sistemas que piensan racionalmente y sistemas que actúan racionalmente).

Algunos ejemplos de sistemas de IA incluyen:

- Software de reconocimiento de voz, que puede comprender y transcribir palabras habladas;
- Sistemas de reconocimiento de imágenes, que pueden identificar objetos o personas en imágenes o vídeos;
- Procesamiento del lenguaje natural (NLP), que permite que las máquinas entiendan y generen un lenguaje similar al humano;
- Algoritmos de toma de decisiones, que pueden analizar datos y tomar decisiones basadas en ese análisis.

En la educación, la IA se utiliza de muchas maneras para mejorar la experiencia de aprendizaje y mejorar los resultados. Algunos ejemplos incluyen (17):

- **Sistemas de aprendizaje adaptativo:** utilizan IA para personalizar las experiencias de aprendizaje de cada alumno en función de sus puntos fuertes, débiles y estilo de aprendizaje. El sistema puede ajustar el contenido y al ritmo de aprendizaje para adaptarse mejor a las necesidades individuales del estudiante.
- **Sistemas de tutoría inteligente:** utilizan IA para brindar tutela personalizada y retroalimentación a los estudiantes mientras trabajan en tareas y problemas prácticos. Estos sistemas pueden ayudar a los estudiantes a identificar áreas en las que necesitan apoyo adicional y brindar orientación y recursos para ayudarlos a aprender de manera más efectiva.
- **Herramientas de traducción y aprendizaje de idiomas:** la IA se puede utilizar para traducir texto y voz de un idioma a otro, lo que facilita que los estudiantes aprendan y se comuniquen con personas que hablan otros idiomas.
- **Herramientas de calificación y evaluación:** la IA se puede usar para calificar tareas y evaluaciones, brindando a los estudiantes comentarios inmediatos y liberando a los educadores para que se concentren en tareas más complejas.
- **Asistentes virtuales:** los asistentes virtuales impulsados por IA pueden ayudar a los estudiantes a encontrar información, completar tareas y responder preguntas, liberando tiempo para que los educadores se concentren en la enseñanza y el aprendizaje.
- **Recomendaciones personalizadas:** la IA puede analizar el desempeño anterior y las preferencias de un estudiante para proporcionar recomendaciones sobre los materiales y recursos de aprendizaje que tienen más probabilidades de ser útiles y atractivos para ese estudiante.

La IA está moldeando cada vez más el mundo en el que vivimos, y es importante aprender sobre su impacto y sus posibles implicaciones, y cómo ser activo en el desarrollo de la tecnología de una manera que beneficie a toda la sociedad. Aquí hay algunas razones por las que es importante aprender sobre la IA:

- **Preparación para el futuro:** la IA se está integrando rápidamente en muchas industrias, y aprender sobre la IA puede ayudar a las personas a prepararse para futuras oportunidades laborales y garantizar que estén equipadas para prosperar en un ambiente laboral que cambia rápidamente.
- **Comprender su impacto:** la IA tiene el potencial de tener impacto significativo en la sociedad, tanto positivo como negativo. Al aprender sobre la IA, las personas pueden obtener una mejor comprensión de sus posibles implicaciones, incluido el posible cambio en los puestos de trabajo, las preocupaciones sobre la privacidad y las consideraciones éticas.
- **Tomar decisiones informadas:** la IA se usa cada vez más para tomar decisiones que afectan la vida de las personas, desde la atención médica hasta la justicia penal. Aprender sobre IA puede ayudar a las personas a tomar decisiones informadas sobre estas aplicaciones y abogar por un uso responsable y ético.

→ **Fomento de la innovación:** la IA tiene el potencial de impulsar la innovación tecnológica y crear nuevos productos y servicios. Al aprender sobre IA, las personas pueden comprender mejor sus aplicaciones potenciales y contribuir a dar forma a su desarrollo de manera que se alinee con sus valores y prioridades.

→ **Desarrollar habilidades de pensamiento crítico:** aprender sobre IA requiere pensamiento crítico y la capacidad de evaluar información compleja.

Aprender sobre la IA es esencial para que las personas comprendan su impacto y estén equipados para navegar por sus posibles implicaciones. Al invertir en educación sobre IA, las personas pueden asegurarse de estar preparadas para participar en la configuración de un futuro más equitativo y responsable (18).

2.2. Sesgo en inteligencia artificial

Los sistemas automatizados no son inherentemente neutrales. Los ordenadores son codificados por personas y, como consecuencia, reflejan prejuicios y preferencias de quienes codifican los algoritmos de IA.

El sesgo en la IA se refiere a la tendencia de los sistemas a exhibir ciertos prejuicios o estereotipos en sus procesos y resultados de toma de decisiones. Esto puede ocurrir debido a los datos que se usan para entrenar la IA, así como a los algoritmos y modelos que se usan para construir y operar el sistema. El sesgo de IA ocurre porque los seres humanos eligen los datos que usan los algoritmos y también deciden cómo se aplicarán los resultados de esos algoritmos. Sin pruebas exhaustivas y equipos diversos, es fácil que los sesgos inconscientes aparezcan en los modelos de aprendizaje automático. Luego, los sistemas de IA automatizan y perpetúan esos modelos sesgados.

Una forma común en que se introduce el sesgo en tales sistemas es debido al uso de **datos sesgados**. Por ejemplo, suponga que desea automatizar el proceso de decisión de si un solicitante es aceptado en su universidad o facultad. Suponga que en el pasado ha admitido principalmente a hombres. Si utiliza estos datos históricos para entrenar el nuevo sistema, es probable que se replique este sesgo de género.

Cuando se utilizan datos del mundo real, como artículos de noticias o publicaciones en redes sociales, es probable que los sistemas de IA refuerzen los prejuicios existentes. Las aplicaciones de traducción, por ejemplo, tienden a estar sesgadas contra las mujeres cuando se traducen de idiomas que usan pronombres neutros en cuanto al género. Ejemplos típicos son “ella cuida a los niños” y “él invierte” (19).

Al observar la clasificación de los datos, **la forma en que se recopilan los datos es muy importante. Tanto el submuestreo como el sobremuestreo pueden generar sesgos**. Un ejemplo de submuestreo es cuando los sistemas de detección de rostros se entranan principalmente en rostros masculinos blancos y, como consecuencia, no reconocen los rostros femeninos más oscuros tan bien como los rostros masculinos más claros. El sobremuestreo, por otro lado, puede conducir a una representación excesiva de ciertos grupos. Por ejemplo, los delitos cometidos en áreas frecuentadas por la policía

estarán más presentes en el conjunto de datos y el modelo de IA entrenado en los datos probablemente reflejará esta desproporción y estas áreas serán etiquetadas como más peligrosas.

Las personas también pueden **reforzar el sesgo en los modelos de IA ya implementados**. Por ejemplo, es posible que un sistema de IA esté mostrando anuncios de una tarjeta de crédito con un alto interés a personas con un nivel educativo más bajo. Es posible que estas hagan clic en este anuncio sin darse cuenta de que esto reforzará el sesgo ya existente y que seguirán recibiendo tales sugerencias.

El sesgo en la IA puede tener impacto negativo significativo en la sociedad y las personas. Algunas de las consecuencias del sesgo en la IA incluyen:

Resultados discriminatorios

Los sistemas de IA que están sesgados pueden producir resultados que perjudican injustamente a ciertos grupos de personas, como las mujeres o los grupos minoritarios. Esto puede conducir a la discriminación y la desigualdad en áreas como el empleo, los préstamos y la atención médica.

Algoritmo de anuncios de Facebook

En 2019, se descubrió que Facebook (20) contravenía la Constitución de los EE. UU. al permitir que sus anunciantes orientaran deliberadamente los anuncios según el género, la raza y la religión, siendo todas ellas datos protegidos según la legislación del país. Los anuncios de trabajo para puestos de enfermería o trabajo de secretaría se sugirieron principalmente a mujeres, mientras que los anuncios de trabajo para conserjes y taxistas se mostraron a un mayor número de hombres, en particular hombres pertenecientes a minorías. El algoritmo aprendió que era probable que los anuncios de bienes inmuebles lograran mejores estadísticas de participación cuando se mostraban a personas blancas, por lo que ya no se mostraban a otros grupos minoritarios.

Este problema se deriva de cómo aprende la máquina de IA. En este caso la plataforma publicitaria formó un patrón a partir de los datos que se le proporcionaron, pero el patrón reflejaba las desigualdades sociales existentes y, si no se hubiera controlado, habría ayudado a propagarlas aún más. Esta práctica se conoce como “publicidad basada en la discriminación” y se ha descubierto que perpetúa estereotipos dañinos y la discriminación en la sociedad. Este ejemplo destaca la importancia de las consideraciones éticas en el desarrollo y uso de la IA. Es esencial que los algoritmos se diseñen y supervisen de manera que no perpetúen ni amplifiquen los sesgos y la discriminación existentes.

Oportunidades limitadas

El sesgo en los sistemas de IA puede limitar las oportunidades para ciertos grupos de personas. Por ejemplo, si un sistema de IA está sesgado en contra de las mujeres, es posible que no las recomiende para ciertos puestos de trabajo o programas educativos, lo que genera una falta de avance y crecimiento para estas personas.

La herramienta de reclutamiento sesgada de Amazon

En 2018, Reuters (21) informó de que Amazon había estado trabajando en un sistema de reclutamiento de IA diseñado para agilizar el proceso de reclutamiento mediante la lectura de currículos y la selección del candidato mejor calificado. Desafortunadamente, la IA parecía tener un problema serio con las mujeres y resultó que el algoritmo había sido programado para replicar las prácticas de contratación existentes, lo que significa que también replicaba sus sesgos. Esto se debe a que los modelos informáticos de Amazon fueron entrenados para examinar a los solicitantes mediante la observación de patrones en los currículos enviados a la empresa durante un período de 10 años. La mayoría provino de hombres, un reflejo del dominio masculino en la industria tecnológica.

En efecto, el sistema de Amazon aprendió que los candidatos masculinos eran preferibles. Penalizaba los currículos que incluían la palabra “femenino”, como en “capitana del club de ajedrez femenino”. Y rebajó la calificación de las graduadas de dos universidades exclusivamente para mujeres. En lugar de ayudar a eliminar los sesgos presentes en el proceso de reclutamiento, el algoritmo simplemente los automatizó. Amazon confirmó que habían desechado el sistema.

El sesgo en la contratación y el reclutamiento puede ocurrir cuando ciertos grupos de personas están injustamente en desventaja o excluidos de las oportunidades laborales en función de factores como su raza, género, edad, religión, orientación sexual o discapacidad. Esto puede conducir a una falta de diversidad en el lugar de trabajo y puede tener consecuencias negativas tanto para las personas afectadas como para la empresa en su conjunto.

Es importante que las empresas sean conscientes del potencial de sesgo en sus procesos de contratación y tomen medidas para mitigarlo. Esto puede incluir el uso de diversas fuentes de contratación, la implementación de criterios de evaluación justos y objetivos y la capacitación de los empleados en diversidad, equidad e inclusión.

Información errónea

El sesgo en los sistemas de IA puede dar lugar a la difusión de información falsa o engañosa. Por ejemplo, si un sistema de IA está sesgado contra un determinado grupo de personas, puede difundir información errónea sobre ese grupo, lo que genera estereotipos negativos y discriminación.

Los sistemas de recomendación de noticias que utilizan algoritmos de IA para personalizar el contenido en función de los hábitos y preferencias de lectura anteriores de un usuario también pueden perpetuar la información errónea. Si estos algoritmos están sesgados hacia perspectivas particulares, pueden reforzar las creencias existentes y difundir información falsa o engañosa.

Toma de decisiones ineficaz

El sesgo en los sistemas de IA puede conducir a una mala toma de decisiones, ya que es posible que el sistema de IA no esté considerando con precisión todos los factores relevantes. Esto puede tener consecuencias negativas en áreas como la atención médica, la justicia penal y las finanzas.

Racismo en la provisión de atención médica en EE. UU.

En 2018, un equipo de la Universidad de California (22), Berkeley, descubrió un problema con un sistema de IA que se estaba utilizando para asignar atención a 200 millones de pacientes en los EE. UU., lo que provocó que los pacientes de color recibieran un estándar de atención más bajo. El problema surgió del hecho de que el sistema estaba asignando valores de riesgo utilizando el costo previsto de la atención médica como variable determinante, y debido a que los pacientes de color a menudo tenían menos capacidad de pago o eran percibidos como menos capaces de pagar por el estándar de atención más alto, el sistema IA esencialmente aprendió que no tenían derecho a tal estándar.

En general, a las personas de esta raza se les asignaron puntajes de riesgo más bajos que a las personas de raza blanca, a pesar de que los pacientes de color también tenían estadísticamente más probabilidades de tener condiciones comórbidas y, por lo tanto, experimentar niveles más altos de riesgo. Esto, a su vez, significó que los pacientes de color tenían menos probabilidades de acceder al estándar de atención necesario y más probabilidades de experimentar efectos adversos como resultado de haberles negado la atención adecuada. Habiendo hecho este descubrimiento, el equipo de UC Berkeley trabajó con la empresa responsable del desarrollo de la herramienta para encontrar variables distintas al costo a través de las que asignar puntajes de riesgo, reduciendo el sesgo en un 84%.

Ha habido numerosos casos en los que el sesgo en la IA ha llevado al racismo en la prestación de atención médica. Un ejemplo es el uso de algoritmos para predecir

qué pacientes están en riesgo de desarrollar ciertas condiciones o enfermedades. Estos algoritmos a menudo se entrenan con datos que reflejan las experiencias de atención médica de ciertas poblaciones, lo que significa que puede no predecir con precisión los riesgos de las personas de otros grupos raciales o étnicos. Esto puede resultar en que los proveedores de atención médica asignen recursos y tratamientos de manera desproporcionada a ciertos grupos raciales o étnicos, en lugar de brindar la misma atención a todos los pacientes.

En general, es crucial que los proveedores de atención médica y las empresas de tecnología sean conscientes del potencial de sesgo en la IA y tomen medidas para mitigarlo. Esto puede implicar el uso de diversos conjuntos de datos para entrenar algoritmos, implementar medidas de seguridad para evitar la discriminación y garantizar que las personas tengan la oportunidad de proporcionar información y comentarios sobre el uso de la IA en la atención médica.

El sesgo en la IA puede tener graves consecuencias, ya que puede conducir a un trato injusto de ciertos grupos y puede perpetuar las desigualdades sociales existentes. Es importante reconocer y abordar el sesgo en los sistemas de IA.

¿Qué podemos hacer con los sesgos en la IA? [23]

Cambiar la forma en que las personas se educan sobre ciencia y tecnología

Una forma de reducir el sesgo en la IA es cambiar la forma en que se educa a las personas sobre ciencia y tecnología. Esto podría incluir proporcionar una representación más diversa en los planes de estudios de ciencia y tecnología, destacando el impacto de los sesgos históricos y actuales en el campo y enfatizando la importancia de las consideraciones éticas en el desarrollo y uso de la IA.

Otro enfoque podría ser incorporar más oportunidades prácticas de aprendizaje experiencial que permitan a los estudiantes involucrarse directamente y comprender los posibles sesgos e implicaciones de la IA. Esto podría incluir proyectos o estudios de casos que desafíen a los estudiantes a identificar y abordar los sesgos en los sistemas de IA al considerar las implicaciones éticas de su uso.

Además, brindar más recursos y apoyo a los educadores para comprender y abordar mejor estos problemas puede ayudar a crear un entorno de aprendizaje más inclusivo y equitativo para todos los estudiantes. Esto podría incluir oportunidades de desarrollo profesional, acceso a investigación y recursos, y apoyo y orientación continuos de expertos en el campo.

Ser consciente de los contextos en los que la IA pueden ayudar a corregir el sesgo, así como en los que existe un alto riesgo de que la IA pueda incrementar el sesgo

Al implementar IA, es importante anticipar dominios potencialmente propensos a sesgos injustos, como en los ejemplos anteriores de sistemas sesgados o con datos sesgados. Dado que tanto las personas como las empresas tienen algún tipo de responsabilidad social, tenemos la obligación de regular nuestros procesos de modelado para garantizar que seamos éticos en nuestras prácticas. Las organizaciones deberán mantenerse actualizadas para ver cómo y dónde la IA puede mejorar la equidad y dónde los sistemas de IA han tenido problemas. Las empresas deben establecer una estrategia de eliminación de sesgo dentro de su estrategia general de IA que contenga una cartera de acciones técnicas, operativas y organizativas:

- **La estrategia técnica** incluye herramientas que pueden ayudarlos a identificar posibles fuentes de sesgo y revelar las características de los datos que afectan la precisión del modelo;
- **La estrategia operativa** comprende la mejora de los procesos de recopilación de datos utilizando internos y auditores externos;
- **La estrategia organizacional** incluye establecer un lugar de trabajo donde las métricas y los procesos se presenten de manera transparente.

Considere los sistemas Humano en el circuito

El objetivo de la tecnología humano en el circuito (Human-in-the-Loop en inglés) es hacer lo que ni un ser humano ni una computadora pueden lograr por sí solos. Cuando una máquina no puede resolver un problema, los humanos deben interferir y resolver el problema. Como consecuencia de este procedimiento, se crea un ciclo de retroalimentación continuo. Con retroalimentación continua, el sistema aprende y mejora su rendimiento con cada ejecución posterior. Como resultado, human-in-the-loop conduce a conjuntos de datos más precisos y a una mayor seguridad y precisión.

Los algoritmos de aprendizaje automático más populares requieren grandes cantidades de datos etiquetados para producir resultados precisos. Sin embargo, hay muchos casos en los que ni siquiera hay una gran cantidad de datos sin etiquetar para extraer. Por ejemplo, si está buscando ejemplos de noticias falsas en un idioma con solo unos pocos miles de hablantes, es posible que todavía no haya ningún ejemplo de noticias falsas en ese idioma. Por lo tanto, el algoritmo no tendrá nada de qué aprender. En este caso, mantener a los humanos informados puede garantizar el mismo nivel de precisión incluso para diferentes tipos de datos. Ese es el caso de Facebook, que aún mantiene a los humanos informados cuando se trata de monitorear y abordar la actividad en la plataforma de redes sociales.

2.3. ¿Por qué el género, la diversidad y la inclusión son importantes para la IA?

Los sistemas de IA están diseñados y construidos por humanos y, a menudo, reflejan los sesgos, los prejuicios y los estereotipos de las personas que los diseñan. Si el equipo de desarrollo de IA no es diverso es más probable que el sistema de IA perpetúe prejuicios y estereotipos dañinos. Los equipos que logran la diversidad en cuanto a raza, género, orientación sexual, edad o condiciones económicas desarrollan mejores sistemas porque las personas diversas razonan de manera diferente y pueden generar más ideas nuevas, son mejores para detectar prejuicios e impulsan un pensamiento más creativo. Es una cuestión de diversidad cognitiva.

Cuando YouTube (24) lanzó su aplicación de carga de vídeos para iOS, entre el 5 y el 10 por ciento de los vídeos subidos por los usuarios estaban al revés. ¿Estaba la gente grabando vídeos incorrectamente? No. Su diseño inicial fue el problema. Fue diseñado para usuarios diestros, pero los teléfonos generalmente giran 180 grados cuando se sostienen con la mano izquierda. Sin darse cuenta, YouTube creó una aplicación que funcionaba casi exclusivamente para diestros.

Los datos nos brindan información para tomar decisiones informadas. Pero la realidad es que los datos se pueden interpretar dependiendo de quién los manipule. La interpretación de los datos por parte de la IA puede estar tan sesgada como la inteligencia humana detrás de ella. ¿Qué puede pasar cuando tienes un excelente equipo de IA pero no es cognitivamente diverso? Como se muestra en el ejemplo anterior, el algoritmo de reclutamiento de Amazon tenía sesgo de género, descartando candidatas competitivas para puestos después de enterarse de que el porcentaje de mujeres en esos puestos era menor.

La promoción del género, la diversidad y la inclusión en la IA puede ayudar a garantizar que las tecnologías de IA se diseñen y desarrollen con un enfoque centrado en el ser humano. Las mujeres y otros grupos subrepresentados aportan perspectivas, experiencias y habilidades únicas a la mesa, lo que puede conducir al desarrollo de tecnologías de IA más creativas, innovadoras y efectivas que satisfagan mejor las necesidades de la sociedad. Cuando los sistemas de IA no son diversos e inclusivos es posible que no satisfagan adecuadamente las necesidades e intereses de todos los miembros de la sociedad. Esto puede generar una falta de confianza en los sistemas de IA y una falta de adopción por parte de ciertos grupos.

2.4. Modelos femeninos a seguir que cambian el campo de la IA

Las mujeres juegan un papel importante en la configuración del futuro de la tecnología y, en particular, en la IA. Al aumentar la visibilidad de los modelos femeninos a seguir en campos tecnológicos, como la IA, y promoverlos, nuestro objetivo es inspirar a los profesionales de la educación no formal a crear programas educativos más inclusivos donde el género, la diversidad y la inclusión juegan un papel central en el desarrollo de la educación, actividades y ambientes.

Cabe señalar que la lista que se muestra a continuación no es exhaustiva; es un intento de brindar una descripción general de mujeres increíbles que son una fuerza impulsora en el campo de la IA en todo el mundo. Al empujar los límites de lo que es posible y al expandir nuestro conocimiento del potencial en número y en diversidad de la IA en salud, infraestructura, clima, energía e idiomas (entre otros), estas mujeres juegan un papel importante como modelos a seguir para la generación futura de mujeres científicas, ingenieras y empresarias.

Carolina Barcenas [25]

Vicepresidente senior de investigación en Visa

Lidera un equipo que impulsa la innovación en cuatro áreas: IA y aprendizaje automático, donde es responsable de explorar y desarrollar formas avanzadas de aprovechar los datos para crear valor comercial para Visa a través de técnicas de IA; cadena de bloques; Seguridad y Comercio del Futuro. El equipo se enfoca tanto en la investigación fundamental como en la impulsada por el producto. Carolina se ha desempeñado como co-líder de Visa Women in Technology en Austin y con frecuencia participa en eventos que apoyan el avance de las mujeres, como el Congreso de Mujeres de Texas, las cumbres Latinas in Tech y Girlstart. Ella es la fuerza organizadora detrás del programa de pasantes del colegio comunitario que se enfoca en candidatos tradicionales.

Ha trabajado tanto en la industria como en la academia y tiene más de 20 años de experiencia en soluciones de aprendizaje automático en fintech. Antes de unirse a Visa, pasó 7 años en PayPal, donde fue responsable de administrar el riesgo de los pequeños y medianos vendedores de comercio electrónico. Su experiencia es en técnicas profundas de minería de datos. Ella tiene un doctorado en Estadística Aplicada del Instituto de Tecnología de Georgia como becaria Fulbright.

En mi caso, comencé a pensar que tendría una carrera como físico; sin embargo, descubrí que mi pasión era resolver problemas y crear abstracciones matemáticas para utilizarlas en la toma de decisiones. Después de darme cuenta de mi pasión profesional, reajusté mis objetivos y seguí una carrera en aprendizaje automático. Los objetivos cambian y, a medida que avanzas, no debes tener miedo de adaptarte a nuevas situaciones. Resulta clave trabajar en algo que te apasione.

Stephanie Lampkin [26]

Fundador técnico y CEO de Blendoor

En su papel como fundadora técnica y directora ejecutiva de Blendoor, a Stephanie Lampkin le apasiona aprovechar la inteligencia aumentada y el análisis de personas para unir una fuerza laboral diversa de candidatos a empresas, asegurando el anonimato de los candidatos en el proceso. Ella cree que esto puede mitigar el sesgo inconsciente, dando a las personas cualificadas más oportunidades y creando mejores empresas y economías como resultado. Stephanie también recopila datos para BlendScore, que es un informe publicado por Blendoor. BlendScore califica a las empresas tecnológicas por su reclutamiento, compensación, retención y promoción de mujeres y minorías subrepresentadas.

Comencé Blendoor tres meses después de tener una experiencia de entrevista bastante sorprendente con Google. Yo vivía en Nueva York en ese momento y acababa de graduarme de MIT Sloan. Estaba en una entrevista para un puesto de líder analítico en Google, y el equipo de contratación me respondió que sería más adecuada para un puesto de ventas o marketing porque no me consideraban lo suficientemente técnica para el puesto. Antes de esta entrevista, había estado programando y mostrando buenas aptitudes para la informática desde una edad temprana, obtuve un título de ingeniería de Stanford, trabajé en Microsoft durante 5 años y medio y acababa de completar un MBA. Algo acerca de sus comentarios no me pareció del todo correcto, pero rechacé cortésmente el puesto de Ventas/Marketing y seguí adelante. Más tarde, descubrí que Google en ese momento tenía sólo un 2% de afroamericanos, un 3% de latinos y alrededor de un 25% de mujeres. El pretexto que comunicaban a los medios era que era un problema diferente; consistía en argumentar que no encontraban suficientes mujeres cualificadas y personas de color. Esto me llevó a un momento de sorpresa. Por eso, creé una plataforma que permite conectar a candidatos y empresas fácilmente. Así las empresas puedan acceder a un grupo más amplio de talento diverso. Así fue como nació Blendoor.

Freyja Jørgensen [27]

Gerente de Innovación en Simula/Gründergarasjen y Líder del Grupo Directivo de NORA.startup

Freyja Jørgensen tiene mucha energía y es el tipo de persona que simplemente hace que las cosas sucedan. Uno de los grupos más activos dentro del Consorcio de Investigación de IA de Noruega (NORA) es NORA.startup. La iniciativa NORA.startup es una puerta de entrada a una mayor cooperación entre la academia, las incubadoras y las empresas emergentes en el campo de la IA, el aprendizaje automático y la robótica. Jørgensen tiene una formación académica que combina biología molecular y psicología. Tiene una amplia

experiencia trabajando en la intersección entre la academia y las startups, y ha estado involucrada en varios proyectos educativos y de innovación. Anteriormente, Jørgensen ha trabajado para promover la ciencia y la tecnología en la educación, pero actualmente dedica su tiempo a apoyar a las empresas de tecnología en la etapa inicial en la incubadora Gründergarasjen. También está trabajando para desarrollar aún más la red NORA.startup y apoyar a las nuevas empresas de IA allí. Jørgensen es responsable de gran parte de la organización interna y la gestión de eventos que enriquecerán el ecosistema de IA de Noruega y más allá.

Desde muy joven supe que, de alguna manera, quería trabajar con personas y ciencia. Soñaba con ser médico de niños o investigador. Tener un abuelo con un doctorado en química y que era profesor de biología en la escuela secundaria me inspiró aún más. Más tarde, al estar involucrada en varios proyectos educativos y de innovación en Simula, rápidamente sentí muchas ganas de trabajar en un entorno dinámico e interdisciplinario. Eso eventualmente me llevó al campo extremadamente inspirador y de rápido crecimiento de las nuevas empresas tecnológicas y la IA. La IA y otras tecnologías seguirán desempeñando un papel central en nuestras vidas, y me parece importante apoyar y comprender este campo. Cada día aprendo algo nuevo y enfrento nuevos desafíos y posibilidades en el trabajo que realizo; esa es una de las principales cosas que me motiva en mi trabajo.

Rebekah Agwunobi [28]

Estudiante Choate Rosemary Hall y ganadora de la beca Caroline D. Bradley

El interés de Rebekah por los ordenadores fue apoyado por su familia desde una edad temprana, comenzando con una clase de JavaScript que tomó en tercer grado. Desde entonces, descubrió la IA y sus aplicaciones para el bien social en Stanford AI4ALL, pasó dos veranos como pasante de investigación en la USC, creó un estudio dirigido al aprendizaje automático en su escuela secundaria y más. Además de su interés por la tecnología, es una activista apasionada que se preocupa por temas que incluyen el encarcelamiento masivo, la reforma penitenciaria y la diversidad en STEM. Tiene la esperanza de usar la IA y la informática como herramientas efectivas para resolver problemas del mundo real.

Quería algo integrado en mi programa de cursos que me empoderara y me permitiera explorar cosas que me interesaban. Es por eso que el trimestre pasado hice un estudio dirigido al aprendizaje automático. En la clase, hablamos sobre algunas de las técnicas generales que aprendí en Stanford AI4ALL, y también hablamos sobre algunos conceptos más avanzados. Participamos en concursos de Kaggle y aprendí mucho sobre las diferentes aplicaciones del aprendizaje automático en áreas como la generación de arte y la música. Aprendí muchos conceptos que subyacían bajo las matemáticas que ya conocía.

Abbey Lin [27]

Líder de controles y cumplimiento en Microsoft Noruega, cofundadora y miembro de la junta Oslo.ai, fundadora de Ethical AI Resources

Abbey Lin ha trabajado en el campo de la ética durante los últimos 15 años, más recientemente con Microsoft. Fue a través de estas experiencias que vio los desafíos éticos emergentes y las oportunidades que dan forma a la IA y quería que más personas fueran conscientes de ello. Durante su baja por maternidad hace algunos años, creó Ethical AI Resources, un sitio web y una fuente de noticias que recopila información y las últimas noticias sobre IA responsable, que se ha convertido en un recurso popular en Twitter [@aithical](https://www.twitter.com/Alethical). Además, Abbey es cofundadora de Oslo.ai, que está ayudando a construir y apoyar una comunidad de entusiastas de la IA en Oslo a través de eventos y oportunidades para establecer contactos. Abbey también es miembro de la junta de Visualyst.co, una empresa emergente de IA centrada en el cumplimiento de las normas de vídeo.

Si te preocupas por la diversidad y la inclusión, también debes preocuparte por la IA responsable. La IA tiene la capacidad de tener un impacto positivo en el mundo, pero también puede, y ya lo ha hecho, tener un impacto negativo adicional en comunidades ya marginadas. La IA responsable debe tenerse en cuenta en todas las etapas del desarrollo del producto y los consumidores de IA deben cuestionar y exigir transparencia sobre cómo se creó el producto.



3

**Directrices para
actividades que promueven
un enfoque inclusivo de la
inteligencia artificial**

Esta sección incluye una colección de actividades prácticas que las instituciones de educación no formal pueden incluir en sus programas educativos para enseñar a los estudiantes sobre IA.

Las actividades propuestas introducen diferentes conceptos de IA e incluyen descripciones paso a paso de cada actividad, materiales que se utilizarán y recursos educativos complementarios para ayudar a los educadores a planificar y ofrecer actividades atractivas y significativas a sus estudiantes.

Antes de comenzar, aquí hay algunos consejos para crear un entorno de aprendizaje acogedor e inclusivo para todos los estudiantes, independientemente de sus antecedentes o conocimientos previos de IA:

- Comience por presentar el concepto de IA a sus estudiantes y por explicar sus diversas aplicaciones. Puede usar videos, textos y otros materiales para proporcionar una visión general de la IA y sus posibles impactos en la sociedad.
- Asegúrese de enfatizar la importancia de la diversidad, la equidad y la inclusión en el campo de la IA. Esto puede incluir discutir las formas en que la IA se ha utilizado para perpetuar los sesgos y discutir estrategias para crear sistemas de IA más inclusivos y equitativos.
- Utilice estudios de casos y ejemplos del mundo real para ilustrar las diversas formas en que se utiliza la IA en diferentes industrias y sectores. Esto puede ayudar a los estudiantes a ver las aplicaciones prácticas de la IA y comprender cómo está afectando al mundo que los rodea.
- Anime a los estudiantes a pensar críticamente sobre las implicaciones éticas de la IA y discutir las posibles consecuencias del uso de la IA en diferentes contextos. Puede usar indicaciones y preguntas de discusión para facilitar esta conversación.
- Considere incorporar actividades prácticas, como ejercicios o proyectos de codificación, para dar a los estudiantes la oportunidad de aplicar lo que han aprendido y explorar los conceptos de IA de una manera más interactiva.
- Anime a los estudiantes a mantenerse al día sobre los últimos desarrollos en IA y a continuar aprendiendo sobre el campo incluso después de que termine el curso. Esto puede incluir mantenerse informado sobre noticias relevantes, asistir a conferencias y talleres, y unirse a comunidades y organizaciones en línea relevantes.

3.1 Actividad 1: El juego de etiquetas [29]

Objetivos de aprendizaje

- Mejorar la comunicación no verbal:** Los participantes practicarán señales y gestos no verbales para formar grupos sin usar la comunicación verbal, fomentando habilidades de comunicación efectivas.
- Fomentar el pensamiento crítico y la toma de decisiones:** Los participantes considerarán varios factores, como los colores, formas, patrones y tamaños de las etiquetas, para tomar decisiones sobre la formación de grupos, promoviendo el pensamiento crítico y las habilidades de toma de decisiones.
- Promover la colaboración y el trabajo en equipo:** Los participantes trabajarán juntos para formar grupos basados en características comunes, fomentando la cooperación y el trabajo en equipo sin depender de la comunicación verbal.
- Explorar la inclusión y la diversidad:** Los participantes reflexionarán sobre sus formaciones grupales y discutirán si consideraron factores más allá de las etiquetas, fomentando la comprensión y la apreciación de la diversidad.
- Facilitar la discusión reflexiva y abierta:** La discusión final facilitada por el líder alienta a los participantes a compartir sus procesos de toma de decisiones y puntos de vista, promoviendo la reflexión, la escucha activa y el intercambio de ideas entre los participantes.
- Crear conciencia sobre el sesgo:** Los participantes se darán cuenta de sus propios prejuicios, principalmente lo que es el sesgo de afinidad y lo que todos comparten en común a pesar de que todos somos diferentes.

Materiales:

Papel de desecho

Cinta adhesiva

Bolígrafos o lápices

Papel de colores

Descripción de la actividad

Prepare etiquetas o pegatinas para cada participante. Tiene la opción de crear las suyas propias recortando formas de papel de color o utilizando pegatinas prefabricadas. Asegúrese de que las etiquetas vienen en una variedad de colores, formas, patrones y tamaños. Por ejemplo, puede tener varias etiquetas moradas, pero podrían estar en forma de círculos, cuadrados y triángulos, con patrones como cuadrados, puntos y rayas.

El facilitador/líder comienza distribuyendo etiquetas individuales a todos los participantes. Cada persona debe pegar su etiqueta a su ropa, colocándola entre su cintura y hombros. Una vez que todos están etiquetados, los participantes comienzan a moverse por el espacio designado, teniendo en cuenta a los demás a su alrededor.

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

Luego, el líder introduce variedad instruyendo a todos a moverse de una manera diferente, como saltar. Después de unos minutos de movimiento, el líder pide a los participantes que formen pequeños grupos sin usar la comunicación verbal. Observe cómo se dividen en grupos porque es crucial. Los participantes pueden considerar sus etiquetas y organizarse en función del color o la forma. Una vez que los grupos se han formado, todos se extienden nuevamente, continuando moviéndose por el espacio.

El proceso se repite varias veces, con los participantes reagrupándose sin hablar. A medida que avanza el juego y se forman más grupos, los participantes finalmente se reúnen en un gran círculo. En este punto, el líder inicia una discusión, alentando a todos a compartir cómo decidieron las formaciones de sus grupos. La conversación explora si alguien miró más allá de las etiquetas o creó intencionalmente diversos grupos que abarcan diferentes formas, colores, patrones y tamaños.

Una guía paso a paso

Paso 1 La persona que lidera el juego debe dar a todos su propia etiqueta. Todos deben pegar su etiqueta a su parte superior entre su cintura y sus hombros.

Paso 2 Todos deben moverse por el espacio y mantenerse a una distancia segura de cada uno.

Paso 3 Despues de un minuto o dos, la persona que lidera el juego debe pedirles a todos que se reúnan en grupos pequeños sin hablar.

Paso 4 Una vez que los grupos se han formado, todos deben separarse nuevamente y continuar moviéndose por el espacio.

Paso 5 Todos deben repetir los pasos tres y cuatro para que hayan hecho tres o cuatro grupos diferentes sin hablar.

Paso 6 Todos deben reunirse en un gran círculo a una distancia segura entre sí.

Comentario

La persona que dirige la actividad debe preguntar:

- ¿Cómo decidieron los grupos en los que se dividirían? Tómense por turnos para compartir sus pensamientos.
- ¿Alguien miró más allá de las etiquetas?
- ¿Alguien formó intencionalmente un grupo diverso con una variedad de formas, colores, patrones y tamaños?

Reflexión

A través de esta actividad, los participantes fueron introducidos al concepto de diversidad y exploraron específicamente el sesgo de afinidad, que es una forma de sesgo inconsciente.

- ¿Se sorprendieron las personas al enterarse del sesgo de afinidad?
- ¿Cómo se sintieron al reflexionar sobre sus formaciones grupales durante el juego de etiquetas y su lista de personas de confianza?
- ¿Cuáles son algunas de las ventajas y desafíos asociados con la colaboración en grupos que comprenden individuos similares y diversos?

Se anima a los participantes a reflexionar sobre los diversos grupos a los que pertenecen, como una clase escolar, un equipo deportivo o un círculo de amigos. Deben considerar si estos grupos consisten principalmente en individuos que son similares o diferentes a ellos mismos. Posteriormente, los participantes deben formar parejas y participar en conversaciones sobre estrategias para mejorar su comprensión de la diversidad en su vida cotidiana e interacciones. Tomando turnos, los participantes pueden compartir sus ideas si así lo desean.

3.2 Actividad 2: Dibujar la descripción [30]

Objetivos de aprendizaje

- Desarrollar habilidades de observación:** Los participantes mejorarán su capacidad para observar y prestar atención a los detalles.
- Mejorar la creatividad:** Los participantes ejercitarán su creatividad e imaginación interpretando las descripciones en forma visual.
- Mejorar las habilidades de comunicación:** Los participantes practicarán una comunicación efectiva al transmitir con precisión las descripciones a través de sus dibujos.
- Fomentar el pensamiento crítico:** Los participantes participarán en el pensamiento crítico analizando descripciones y traduciéndolas en representaciones visuales.
- Fomentar el trabajo en equipo:** Los participantes colaborarán con otros compartiendo sus interpretaciones y trabajando juntos para resolver desafíos.

Materiales

Bolígrafos o lápices de colores

Papel de desecho

Impresiones de las fotos y descripciones del objeto, escena, persona o concepto elegido

Descripción de la actividad

Para comenzar la actividad, los participantes se reúnen en un grupo y se les proporcionan materiales de dibujo como papel y lápices. El facilitador o un participante designado asume el papel de descriptor.

Leerán una descripción de un objeto, escena, persona o concepto, sin revelar lo que es. Las descripciones pueden ser específicas o abstractas, lo que permite una variedad de interpretaciones. A medida que el facilitadordescriptor lee la descripción, los participantes escuchan atentamente y visualizan los detalles en sus mentes. Luego tienen una cantidad fija de tiempo para traducir sus imágenes mentales en dibujos en sus papeles. Los dibujos deben tener como objetivo capturar la esencia del objeto o escena descrito.

Una vez que se acaba el tiempo de dibujo, los participantes comparten sus dibujos con el grupo. El descriptor revela el objeto o escena original, y los participantes comparan sus dibujos para ver con qué precisión capturaron la descripción. Esto fomenta la reflexión y la discusión, permitiendo a los participantes explorar las diferencias y similitudes en sus interpretaciones.

Después de cada ronda, los roles de descriptor y participantes se pueden rotar, asegurando que todos tengan la oportunidad de practicar la descripción y el dibujo. La actividad se puede personalizar variando la complejidad de las descripciones, introduciendo diferentes temas o incorporando límites de tiempo para agregar un elemento de desafío.

Una guía paso a paso

Paso 1

La persona que dirige la actividad debe leer la descripción de objeto, persona o concepto. Puede usar el ejemplo a continuación, de Katherine Johnson, quien fue una de las primeras mujeres afroamericanas que trabajaron para la NASA, recibiendo su premio presidencial [31].

Ejemplo

Descripción de la imagen: En esta escena, vemos a un matemático de la NASA recibiendo un premio presidencial en 2015. Este matemático nació en 1918 e hizo importantes contribuciones en los primeros tiempos de la NASA, incluyendo el uso pionero de computadoras para muchas tareas.



Paso 2

Todos deben dibujar el objeto, escena, persona o concepto basado en la descripción.

Paso 3

Dile a todos unos minutos para completar sus dibujos.

Paso 4

Anime a todos a compartir lo que dibujaron. Compare las similitudes y diferencias entre los dibujos. Para la imagen de ejemplo utilizada algunas de las preguntas pueden ser:

- ¿Qué llevan puesto?
- ¿De qué tamaño son (altos, bajos, gordos, delgados)?
- ¿De qué color es su piel?
- ¿Tienen género?
- ¿Qué edad tienen?

Paso 5

Muestre a todos la foto de la primera persona y revele su nombre, tal vez algunos detalles sobre quiénes son o qué hacen / hicieron. Conversen sobre qué dibujo es más similar a la imagen real de la persona. ¿Cuáles son las diferencias más significativas en los dibujos de las personas y la imagen real?

Comenta

Pregúntele a la gente por qué dibujaron la descripción de la manera en que lo hicieron. ¿Qué pasa con la descripción que les hizo pensar que el objeto, escena, persona o concepto se veía de cierta manera? ¿Qué fue lo que les hizo elegir cierta edad, raza, género, tamaño o aspecto para la persona? ¿Qué les hizo elegir los colores o formas que eligieron?

Reflexión

En esta actividad, todos dibujaron escenas, personas o conceptos basados en descripciones del trabajo que han realizado y el impacto que han tenido. Pida a todos que piensen en los dibujos que difieren mucho de la imagen real. A menudo asociamos ciertos rasgos con ciertos trabajos, personalidades, logros, colocación o usabilidad. Esto es un sesgo inconsciente y es causado por muchas cosas, como solo aprender sobre la historia desde una perspectiva, o ver programas de televisión y películas donde los personajes y los escenarios se basan en estereotipos ignorantes o dañinos.

1. Reflexión individual: Despues de completar sus dibujos, los participantes pueden tomarse un momento para reflexionar sobre sus propias interpretaciones, considerando preguntas tales como:

- ¿Cómo interpreté la descripción?
- ¿En qué elementos me centré y por qué?
- ¿Capturé la esencia del objeto, escena, persona o concepto descrito?
- ¿A qué desafíos me enfrenté durante el proceso de dibujo?
- ¿Qué podría haber hecho diferente para mejorar mi dibujo?

Alentar a los participantes a reflexionar individualmente les permite obtener información sobre su propio proceso de pensamiento, elecciones artísticas y áreas de mejora.

2. Reflexión grupal: Despues de compartir sus dibujos con el grupo, los participantes pueden participar en una sesión de reflexión grupal facilitada por el descriptor o un líder designado. Esta sesión puede incluir la discusión y exploración de las diferentes interpretaciones y variaciones en los dibujos. Algunas indicaciones de reflexión para la discusión grupal pueden incluir:

- ¿Qué similitudes y diferencias notamos en nuestros dibujos?
- ¿Qué aspectos de la descripción fueron difíciles de capturar visualmente?

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

- ¿Cómo contribuyeron las variaciones en nuestras interpretaciones a una comprensión más rica de la descripción?
- ¿Algún dibujo proporcionó perspectivas o ideas únicas?

La reflexión grupal permite a los participantes obtener una perspectiva más amplia sobre las diversas interpretaciones y fomenta el diálogo abierto y la apreciación de los diferentes enfoques. También fomenta un sentido de aprendizaje colectivo y experiencias compartidas.

Explore varios enfoques para reducir el sesgo inconsciente. Pregunte si alguien puede proponer sugerencias. Por ejemplo, en lo que respecta a la imagen de ejemplo, puede hacer lo siguiente:

- Sumérjase en representaciones positivas de individuos de diversos géneros, razas o habilidades al involucrarse con literatura como "Pequeños líderes: hombres excepcionales en la historia negra" y "Mujeres excepcionales en la historia negra" de Vashti Harrison o películas como "El halcón de mantequilla de maní".
- Inicie discusiones sobre casos en los que las personas han enfrentado discriminación basada en factores como raza, género, sexualidad, capacidad y edad.
- Comprender el problema permite resolverlo mejor
- Acostúmbrese a buscar múltiples perspectivas sobre los acontecimientos históricos. Las personas poseen puntos de vista y experiencias diferentes, lo que influye en la forma en que perciben y relatan acontecimientos pasados.

Nota: La imagen utilizada en esta actividad actúa como ejemplo. Siéntase libre de usar cualquier tipo de imagen que se ajuste a su curso o plan de estudios

3.3 Actividad 3: Adivina a la persona

Objetivos de aprendizaje

- 1. Reconocer sesgos:** Los participantes se darán cuenta de la existencia de sesgos, específicamente el efecto halo, que influye en nuestras percepciones y juicios de los individuos basados en información limitada o nociones preconcebidas.
- 2. Comprender el efecto halo:** Los participantes obtendrán una comprensión del efecto halo como un sesgo cognitivo que nos lleva a atribuir cualidades positivas a los individuos en función de una sola característica o rasgo, como la apariencia física o la fama.
- 3. Suposiciones de desafío:** Se alentará a los participantes a suposiciones y prejuicios examinando sus percepciones iniciales de la mujer famosa.
- 4. Fomentar la concienciación sobre la diversidad en las profesiones:** Los participantes ampliarán su comprensión de la amplia gama de profesiones al darse cuenta de que las personas pueden sobresalir en campos que pueden no alinearse con los estereotipos o expectativas comunes. Apreciarán que los talentos y las habilidades no están limitados por el género o la apariencia.
- 5. Fomentar el pensamiento crítico:** Los participantes desarrollarán habilidades de pensamiento crítico mediante el análisis de sus propios procesos de pensamiento y sesgos al hacer juicios. Aprenderán a cuestionar la influencia del efecto halo y se esforzarán por realizar evaluaciones más objetivas e imparciales de los individuos.
- 6. Fomentar el diálogo abierto:** La actividad promoverá discusiones abiertas entre los participantes, permitiéndoles compartir sus suposiciones iniciales, reflexionar sobre sus sesgos y explorar el impacto del efecto halo. Este diálogo fomentará una comprensión más profunda de los sesgos y sus implicaciones en diversos contextos.
- 7. Empoderar a los participantes para desafiar los prejuicios:** A través de esta actividad, los participantes estarán facultados para desafiar los prejuicios y estereotipos que encuentran en su vida diaria. Desarrollarán estrategias para reconocer y superar el efecto halo, promoviendo juicios y procesos de toma de decisiones más justos y equitativos.
- 8. Mejorar la empatía y la inclusión:** Al explorar los sesgos y el efecto halo, los participantes desarrollarán empatía y una mayor apreciación de la diversidad. Comprenderán la importancia de crear entornos inclusivos que valoren a las personas por sus habilidades y contribuciones únicas en lugar de atributos superficiales.
- 9. Aplicar conocimientos a situaciones del mundo real:** Se alentará a los participantes a transferir sus aprendizajes de la actividad a situaciones del mundo real, como interacciones en el lugar de trabajo, procesos de contratación o relaciones personales. Se esforzarán por hacer juicios y decisiones más informados basados en una evaluación exhaustiva de los individuos en lugar de confiar únicamente en el efecto halo.

10. Fomentar una cultura de inclusión: La actividad tiene como objetivo contribuir a la creación de una cultura de inclusión, donde se cuestionen los prejuicios, se celebre la diversidad y se valore a los individuos por sus habilidades y logros en lugar de características superficiales. Se alentará a los participantes a ser defensores de la justicia y la igualdad en sus vidas personales y profesionales.

Materiales

Imagen de personas de las que le gustaría hablar que son “víctimas” del efecto halo

Descripción de la actividad

En esta actividad, el facilitador presenta a los participantes una imagen de una persona famosa sin revelar su profesión. Luego se les pide a los participantes que adivinen qué profesión creen que ejerce basándose únicamente en la imagen.

El propósito de esta actividad es crear conciencia sobre los sesgos, centrándose específicamente en el efecto halo. El efecto halo es un sesgo cognitivo en el que los individuos forman juicios generales o impresiones de los demás basados en información limitada o un solo rasgo positivo. En este caso, los participantes pueden verse influenciados por la imagen de una persona famosa, asumiendo una profesión que se alinea con estereotipos comunes o nociones preconcebidas.

Después de que los participantes hayan compartido sus conjeturas iniciales, el facilitador inicia una discusión sobre los sesgos y el impacto del efecto halo. Se anima a los participantes a reflexionar sobre sus suposiciones y considerar cómo el efecto halo puede conducir a juicios inexactos sobre la profesión o las habilidades de alguien. El facilitador guía la conversación para explorar la diversidad de profesiones y la importancia de reconocer los talentos y habilidades individuales más allá de las características superficiales.

Una guía paso a paso

Paso 1

La persona que dirige la actividad debe mostrar la imagen de la persona famosa (por ejemplo, Hedy Lamarr) y pedir a los participantes que adivinen su profesión.

Ejemplo

¿Quién es Hedy Lamarr? (31)

- Actriz en películas de Hollywood;
- Inventora que fue pionera en la tecnología que algún día formaría la base de los sistemas de comunicación WiFi, GPS y Bluetooth de hoy;
- Recibió el premio Bulbie Gnass Spirit of Achievement Award de la Convención de Invenciones.



Paso 2

Dele a todos unos minutos para pensar en su respuesta.

Paso 3

Pida a los participantes que compartan sus respuestas. En este paso puede ser tan creativo como quiera. Por ejemplo, si se encuentra en un lugar físico, como un aula, puede pedir a los participantes que usen post-it's para escribir su respuesta y luego usar una pizarra blanca o una pared para mostrar los resultados, o si está en una sesión en línea, puede usar Mural o Miro board.

Paso 4

Anime a todos a compartir lo que escribieron. Compare las similitudes y diferencias entre las respuestas:

- ¿Cuál fue tu suposición inicial sobre la profesión de esa persona en la imagen? ¿Qué factores influyeron en tu percepción?
- ¿La fama o la apariencia de esa persona en la imagen afectó su juicio sobre su profesión? ¿Cómo?
- ¿Alguna vez te has encontrado con situaciones en las que hiciste suposiciones sobre las habilidades o la profesión de alguien basadas en características superficiales? ¿Puedes compartir algún ejemplo?

Comentario

Haga preguntas a las personas sobre sus propios prejuicios basados en el ejemplo anterior, como ¿Puedes pensar en algún caso de la vida real en el que el efecto halo pueda haber llevado a juicios sesgados u oportunidades perdidas? ¿Cómo crees que el efecto halo puede afectar los procesos de toma de decisiones, como la contratación o las promociones? ¿Qué estrategias podemos adoptar para fomentar la diversidad y la inclusión, tanto en nuestra vida personal como en contextos sociales más amplios?

Reflexión

A través de esta actividad, los participantes se dan cuenta de lo fácil que podemos caer presa de los prejuicios, haciendo juicios rápidos sobre la profesión o las habilidades de alguien sin considerar el alcance total de sus habilidades y calificaciones. La suposición inicial que hicieron sobre la profesión de la persona famosa en la imagen fue influenciada por su apariencia, lo que les hizo cuestionar el impacto potencial del efecto halo en su propio pensamiento.

Participar en una discusión abierta con otros participantes profundizó aún más su comprensión de los sesgos y sus implicaciones. Escuchar a otros compartir sus suposiciones y experiencias iniciales reforzó la omnipresencia de los prejuicios en varios aspectos de la vida. Fue revelador reconocer que ellos también a veces han hecho juicios basados en características superficiales, pasando por alto las verdaderas habilidades y el potencial de los individuos.

La actividad los lleva a reflexionar sobre el impacto social más amplio de los prejuicios y la importancia de crear entornos que valoren la diversidad y la inclusión. Les hizo darse cuenta de que los sesgos pueden obstaculizar los procesos justos de toma de decisiones, afectando áreas como la contratación, las promociones y la igualdad en general.

Un aspecto que destaca durante la discusión es la necesidad de autoconciencia e introspección. Reconocer nuestros propios prejuicios es crucial para desafiarlos y superarlos. La actividad nos animó a cuestionar nuestras suposiciones, a ser conscientes de la influencia potencial del efecto halo y a esforzarnos por obtener juicios más objetivos e inclusivos.

Nota: La imagen utilizada en esta actividad actúa como ejemplo. Siéntase libre de usar cualquier tipo de imagen que se ajuste a su curso o plan de estudios

3.4 Actividad 4: Sesgo de género usando LearningML con conjunto de datos 1

Objetivos de aprendizaje

- 1. Practicar el reconocimiento y la mitigación del sesgo de género:** Los participantes practicarán activamente la identificación y el tratamiento del sesgo de género mediante la participación en el sitio web de LearningML y la utilización de actividades informáticas. Desarrollarán habilidades para analizar críticamente algoritmos, modelos y datos para minimizar el impacto del sesgo en los procesos impulsados por la tecnología.
- 2. Aumentar la conciencia del sesgo inconsciente:** Los participantes obtendrán una comprensión más profunda del sesgo inconsciente y sus implicaciones. Explorarán cómo los sesgos pueden influir en la toma de decisiones y aprenderán estrategias para reconocer y contrarrestar los sesgos inconscientes en diversos contextos.
- 3. Reflexionar sobre el uso de software automatizado para la contratación:** Los participantes reflexionarán sobre las consideraciones éticas y los posibles sesgos asociados con el uso de software automatizado para fines de contratación. Analizarán críticamente las ventajas, limitaciones y riesgos potenciales de confiar en sistemas impulsados por la tecnología para las decisiones de contratación. Los participantes desarrollarán una comprensión refinada del papel del juicio humano en el proceso de reclutamiento.
- 4. Mejorar el pensamiento crítico y las habilidades de toma de decisiones:** Participar en actividades basadas en computadoras y reflexionar sobre el sesgo en la tecnología y los sistemas automatizados fomentará las habilidades de pensamiento crítico de los participantes. Desarrollarán habilidades para evaluar la confiabilidad y validez del software automatizado y tomar decisiones informadas con respecto a su implementación en varios dominios.
- 5. Fomentar la conciencia de las implicaciones éticas de la tecnología:** Los participantes obtendrán información sobre las consideraciones éticas que rodean el uso de la tecnología, particularmente en relación con el sesgo y la discriminación. Reflexionarán sobre la responsabilidad de los individuos y las organizaciones para garantizar la equidad, la transparencia y la inclusión al utilizar sistemas automatizados.
- 6. Promover la alfabetización digital y el uso responsable de la tecnología:** Los participantes desarrollarán habilidades de alfabetización digital y comprenderán el impacto potencial de la tecnología en la sociedad. Explorarán la importancia del uso responsable de la tecnología, teniendo en cuenta las posibles consecuencias de los algoritmos sesgados y la necesidad de una evaluación y mejora continuas.
- 7. Fomentar la autorreflexión y el aprendizaje continuo:** Se alentará a los participantes a reflexionar sobre sus propios prejuicios, suposiciones y perspectivas a lo largo del proceso de aprendizaje. Desarrollarán una mentalidad de crecimiento, abrazando el valor del aprendizaje continuo y buscando activamente oportunidades para desafiar y ampliar su comprensión del sesgo y la tecnología.

Materiales

Ordenador portátil con acceso a Internet

[LearningML](#)

Descripción de la actividad

En los procesos de reclutamiento y selección actuales, sitios web como LinkedIn se han convertido en el paso inicial. Debido a que estas plataformas utilizan modelos de IA para la selección de candidatos, es crucial garantizar que estos modelos estén diseñados de manera que brinden igualdad de oportunidades para personas de todas las razas, géneros, religiones y nacionalidades. Sin embargo, si los modelos de IA se crean con un sesgo inconsciente, existe el riesgo de introducir discriminación durante el proceso de contratación.

En esta actividad, mostraremos un ejemplo de cómo se puede introducir el sesgo de género en un modelo de IA diseñado específicamente para ayudar a las startups a contratar expertos en informática. El modelo se desarrollará utilizando datos sesgados basados en el “género” y clasificará a los candidatos en dos grupos: “Contratados” o “No contratados”. A través de esta demostración, los participantes serán testigos de cómo el modelo toma decisiones de contratación y observarán el impacto de este sesgo introducido en los resultados.

El conjunto de datos utilizado para entrenar este modelo es relativamente pequeño, consta de 10 imágenes para la clase “Contratado” y 10 imágenes para la clase “No Contratado”. Las imágenes que representan a la clase “Contratado” representan predominantemente a hombres que trabajan con computadoras portátiles, como se espera de los informáticos. En contraste, para la clase “No contratado”, introdujimos intencionalmente el sesgo al seleccionar imágenes de mujeres involucradas en actividades no relacionadas, como cocinar o cuidar niños. Solo hay unas pocas imágenes de mujeres trabajando con computadoras portátiles, lo que refuerza los roles de género estereotipados.

A través de esta actividad, los participantes obtendrán una comprensión de primera mano de las posibles consecuencias de los datos sesgados y su influencia en los modelos de IA. Reconocerán la importancia de crear modelos de IA justos e imparciales para garantizar prácticas de contratación equitativas y desafiar los sesgos existentes presentes en los procesos de reclutamiento.

Una guía paso a paso

Paso 1

- Haga clic [aquí](#) para navegar por el sitio web que contiene los recursos (imágenes) que necesitará para la actividad
- Descargue el archivo: [source_activity1.zip](#) en su escritorio o carpeta de trabajo
- Descomprima el archivo descargado. El archivo source_activity1.zip contiene dos archivos:
 1. activity1_gender_ub.json
 2. test_image.jpeg

Braga FAIas LTIA Activities

Activity guides:

- [Activity 1 Guide](#)
- [Activity 2 Guide](#)
- [Activity 3 Guide](#)

Activity resources:

- 1 • [Resources for Activity 1](#)
• [Resources for Activity 2](#)

Paso 2 Lance el sitio web de [LearningML](#)

Paso 3 Haga clic en: [LearningML v1.3 \(avanzado\)](#)

The screenshot shows the LearningML website homepage. At the top, there is a navigation bar with links: QUÉ ES, POR QUÉ, APRENDER, DESARROLLO, BLOG, DESCARGAS, ACERCA DE, CONTACTO, and a language switcher. Below the navigation is a large green banner featuring a cartoon character of a green alien with purple hair and a purple vest, sitting at a control panel with various buttons and screens. In the center of the banner, there are three green buttons: 'LearningML v1.2 (la original)', 'LearningML v1.3 (avanzada)' (which is highlighted with a red box and a blue number '3'), and 'Iml-Snap!'. Below the banner, there are four descriptive boxes: 1. 'LearningML v1.2 (la original)' - 'Ideal para introducirse en el mundo del Machine Learning. Se puede usar desde los últimos cursos de primaria.' 2. 'LearningML v1.3 (avanzada)' - 'En esta versión se añade la clasificación de conjuntos de números y el modo avanzado, con el que podrás explorar el comportamiento de los algoritmos de ML.' 3. 'Iml-Snap!' - 'Para los que quieren más potencia programando aplicaciones. Todas las fases del ML se hacen programando. Ideal para bachillerato, formación profesional y primeros cursos universitarios.' 4. 'LearningML-Desktop' - 'Para los que prefieren tener instalado LearningML v1.3 en su ordenador (Linux, Windows, Mac) y pasar de Internet. Ideal para ser incorporado en distribuciones educativas de Linux y para colegios que tengan problemas de conexión a internet.'

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

Paso 4 Seleccione su idioma: haga clic en el círculo

Paso 5 Haga clic en reconocer imágenes

Paso 6 Suba el archivo: **activity1_gender_ub.json**



Nota: Como puede ver en la siguiente imagen, se han añadido las dos clases (**Contratado** y **No contratado**) con sus respectivas imágenes.

1. Entrenar **Entrenar**
Primero necesito algunas imágenes de ejemplo
+ Añadir nueva clase de imágenes
No contratado
Hired (10)
Contratado

2. Aprender **Aprender**
Llegó el momento de aprender a clasificar imágenes
Aprender a reconocer imágenes

3. Probar **Probar**
Introduce términos nuevos y comprueba si se clasifican correctamente

LearningML tiene 3 etapas como cualquier sistema de IA:

- 1. Entrenar:** En esta etapa alimentamos nuestro sistema o modelamos los datos que tenemos.
- 2. Aprender:** En esta etapa, el sistema está trabajando en el backend y encuentra patrones entre los datos y une las cosas que son similares para que pueda reconocer cualquier información desconocida.
- 3. Probar:** En esta etapa, el sistema está entrenado y listo para reconocer y categorizar la nueva imagen. Cuidado: su imagen de prueba no debe ser parte de sus datos de entrenamiento.

Paso 7 Haga clic en aprender a reconocer imágenes:

2. Aprender

Llegó el momento de aprender a clasificar imágenes

7

Aprender a reconocer imágenes



Paso 8 Suba la imagen de prueba

3. Probar

Introduce términos nuevos y comprueba si se clasifican correctamente

8



Subir Imagen de prueba

Resultados

En esta actividad, en el test_image hay dos mujeres desarrolladoras. Como resultado, el modelo clasifica al candidato de la siguiente manera: **No contratado**

3. Probar

Introduce términos nuevos y comprueba si se clasifican correctamente



- Not hired (98.03 %)
- Hired (1.97 %)

Para discutir

- ¿Crees que hay algún sesgo adicional en el conjunto de datos de imagen utilizado para entrenar el modelo?
- ¿Por qué crees que el resultado es No contratado?
- ¿Qué cambiarías en los datos para cambiar el resultado?

Nota para la discusión

Para que el modelo clasifique a la candidata como contratada, debe capacitarse sin sesgo de género. Si agregamos el mismo número de imágenes masculinas y femeninas en circunstancias similares, el modelo clasificará a un candidato por igual: en otras palabras, el género de esa persona sería irrelevante para su elegibilidad para el puesto.

Reflexión

La actividad que explora el sistema de contratación con sesgo de género ha sido reveladora e invita a la reflexión. Arrojó luz sobre las posibles implicaciones del sesgo inconsciente en procesos de reclutamiento, específicamente relacionados con el género.

Al examinar el modelo de IA diseñado para ayudar a las nuevas empresas a contratar expertos en informática, observamos cómo se puede introducir el sesgo a través del conjunto de datos utilizado para el entrenamiento. La limitada representación de mujeres en la clase “contratada”, en comparación con las imágenes estereotipadas de mujeres en roles no relacionados con la informática en la clase “No contratado”, propiciaba la perpetuación de los estereotipos de género y la desigualdad de oportunidades.

Subraya la necesidad de una mayor diversidad e inclusión en los conjuntos de datos de capacitación para garantizar una toma de decisiones justa. Al confiar en datos sesgados, corremos el riesgo de perpetuar la discriminación práctica y obstaculizar el progreso hacia una sociedad más inclusiva.

Además, esta actividad enfatizó la importancia del pensamiento crítico y ético. Para ello, se hizo uso de consideraciones en el desarrollo y uso de modelos de IA para fines de contratación. La actividad reveló las posibles consecuencias de confiar únicamente en sistemas automatizados sin intervención y supervisión.

Referencia: <https://fosteringai.github.io/project/result3/>

3.5 Actividad 5: Sesgo de género usando LearningML con el conjunto de datos 2

Objetivos y materiales de aprendizaje

Igual que en la actividad 4

Descripción detallada

En la actividad 4, creamos un modelo de clasificación para ayudar a una startup a contratar a un experto en informática. El modelo de la Actividad 4 fue entrenado con datos sesgados de “género” y, como resultado, el modelo clasificó a los hombres como más propensos a ser contratados para el trabajo que a las mujeres.

En la actividad 5, vamos a crear un nuevo modelo de clasificación para ayudar a las startups a contratar a un experto en Informática, pero en este caso entrenaremos el modelo sin sesgo de “género” en los datos. Agregamos el mismo número de imágenes masculinas y femeninas, y ambas en situaciones similares.

Como resultado, puede ver cuán importante es el tipo y la cantidad de datos que utilizamos para crear un modelo de IA para evitar sesgos raciales y de género, o incluso para crear involuntariamente sesgos hacia otro grupo al tratar de corregir el sesgo.

Una guía paso a paso

Paso 1

- Haga clic [aquí](#) para navegar por el sitio web que contiene los recursos (imágenes) que necesitará para la actividad
- Descargue el archivo: **source_activity2.zip** en su escritorio o carpeta de trabajo
- Descomprima el archivo descargado. El archivo source_activity2.zip contiene dos archivos:
 1. activity2_gender_ub.json
 2. test_image.jpeg

Braga FAIaS LTIA Activities

Activity guides:

- [Activity 1 Guide](#)
- [Activity 2 Guide](#)
- [Activity 3 Guide](#)

Activity resources:

- 1. [Resources for Activity 1](#)
- 2. [Resources for Activity 2](#)

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

Paso 2 Haga click [aquí](#) para visitar el sitio web de learningML

Paso 3 Haga clic en: LearningML v1.3 (avanzado)

The screenshot shows the homepage of the LearningML website. At the top, there is a navigation bar with links: QUÉ ES, POR QUÉ, APRENDER, DESARROLLO, BLOG, DESCARGAS, ACERCA DE, CONTACTO, and a language dropdown set to Spanish. Below the navigation is a large cartoon illustration of a green alien-like character with a purple vest, working on a control panel with various buttons and screens. A blue circle with the number '2' is positioned in the top left corner of the page. Below the illustration, there are four main sections: 'LearningML v1.2 (la original)', 'LearningML v1.3 (avanzada)' (which is highlighted with a blue circle containing the number '3'), 'Iml-Snap!', and 'LearningML-Desktop'. Each section has a brief description. The 'LearningML v1.3 (avanzada)' section is described as ideal for introducing machine learning, while the 'LearningML v1.3 (avanzada)' section is described as advanced, for programming applications, and the 'LearningML-Desktop' section is described as for those who prefer to install it on their computer.

Paso 4 Seleccione su idioma: haga clic en el círculo

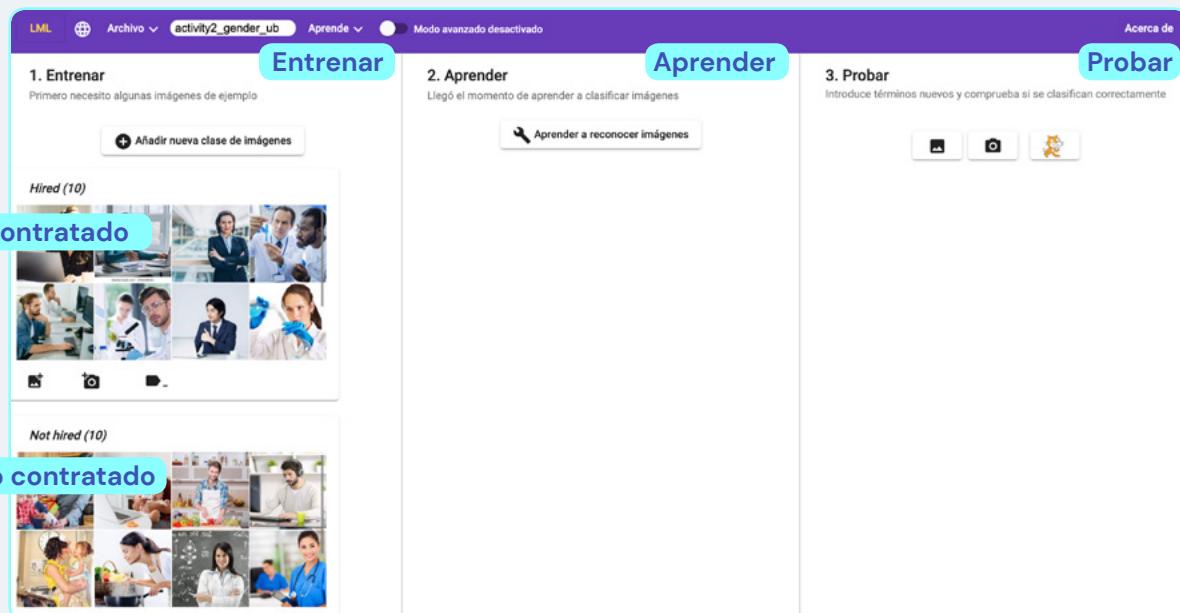
Paso 5 Haga clic en reconocer imágenes

Paso 6 Suba el archivo: **activity2_gender_ub.json**

The screenshot shows the 'Aprende' section of the LearningML interface. At the top, there is a toolbar with 'Archivo', 'sin nombre', 'Aprende', and a toggle switch for 'Modo avanzado desactivado'. Below the toolbar, there are two main sections: 'Textos' and 'Imagenes'. The 'Imagenes' section is highlighted with a blue circle containing the number '4'. It contains the text 'Enseña al ordenador a reconocer imágenes' and a button labeled 'Reconocer imágenes'. To the right of the 'Imagenes' section, there is a small preview window showing a green alien character. A blue circle with the number '5' is placed over the 'Reconocer imágenes' button. In the bottom right corner of the screen, there is a floating menu with options: 'Nuevo', 'Texto' (highlighted with a blue circle containing the number '6'), 'Cargar desde tu ordenador' (which is highlighted with a red rectangle), and 'Guardar en tu ordenador'. A blue circle with the number '7' is placed over the 'Cargar desde tu ordenador' option.

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

Nota: Como puede ver en la siguiente imagen, se han añadido las dos clases (**Contratado** y **No contratado**) con sus respectivas imágenes.



LearningML tiene 3 etapas como cualquier sistema de IA:

- 1. Entrenar:** En esta etapa alimentamos nuestro sistema o modelamos los datos que tenemos.
- 2. Aprender:** En esta etapa, el sistema está trabajando en el backend y encuentra patrones entre los datos y une las cosas que son similares para que pueda reconocer cualquier información desconocida.
- 3. Probar:** En esta etapa, el sistema está entrenado y listo para reconocer y categorizar la nueva imagen. Cuidado: su imagen de prueba no debe ser parte de sus datos de entrenamiento.

Tenemos dos categorías para clasificar a un candidato: Contratado y No Contratado. Cada clase necesita 10 imágenes, pero agregamos imágenes sin sesgo de género en este ejemplo. Añadimos el mismo número de imágenes de mujeres y hombres en el mismo tipo de situaciones.

Paso 7 Haga clic en aprender a reconocer imágenes:

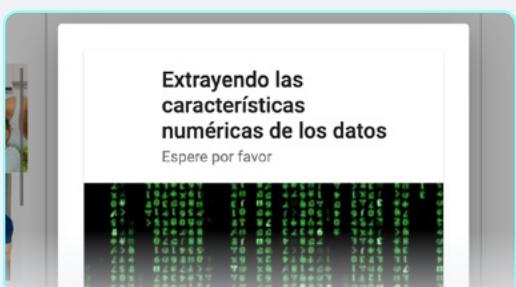
2. Aprender

Llegó el momento de aprender a clasificar imágenes

7

Aprender a reconocer imágenes

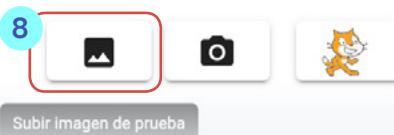
Extrayendo las características numéricas de los datos
Espere por favor



Paso 8 Suba la imagen de prueba

3. Probar

Introduce términos nuevos y comprueba si se clasifican correctamente



Resultados En esta actividad, en el test_image hay una mujer desarrolladora. Como resultado, el modelo clasifica a la candidata de la siguiente manera: **Contratado**

3. Probar

Introduce términos nuevos y comprueba si se clasifican correctamente



- Hired (98.59 %)
- Not hired (1.41 %)

Para Discutir

¿Crees que hay algún sesgo adicional en el conjunto de datos de imagen utilizado para entrenar el modelo?

Reflexión

Esta actividad presentó una valiosa lección al enfatizar la importancia de entrenar modelos de IA sin sesgos. Al incorporar conscientemente una representación equilibrada de ambos géneros en los datos, en situaciones similares, el objetivo era crear un modelo más justo para ayudar a las startups en su proceso de contratación de expertos en Informática.

Subraye la realidad de que el tipo y la cantidad de datos utilizados en el entrenamiento de modelos de IA desempeñan un papel fundamental en la determinación del comportamiento del modelo y los posibles sesgos. Los datos sesgados pueden perpetuar e incluso amplificar los sesgos sociales existentes y las prácticas discriminatorias, mientras que los datos imparciales pueden ayudar a contrarrestar tales sesgos y promover la equidad.

Es esencial reconocer que abordar el sesgo en los modelos de IA va más allá de los sesgos de género. Debemos permanecer vigilantes y asegurarnos de que también se tengan en cuenta los sesgos raciales, étnicos y de otro tipo. El objetivo debe ser crear modelos que no solo eviten preservar los sesgos existentes, sino que trabajen activamente para mitigar la discriminación de cualquier tipo.

Actividades como las anteriores sirven como un poderoso recordatorio de la responsabilidad ética que tenemos al desarrollar sistemas de IA. Destacan la necesidad de conjuntos de datos diversos y representativos, técnicas meticulosas de preprocesamiento y evaluación continua para identificar y rectificar los sesgos que puedan surgir.

Para fomentar modelos de IA imparciales, es crucial fomentar la colaboración interdisciplinaria entre científicos de datos, especialistas en ética, científicos sociales y expertos en el dominio. Al trabajar juntos, podemos desarrollar metodologías y directrices sólidas que prioricen la equidad, la transparencia y la rendición de cuentas en el desarrollo de la IA.

En conclusión, la reflexión sobre estas actividades enfatiza el papel significativo que juegan los datos y el sesgo en la creación de modelos de IA. Al reconocer y abordar activamente los sesgos, podemos esforzarnos por construir sistemas de IA que promuevan la igualdad, la justicia y la inclusión, ayudándonos a crear un futuro en el que las tecnologías de IA realmente sirvan y beneficien a todos los miembros de la sociedad.

3.6. Actividad 6: Identificar el sesgo en la IA

Objetivos de aprendizaje

1. **Comprender el concepto de sesgo en IA:** Los participantes obtendrán una comprensión integral del sesgo en la inteligencia artificial y su impacto potencial en los procesos de toma de decisiones. Explorarán diferentes formas de sesgo, incluido el sesgo de datos, el sesgo algorítmico y el sesgo social.
2. **Analizar técnicas de identificación de sesgos:** Los participantes aprenderán varios métodos y técnicas utilizados para identificar el sesgo en los sistemas de IA. Explorarán enfoques como el análisis de datos, las métricas de evaluación y las medidas de equidad para evaluar y detectar sesgos en los modelos de aprendizaje automático.
3. **Evaluar las implicaciones éticas de la IA sesgada:** Los participantes examinarán críticamente las consideraciones éticas asociadas con los sistemas de IA sesgados. Reflexionarán sobre las posibles consecuencias de la toma de decisiones sesgada en aplicaciones del mundo real y comprenderán la importancia de la equidad, la transparencia y la responsabilidad en el desarrollo y despliegue de la IA.
4. **Mejorar las habilidades de análisis de datos:** A través del ejercicio, los participantes mejorarán sus habilidades de análisis de datos mediante la exploración de conjuntos de datos y la identificación de posibles sesgos. Desarrollarán competencia en preprocesamiento de datos, selección de características y técnicas de visualización de datos para analizar e interpretar los sesgos presentes en los sistemas de IA.
5. **Fomentar la conciencia del impacto social:** Los participantes desarrollarán una comprensión de cómo los sistemas de IA sesgados pueden perpetuar los sesgos sociales existentes y reforzar la desigualdad. Explorarán casos y ejemplos para comprender el impacto social más amplio de una IA sesgada en áreas como la contratación, la justicia penal y la atención médica.
6. **Promover el desarrollo responsable de la IA:** Los participantes aprenderán la importancia del desarrollo responsable de la IA y la necesidad de incorporar la equidad, la diversidad y la inclusión en el diseño e implementación de los sistemas de IA. Considerarán estrategias para mitigar el sesgo y garantizar que las tecnologías de IA beneficien a todos los individuos y comunidades.
7. **Fomentar el pensamiento crítico y la resolución de problemas:** A través del ejercicio, los participantes mejorarán sus habilidades de pensamiento crítico y resolución de problemas. Aprenderán a abordar cuestiones complejas relacionadas con el sesgo en la IA con una mentalidad sistemática y analítica, evaluando diferentes perspectivas y proponiendo soluciones efectivas.
8. **Fomentar la colaboración y el pensamiento interdisciplinario:** Los participantes participarán en discusiones y actividades de colaboración para explorar el sesgo en la IA desde múltiples puntos de vista. Reconocerán la naturaleza interdisciplinaria de abordar el sesgo, fomentando la colaboración entre científicos de datos, especialistas en ética, formuladores de políticas y otras partes interesadas.

Materiales

Ordenador portátil con acceso a Internet

Cuenta de Google para usar Google Colab

No se requiere experiencia previa de programación

Descripción de la actividad

A lo largo del ejercicio, los participantes trabajarán con un conjunto de datos de Kaggle, una reconocida plataforma para la ciencia de datos y el aprendizaje automático. Se les presentarán diferentes formas de sesgo, como el sesgo de datos, el sesgo algorítmico y el sesgo social, y aprenderán cómo estos sesgos pueden manifestarse en los sistemas de IA. Los participantes tendrán la oportunidad de analizar el conjunto de datos, preprocesar los datos y emplear técnicas de análisis de datos para identificar posibles sesgos.

Utilizando métricas de evaluación y medidas de equidad, los participantes evaluarán el rendimiento y la equidad de los modelos de aprendizaje automático entrenados en el conjunto de datos. Obtendrán información sobre los desafíos de identificar y medir el sesgo en los sistemas de IA y aprenderán a interpretar los resultados de la evaluación del sesgo.

El ejercicio también enfatiza las implicaciones éticas de los sistemas de IA sesgados. Los participantes explorarán casos y ejemplos del mundo real que demuestran el impacto social de una IA sesgada en áreas como la contratación, la justicia penal y la atención médica. Participarán en discusiones sobre la importancia de la equidad, la transparencia y la rendición de cuentas en el desarrollo y despliegue de la IA.

Una guía paso a paso

- Paso 1** Descargue los recursos para este ejercicio visitando [el sitio web](#) y descargando los ficheros **Identifying_bias_in_AI.ipynb**, **toxic_data.csv** e **Instructions_notebook_bias.pdf**. Guárde los en una carpeta dedicada a esta actividad.

IO4 Handbook Activities ([Download](#))

- 1 • **Identifying_bias_in_AI.ipynb:** a jupyter notebook
- **toxic_data.csv:** a csv file containing the data to be used in the notebook
- **Instructions_notebook_bias.pdf:** a pdf containing instructions on how to use the notebook

Nota sobre los datos

A finales de 2017, la plataforma [Civil Comments](#), una aplicación de comentarios con todas las para sitios de noticias independientes, cerró y decidió hacer que sus ~ 2 millones de comentarios públicos de su plataforma estuvieran disponibles en un archivo abierto y duradero para que los investigadores pudieran comprender y mejorar el civismo en conversaciones en línea en los próximos años. Jigsaw, una unidad dentro de Google que explora las amenazas en sociedades abiertas y construye tecnología que inspira soluciones escalables, patrocinó este esfuerzo y extendió la anotación de estos datos por parte de evaluadores humanos con varios atributos conversacionales tóxicos. En este ejercicio, trabajará con un pequeño subconjunto de los datos que se utilizaron en el concurso [Jigsaw Unintended Bias in Toxicity Classification](#).

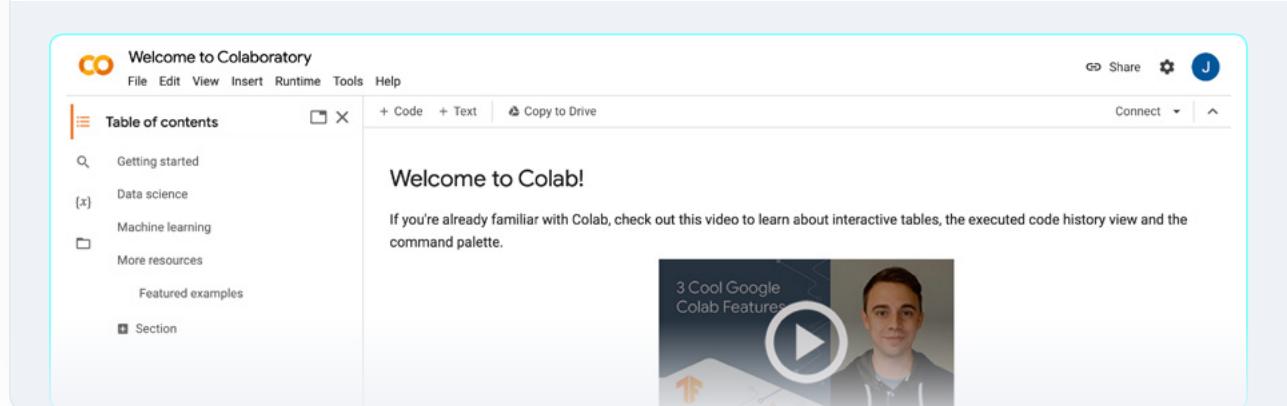
Paso 2 Conéctese en su cuenta de gmail (si aún no lo ha hecho).

Paso 3 Abra cualquier navegador de Internet que esté utilizando y escriba [Google Colab](#).

Nota: Google Colab es una plataforma en línea fácil de usar que le permite escribir y ejecutar código sin necesidad de ninguna experiencia previa en programación. Es como tener un cuaderno virtual donde puedes escribir y ejecutar código usando Python, un lenguaje de programación popular.

Con Google Colab, puede crear y editar celdas de código, que son como secciones individuales donde puede escribir y ejecutar su código. Proporciona una interfaz simple e interactiva donde puede escribir su código y ver la salida de inmediato.

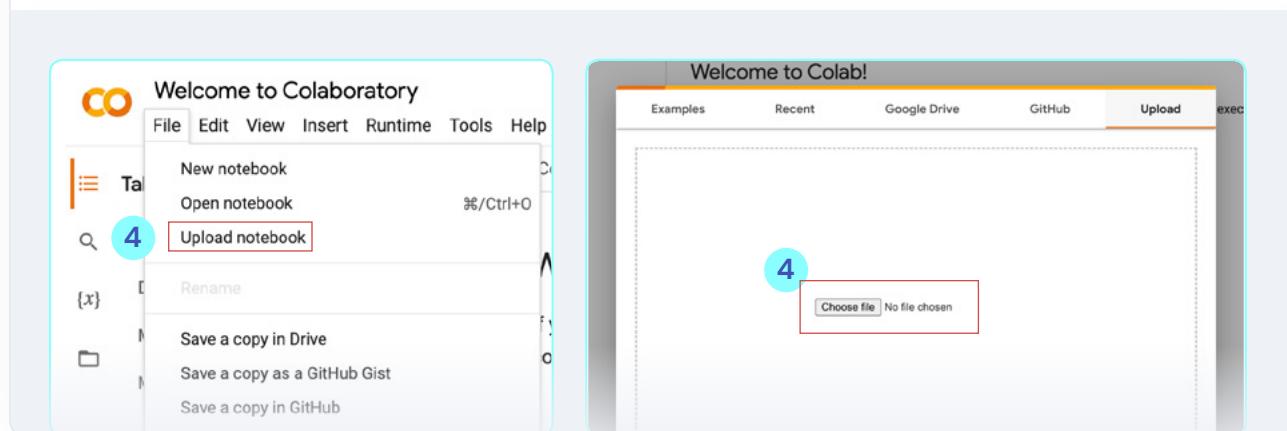
Una de las grandes ventajas de Google Colab es que viene preinstalado con muchas bibliotecas y frameworks populares, como TensorFlow y PyTorch, que se usan comúnmente en análisis de datos, aprendizaje automático y otros campos. Esto significa que puede explorar y trabajar fácilmente con funcionalidades avanzadas sin la necesidad de configuraciones o instalaciones complejas.



The screenshot shows the 'Welcome to Colaboratory' page in Google Colab. The left sidebar contains a 'Table of contents' with sections like 'Getting started', 'Data science', 'Machine learning', 'More resources', 'Featured examples', and 'Section'. The main area displays the text 'Welcome to Colab!' and a note: 'If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view and the command palette.' Below this is a video thumbnail with the title '3 Cool Google Colab Features' and a play button.

3. Directrices para actividades que promueven un enfoque inclusivo de la IA

Paso 4 Sube el archivo **Identifying_bias_in_AI.ipynb** que descargó inicialmente.



Paso 5 Una vez que esté cargado, esto es lo que verá. En el lado izquierdo está la lista de contenido del cuaderno para facilitar la navegación y en la ventana principal (centro derecho de la pantalla) puede ver el cuaderno real y su contenido desplazándose hacia abajo.

The image shows a screenshot of a Google Colab notebook titled 'Identifying_bias_in_AI.ipynb'. On the left, there is a 'Table of contents' sidebar with several sections expanded, including 'Identifying Bias in AI', 'Prepare data and train model', 'Importing all packages and load data', 'Data exploration', 'Data preprocessing', 'Working with words and vocabularies', 'Train model', 'How does the model make decisions?', 'Deep dive', and 'Other possible types of biases'. On the right, the main content area displays the first section of the notebook, titled 'Identifying Bias in AI'. The text describes the activity as a programming exercise based on 'identifying bias in AI', which was released under the Apache adaption has been development in the context of the erasmus+ project FAIAS. It explains that the exercise uses labeled data to train a model to decide whether a comment or a text is toxic or not. The text also mentions that the easiest way to get started is to open the notebook in Google Colab and run the code. The content area continues with information about code cells and package imports.

Paso 6 Siga las instrucciones del cuaderno para obtener la salida.

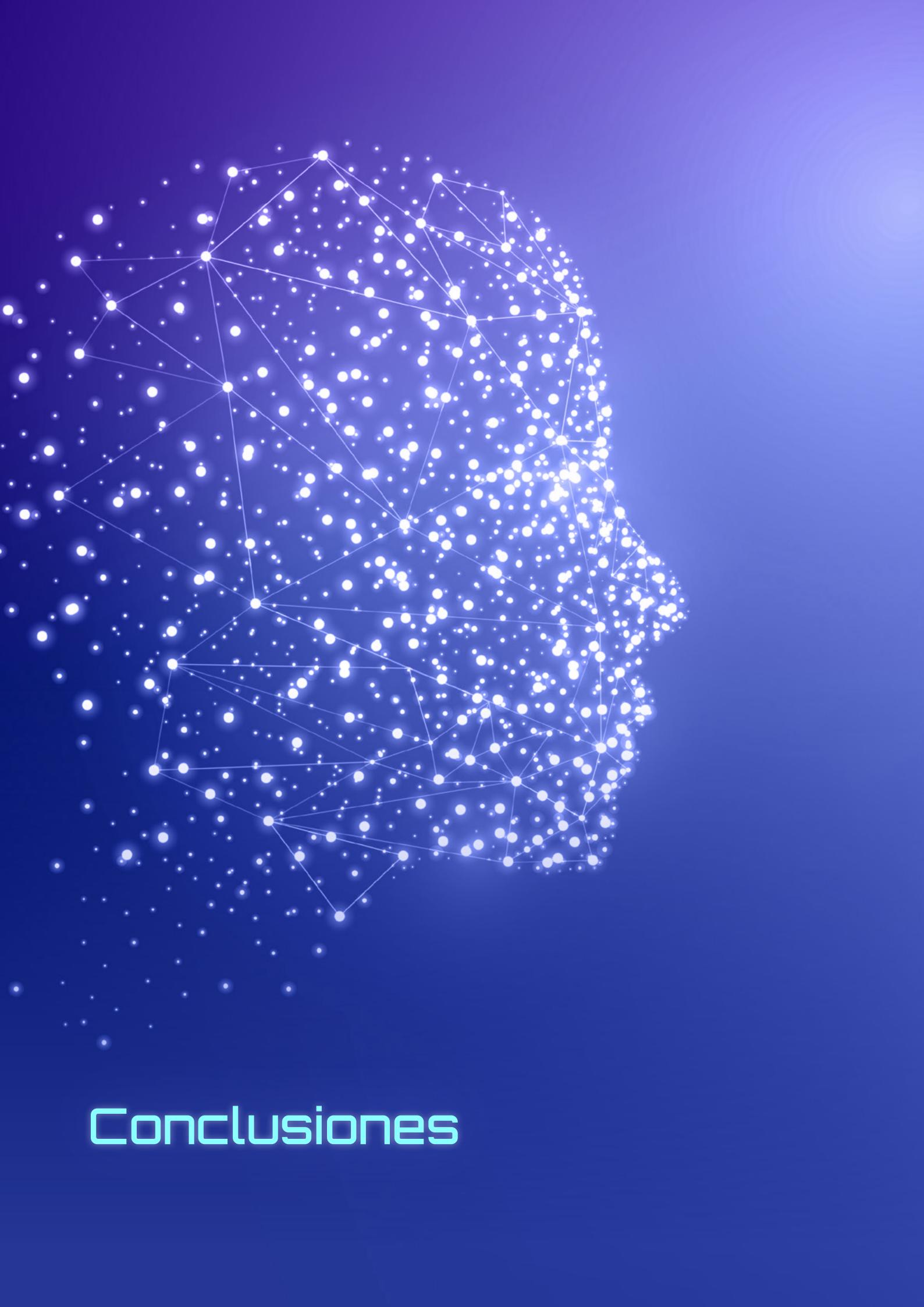
Reflexión

El análisis de conjuntos de datos del mundo real y la exploración de diversas técnicas para detectar sesgos en los modelos de IA permiten a los participantes comprender la complejidad y los desafíos asociados con este problema. Destacó la importancia de evaluar críticamente los datos utilizados para entrenar la actividad de IA y considerar las posibles consecuencias de los resultados sesgados.

La actividad también enfatizó las consideraciones éticas que rodean los sistemas de IA sesgados. Llevó a reflexionar sobre el impacto social de tales sesgos, particularmente en áreas como la contratación, la justicia penal y la atención médica. Sirvió como un recordatorio de la responsabilidad que tenemos como desarrolladores y usuarios de IA para garantizar la equidad, la transparencia y la rendición de cuentas en su diseño e implementación.

A través de este ejercicio, los participantes desarrollan aún más sus habilidades de análisis de datos, aprendiendo técnicas valiosas para el preprocesamiento de datos, la selección de características y la visualización de datos. Estas habilidades les permitieron identificar e interpretar posibles sesgos dentro de los modelos de IA, fortaleciendo su capacidad para analizar y abordar el sesgo en proyectos futuros.

Colaborar con otros participantes y participar en debates sobre el sesgo en la IA proporcionó una perspectiva más amplia sobre el tema. Escuchar diferentes puntos de vista y considerar diversas perspectivas mejoró mi comprensión de la naturaleza multifacética del sesgo y sus implicaciones en los sistemas de IA.



Conclusiones

En conclusión, el proyecto Fomentando la Inteligencia Artificial en las Escuelas (FAIaS) y la guía adjunta para educadores no formales proporcionan valiosos recursos y conocimientos para navegar por el complejo panorama de la ética en IA, el desarrollo responsable y las prácticas docentes inclusivas.

A medida que la IA continúa dando forma a nuestra sociedad, es crucial priorizar el despliegue ético y responsable de los sistemas de IA al tiempo que se garantiza la diversidad, la equidad y la inclusión.

La guía reconoce las ventajas de la educación no formal en la enseñanza de la IA y destaca su flexibilidad y adaptabilidad para satisfacer las necesidades de diversos estudiantes. Enfatiza la importancia de incorporar aspectos técnicos y éticos de la IA en los programas educativos, equipando a las personas con las habilidades y el conocimiento necesarios para comprender y navegar por las implicaciones de los sistemas de IA. Al promover experiencias prácticas, fomentar la creatividad y la innovación, apoyar a las comunidades subrepresentadas, fomentar el aprendizaje permanente y desarrollar habilidades prácticas, la educación no formal puede empoderar a las personas para prosperar en la era digital.

Además, la guía subraya la importancia de promover el género, la diversidad y la inclusión en los programas de educación no formal sobre IA. Al crear entornos de aprendizaje inclusivos que valoran y respetan las diversas identidades y experiencias de los estudiantes, los educadores pueden fomentar un sentido de pertenencia y apoyar el éxito académico y personal de todos los estudiantes. Las prácticas docentes inclusivas, como ser consciente de los prejuicios, usar un lenguaje inclusivo, abordar las necesidades individuales, crear espacios seguros y buscar apoyo cuando sea necesario, son cruciales para crear entornos de aprendizaje de apoyo y atractivos.

Al aprovechar el potencial de la educación no formal y adoptar prácticas docentes inclusivas, podemos construir un futuro más equitativo e inclusivo para la IA. El proyecto FAIaS y la guía para educadores no formales sirven como recursos importantes en este viaje, proporcionando a los educadores las herramientas y el conocimiento para empoderar a la próxima generación de profesionales y usuarios de IA. Al priorizar la ética, el desarrollo responsable y la inclusión, podemos garantizar que la tecnología de IA sirva a los mejores intereses de las personas y las comunidades, al tiempo que defiende los principios de equidad y justicia. Juntos, podemos dar forma a un futuro en el que la IA se aproveche para el beneficio de todos.

Definiciones y glosario [33], [34], [35]

Diversidad	Socialmente, se refiere a la amplia gama de identidades. En términos generales, incluye raza, etnia, género, edad, origen nacional, religión, discapacidad, orientación sexual, nivel socioeconómico, nivel educativo, estado civil, idioma, estado de veterano, apariencia física, etc. También involucra diferentes ideas, perspectivas y valores.
Sesgo	Prejuicio a favor o en contra de una cosa, persona o grupo en comparación con otra, generalmente de manera injusta o negativa. El sesgo inconsciente, también conocido como sesgo implícito, se define como "actitudes y estereotipos que influyen en el juicio, la toma de decisiones y el comportamiento de maneras que están fuera de la conciencia y / o control consciente".
Inclusión	El acto de crear un entorno en el que cualquier individuo o grupo será bienvenido, respetado, apoyado y valorado como miembro plenamente participante. Un clima inclusivo y acogedor abraza y respeta las diferencias.
Educación no formal	Educación institucionalizada, intencional y planificada por un proveedor de educación, con la característica definitoria de que es una adición, alternativa y/o complemento de la educación formal dentro del proceso de aprendizaje permanente de los individuos. A menudo se proporciona para garantizar el derecho de acceso a la educación para todos. Atiende a personas de todas las edades, pero no necesariamente aplica una estructura de vía continua. Puede ser de corta duración y / o baja intensidad, y generalmente se proporciona en forma de cursos cortos, talleres o seminarios. La educación no formal conduce principalmente a cualificaciones que no son reconocidas como cualificaciones formales por las autoridades educativas nacionales pertinentes o a ninguna cualificación en absoluto. La educación no formal puede abarcar programas que contribuyen a la alfabetización y educación de adultos y jóvenes para niños que no asisten a la escuela, así como programas sobre preparación para la vida, habilidades laborales y desarrollo social o cultural.
Prejuicio	Un juicio o preferencia preconcebida, especialmente una que interfiere con el juicio imparcial y puede estar arraigada en estereotipos, que niega el derecho de los miembros individuales de ciertos grupos a ser reconocidos.
Estereotipo	Una forma de generalización arraigada en creencias generales y suposiciones falsas, producto de procesos de categorización que pueden resultar en una actitud prejuiciosa, juicio crítico y discriminación intencional o no intencional. Los estereotipos son típicamente negativos, basados en poca información y no reconocen el individualismo y la agencia personal.

Referencias

1. Rosenfeld, Hannah, Danni Yu, and Abhishek Gupta. 2023. "Ethical AI requires future-ready, inclusive education system | World Economic Forum." The World Economic Forum, June 5, 2023. <https://www.weforum.org/agenda/2023/06/ethical-ai-future-ready-inclusive-education-system/>
2. Blackman, Reid. 2020. "A Practical Guide to Building Ethical AI." Harvard Business Review, October 15, 2020. <https://hbr.org/2020/10/a-practical-guide-to-building-ethical-ai>
3. Burton, Emanuelle, Judy Goldsmith, Sven Koenig, Benjamin Kuipers, Nicholas Mattei, and Toby Walsh. 2017. "Ethical Considerations in Artificial Intelligence Courses." AI Magazine 38(2) (July): pp22–34. <https://doi.org/10.1609/aimag.v38i2.2731>
4. Norqvist, Lars, and Eva Leffler. 2017. "Learning in non-formal education: Is it "youthful" for youth in action?" Springer link 63 (March): pp235–256. <https://doi.org/10.1007/s11159-017-9631-8>
5. Zowghi, Didar, and Francesca da Rimini. 2023. Responsible AI: Best Practices for Creating Trustworthy AI Systems. N.p.: Pearson Addison Wesley. <https://doi.org/10.48550/arXiv.2305.12728>
6. Lang, Michael, Mark Freeman, Gaye Kiely, and Amy B. Woszczynski. 2022. "Special Issue Editorial: Equality, Diversity, and Inclusion in IS Education." Journal of Information Systems Education 33, no. 1 (February): pp1–6.
7. Perry, Sylvia P., Mary C. Murphy, and John F. Dovidio. 2015. "Modern prejudice: Subtle, but unconscious? The role of Bias Awareness in Whites' perceptions of personal and others' biases." Journal of Experimental Social Psychology 61 (November): pp64–78.
8. Warikoo, Natasha, Stacey Sinclair, Jessica Fei, and Drew Jacoby-Senghor. n.d. "Examining Racial Bias in Education: A New Approach." Educational Researcher 45 (9). <https://doi.org/10.3102/0013189X16683408>
9. Zittleman, Karen, and David Sadker. 2022. "Gender Bias in Teacher Education Texts: New (and Old) Lessons." Journal of Teacher Education 53, no. 2 (March): 168–180. <https://doi.org/10.1177/0022487102053002008>
10. Spencer, Barbara. 1980. "Overcoming the Age Bias in Continuing Education." New Directions for Continuing Education 8: pp71–86.
11. Willis, Robert J., and Sherwin Rosen. 1979. "Education and Self-Selection." Journal of Political Economy 87, no. 5 (October). <https://doi.org/10.1086/260821>
12. Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." Journal of Economic Literature 40, no. 2 (June): pp436–482. <https://doi.org/10.1257/002205102320161258>.
13. Derman-Sparks, Louise, and Julie O. Edwards. 2019. "Understanding Anti-Bias Education." National Association for the Education of Young Children 74, no. 5 (November): pp6–13. <https://www.jstor.org/stable/26842300>
14. Stubbs, Sue. 2008. Inclusive Education. N.p.: The Atlas Alliance, pp69–106.

15. Mohammed, Phaedra S., and Eleanor N. Watson. 2019. "Towards Inclusive Education in the Age of Artificial Intelligence: Perspectives, Challenges, and Opportunities." In *Artificial Intelligence and Inclusive Education: Speculative Futures and Emerging Practices*, edited by Jeremy Knox, Yuchen Wang, and Michael Gallagher, 17–37. N.p.: Springer Nature Singapore. https://doi.org/10.1007/978-981-13-8161-4_2
16. Xu, Feiyu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. 2019. "Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges." In *Lecture Opmerkings in Computer Science*, pp563–574. Vol. 11839. N.p.: Springer. https://doi.org/10.1007/978-3-030-32236-6_51
17. Gururaj, Tejasri. 2023. "10 examples of how artificial intelligence is improving education." *Interesting Engineering*. <https://interestingengineering.com/lists/examples-how-artificial-intelligence-improving-education>
18. Morron, Mirriam. 2023. "The Impact of AI on Learning and Creativity: A Critical Look at the Future." LinkedIn. <https://www.linkedin.com/pulse/impact-ai-learning-creativity-critical-look-future-mirriam-morron/>
19. Ullman, Stefanie. 2022. "Gender Bias in Machine Translation Systems." In *Artificial Intelligence and Its Discontents*, pp123–144. 10.1007/978-3-030-88615-8_7.
20. Hao, Karen. 2019. "Facebook's ad-serving algorithm discriminates by gender and race." MIT Technology Review. <https://www.technologyreview.com/2019/04/05/1175/facebook-algorithm-discriminates-ai-bias/>
21. Dastin, Jeffrey. 2018. "Amazon scraps secret AI recruiting tool that showed bias against women." Reuters. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
22. Haag, Christian, and Alexander Smith. 2019. "5 Examples of Biased Artificial Intelligence." Logically. <https://www.logically.ai/articles/5-examples-of-biased-ai>
23. Manyika, James, Jake Silberg, and Britanny Presten. 2019. "What Do We Do About the Biases in AI?" Harvard Business Review. <https://hbr.org/2019/10/what-do-we-do-about-the-biases-in-ai>
24. Gonzalez, Glen. 2020. "Breaking Bias." Smith Communication Partners. <https://smithcp.com/breaking-bias/>
25. "Visa SVP Carolina Barcenas: "Educating the general public on AI is key to ethical use and smart" 2019. Medium. <https://medium.com/authority-magazine/visa-svp-carolina-barcenas-educating-the-general-public-on-ai-is-key-to-ethical-use-and-smart-ea080f37966c>
26. "Role Models in AI: Pasohanie Lampkin | by AI4ALL Team | AI4ALL." 2019. Medium. <https://medium.com/ai4allorg/role-models-in-ai-Pasohanie-lampkin-f50e6885afa4>
27. "Female role models changing the field of Artificial Intelligence in Norway – NORA." n.d. Norwegian Artificial Intelligence Research Consortium. Accessed June 16, 2023. <https://www.nora.ai/news-and-events/news/30-women-in-norway-changing-the-field-of-ai.html>
28. "Role Models in AI: Rebekah Agwunobi | by AI4ALL Team | AI4ALL." 2018. Medium. <https://medium.com/ai4allorg/role-models-in-ai-rebekah-agwunobi-f3bc438580da>
29. "The tag game | Activities." n.d. Scouts. Accessed June 16, 2023. <https://www.scouts.org.uk/activities/the-tag-game/>

30. "Describe and draw | Activities." n.d. Scouts. Accessed June 16, 2023. <https://www.scouts.org.uk/activities/draw-the-description/>
31. Katherine Johnson, https://en.wikipedia.org/wiki/Katherine_Johnson
32. "Inspirational women: Hedy Lamarr - Code Week." 2022. EU Code Week blog. <https://blog.codeweek.eu/inspirational-women-hedy-lamarr/>
33. UNESCO Institute for Statistics, Glossary, <http://uis.unesco.org/en/glossary>
34. College of the Environment, University of Washington, Diversity, Equity and Inclusion Glossary ,<https://environment.uw.edu/about/diversity-equity-inclusion/tools-and-additional-resources/glossary-dei-concepts/>
35. Harvard Human Resources, Glossary of Diversity, Inclusion and Belonging Terms, https://edib.harvard.edu/files/dib/files/dib_glossary.pdf

Créditos de imágenes

Portada, contraportada – Diseño propio, utilizando imagen diseñada por [shuvojit11 / Freepik](#)

Paginas 6, 14 – Diseño propio, utilizando imagen diseñada por [rawpixel.com / Freepik](#)

Pagina 28 – Imagen diseñada por [rawpixel.com / Freepik](#)

Pagina 34 – Imagen “[Katherine Johnson Receives Presidential Medal of Freedom \(NHQ201511240003\)](#)” por [NASA HQ PHOTO](#). licenciado bajo [CC BY-NC-ND 2.0](#)

Pagina 38 – Imagen “[Hedy Lamarr](#)” de [ReznorH](#). está marcado con [Domínio Público 1.0](#)

Paginas 43 a 61 – Capturas de pantalla tomadas del sitio web [learningml.org](#)

Pagina 57 – Imagen de sección diseñada por [Freepik](#)

Pagina 63 – Fuente propia

Fomentando la Inteligencia Artificial en las Escuelas

Guía de inclusión para educadores (no formales) en el contexto de la educación en Inteligencia Artificial



Co-funded by the
Erasmus+ Programme
of the European Union



Universidad
Rey Juan Carlos

Teatro Circo de Braga
EM, S.A.

