

Programa de Inteligencia Artificial para 1° y 2° medio:

FUNDAMENTOS CIUDADANÍA Y ÉTICA



ÍNDICE

CRÉDITOS	3
INTRODUCCIÓN	4
ACTIVIDAD 1: ¿QUÉ ES Y PARA QUÉ USAMOS LA IA?	10
ACTIVIDAD 2 ALGORITMOS COMO OPINIONES	27
ACTIVIDAD 3: MATIZ ÉTICA	35
ACTIVIDAD 4: INTRODUCCIÓN AL APRENDIZAJE AUTOMATIZADO SUPERVISADO Y SESGO ALGORÍTMICO	44
ACTIVIDAD 5: CASO I: “FALSIFICACIÓN DE INFORMACIÓN” (PARTE 1)	94
ACTIVIDAD 6: CASO I: “FALSIFICACIÓN DE INFORMACIÓN” (PARTE 2)	109
ACTIVIDAD 7: CASO II: ¿CÓMO ENSEÑAREMOS A LA IA A MANEJAR?	117
ACTIVIDAD 8: BÚSQUEDA DEL TESORO EN YOUTUBE	129
ACTIVIDAD 9: REDISEÑO DE YOUTUBE	137
ACTIVIDAD 10: SEMINARIO SOCRÁTICO DE YOUTUBE	148
ACTIVIDAD 11: CASO III: ¿CÓMO NOS PREPARAMOS PARA LA SIGUIENTE PANDEMIA?	155
ACTIVIDAD 12: CASO IV: ¿DEBEMOS ADOPTAR LAS TECNOLOGÍAS BASADAS EN IA?	168
AGRADECIMIENTOS	182

CRÉDITOS

Este programa fue desarrollado por el Centro de Innovación del Ministerio de Educación de Chile, en conjunto con el Portal Educarchile, proyecto de colaboración entre Fundación Chile y el ente gubernamental.

Este material está licenciado con Creative Commons CC-BY-NC, que permite adaptar los contenidos libremente, de manera no comercial y siempre que se incluya el reconocimiento a los creadores.

Utilice el siguiente texto para reconocer este trabajo: “El Programa de inteligencia artificial para 1° y 2° medio: fundamentos, ciudadanía y ética fue desarrollado por el Centro de Innovación del Ministerio de Educación chileno en conjunto con Fundación Chile, en el marco del proyecto Portal Educarchile”.

Puede encontrar más información sobre esta licencia en <https://creativecommons.org/licenses/by/4.0/>.

Las actividades 1-4 y 8-10 son una adaptación de los recursos del programa *An Ethics of Artificial Intelligence Curriculum for Middle School Students*¹, ampliados y reinterpretados para el contexto nacional.

Recibiremos sugerencias, consultas, experiencias de uso y reportes de errores en innovacion@mineduc.cl.

Última actualización: abril de 2021.

CENTRO DE INNOVACIÓN

MINISTERIO DE EDUCACIÓN



¹ Programa creado por Blakeley H. Payne con el apoyo del Personal Robots Group del MIT Media Lab, dirigido por Cynthia Breazeal.

INTRODUCCIÓN

Cada día es más común ver servicios de entretenimiento potenciados por inteligencia artificial (IA) que permiten reemplazar rostros en tiempo real, recibir recomendaciones personalizadas de contenido (sugerencias de noticias, videos, rutas de transporte, etc.) o añadir color y resolución a un video antiguo. Sin embargo, el campo de aplicación de la inteligencia artificial es mucho más amplio y transversal y tiene el potencial de comprometer derechos esenciales cuando no se definen marcos éticos de operación, pudiendo, por ejemplo, vulnerar la privacidad de datos, favorecer la discriminación arbitraria, la manipulación o comprometer la seguridad y la honra de las personas. Una IA correctamente implementada, por otro lado, puede jugar un rol fundamental en la planificación de ciudades, la anticipación de desastres y conductas de riesgo, la investigación científica, la eficiencia de las industrias o la apertura de nuevos caminos para la creatividad artística.

Este programa busca entregar herramientas básicas para que docentes y estudiantes comprendan, por un lado, **el funcionamiento básico de las aplicaciones basadas en inteligencia artificial** y, por el otro, las **implicancias éticas y ciudadanas** que tienen las decisiones de diseño e implementación de estas tecnologías, en cuanto involucran competencias fundamentales para las y los futuros ciudadanos, no sólo como desarrolladores de tecnología sino como profesionales y técnicos de las más diversas áreas del conocimiento.

Estructura del programa

El programa consta de 12 actividades “autoexplicativas” de desarrollo teórico, práctico y de análisis, compuestas por una guía de contenidos (incluida como capítulo en este documento) y, en la mayoría de los casos, por una presentación de diapositivas. La guía de contenidos está conformada por una ficha introductoria, una guía docente con propuestas de interacción en aula y una o más hojas de trabajo para las/los estudiantes. La actividad 1 integra un capítulo adicional con material de profundización, que entrega conceptos y distinciones de corte introductorio. Todas las actividades incluyen una parte expositiva y una actividad práctica.

El material fue desarrollado para implementarse de manera secuencial, dada la incorporación gradual de conceptos, reflexiones e instancias de aplicación. Sin embargo, hay grupos de actividades que se pueden implementar de manera modular, como los estudios de caso y algunas actividades de experimentación con tecnología, siempre que el/la docente se familiarice antes con los conceptos y distinciones de las actividades precedentes y resguarde los espacios para entregar los contenidos necesarios. Las actividades se pueden ordenar funcionalmente de la siguiente manera:

	1	2	3	4	5	6	7	8	9	10	11	12
Actividades con componentes teóricos clave	X	X	X	X								
Actividades con experimentación de tecnologías de IA				X	X	X						
Actividades de estudio de casos					X	X	X				X	X
Actividades de aplicación de aprendizajes							X	X	X	X		X

Propuesta curricular y extracurricular

Las actividades se encuentran localizadas curricularmente en 1º y 2º Medio, en las asignaturas troncales de **Tecnología** (Ejes “Resolución de Problemas Tecnológicos” y “Tecnología, Ambiente y Sociedad”) e **Historia, Geografía y Ciencias Sociales** (Ejes “Formación ciudadana” y “Geografía”), con oportunidades de integración curricular con **Matemática, Inglés y Lengua y Literatura**. La siguiente tabla muestra la propuesta curricular global del programa:

	Tecnología		Historia		Matemática		Inglés		Lengua		Duración
	I M	II M	I M	II M	I M	II M	I M	II M	I M	II M	
1	OA1	OA5			U2	OA2					90 minutos
2	OA2 OA3	OA5									50 minutos
3	OA2 OA3	OA5	OA19	OA24							60 minutos
4	OA2 OA3	OA2		OA24							150 a 160 minutos
5	OA5	OA5 OA6		OA22			OA1 OA9	U1			75 minutos
6	OA5	OA5 OA6		OA22							60 minutos
7	OA5	OA5		OA22							90 minutos

	OA6	OA6		OA24						
8	OA3	OA6	OA22							60 minutos
9	OA2 OA3	OA3 OA5	OA23	OA22						90 minutos
10	OA4	OA3 OA6	OA23	OA22				OA22	OA22	60 a 80 minutos
11	OA2 OA4 OA5 OA6	OA5 OA6	OA22 OA25	OA22 OA23				OA10 OA21	OA10 OA21	90 minutos
12	OA5 OA6	OA3 OA5 OA6	OA25	OA23				OA10 OA21	OA10 OA21	90 minutos

El programa está planteado como una oportunidad de “codocencia”, mediante la alternancia de las dos asignaturas troncales y la integración de otras especialidades, en un diálogo interdisciplinario y flexible. En este sentido, la propuesta curricular es una guía general que deberá ser revisada y adaptada a los requerimientos de cada equipo profesional, pudiendo integrar nuevas asignaturas.

Dado el interés general de las actividades y su enfoque práctico, es posible implementar el programa en talleres extraprogramáticos de tecnología, ciencias o Formación Ciudadana, entre otros, así como adaptar el material a diferentes niveles de Educación Básica o Media. Puede utilizarse también para complementar programas de pensamiento computacional o como módulo introductorio para el uso de herramientas de inteligencia artificial en talleres de robótica educativa (sobre todo la actividad 4).

Conocimientos previos

El programa no requiere contar con conocimiento técnico ni experiencia previa en implementación de talleres de programación o ética de la tecnología, toda vez que la información necesaria se encuentra contenida en el programa, contextualizada y dosificada. En cualquier caso, es necesario explorar y familiarizarse con los conceptos y herramientas propuestas en cada actividad.

Es importante señalar que la inteligencia artificial sigue un camino paralelo —y complementario— al de la programación “convencional”. En términos simplificados, las herramientas de pensamiento computacional (por ejemplo, la programación por bloques), están ligadas a desafíos que requieren

soluciones (algoritmos) precisas (y no ambiguas), tales como llevar a un personaje de un punto A a un punto B. La IA, por su parte, está generalmente asociada al “aprendizaje” de modelos algorítmicos a partir de ejemplos previos, lo que le permite encontrar estrategias variadas y flexibles en contextos complejos, como lograr que un vehículo real llegue de un punto de la ciudad a otro sin conductor.

Acceso a Internet y requerimientos del sistema

Casi la mitad de las actividades del programa se pueden implementar sin necesidad de conexión a Internet, es decir, empleando únicamente los contenidos descargados (hojas de trabajo, presentaciones de diapositivas, entre otros) y los materiales que se especifican en las fichas iniciales. Sin embargo, por la naturaleza de la temática, existe un segundo grupo de actividades que requieren experimentar con plataformas a través del navegador. Un tercer conjunto corresponde a actividades que cuentan con adaptaciones para trabajar sin conexión, como se muestra en la siguiente tabla:

	1	2	3	4	5	6	7	8	9	10	11	12
Actividades sin conexión a Internet (<i>offline</i>)	X	X	X							X	X	
Actividades con conexión a Internet				X	X	X		X	X			
Actividades con variantes sin conexión							X					X

El ancho de banda de Internet, la memoria RAM y la capacidad de la CPU requeridos para descargar las actividades y visualizar los contenidos no son diferentes a los necesarios para realizar tareas cotidianas, tales como usar Google Drive, correo electrónico y YouTube. La actividad 4, por su parte, presenta la plataforma web Teachable Machine, que si bien no exige ancho de banda y memoria RAM mayores a los anteriormente descritos, requiere una capacidad de CPU superior al promedio de los usos del navegador, ya que realiza el procesamiento de los modelos de manera local.

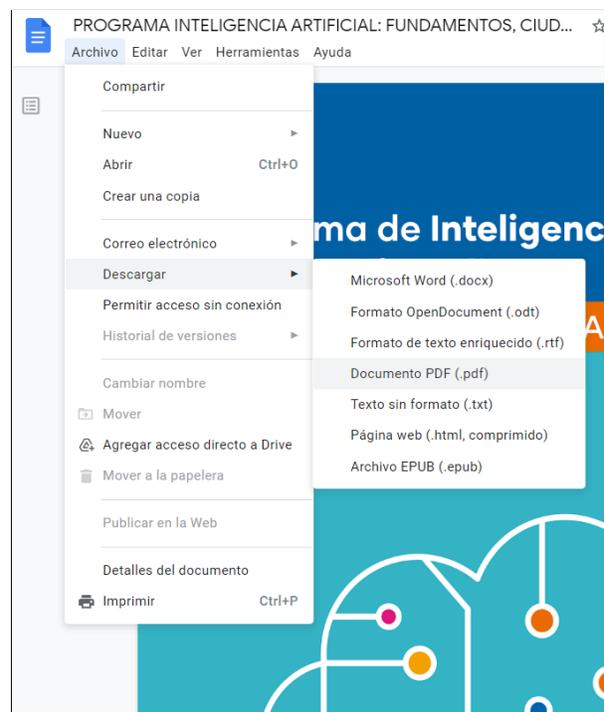
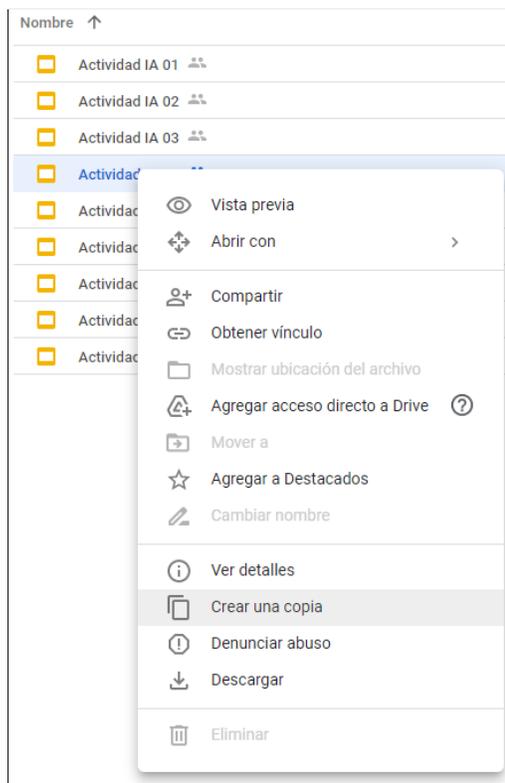
Modo de uso

Este programa fue desarrollado en formato Google Suite o Google for Education, que permite la actualización permanente e inmediata del material, así como una amplia variedad de formatos de descarga y uso, además de ser una de las plataformas más extendidas en el ámbito escolar.

Algunas actividades contienen videos y/o plataformas que se encuentran originalmente en idioma inglés, en cuyo caso se entregan las instrucciones para traducir de manera automática, así como la previsión de los errores que se pueden encontrar. En cualquier caso, recomendamos usar el navegador de Google Chrome para trabajar.

Existen tres formas para explorar, guardar y gestionar el material:

1. **Crear una copia en línea en su Drive.** Para ello, ingrese a [la carpeta del programa](#), haga clic derecho sobre el material que desee guardar en su unidad y seleccione “crear una copia” (imagen izquierda). Esto le permitirá contar con su propia versión del programa, compartirla con otras personas y editarla para que se ajuste a los requerimientos de su asignatura, equipo docente o grupo curso. Recomendamos esta opción, si cuenta con la Suite de Google (institucional) o con una cuenta en Gmail.
2. **Descargar en formato de Microsoft Office** (Word y PowerPoint) para guardar y editar de manera local. Para ello, abra el recurso que quiere descargar, vaya a *Archivo > Descargar* y seleccione Microsoft Word o PowerPoint (imagen derecha). Recuerde que pueden perderse y/o deformarse algunas características del diseño en el proceso de conversión.
3. **Descargar en formato PDF.** Siga los mismos pasos del punto anterior, pero seleccione .pdf como formato de salida. Esta fórmula minimiza errores de diseño en la conversión y genera documentos listos para imprimir, aunque no se pueden editar.



Accesos útiles

- Carpeta en línea con todo el material del programa:
<https://drive.google.com/drive/folders/1QYsQhVci0LRLLoDuLH4a7vUSUxNgqkGEm?usp=sharing>
- Carpeta en línea con presentaciones de diapositivas:
<https://drive.google.com/drive/folders/1kgva9NXatKTpNI26TyuDGKQjUMRUQARB?usp=sharing>
- Programa en versión descargable (ZIP):
<https://drive.google.com/drive/folders/1yb638D6cqwCUBTzxRjqK6gc8lq5kzLkN?usp=sharing>
- Sitios de la iniciativa:
<https://www.innovacion.mineduc.cl/iniciativas/plan-nacional-de-lenguajes-digitales/inteligencia-artificial>
<https://www.educarchile.cl>

ACTIVIDAD 1

¿QUÉ ES Y PARA QUÉ USAMOS LA IA?

Oportunidades curriculares

Tecnología

- **I° Medio - OA 1:** Identificar oportunidades o necesidades personales, grupales o locales que impliquen la creación de un servicio, utilizando recursos digitales u otros medios.
- **II° Medio - OA 5:** Usar software de dibujo para crear y representar diferentes ideas por medio de imágenes.

Matemática

- **I° Medio - Unidad 2:** Incorporar la noción de linealidad en dos variables (OA 4, OA 5 y OA 6)
- **II° Medio - OA 2:** Mostrar que comprenden las relaciones entre potencias, raíces enésimas y logaritmos (aplicado a noción de algoritmo)

Visión general

La actividad introduce elementos centrales de los sistemas de inteligencia artificial (IA), a partir de conocimientos previos, conceptos básicos y casos de uso diario. Las y los estudiantes indagan en el rol que juegan los datos que entregamos a diferentes plataformas y dispositivos, finalizando con un juego grupal, en el cual deben establecer cuáles son las predicciones y conjuntos de datos que utilizan diferentes sistemas de IA.

Objetivos

- Comprender las aplicaciones y partes de la inteligencia artificial.
- Comprender, mediante la exploración de casos, el concepto de predicción y conjuntos de datos como objetivos de la inteligencia artificial.

Agenda (90 minutos)

- Introducción (base teórica) (50 min)
 - ¿Qué es un algoritmo?
 - Problemas irresolubles para los algoritmos "tradicionales"
 - Aplicaciones comunes de la IA
 - Partes de la IA
- Desarrollo (30 min)
- Cierre (10 min)

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Hoja del estudiante (una por grupo)
- Hoja de trabajo "Bingo de IA" (una por grupo)
- Material de profundización para el/la docente

Palabras clave

Algoritmos, problemas irresolubles para algoritmos "convencionales", sistemas de inteligencia artificial (IA), partes de la IA.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1bV2tZDRJLUNb-iL2gtbFQc5fhedQx0cOJvc_UBAzcH8/edit?usp=sharing

GUÍA DOCENTE

Introducción (50 minutos)

Esta actividad introduce a los y las estudiantes en el léxico esencial para comprender el funcionamiento y las implicancias de los sistemas de inteligencia artificial en servicios de uso cotidiano. Por su carácter introductorio y densidad teórica, sugerimos estudiar el material anexo que se incluye al final de la actividad, en donde encontrará las explicaciones necesarias para comprender lo señalado y profundizar en lo que considere más pertinente.

Una vez revisado el material, dispóngase a preparar su clase, considerando entre 40 y 60 minutos para la parte expositiva. Para ello, se ha elaborado una presentación de diapositivas con una síntesis de los conceptos revisados en el material de estudio (puede utilizar los ejemplos y planteamientos sugeridos en el mismo). Revise las recomendaciones de interacción pedagógica en la guía de la actividad y las notas a pie de página de las láminas.

En la mitad de la presentación, se incluye un espacio optativo de evaluación formativa que le permitirá constatar si los y las estudiantes han aprendido a diferenciar un algoritmo “convencional” de un sistema de inteligencia artificial, distinción fundamental para seguir avanzando en el programa. El o la docente pregunta de forma oral por cinco ejemplos de servicios y las y los estudiantes contestan de qué tipo se tratan.

Desarrollo (30 minutos)²

Una vez finalizada la presentación, y habiendo explicado las partes de la IA, estará en condiciones de presentar a las y los estudiantes el juego “Bingo de IA”. Para ello, forme grupos de dos a cuatro integrantes, lo que les permitirá compartir y analizar sus respuestas. Cada equipo deberá contar con su hoja del estudiante y su hoja de trabajo con el bingo.

Una vez definido el tamaño de los grupos, reparta las hojas, explique las instrucciones contenidas en la guía del estudiante, lea los encabezados de la plantilla del bingo y comience el juego. Las hojas presentan 24 aplicaciones de inteligencia artificial en plataformas tecnológicas de uso cotidiano, tales como redes sociales, apps para celular y correo electrónico. Los grupos tendrán que deducir los conjuntos de datos sobre los que

² Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

estas aplicaciones operan, así como la predicción que están intentando hacer (pueden investigar los fundamentos de estos sistemas haciendo uso de computadores). La actividad durará hasta que uno de los grupos complete una de las cinco líneas horizontales o verticales o una diagonal y puedan demostrar la relación que hay entre cada sistema de inteligencia artificial, el conjunto de datos y la predicción que propusieron.

Debe esperar a que un grupo diga “Bingo” para revisar las respuestas frente al curso y verificar que la información entregada sea la correcta. Si así lo considera, lleve algún premio para estimular la actividad.

Cierre (10 minutos)

Para concluir, elabore con los y las estudiantes una síntesis con los principales conceptos aprendidos. Utilice la pizarra y tome nota. Revise las respuestas, reforzando las correctas y corrigiendo las erróneas. Enfatice la importancia de comprender las aplicaciones de IA que se utilizan en la vida cotidiana y que producen un efecto en nuestras vidas.

Algunas preguntas de cierre que puede realizar son:

- ¿Qué es la inteligencia artificial?
- ¿Para qué la usamos?
- ¿Por qué pueden afectar nuestra vida de manera positiva o negativa?
- ¿Son neutrales los algoritmos y los sistemas de IA?

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Juego de bingo IA**

Asignatura: Tecnología / Matemática

Fecha: _____ **Curso:** _____

Objetivos:

- Comprender las aplicaciones y partes de la inteligencia artificial.
- Comprender, mediante la exploración de casos, el concepto de predicción y conjuntos de datos como objetivos de la inteligencia artificial.

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____
4. _____

Nombre de la actividad: Algoritmos como opiniones

Introducción

Ahora que han aprendido sobre algoritmos y aplicaciones de inteligencia artificial, vamos a poner a prueba esos conocimientos.

Este juego busca que sean capaces de identificar los datos que necesitan diferentes sistemas de IA para funcionar y el resultado o predicción que están intentando hacer. Piensen con calma: ¿con qué información está trabajando cada uno de los sistemas de IA que se presentan en la hoja del bingo? ¿Qué es lo que están tratando de obtener, a partir de los datos que conocen, para realizar predicciones en escenarios nuevos? ¿Para qué fueron entrenados esos algoritmos? ¿Qué buscan reconocer o predecir?

En el siguiente ejemplo: “Reconocimiento de la placa patente de un vehículo”, los **datos** (con los que se entrenó el modelo) serían fotos de numerosas placas patente con sus respectivos datos de lectura y la **predicción** la identificación de una placa específica, dada una imagen de la patente.

Instrucciones del juego de Bingo IA

1. La o el profesor entregará una hoja con el bingo por equipo. Estas hojas tienen una serie de recuadros, cada uno de los cuales presenta un sistema de inteligencia artificial que pueden encontrar en un celular, computador u otro dispositivo.
2. El juego comienza cuando la o el profesor lee frente a todo el curso los encabezados de cada recuadro del bingo.
3. El equipo debe pensar y anotar, en cada caso, el conjunto de datos con el que trabaja el sistema de IA y la predicción que está intentando realizar.
4. Cuando un equipo complete los cinco recuadros de una línea de su hoja de juego, o bien una diagonal, debe levantar la mano y en voz alta señalar **¡¡¡BINGO!!!**
5. El juego se detiene y se revisan los resultados en conjunto con todo el curso.
6. Si las respuestas están correctas, **¡el equipo es el ganador!** En caso contrario, se reanuda el juego hasta que un grupo responda todo de manera correcta.

Tiempo estimado de actividad: 30 minutos.

BINGO DE I.A.

<p>Obtuvo un pronóstico del tiempo de un sitio web o usó una aplicación del clima</p> <p>Conjunto de datos: Predicción:</p>	<p>Envió un mensaje de reconocimiento de voz (<i>voice-to-text</i>)</p> <p>Conjunto de datos: Predicción:</p>	<p>Usó un motor de búsqueda en línea como Google o Bing</p> <p>Conjunto de datos: Predicción:</p>	<p>Vi un resultado autocompletable en la búsqueda de Google</p> <p>Conjunto de datos: Predicción:</p>	<p>Un computador calificó mi tarea de escritura</p> <p>Conjunto de datos: Predicción:</p>
<p>Usó la "búsqueda segura" en Google</p> <p>Conjunto de datos: Predicción:</p>	<p>Vi una respuesta sugerida a un correo electrónico en Gmail</p> <p>Conjunto de datos: Predicción:</p>	<p>Usó un filtro de Snapchat (¿cuál es tu favorito?)</p> <p>Conjunto de datos: Predicción:</p>	<p>Jugué un videojuego con detección de movimiento, como Mario Party, Nintendo, Wii U, etc.</p> <p>Conjunto de datos: Predicción:</p>	<p>Recibí una sugerencia de Emoji para reemplazar una palabra (ej. "lol" es reemplazado por un Emoji de cara sonriente)</p> <p>Conjunto de datos: Predicción:</p>
<p>Vi un producto patrocinado en Google o Amazon. Ej. "dado que compraste __, pensamos que te gustaría ..."</p> <p>Conjunto de datos: Predicción:</p>	<p>Recibí un correo electrónico en la carpeta de <i>spam</i> (¿era realmente spam?)</p> <p>Conjunto de datos: Predicción:</p>	<p>LIBRE</p>	<p>Hice clic en una publicidad de Instagram (¿qué tipo de publicidad ves normalmente en la aplicación en comparación a tu compañero/a?)</p> <p>Conjunto de datos: Predicción:</p>	<p>Vi un artículo sugerido en una aplicación de noticias (¿qué tipo de artículos normalmente ves en comparación a tu compañero/a?)</p> <p>Conjunto de datos: Predicción:</p>
<p>Marqué un correo electrónico como "importante"</p> <p>Conjunto de datos: Predicción:</p>	<p>Vi publicidad sugerida en Snapchat (de ser así, ¿de qué? ¿Cómo se diferencia de la publicidad que ve tu compañero/a?)</p> <p>Conjunto de datos: Predicción:</p>	<p>Utilicé la función de autocorregir texto o usé el autocorrector</p> <p>Conjunto de datos: Predicción:</p>	<p>Escuché una canción recomendada en Spotify (¿qué tipo de música usualmente te recomienda en comparación a tu compañero/a?)</p> <p>Conjunto de datos: Predicción:</p>	<p>Vi un producto recomendado en Facebook (si así fue, ¿para qué?)</p> <p>Conjunto de datos: Predicción:</p>
<p>Vi un <i>nudge</i> recordatorio en Gmail para responder un correo electrónico</p> <p>Conjunto de datos: Predicción:</p>	<p>Utilicé mi huella digital para desbloquear un dispositivo o usé el reconocimiento facial</p> <p>Conjunto de datos: Predicción:</p>	<p>Usé una aplicación de mapas para encontrar un camino o destino</p> <p>Conjunto de datos: Predicción:</p>	<p>Usé una aplicación para reconocer una canción</p> <p>Conjunto de datos: Predicción:</p>	<p>Me comunicué con un <i>bot</i> de servicio al cliente</p> <p>Conjunto de datos: Predicción:</p>

Material de profundización (para el/la docente)

En este apartado va a encontrar los conceptos y definiciones esenciales para entender los fenómenos que se describen en la actividad, con ejemplos y casos cotidianos. Puede utilizar esta información como material de estudio o para contextualizar el contenido de la presentación de diapositivas incluida con esta actividad, profundizando en lo que le parezca más pertinente para el grupo curso.

Veremos los siguientes puntos:

1. Algoritmos.
2. Problemas irresolubles para algoritmos “convencionales”.
3. Aplicaciones de la inteligencia artificial.
4. Partes de la inteligencia artificial.

1. ¿Qué es un algoritmo?

Un algoritmo es una **secuencia finita de instrucciones diseñada para realizar una tarea o resolver un problema**. Un ejemplo de algoritmo es una receta de cocina, la cual paso a paso nos indica las instrucciones para realizar una preparación. Una receta de cocina es un algoritmo que genera una salida (*output*), que se manifiesta en el plato que nos permite realizar. La receta de cocina también tiene una entrada (*input*), en la forma de los ingredientes necesarios para la preparación. De esta manera, podemos entender un algoritmo como un proceso compuesto de 3 elementos:



Distintos tipos de entradas y salidas

Algunos algoritmos no tienen una salida concreta. Por ejemplo, un algoritmo que nos enseña un baile no genera un producto tangible o una “salida observable”, aunque podríamos considerar la habilidad de ejecutar el baile como la “salida” del algoritmo.

Algunos algoritmos también pueden no tener una entrada concreta. Por ejemplo, considere las instrucciones para patear un penal en un partido de fútbol. Aunque no existen “ingredientes”, podríamos considerar elementos como la pelota, el arco, el arquero, el punto de penal y el pateador como los elementos de entrada del algoritmo, necesarios para patear el penal.

Los algoritmos son herramientas sumamente importantes, pues nos permiten “empaquetar” las soluciones que encontremos a nuestros problemas y traspasarlos a otros miembros de la sociedad que puedan estar afectados por ellos. De esa forma, las sociedades son capaces de aprender y acumular aprendizaje en el tiempo. Por este motivo, una de las principales características de un algoritmo es su **claridad o precisión**. Un algoritmo debe ser **muy preciso** (no ambiguo) explicando cómo ejecutar una tarea o, de lo contrario, puede terminar generando resultados inesperados y erróneos. Por ejemplo:

Un niño fue a comprar, pues su madre le dijo “tráeme una caja de leche y, si tienen huevos, tráeme seis”. El niño llegó entonces con seis cajas de leche y cuando la mamá le preguntó por qué, él respondió “porque sí tenían huevos”.

En el ejemplo anterior, la secuencia finita de instrucciones (o el algoritmo) definido por la madre es ambiguo y poco claro. Normalmente, esperamos que nuestros interlocutores puedan comprender algoritmos ambiguos a través del contexto. Así, entendemos que es extrañísimo que el niño haya entendido que debía realizar una prueba lógica sobre el inventario de huevos en la tienda para inferir la cantidad de cajas de leche que debe comprar (ahí está el chiste). Sin embargo, cuando no hay contexto, un algoritmo no preciso puede inducirnos a errores.

Imaginemos que el curso tiene a su cuidado un acuario con peces y que es momento de cambiar el agua y acondicionarla para que sea apta para su subsistencia. Para ello compra un producto químico donde vienen las instrucciones: *"Añada una tapita del producto y, si el acuario no tiene suficiente agua, añade un poco más"*. En el ejemplo, no queda claro si cuando el acuario no tiene suficiente agua debemos agregar más agua o más del producto que utilizamos para tratarla. La falta de precisión del algoritmo generará distintos resultados, algunos de los cuales serán simplemente incorrectos, mientras que otros resultarán potencialmente peligrosos, causando en este caso la muerte de los peces del acuario. Observe además que un aficionado a los acuarios puede tener suficiente contexto como para ejecutar este algoritmo, a pesar de su ambigüedad.

Finalmente, un tipo particular de algoritmo con el que estamos más familiarizados hoy en día es el algoritmo computacional. Como el resto de los algoritmos, el **algoritmo computacional** es también un conjunto finito de instrucciones precisas que le indican a un computador cómo resolver una tarea. Los algoritmos computacionales son especificados en algún lenguaje especial que llamamos **lenguaje de programación**. De hecho, la programación de computadores consiste precisamente en el diseño, definición e implementación de algoritmos que le permitan a un computador resolver una tarea o problema en particular. Dado que estamos tratando con computadores, los algoritmos computacionales (y, por lo tanto, los lenguajes de programación) tienen una fuerte base matemática, aunque sus aplicaciones varían en complejidad y magnitud. De esta manera, lo que le permite a un teléfono móvil reproducir un video desde YouTube es un algoritmo, como también lo es el que permitió generar la primera imagen de un agujero negro, aunque presentan diferentes niveles de complejidad, como se explica a continuación.

2. Problemas irresolubles para algoritmos "convencionales"

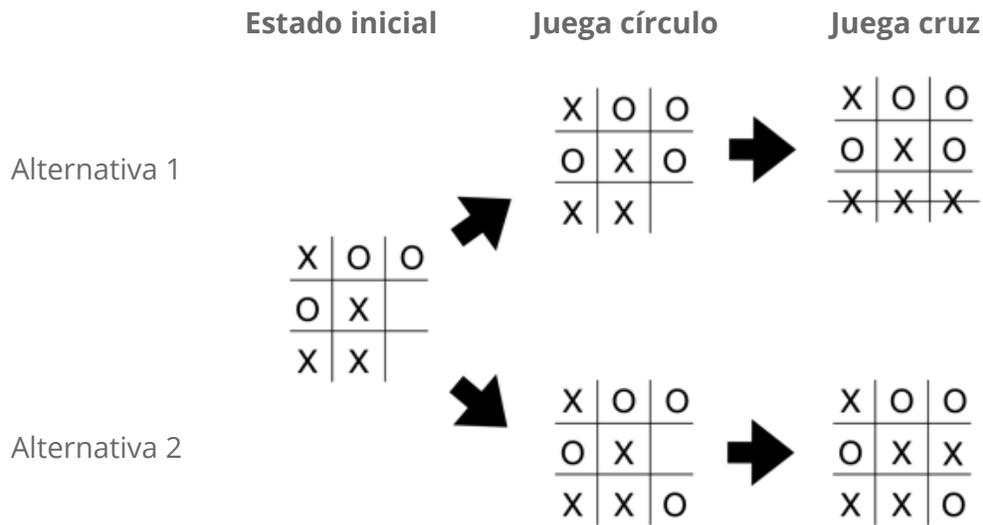
Como comentamos en la sección anterior, el programa que le permite a un teléfono móvil ver videos desde Internet es un algoritmo, como también lo es aquel que generó la primera imagen de un agujero negro. Intuitivamente podríamos pensar que ambos problemas son distintos en magnitud y dificultad. Parece claro que ver un video desde la red (por ejemplo, YouTube) es una tarea más sencilla que generar la imagen de un agujero negro a partir de

miles de imágenes generadas por decenas de distintos observatorios alrededor del mundo. Esto, incluso si sólo consideramos la cantidad de entradas (*inputs*) para cada uno de los algoritmos (un video versus miles de imágenes de distintos observatorios) **¿Cómo podemos evaluar si una tarea es más difícil que otra?**

En lo que sigue, presentaremos un ejemplo que nos permitirá encontrar una medida de dificultad para una tarea.

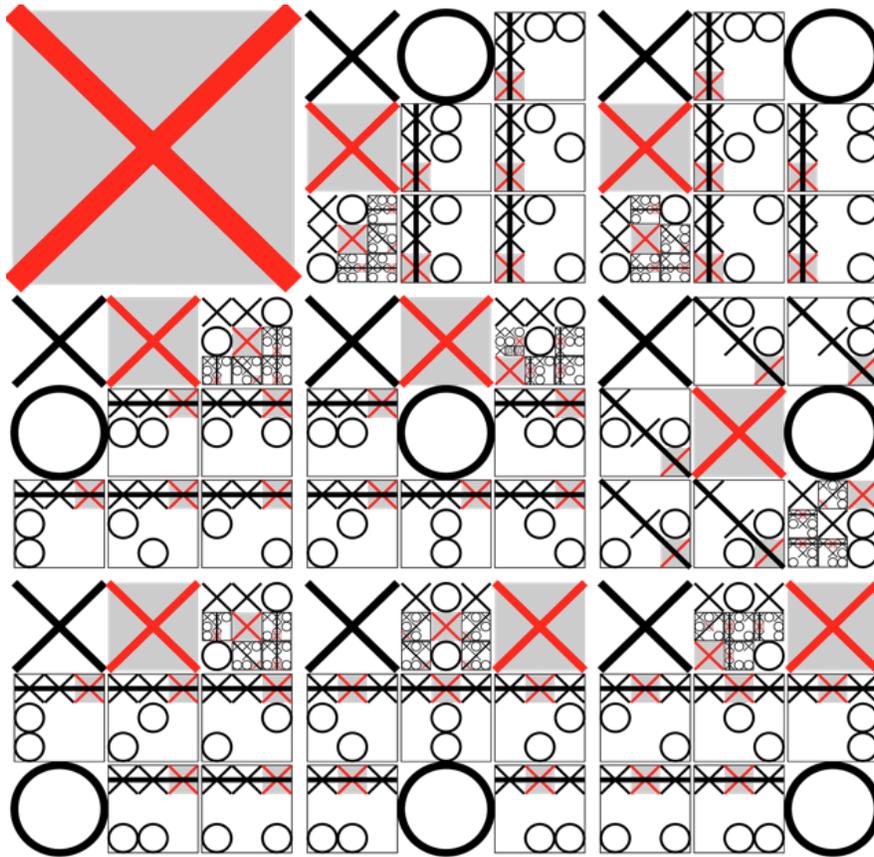
Pensemos en el algoritmo para jugar el juego “gato” o como también es conocido: “Tres en línea”. Como sabemos, este es un juego simple, donde los jugadores se van turnando para rellenar una cuadrícula de 3x3 casillas. El primer jugador empieza siempre con cruces, mientras que el segundo con círculos. El primero en crear una línea de 3 cruces o círculos (horizontal, vertical o diagonal) gana la partida. El “gato” es un juego muy simple y que con pocas instrucciones es posible dominar. Sin embargo, ¿Cómo podríamos crear un algoritmo que no tan solo juegue, sino que gane?

Al jugar “gato”, usualmente no jugamos de manera aleatoria. Además, también tomamos en cuenta que jugamos con otras personas que también intentan ganar. Por lo tanto, nuestras acciones apuntan tanto a ganar como a evitar que la otra persona gane. Para ello, evaluamos distintas alternativas y seleccionamos aquella que más nos convenga. Por ejemplo, en la siguiente figura, considere que usted debe jugar con un círculo, a partir de la situación dada en el estado inicial. ¿Cuál de las dos opciones posibles elegiría (tableros en el centro)? Entre ambas alternativas, claramente la mejor es la segunda, pues evita que el otro jugador gane. En cualquiera de los dos casos, es posible saber que usted no ganará esta partida.



Esta estrategia de juego se llama “pensar hacia adelante” y es una virtud que jugadores expertos de cualquier juego desarrollan. En el ejemplo anterior, somos capaces de pensar 2 pasos hacia adelante en el juego, evaluando cuatro tableros y dos resultados finales. Un posible algoritmo para jugar al gato puede hacer lo mismo: evaluar jugadas hacia adelante y elegir la jugada que permita ganar o empatar el juego, en el peor de los casos. En este caso, nos preguntamos ¿Cuántos posibles tableros y resultados distintos hay para el juego gato?

La respuesta es sorprendente. Un cálculo simple nos permite saber que existen casi veinte mil tableros del juego gato distintos (juegos parciales), aunque al simplificarlos nos podemos quedar con solo unos pocos. De hecho, las figuras siguientes muestran los algoritmos óptimos para jugar al gato. Con ellos **es imposible perder** —en el peor de los casos, el juego termina en empate—. No es necesario que el curso comprenda a cabalidad esta figura, sino que entienda que el gato ofrece un número de estrategias suficientemente manejable y predecible como para graficarlo en un esquema de este tamaño.



El esquema muestra la estrategia óptima para comenzar jugando con cruz desde el casillero superior izquierdo, dependiendo de las jugadas que haga el círculo en cada movimiento. El tamaño decreciente de las figuras representa jugadas más avanzadas en el tiempo.

Fuente: Wikimedia Commons³

La cantidad de escenarios a evaluar es una medida de cuán difícil es una tarea en particular. Mientras más escenarios debemos evaluar para encontrar la solución a una tarea, más difícil es ese problema.

Como vimos en el ejemplo, el juego gato es un problema fácil. Aunque la cantidad de tableros que debemos evaluar tal vez nos parezca sorprendentemente alta, para un computador, la cantidad de escenarios a evaluar en las figuras anteriores es muy baja (un computador moderno puede ejecutar millones de instrucciones por segundo).

Consideremos ahora que queremos hacer lo mismo que antes, pero esta vez para el **ajedrez**. ¿Cómo podríamos crear un algoritmo para jugar ajedrez? Utilizando la misma estrategia que antes, podemos evaluar las jugadas y “pensar hacia adelante”, eligiendo

³ <https://en.wikipedia.org/wiki/Tic-tac-toe#/media/File:Tictactoe-X.svg>; https://es.wikipedia.org/wiki/Tres_en_l%C3%ADnea

aquellas jugadas que nos lleven a ganar o empatar en el peor de los casos. Hay un pequeño problema con esta estrategia. Solo para los primeros 10 movimientos en el tablero hay 69 billones de alternativas distintas (millones de millones). En realidad, la cantidad de jugadas en el ajedrez es tan grande que ni siquiera un computador podría evaluarlas.

A pesar de lo anterior, sabemos que tanto humanos como computadores juegan ajedrez, sin necesariamente tener presente el universo completo de alternativas disponible. ¿Cómo es esto posible? Los mejores jugadores de este juego evalúan sólo algunas jugadas a futuro y **utilizan su experiencia para saber cuáles de ellas vale la pena evaluar**. Esto hace que el ajedrez sea un juego entretenido, pues la mayor parte de las partidas son distintas unas de las otras.

Algo similar sucede desde el campo de la informática. **En general, los problemas muy “difíciles”,** incluso para las computadoras, **los abordamos actualmente con inteligencia artificial (IA)**. Los algoritmos de IA son distintos a los algoritmos convencionales pues, en vez de requerir un conjunto de instrucciones precisas para resolver una tarea, **necesitan un conjunto de “ejemplos” desde los cuales el algoritmo “aprende”,** como veremos en el siguiente punto y a lo largo de este programa.

En el caso del ajedrez, en vez de definir exactamente el conjunto de instrucciones para ganar una partida (lo cual, ya vimos, es imposible por la cantidad existente de distintas partidas de ajedrez), una IA diseñada para jugar ajedrez recibe “ejemplos” de partidas en las cuales un jugador ganó. En tal caso, el algoritmo de IA **aprende** patrones de juego que le permitan ganar. La mecánica a través de la cual el algoritmo de IA aprende usualmente se basa en el cálculo de estadísticas desde los ejemplos a los cuales es expuesto (por ejemplo, qué jugadas tienen la probabilidad más alta de ganar en una circunstancia determinada).

3. Aplicaciones de la IA

Como hemos visto, la inteligencia artificial tiene aplicaciones en problemas donde los algoritmos convencionales no son adecuados, dada la dificultad de una tarea. Considere, por ejemplo, el reconocimiento facial en imágenes. ¿Cómo podríamos diseñar un algoritmo convencional para realizar esta tarea? En este caso, deberíamos utilizar el **pensamiento computacional** (que es una de las bases de la programación⁴) para diseñar un conjunto de instrucciones específico que nos permita reconocer una cara. Podríamos partir reconociendo los ojos: la forma (¿qué forma de los ojos tiene la persona? ¿Almendrados,

⁴ Los y las estudiantes que hayan participado en talleres de robótica y/o hayan incursionado en proyectos y lenguajes como Scratch, Code.org o la Hora del Código ya deberían familiarizarse con el pensamiento computacional.

redondos, profundos?), el color (café, negro, verde, azul), la distancia a otros rasgos en la cara (ojos más juntos, ojos más separados, etc.).

Sin embargo, ¿qué ocurre si la persona en la imagen es más joven? ¿Debería nuestro algoritmo tomar en cuenta entonces que los rasgos pueden ser más pequeños si es un niño? ¿Y si es un anciano? ¿Y si aparece de perfil o desde arriba? ¿Y si la persona usa lentes? ¿Y si el rostro no está completo? Como podemos ver, la cantidad de escenarios que debemos evaluar para reconocer a una cara "promedio" es muy alto e impredecible, por lo que ésta es una tarea difícil. No es extraño entonces que los sistemas de reconocimiento facial utilicen inteligencia artificial.

En la actualidad, las aplicaciones de la IA son muy amplias, tanto en la industria como en tareas cotidianas. En general, podemos reconocer las principales aplicaciones como ejemplos de algunas de estas áreas:

- **Procesamiento y generación de texto, imágenes, video y audio** (por ejemplo, agregar bigote a un actor en una película, crear una canción con el estilo de un/a artista o generar personajes de ficción que no existen a partir de rasgos previamente estudiados)
- **Control de vehículos, robots y maquinarias** (por ejemplo, la conducción de vehículos autónomos o el aterrizaje de misiones en Marte)
- **Planificación** (por ejemplo, organizar el calendario de un torneo de fútbol, minimizando viajes largos y alternando adecuadamente partidos difíciles y citas de local y visita)
- **Coordinación** (por ejemplo, de drones dentro de un enjambre, semáforos en una ciudad automatizada o vehículos en bodegas robotizadas)
- **Aprendizaje desde comportamiento humano** (por ejemplo, los sistemas de recomendación en plataformas de video o mecanismos de optimización de la publicidad en redes sociales)

4. Partes de la IA

Tal como los estudiantes, los algoritmos de IA deben "estudiar" para "pruebas" o "tareas" en las cuales son evaluados. Si en las "pruebas" los algoritmos aprueban, estos son utilizados. De otra manera, "repiten" y deben ir a estudiar nuevamente. Más específicamente, los algoritmos de IA tienen 2 fases esenciales: El **entrenamiento** (estudio) y la **evaluación** (las pruebas). Este entrenamiento puede ser supervisado o no supervisado, pero, por el momento, es importante que nos quedemos con este proceso general.

Durante la fase de entrenamiento, los algoritmos son expuestos a un conjunto de datos que contienen los ejemplos desde los cuales deben aprender. El conjunto de datos varía con respecto al objetivo que tenga el algoritmo de IA. Por ejemplo, si este requiere aprender a jugar ajedrez, el **conjunto de datos** tendrá miles de ejemplos de partidas de ajedrez. Si el algoritmo requiere aprender a recomendar una película a un usuario en base a su perfil, el conjunto de datos tendrá miles de perfiles de personas y las películas que les gustan.

En la fase de evaluación, el algoritmo es expuesto a muestras para las cuales debe demostrar su aprendizaje, usualmente a través de una predicción (que es el propósito que vemos en esta actividad), pero, alternativamente, también a través de la generación de un nuevo ejemplo.

Nuevamente, el tipo de muestras varía con respecto al objetivo que tenga el algoritmo. Por ejemplo, si requiere aprender a jugar ajedrez, las muestras son nuevas partidas de ajedrez para los cuales el algoritmo debe generar nuevas jugadas. Si el algoritmo requiere aprender a recomendar películas, las muestras son nuevos perfiles para los cuales debe predecir nuevas recomendaciones.

Al final de la fase de evaluación, se mide cómo lo hizo el algoritmo (¿cuántas partidas de ajedrez ganó? o ¿fueron “buenas” recomendaciones?). En base a esta evaluación, el algoritmo puede ser descartado (si lo hizo mal) o desplegado a producción (si lo hizo bien), es decir, puesto a disposición de las personas.

Es importante entender que, tal como las personas pueden aprobar un examen equivocándose en algunas pocas preguntas, una IA también puede aprobar su evaluación equivocándose en algunas muestras. **Ningún algoritmo de IA es infalible**, aunque los estándares que utilizamos para aprobar una IA son mucho más altos que los que ponemos a las personas: considere que la nota de aprobación 4 representa un 50% de logro con respecto a la nota 7 (siendo la nota 1 la más baja), mientras que los algoritmos de IA usualmente requieren porcentajes de logro por sobre el 95% y a veces por sobre el 99%. Los estándares que imponemos a las IA dependen del tipo de tarea que deban cumplir. Por ejemplo, una IA que maneje un automóvil tendrá estándares muy altos, mientras que una que recomiende películas puede tener estándares más bajos.

Puede ser que usted haya recibido una recomendación de una película mala por parte de una IA o una dirección incorrecta por parte del GPS. Los algoritmos de IA pueden equivocarse y es importante que entendamos que no son infalibles.

Las dos fases antes mencionadas tienen que ver con la "construcción" de un algoritmo de IA. Sin embargo, a no ser que participemos de su construcción, normalmente sólo observamos la IA "en producción", es decir, interactuamos con ella como usuarios, sea jugando ajedrez o recibiendo recomendaciones. Por lo tanto, es importante que podamos reconocer **tres partes principales de una IA**:

- A. **Conjunto de Datos:** ¿qué datos habrá sido necesario para entrenar la IA?
- B. **Algoritmo de aprendizaje:** es el conjunto de operaciones que sustenta el sistema de IA
- C. **Predicción:** ¿Cuál es el propósito del algoritmo? ¿Qué acción intenta predecir (o qué tipo de muestra intenta generar)?

ACTIVIDAD 2

ALGORITMOS COMO OPINIONES

Oportunidades curriculares

Tecnología

- **I°Medio - OA 2:** Desarrollar un servicio que implique la utilización de recursos digitales u otros medios, considerando aspectos éticos, sus potenciales impactos y normas de cuidado y seguridad.
- **I°Medio - OA 3:** Evaluar el servicio desarrollado considerando criterios propios, técnicos y valóricos, y proponer mejoras asociadas tanto a los procesos como al producto final.
- **II°Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y al ambiente, considerando criterios éticos, económicos, ambientales y sociales.

Visión general

Esta actividad introduce el concepto de “optimización” de algoritmos y su relación con los intereses y preferencias de quienes los desarrollan. Para ello, se pide a las y los estudiantes elaborar un algoritmo “convencional” para hacer “el mejor completo”, lo que da pie a una reflexión en torno al componente subjetivo de todo sistema informático que opera con lenguajes de programación y/o inteligencia artificial.

Objetivos

- Comprender el concepto de algoritmo “convencional”.
- Identificar y diseñar algoritmos simples.
- Comprender el concepto de “optimización” de un algoritmo y cómo esto está incidido por opiniones.

Agenda (50 minutos)

- Introducción (10 min)

- Desarrollo (30 min)
- Cierre (10 min)

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Hoja del estudiante (una por grupo)

Palabras clave

Algoritmo, pensamiento computacional, optimización, opiniones.

Acceso a presentación de diapositivas

<https://docs.google.com/presentation/d/1tah4xiTFD1c95r5aNq35PE8xukvivNwgkjtXKiKPns/edit?usp=sharing>

GUÍA DOCENTE

Introducción (10 minutos)

Esta actividad recupera el concepto de algoritmo en el sentido “convencional” del término, es decir, como un conjunto de instrucciones precisas (no ambiguas) en las que no interviene la inteligencia artificial⁵. Esto, con el objetivo de simplificar la comprensión de las nociones de optimización y “opinión” en los algoritmos.

Comience la clase con la presentación de diapositivas, que retoma y refuerza algunas de las definiciones entregadas en la actividad anterior, para introducir la actividad. Use la misma presentación para entregar las instrucciones y fomentar la reflexión en torno a la actividad.

Las láminas incluyen preguntas que se pueden ir haciendo a los/las estudiantes para profundizar la reflexión o abrir nuevas aristas de la conversación, entregando detalles y sugerencias para el/la docente en los comentarios. El curso comprenderá que existen múltiples formas de cumplir un mismo requerimiento (las recetas variarán, aunque haya grupos que elijan los mismos ingredientes).

Desarrollo (30 minutos)⁶

La actividad se centra en la instrucción “Vamos a elaborar el mejor completo” y persigue dos objetivos: lograr que las y los estudiantes construyan algoritmos en base a instrucciones precisas y que comprendan que todo algoritmo lleva implícita una opinión, que se manifiesta en elección de la entrada (*input*), que para esta actividad entenderemos como los “ingredientes”, así como en el procedimiento o algoritmo que permite obtener una salida (*output*), entendida como el “producto final”. Finalmente, comprenderán que los algoritmos se encuentran “optimizados” para cumplir con ciertos objetivos.

⁵ Para esta actividad, utilizaremos el concepto simplemente como “algoritmo”, para no entorpecer la fluidez de la lectura. La noción de “algoritmos de aprendizaje” en el contexto de la inteligencia artificial se retomará en la actividad 4.

⁶ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

Separe el curso en grupos (recomendamos 4 integrantes por equipo) y dígales que realizarán un ejercicio en donde construirán un algoritmo. Use la siguiente definición: Los algoritmos son conjuntos de instrucciones secuenciales, claras y precisas que nos permiten resolver un problema o ejecutar una tarea. A continuación, puede hacer preguntas tales como, **¿A quién le gustan los completos? ¿Cómo los prefieren?, ¿con palta o con chucrut? ¿Algún vegetariano por acá?** La idea es empezar a visualizar y reconocer diferencias.

Explique al curso que tendrán que **crear las instrucciones para hacer el mejor completo**, sin especificar a qué se refiere con “mejor”. Es importante recalcar esta palabra, para que entren en juego las discrepancias. Entregue una copia de la hoja del estudiante por grupo. Para ayudar a las/los estudiantes a escribir algoritmos precisos, se sugieren las siguientes condiciones:

1. Pensar en los materiales y utensilios para preparar los ingredientes.
2. Tener presente dónde cocinar y pensar en los tiempos y temperaturas de cocción.
3. Agregar los pasos para armar el mejor completo de su gusto.

Para el desarrollo de esta parte de la actividad, sugerimos de 10 a 15 minutos, para que puedan pensar en instrucciones precisas y claras. Una vez que todos los grupos hayan elaborado sus algoritmos, solicíteles exponer su trabajo.

Entre las cosas que puede destacar de los trabajos presentados están la precisión con la que desarrollen su algoritmo o las diferencias que existan entre los distintos completos (por ejemplo, algunos habrán pensado en un completo italiano, otros en uno sin palta, con sucedáneos vegetales o con salsa americana y mostaza).

Haga notar que los algoritmos no sólo dependen de instrucciones precisas, sino que fundamentalmente de **la intención de quien construye el algoritmo** (en este caso, su gusto). Es por ello que decimos que los algoritmos están **“optimizados para determinados fines”**. De este modo, el objetivo de un equipo pudo haber sido el de “optimizar la apariencia del completo”, “optimizar el orden de los utensilios” o “la presencia de verduras” (en realidad puede haber más de un fin).

Cierre (10 minutos)

Sugerimos iniciar la reflexión con algunas preguntas:

—**¿Cuál es el mejor completo? ¿Cómo podemos establecer esa respuesta? ¿De qué depende?**

Las respuestas deberían estar relacionadas con la persona que construye las instrucciones. Resalte entonces los conceptos clave de **optimización** y **opiniones**, ejemplificando con algunas de las propuestas de los grupos.

Para cerrar, plantee la siguiente problemática:

—**¿Han pensado cómo funcionan los buscadores como Google para mostrar los resultados? ¿Se puede aplicar en este caso el concepto de optimización? De ser así, ¿para qué están optimizadas las búsquedas?**

Escuche algunas reflexiones e indique que Google, entre otras cosas, optimiza su algoritmo para ofrecernos resultados “adecuados” a nuestro perfil (con el objetivo de ofrecernos publicidad personalizada). Puede mencionar (y hacer la prueba, si se cuenta con computadores con conexión a Internet) que dos personas que introducen la misma búsqueda en Google o YouTube obtendrán resultados distintos y preguntar a la clase por qué creen que ello sucede. Si tiene tiempo y considera oportuno profundizar, pregúnteles qué pasa cuando no iniciamos sesión en YouTube, cuando tenemos historial previo de búsquedas o cómo cambian los resultados según el país en el que nos encontramos. Todos estos son ejemplos de optimización del algoritmo.

Finalice la clase con la pregunta:

—**¿Cómo decidimos para qué debería optimizarse un algoritmo?**

Esta pregunta dará pie a la actividad “Matriz ética”, en la que se introduce al ecosistema de partes interesadas en la construcción de un algoritmo.

Es importante que, al final de la actividad, los y las jóvenes entiendan los siguientes puntos:

- **Los algoritmos deben ser precisos y no ambiguos**, en tanto las máquinas lo son. Esto está relacionado con lo que conocemos como “pensamiento computacional”, que se refiere al planteamiento de tareas o soluciones en formas que un computador (o un agente de procesamiento⁷) pueda ejecutar. Usualmente, las y los estudiantes presentan pasos incompletos o imprecisos que, al ser seguidos al pie de la letra, terminan con un resultado distinto al esperado.
- **Los algoritmos son desarrollados por personas**. Distintos algoritmos para realizar una misma tarea resultarán en diferentes resultados.
- **Distintas personas pueden tener diferentes motivaciones y objetivos al diseñar un algoritmo**. Con esto se ilustra los conceptos de “opinión” y “optimización” del algoritmo para lograr resultados con características favorables a determinadas “partes interesadas”.

⁷ Usamos esta distinción para incluir en el espectro a otros actores que pueden procesar el lenguaje en que está escrito el algoritmo, como los seres humanos. De este modo, una receta de cocina, las instrucciones para llegar de un punto A a un punto B o las disposiciones para construir el encabezado de un sitio web son algoritmos que podrían procesar y ejecutar ambos agentes.

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Algoritmos como opiniones**

Asignatura: Tecnología

Fecha: _____ **Curso:** _____

Objetivos:

- Comprender el concepto de algoritmo.
- Identificar y diseñar algoritmos simples.
- Comprender el concepto de "optimización" de un algoritmo simple y cómo esto está incidido por opiniones.

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____
4. _____

Introducción

Escuchen atentamente las instrucciones del profesor o profesora y la definición de "algoritmo" que utilizaremos. Si ya lo comprendieron, están listas/os para desarrollar la siguiente actividad, que pone a prueba su capacidad para elaborar un algoritmo.

Seguramente, alguna vez han visto una receta o las instrucciones para elaborar alguna comida, como las que salen en la parte de atrás de los paquetes de tallarines. Esas instrucciones (con pasos a seguir en orden y específicamente definidas) son **un algoritmo**.

En esta actividad, vamos a hacernos cargo de crear nuestras propias instrucciones, para elaborar **el mejor completo**. Esperen las instrucciones para formar los grupos y comiencen a pensar en cómo prepararían este completo:

1. Comiencen anotando los ingredientes y los utensilios que necesitarán para el proceso.
2. No olviden considerar el lugar que necesitan, los tiempos y temperaturas de cocción, así como la preparación de los ingredientes individuales que se necesitan para armar el completo.

Creemos “el mejor completo”

Ingredientes:

Utensilios:

Lugar:

Pasos para elaborar el mejor completo (nuestro algoritmo):

ACTIVIDAD 3

MATIZ ÉTICA

Oportunidades curriculares

Tecnología

- **I°Medio - OA 2:** Desarrollar un servicio que implique la utilización de recursos digitales u otros medios, considerando aspectos éticos, sus potenciales impactos y normas de cuidado y seguridad.
- **I°Medio - OA 3:** Evaluar el servicio desarrollado, considerando criterios propios, técnicos y valóricos, y proponer mejoras asociadas tanto a los procesos como al producto final.
- **II°Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y al ambiente, considerando criterios éticos, económicos, ambientales y sociales.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 19:** Explicar el problema económico de la escasez y las necesidades ilimitadas con ejemplos de la vida cotidiana, y de las relaciones económicas (por ejemplo, compra y venta de bienes y servicios, pago de remuneraciones y de impuestos, importaciones-exportaciones) que se dan entre los distintos agentes (personas, familias, empresas, Estado y resto del mundo).
- **II°Medio - OA 24:** Analizar y debatir sobre los desafíos pendientes para el país, por ejemplo, reducir la pobreza y la desigualdad, garantizar los derechos de los grupos discriminados, lograr un desarrollo sustentable, perfeccionar el sistema político y fortalecer la relación con los países vecinos, y reconocer los deberes del Estado y la responsabilidad de todos los miembros de la sociedad para avanzar en ellos.

Visión general

En esta actividad, las y los estudiantes identifican las “partes interesadas” en el algoritmo del completo desarrollado en la actividad anterior, junto con los intereses que éstas tienen en el mismo. A través de este ejercicio, comprenderán los algoritmos como soluciones a problemas sociotécnicos. Para ello, construyen una matriz ética que evidencia los conflictos de interés y la superposición de intereses en la construcción del algoritmo.

Objetivos

- Comprender que todos los sistemas técnicos (como la IA y aquéllos basados en programación) son sistemas sociotécnicos no neutrales, en tanto están al servicio de intereses y agendas particulares.
- Reconocer las partes interesadas en un algoritmo y su interacción por medio de la construcción de una matriz ética.
- Dimensionar el fenómeno y las implicancias de las tecnologías digitales en la vida cotidiana.

Agenda (60 minutos)

- Introducción (20 min)
- Desarrollo (30 min)
- Cierre (10 min)

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Hoja del estudiante (una por grupo)
- Papel *kraft*
- Notas autoadhesivas tipo *Post-it*

Palabras clave

Matriz ética, sistemas sociotécnicos, partes interesadas (stakeholders), intereses o valores, optimización.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1sQV96BUVd7D_Tez8-abeXZKvOgBc6TFY9u7M2590fjo/edit?usp=sharing

GUÍA DOCENTE

En la actividad anterior, las y los estudiantes trabajaron en el desarrollo de un algoritmo para preparar el “mejor” completo. Aprendieron entonces que en la elaboración de un algoritmo (en tanto lenguaje computacional) había que ser muy claros y precisos, comprendiendo que todo algoritmo está incidido por la opinión de quien lo construye y “optimizado” para algún fin.

Para finalizar aquella actividad el o la docente les planteó la siguiente problemática: “¿Han pensado para qué están optimizados los buscadores como el de Google?”

Si bien todavía no atenderemos directamente el problema específico de este tipo de buscadores, pondremos atención en el tema de fondo, que son los intereses que subyacen a los algoritmos y cómo éstos se superponen y entran en conflicto. Para ello, elaboraremos matrices éticas, tomando como base los algoritmos construidos en la actividad anterior.

En grupos, los y las estudiantes identificarán las partes interesadas en el algoritmo de su completo y sus objetivos particulares. Posteriormente, trabajarán matrices éticas para un almuerzo escolar, en donde representarán los valores de diferentes actores interesados en su composición. Junto con la guía de trabajo, los grupos de estudiantes deben contar con papel *kraft* (2 o 3 pliegos) y hojas autoadhesivas tipo *Post-it*.

Introducción (20 minutos)

Sugerimos desarrollar la actividad considerando la siguiente pauta y complementar con la presentación de diapositivas:

—**¿Cuáles fueron los objetivos de su algoritmo para la elaboración del completo?**

Los/las estudiantes deberían decir “el más rico”, “el más sano”, “el más llenador”, etcétera.

—**Ahora tómense cinco minutos con los mismos grupos de la actividad anterior para hacer una lluvia de ideas con posibles objetivos para su algoritmo y escríbanlos en notas autoadhesivas (*Post-it*). Intenten identificar al menos diez. Por ejemplo: el más rico, el más saludable, el más inclusivo, el más rápido de preparar, etcétera.**

Terminado el tiempo, pida al curso compartir los objetivos identificados. La idea es que den cuenta de las personas que puedan estar interesadas en el algoritmo del completo. En ese sentido, se plantea la siguiente secuencia:

—Con el mismo grupo, ahora vamos a hacer una lluvia de ideas sobre 5 a 10 personas diferentes que podrían interesarse en su algoritmo del completo. ¿A quién le importa lo que comes? ¿A quién le importa lo que almuerzas, por ejemplo, en el colegio?

Espere algunas respuestas.

—A estas personas que se preocupan por sus algoritmos las llamaremos “partes interesadas”. Para cada una de ellas, indiquen una o dos razones por las cuales podrían interesarse.

Brinde 10 min para que puedan proponer personas interesadas. Si evalúa que las y los estudiantes se han complicado con esto, le sugerimos hacer las siguientes preguntas: ¿de dónde sacan su comida? ¿Quién gana dinero con los alimentos que compran? ¿Qué familiares o personas con las que comparten o supervisan su dieta podrían estar interesados en lo que comen?

En este caso, es importante que piensen en los agentes e instituciones indirectos, que puedan estar detrás de la cadena de producción, distribución y venta de los ingredientes. Por ejemplo: almacenes, empresas agricultoras, transportistas, las compañías que producen los ingredientes, etcétera.

—Por último, para compartir entre todos, ¿Qué partes interesadas incluyeron? ¿Por qué les importa su algoritmo?

Desarrollo (30 minutos)⁸

—Ahora quiero presentarles la matriz ética [lámina 5 de la presentación de diapositiva]. En la primera columna detallamos las partes interesadas y en la primera fila los valores, es decir, el interés que cada una tiene en su algoritmo.

Guíese por los ejemplos que se muestran en las láminas de la presentación y explique que deberán elaborar matrices éticas para un almuerzo escolar, empezando por las más pequeñas y aumentando el nivel de complejidad (1x1, 3x3, 2x3 y, finalmente, 5x5). Si lo considera pertinente, plantee situaciones diferentes para cada matriz.

—Ahora es su turno. Con su grupo, completen las matrices éticas que propone la guía. Para cada matriz, discutan cuál debería ser el objetivo del algoritmo para esas partes interesadas. Además, justifiquen en cada caso por qué marcaron cada cuadro y cuál fue el objetivo que decidieron para su algoritmo. Utilicen los pliegos de papel *kraft* y los *Post-it* para ir rellenando la matriz, tal como se pide en los ejemplos de la guía.

Cierre (10 minutos)

Una vez finalizada la actividad, sugerimos compartir con el curso algunas de las matrices realizadas por los grupos. Al concluir, puede hacer las siguientes preguntas para la discusión final:

- ¿Decidir un objetivo para su algoritmo fue más fácil o más difícil a medida que aumentó el número de partes interesadas?
- ¿Qué pasa a medida que cambia el número de intereses?
- ¿Existen conflictos entre las partes interesadas?
- ¿Existen intereses transversales a los diferentes grupos?
- ¿Por qué le hemos denominado matriz ética?

⁸ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Matriz ética**

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales

Fecha: _____ **Curso:** _____

Objetivos:

- Dimensionar el fenómeno y las implicancias de las tecnologías digitales en la vida cotidiana.
- Comprender que todos los sistemas informáticos son sistemas “sociotécnicos” no neutrales, que están al servicio de intereses particulares.

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____
4. _____

Introducción

Como vimos en la presentación, para elaborar una matriz ética debemos identificar los valores o factores que van a estar relacionados con el algoritmo y, también, las partes interesadas. Con esa información podemos identificar el objetivo del algoritmo.

Empecemos viendo la matriz que se mostró en la presentación, con respecto al almuerzo escolar:

Interesados	Sabor	Nutrientes	Costo
Estudiante			
Padres			
Doctor/a			

En este caso, sabemos que para optimizar este algoritmo debemos poner atención en el sabor, los nutrientes y, finalmente, el costo, siendo este el menos representado entre las partes interesadas.

Instrucciones de la actividad:

Copien las siguientes matrices éticas en el papel *Kraft*. En la primera columna, incluyan las partes interesadas que se preocupen por el resultado de su algoritmo. En la primera fila, definan un *valor o interés* que la parte interesada querría ver expresado en su algoritmo. Debajo de la matriz, escriban el objetivo que perseguiría el algoritmo en cada caso. En la matriz 1x1 sólo incluyan un valor y un objetivo.

Matriz ética de 1x1

Objetivo:

Objetivo:

Matriz ética de 3x3

Objetivo:

Matriz ética de 2x3

Objetivo:

Matriz ética de 5x5

Objetivo:

ACTIVIDAD 4

INTRODUCCIÓN AL APRENDIZAJE AUTOMATIZADO SUPERVISADO Y SESGO ALGORÍTMICO

Oportunidades curriculares

Tecnología

- **I°Medio - OA 2:** Desarrollar un servicio que implique la utilización de recursos digitales u otros medios, considerando aspectos éticos, sus potenciales impactos y normas de cuidado y seguridad.
- **I°Medio - OA 3:** Evaluar el servicio desarrollado considerando criterios propios, técnicos y valóricos, y proponer mejoras asociadas tanto a los procesos como al producto final.
- **II°Medio - OA 2:** Proponer soluciones que apunten a resolver necesidades de reducción de efectos perjudiciales relacionados con el uso de recursos energéticos y materiales en una perspectiva de sustentabilidad, utilizando herramientas TIC colaborativas de producción, edición, publicación y comunicación.

Historia, Geografía y Ciencias Sociales

- **II°Medio - OA 24:** Analizar y debatir sobre los desafíos pendientes para el país, por ejemplo, reducir la pobreza y la desigualdad, garantizar los derechos de los grupos discriminados, lograr un desarrollo sustentable, perfeccionar el sistema político y fortalecer la relación con los países vecinos, y reconocer los deberes del Estado y la responsabilidad de todos los miembros de la sociedad para avanzar en ellos.

Visión general

En esta actividad, las y los estudiantes experimentan el entrenamiento de sistemas de inteligencia artificial, comprendiendo la importancia de la selección de los datos de entrenamiento y su implicancia. El curso utiliza la herramienta **Teachable Machine** para enseñar a un modelo a distinguir perros de gatos, a partir de un set de fotos dado, observando que el conjunto de datos de entrenamiento está sesgado para identificar con mayor precisión a los gatos. Para finalizar, las y los estudiantes tienen la oportunidad de reentrenar el modelo y reflexionar en torno a un caso real de discriminación en el contexto de tecnologías de reconocimiento facial.

Objetivos

- Conocer el mecanismo por el cual algunos algoritmos de inteligencia artificial son capaces de realizar predicciones sobre determinados contenidos.
- Comprender cómo la elaboración del conjunto de datos utilizados para "entrenar" un algoritmo de IA puede impactar los resultados de sus predicciones.
- Conocer el concepto de "sesgo algorítmico" y sus implicancias en contextos reales.

Agenda (150 a 160 minutos)

- Introducción (40 min.)
 - Aprendizaje automatizado supervisado
 - Clasificación
 - Teachable Machine
 - Partes de la IA
- Desarrollo (80-90 min.)
 - Uso del Teachable Machine con estudiantes
 - Actividad: Aprendizaje automatizado supervisado
 - Actividad: Sesgo Algorítmico
 - Cuestionario
 - Video: Gender Shades (Matices de género)
 - Discusión
- Cierre (30 min.)
 - Actividad: Clasificación de instrumentos musicales
 - Conclusiones

Materiales para esta actividad

- Presentación de diapositivas
- Hojas de trabajo x 3
- Computadores con acceso a Internet
- Tarjetas de perros y gatos (incluidas en la actividad)

- Conjuntos de instrumentos musicales

Palabras clave

Sesgo algorítmico, aprendizaje automatizado supervisado; clasificación, predicción, datos de entrenamiento, datos de prueba.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1_m7EkyJLLY5SFYxa7C5Evpm1_KEPIy6n5k5o8NrGY4/edit?usp=sharing

Acceso a conjuntos de datos de imágenes en versión digital

<https://drive.google.com/drive/folders/1fPhc9V1rwRUj7ARTEi58Oh8z4YwV/k1-T?usp=sharing>

GUÍA DOCENTE

Esta es una actividad larga que puede tomar dos o tres clases. Las y los estudiantes aprenderán varios conceptos ligados con los sistemas de IA, tales como aprendizaje automatizado supervisado, clasificación, predicción y sesgo algorítmico.

Se espera que las y los alumnos aprendan sobre el aprendizaje automatizado supervisado (*supervised machine learning*), experimentando por primera vez con una plataforma que permite entrenar y probar algoritmos de IA. Esto les permitirá conocer en la práctica los principios básicos del funcionamiento de muchos sistemas de IA y reflexionar sobre las consecuencias éticas de un entrenamiento sesgado.

Preparación (antes de la clase)

- Considere al menos dos bloques de clases de 2 horas pedagógicas para implementar esta actividad (idealmente distribuidos en jornadas diferentes). Ambas instancias requieren contar con sala de informática o equipamiento móvil que cuente con conexión a Internet y cámara web,
- Repase los contenidos de la actividad 1, en relación con las fases de la inteligencia artificial (entrenamiento y evaluación)
- Familiarícese con la plataforma *Teachable Machine* (textualmente “máquina a la que se le puede enseñar”), herramienta de Google que permite entrenar modelos de aprendizaje automatizado para reconocer imágenes o sonido, de forma rápida y sin necesidad de contar con conocimiento especializado ni programación.
- Para la Parte II de la actividad práctica, cada pareja debe contar sets de imágenes de perros y gatos (conjunto inicial de entrenamiento, conjunto de datos de prueba y conjunto de datos redefinido). Considere tres alternativas:
 - a. Imprimir un juego de fotos (o tarjetas) por pareja para que puedan presentarlas frente a la cámara web.
 - b. Descargar los sets de imágenes desde la [carpeta en línea](#) y crear una copia en cada computador, de modo que los grupos puedan cargarlas de manera local en versión digital.
 - c. Importar las imágenes directamente desde Google Drive. Recomendamos esta opción sólo si tiene un manejo avanzado de Google Drive y Classroom.
- Realice pruebas con la plataforma *Teachable Machine*. Encontrará más instrucciones sobre cómo usar la plataforma en la sección “desarrollo” y en la presentación de diapositivas.

- Familiarícese con los videos que mostrará al curso (tutoriales y “Matices de género”), que se pueden traducir automáticamente con el sistema de YouTube, como se explica más adelante. Lea las correcciones a la traducción automática del video “Matices de género” que se detallan en los comentarios de la diapositiva 10.

Introducción (40 minutos)⁹

En esta parte, nuestro propósito es que las y los estudiantes comprendan los problemas de clasificación y el papel que juegan los datos de entrenamiento en la precisión de la misma.

—**Para comenzar: ¿Recuerdan cuando discutimos la definición de IA? ¿Cuáles fueron los tres componentes de los que hablamos?**

El curso debería contestar: el conjunto de datos, el algoritmo de aprendizaje y la predicción.

—**¡Muy bien! Hoy vamos a hablar un poco más sobre los conjuntos de datos de entrenamiento y profundizar un poco en el aspecto del algoritmo de “aprendizaje” de la IA. Particularmente, vamos a hablar sobre una forma muy común de inteligencia artificial llamada aprendizaje automatizado supervisado.**

Utilizando la presentación de diapositivas de la actividad, muestre lo siguiente:

—**En un sistema de aprendizaje automatizado supervisado, un computador aprende por medio del ejemplo. ¿Cuántos/as de ustedes han visto a un padre o a una madre enseñar los colores a un bebé o a un/a niño/a pequeño/a? Una posibilidad es sentarse con sus hijos/as y con un libro con formas o animales de colores, señalando una imagen y repitiendo “oso café” o “pájaro rojo”. La idea es que la niña o el niño aprenda con estos ejemplos y que más adelante sea capaz de identificarlos por su cuenta.**

Luego de ver los ejemplos de la presentación de diapositivas y entender cómo se aprende vía clasificación, el o la docente pregunta:

⁹ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

—¿Quién puede decirme algunos ejemplos útiles de clasificación que use en su vida cotidiana o que haya aprendido en el colegio?

Las y los estudiantes deberían responder: géneros musicales, categorización de animales y/o seres vivos, categorías de libros en la biblioteca, etc.

—Muy bien, ahora mencionaremos algunos ejemplos de la tecnología de clasificación:

- ★ Detección de rostro: es o no es un rostro
- ★ Detección de *spam* (correo no deseado): es o no *spam* (en base al estudio de correo clasificado previamente como no deseado)
- ★ Detección de escritura a mano (por ejemplo, para depositar cheques en el banco): es o no una de las 27 letras o uno de los 10 números posibles

—Ahora construyamos nuestros propios clasificadores. Les demostraré cómo funciona la plataforma Teachable Machine de Google (literalmente, máquina a la que puedes “enseñar”).

Desarrollo (80-90 minutos)

Esta actividad está compuesta por tres partes:

1. Parte I: entrenamiento de un modelo de inteligencia artificial para que reconozca lo que se muestra frente a la cámara web (si no se cuenta con cámara, existe una variación a este ejercicio, subiendo directamente las imágenes en formato digital)
2. Parte II: construcción de un clasificador de perros y gatos con un set de imágenes que se entrega
3. Parte III: reflexión en torno al concepto de “**sesgo algorítmico**”, a partir de los resultados del clasificador y un video

Es necesario contar con computadores por lo cual se sugiere utilizar el laboratorio de computación o sala Enlaces, biblioteca CRA o laboratorio móvil.

Parte I: Exploración de la herramienta

Cuando esté todo listo, muestre las láminas 7, 8 y 9 de la presentación y repase las fases de la inteligencia artificial aprendidas en la actividad 1, según lo cual distinguimos los **datos de entrenamiento** y los **datos de prueba**. El primer grupo es el que usamos para entrenar el modelo (buscando que no se produzcan sesgos) y el segundo el que utilizamos para

evaluar si el modelo es capaz de realizar la tarea para la cual se le entrenó, con la probabilidad de éxito deseada.

Haga una demostración de *Teachable Machine* (teachablemachine.withgoogle.com) o presente las láminas en las que se explica cómo funciona (9 a 11). Solicítele ver los videotutoriales “Recopilación” (hasta el segundo 0:52¹⁰) y “Preparación” (hasta el segundo 00:10), que se encuentran bajo el título “¿Cómo se utiliza?”.

—¿Entendieron? ¡Ahora inténtenlo ustedes!

Dé tiempo para probar la plataforma (aprox. 8 a 10 minutos) con un “Proyecto de imagen”¹¹ (esta parte de la actividad está pensada para hacerse con la cámara web). Solicite entrenar dos “clases” con objetos a elección (ver párrafo siguiente). Puede pedirles que cuenten al curso los conjuntos de datos de entrenamiento que están utilizando. Explíqueles que mientras más imágenes entreguen al sistema, mayor será el conjunto de datos de entrenamiento (lo que puede llevar a una mayor precisión), pero también el tiempo requerido para procesar los datos. Indíqueles que suministren al menos 10 imágenes y no más de 30.

Antes de entrenar el algoritmo, sugiera a las y los estudiantes poner un nombre a cada “clase” (*class*, categoría o clasificador), que corresponden a cada uno de los elementos o conjunto de elementos que se presentarán frente a cámara para entrenar el modelo, para su posterior identificación (predicción). Por ejemplo: “estuche” y “cuaderno” o “una mano” y “dos manos”. Una vez finalizado el entrenamiento, los grupos deberían obtener predicciones acertadas sobre el elemento que se presenta frente a cámara (datos de prueba). Guíese con la lámina 9.

Una vez finalizado el ejercicio, algunas preguntas que puede realizar para profundizar la comprensión del proceso son:

—**¿Qué sucede cuando ambas clases con las que entrenaron el modelo se muestran al mismo tiempo? ¿Qué ocurre si cambian de posición los objetos o personas en la imagen?**

A continuación. Promueva la siguiente discusión abierta (las parejas de estudiantes deberán contestar las mismas preguntas en su guía):

¹⁰ Puede entregar las siguientes instrucciones: para activar los subtítulos en español, haga clic sobre el ícono de subtítulos, ubicado en la esquina inferior derecha del video (se activará el subtítulo automático en inglés). Presione el ícono de configuración (engranaje) y luego “subtítulos”. Marque “traducción automática” y seleccione español.

¹¹ La página de referencia es <https://teachablemachine.withgoogle.com/train>

- ¿Qué pasa si sólo entrenan una categoría (*class*)?
- ¿Qué sucede a medida que aumenta el conjunto de datos?
- ¿Qué sucede cuando su conjunto de datos de prueba es diferente de su conjunto de datos de entrenamiento?

—**Excelente. Ahora vamos a construir un clasificador gato-perro usando *Teachable Machine*.**

¡Hemos superado la primera parte de la actividad! (reconocimiento de etapas del entrenamiento y primera prueba de Teachable Machine).

Parte II: Construcción de clasificador gato-perro

—**Con su compañera o compañero, construyan un modelo que clasifique gatos y perros.**

Entregue las láminas impresas a las y los estudiantes o indíqueles cómo acceder a los sets de imágenes. Enfático que en esta parte de la actividad sólo utilicen el “Conjunto de datos de entrenamiento” (conjunto inicial). Entrégueles entre 8 y 10 minutos para construir sus clasificadores y registrar la información en la hoja del estudiante (preguntas 1, 2 y 3 de la Parte II).

—**¿Cómo funcionan sus clasificadores?**

Las preguntas 1, 2 y 3 están orientadas a que las y los estudiantes se den cuenta de que el sistema está realizando mejores predicciones para los gatos que para los perros, dada la mayor calidad (diversidad) de la muestra. Una vez se hayan dado cuenta de lo anterior, motive la reflexión con las siguientes preguntas:

- ¿Es útil este clasificador si sólo funciona bien en los gatos?
- ¿Por qué creen que funciona mejor en gatos que en perros?
- ¿Cómo podríamos mejorarlo con nuestros datos de entrenamiento? (Si las y los estudiantes tienen problemas, pídeles que destaquen las similitudes y diferencias de los conjuntos de datos, en base a las preguntas 1 y 2 de la guía. Por ejemplo, los “perros eran peludos y parecidos a los gatos o no eran tan diversos”).

—**Cuando los algoritmos, y específicamente los sistemas de inteligencia artificial, presentan resultados “injustos” de manera sistemática, llamamos a eso sesgo**

algorítmico. Podríamos afirmar que nuestro clasificador gato-perro está sesgado en favor de los gatos, ya que funciona muy bien para ellos y tiene un sesgo en desmedro de los perros. Es por ello que decimos que las aplicaciones informáticas son sistemas sociotécnicos no neutrales (como vimos en la actividad 3), lo que significa que están sujetos a sesgos (por ejemplo, de origen cultural o en la agenda de grupos de interés, entre otros).

Para corregir este sesgo, las y los estudiantes tendrán la oportunidad de entrenar nuevamente el modelo, esta vez con un “conjunto de datos redefinido”. Dé 8 a 10 minutos a los/las estudiantes para terminar esta tarea. deberán crear un proyecto nuevo o actualizar la página.

—¿Cómo están funcionando sus clasificadores ahora? ¿Qué hicieron para que funcionaran mejor?

Si las y los estudiantes dicen que usaron menos datos de entrenamiento, invítelos/as a pensar si es mejor tener más o menos datos. Invítelos/las también a pensar en el valor de la diversidad de los datos de entrenamiento.

Use el cuestionario de la segunda parte del Punto II de la guía del estudiante para que los equipos ejerciten en torno a escenarios de sesgo algorítmico con la clasificación de perros y gatos.

Parte III: Reflexión en torno al concepto de “sesgo algorítmico”

—Hemos visto de primera mano cómo nuestros sistemas de aprendizaje automatizado supervisado pueden presentar un sesgo algorítmico. Ahora quiero que nos tomemos un tiempo para ver un video sobre un caso real.

Reproduzca el video de reconocimiento facial *Gender Shades* (Matices de género) que se incluye en la presentación de diapositivas, activando los subtítulos automáticos en español (según lo explicado anteriormente). Considere las correcciones a la traducción automática que se entregan en la lámina 10 de la presentación.

Motive la siguiente discusión:

—¿Qué problema identificó Joy?

Espere a que las y los estudiantes digan que el sistema de reconocimiento facial no distinguió los rostros femeninos más oscuros tan bien como los rostros masculinos más claros. Si necesitan un estímulo adicional, pregunte si la tecnología de la que habló Joy funcionó igual para todas las personas.

—¿Por qué es esto un problema?

Espere a que las y los estudiantes mencionen que este tipo de casos suponen situaciones de discriminación o subrepresentación de determinados grupos: “no se puede decir que una tecnología funciona si no lo hace para todo el mundo” (refuerce esa idea).

Si las y los estudiantes necesitan ayuda, pregúnteles si les gusta usar los filtros y efectos de TikTok, Instagram o Snapchat. ¿Estaría mal si no todas las personas tuvieran acceso a esa tecnología o funcionara mal sólo para algunas?

—¿Cómo sugiere Joy que podemos solucionar este problema?

Espere a que las y los estudiantes mencionen una mejor selección del conjunto de datos. Si necesitan un estímulo adicional, pregúnteles cómo mejoraron sus algoritmos de clasificación durante la actividad con Teachable Machine.

—¿Qué valor creen que tiene diseñar sistemas tecnológicos con criterios de diversidad e inclusión? ¿O con una participación más amplia de la ciudadanía?

La idea de esta pregunta es reflexionar brevemente sobre los estándares que debería tener este tipo de tecnologías desde una perspectiva ciudadana.

Cierre (30 minutos)

(Esta parte de la actividad requiere micrófono. Si los equipos no cuentan con esta característica, cierre con la reflexión anterior. Si no tiene instrumentos musicales, utilice otro tipo de objetos que emitan sonido, como las palmas o percusión con objetos).

—**Ahora vamos a construir un clasificador más. Voy a entregar a cada uno/a de ustedes un set de instrumentos musicales. Quiero que construyan clasificadores que operen cuando un determinado instrumento esté sonando. Después de completar el**

proceso de entrenamiento, quiero que prueben su clasificador tocando varios instrumentos y vean qué sucede.

Dé tiempo para construir el clasificador hasta que las y los estudiantes estén probándolo.

—¿Qué pasó cuando tocaron varios instrumentos a la vez? ¿El algoritmo predijo ambos instrumentos?

(No, porque sólo le hemos enseñado a clasificar un instrumento a la vez).

A veces, la clasificación puede ser muy difícil porque es complicado enumerar todas las categorías o clases posibles. Piensen cuántas categorías necesitaríamos para clasificar cada combinación de instrumentos musicales que tienen ustedes.

Este problema también apareció en el trabajo de Joy. ¿Recuerdan el video? El mayor problema con los sistemas de reconocimiento facial que ella estudió es que no funcionaban tan bien en los rostros femeninos más oscuros como lo hacían en los rostros masculinos más claros.

Sumemos a esto que no todas las personas se identifican como hombres o mujeres y los sistemas de reconocimiento facial no pueden capturar eso. Esto es algo que siempre debemos considerar cuando estamos clasificando (sea reconociendo caras, otorgando puestos de trabajo o créditos bancarios, entre otros casos), para asegurarnos de no discriminar a ningún grupo, generando para ello sistemas inclusivos que puedan recoger la diversidad de la sociedad.

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Introducción al aprendizaje automatizado supervisado y sesgo algorítmico**

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales

Fecha: _____ **Curso:** _____

Objetivos:

- Conocer el mecanismo por el cual algunos algoritmos de inteligencia artificial son capaces de realizar predicciones sobre determinados contenidos.
- Comprender cómo la elaboración del conjunto de datos utilizados para "entrenar" un algoritmo de IA puede impactar en los resultados de sus predicciones.
- Conocer el concepto de "sesgo algorítmico" y sus implicancias en contextos reales.

Nombres de las/los integrantes del grupo

1. _____
2. _____

Parte I: Introducción a la actividad sobre aprendizaje automatizado supervisado

Instrucciones

1. Ingresen a <https://teachablemachine.withgoogle.com>. Asegúrense de que el sitio esté cargado en español. Si no, seleccionen el idioma desde el pie de página, como se muestra en la imagen.



2. Vayan hasta donde dice “¿Cómo se utiliza?” y reproduzcan los tutoriales “**Recopilación**” hasta el segundo 00:52 y “**Preparación**”, hasta el segundo 00:10.

¿Cómo se utiliza?

Class 1

Class 2

TRAIN MODEL

1 Recopilación

Recopila y agrupa los ejemplos que quieres que aprenda el ordenador por clases o categorías.

[Video: Recopila muestras](#)

2 Preparación

Prepara el modelo y pruébalo al momento para ver si puede clasificar correctamente los nuevos ejemplos.

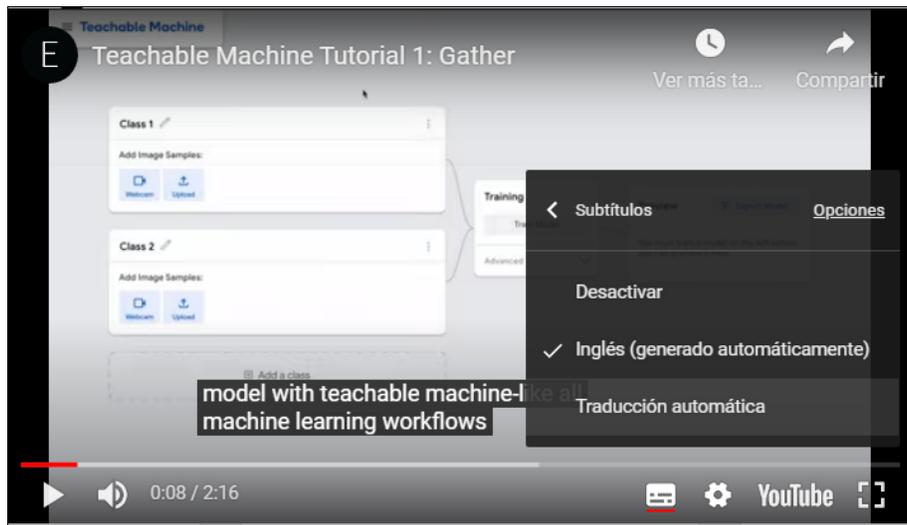
[Video: Prepara tu modelo](#)

3 Exportar

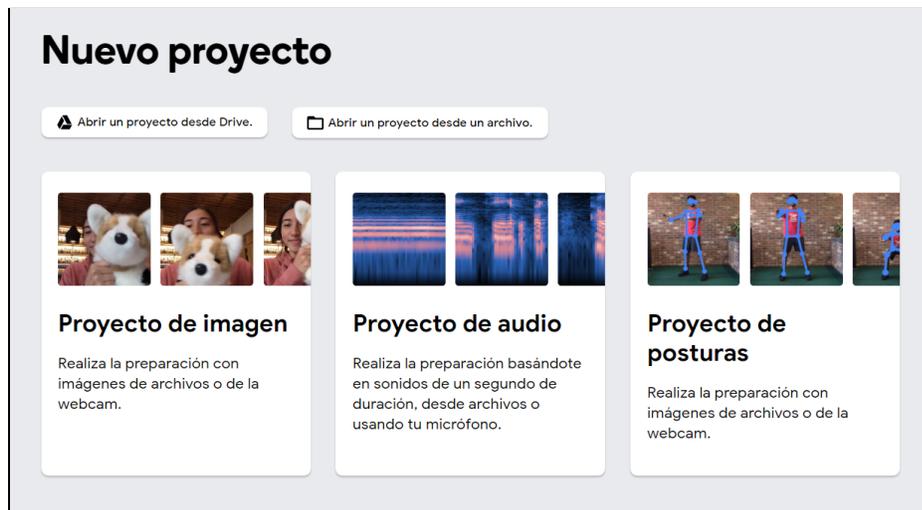
Exporta el modelo para tus proyectos: sitios, aplicaciones y mucho más. Puedes descargar el modelo o alojarlo online de forma gratuita.

[Video: Exportar el modelo](#)

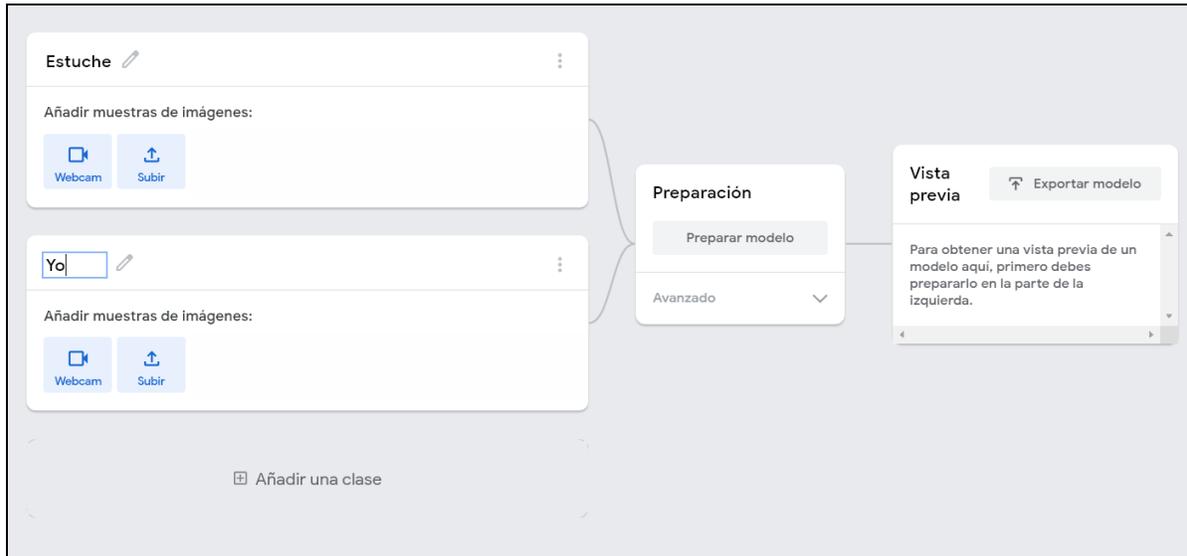
Recuerden activar los subtítulos automáticos de YouTube, haciendo clic sobre la rueda de configuración y seleccionando “inglés (generado automáticamente)”. Repitan el proceso y seleccionen “traducción automática” al español. Si lo necesitan, pongan el tamaño del subtítulo en 100%.



3. Serán redirigidos/as a la página de entrenamiento de modelos. Hagan clic sobre la opción “**Proyecto de imagen**”.



Entrenen un modelo situando dos objetos diferentes frente a la cámara (dos “class” o “clases”). Elijan la opción “Webcam” y tomen muestras de la persona u objeto que hayan elegido, manteniendo presionado el botón de grabación.



La cantidad de tiempo que mantengan el botón presionado determinará el número de muestras que entreguen al sistema. Se recomienda generar entre 10 y 30 imágenes por clase (un segundo de grabación aproximadamente) para evitar tiempos de procesamiento largos. En ese lapso, intenten variar la posición del objeto para lograr una diversidad de muestras. Pulsen “**Preparar modelo**” y esperen a que procese, sin cambiar la pestaña del navegador.

Cuando el modelo esté entrenado, experimenten con los resultados y luego discutan:

1) ¿Qué pasa cuando solo entrenan una clase?

2) ¿Qué sucede cuando aumentan el número de imágenes en su conjunto de datos? Asegúrense de que ambas categorías tengan al menos diez imágenes.

3) Si han entrenado con los objetos en una sola posición, intenten variarla. ¿Qué sucede cuando su conjunto de datos de prueba es diferente a su conjunto de datos de entrenamiento?

Parte II: Introducción al sesgo algorítmico

En este ejercicio aprenderán los tres componentes de un sistema de inteligencia artificial (IA) y el rol de los datos de entrenamiento.

Instrucciones

- Vayan nuevamente a <https://teachablemachine.withgoogle.com>
- Hagan clic en “Primeros pasos”
- Selecciones “Proyecto de imagen”

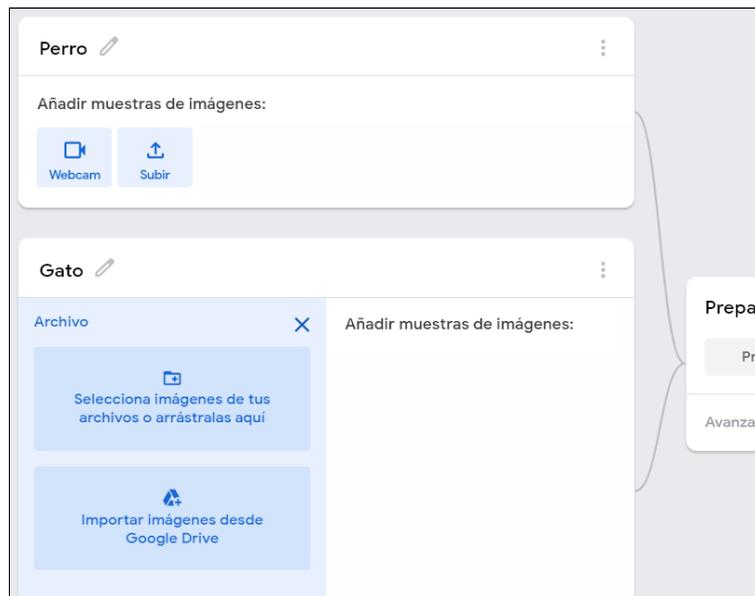
Pasos para agregar un conjunto de datos de entrenamiento:

- 1) Creen dos clases, una para gato y otra para perro
- 2) Esperen a que su profesor o profesora les dé instrucciones (puede que trabajen con imágenes impresas de perros y gatos frente a la webcam o cargando imágenes al sistema)
- 3) Si están trabajando con impresos, hagan clic en “Webcam” y muéstrenlos frente a la cámara. Si usarán archivos de imagen, elijan “Subir”.

Importante: usar sólo el “conjunto inicial de datos de entrenamiento” en esta parte de la actividad.

(Las siguientes instrucciones son sólo para la opción de cargar archivos de imágenes)

- 4) Debajo de "file", elijan si subirán los datos desde el computador o desde una unidad de Google Drive.
- 5) Localicen el conjunto de datos de entrenamiento para las clases gato y perro.



- 6) Hagan clic en "Abrir" o "Select" para incorporar el conjunto de imágenes. **Asegúrense de agregar todas las imágenes y una sola vez. Mantengan "Ctrl" presionado para seleccionar varios ítems.**

¿Listos/as? Ahora vamos a responder algunas preguntas sobre los datos.

Pregunta 1:

a) Para el conjunto de datos de entrenamiento del perro, registren lo siguiente:

¿Cuántas imágenes se incluyeron? _____

¿En qué se parecen las imágenes? _____

¿En qué se diferencian las imágenes? _____

b) Para el conjunto de datos de entrenamiento del gato, registren lo siguiente:

¿Cuántas imágenes se incluyeron? _____

¿En qué se parecen las imágenes? _____

¿En qué se diferencian las imágenes? _____

Pregunta 2: Entrenen a su clasificador con los dos conjuntos de datos de entrenamiento entregados.

Una vez que su clasificador esté terminado, prueben su modelo con las imágenes del **conjunto de datos de prueba**, anotando los resultados de las predicciones del sistema:

Imagen	Clasificación	Porcentaje de confianza	¿Correcto?
			

			
---	--	--	--

Pregunta 3: ¿En qué grupo funcionó mejor su clasificador? (marque uno)

Gatos

Perros

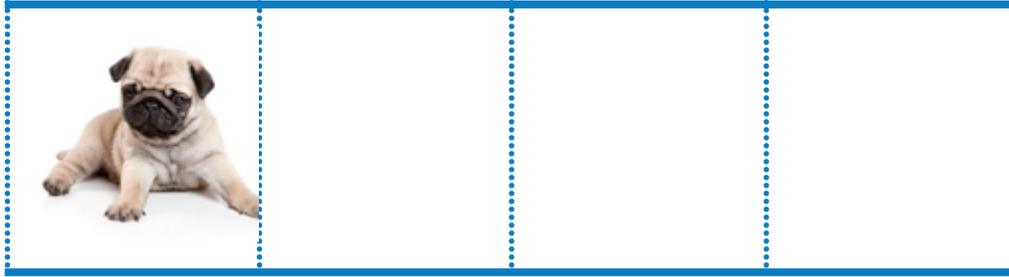
¿Por qué creen que pasó? _____

Pregunta 4: Ahora entrenen un nuevo clasificador con el **conjunto de datos de entrenamiento redefinido**.

Cuando el clasificador esté terminado, prueben los resultados del modelo con el conjunto de datos de prueba, anotando los resultados de las predicciones. Completen la siguiente tabla:

Imagen	Clasificación	Porcentaje de confianza	¿Correcto?
			





Pregunta 5: Su nuevo algoritmo funcionó... (marquen uno)

Mejor para perros

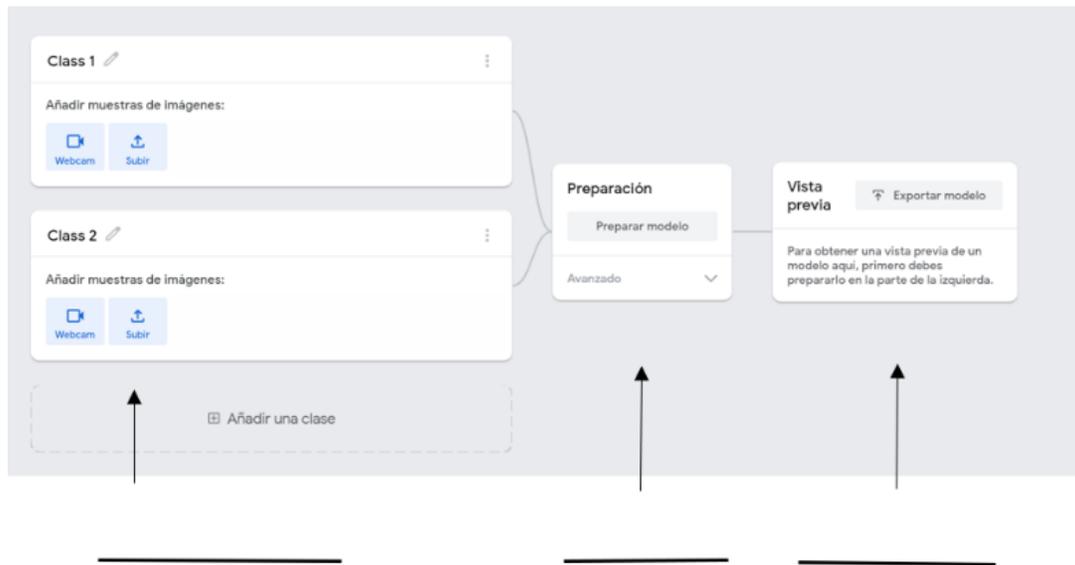
Lo mismo para gatos y perros

Mejor para gatos

Pregunta 6: Entreguen sus observaciones con respecto a la pregunta 5:

Parte III: Aprendizaje automatizado supervisado (cuestionario)

1. ¿Cuáles son los tres componentes de un sistema de IA?



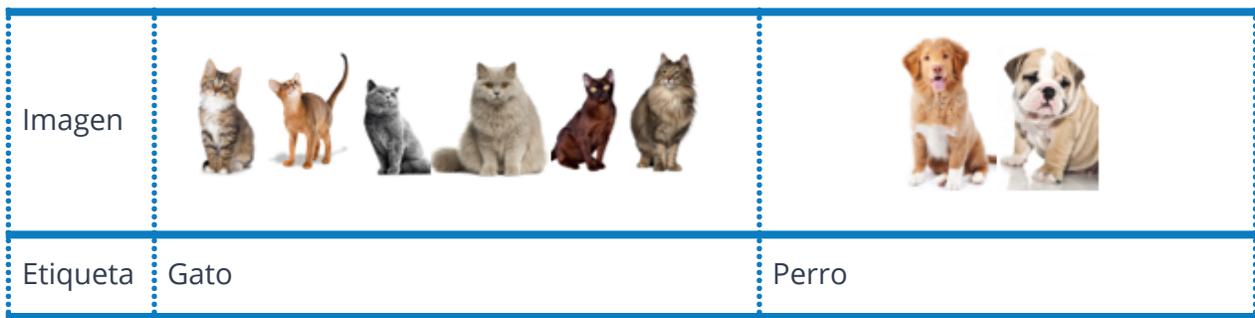
2. Un algoritmo de aprendizaje automatizado supervisado ha sido entrenado con las siguientes imágenes con la etiqueta "gato".

Imagen	 
Etiqueta	Gato

¿Cómo clasificará la siguiente imagen? (marca una):



3. Un algoritmo de aprendizaje automatizado supervisado ha sido entrenado con las siguientes imágenes:



3.1 ¿Esperas que la precisión del algoritmo sea (marca uno):

Mejor para gatos Lo mismo para gatos y perros Mejor para perros

3.2. ¿Qué conjunto de datos de entrenamiento entregaría la clasificación más precisa tanto para gatos como para perros?



B.

	
Gato	Perro

C.

	
Gato	Perro

D.

	
Gato	Perro

Tiempo estimado total de actividad: 2 horas

Conjuntos de datos de imágenes

En esta sección encontrará versiones imprimibles de los conjuntos de imágenes para que las y los estudiantes construyan sus clasificadores gato-perro, en caso de que haya decidido trabajar con las tarjetas impresas frente a la cámara. Tanto en esta versión como en la digital, las imágenes están clasificadas en tres conjuntos de datos:

Conjunto de datos	Descripción
Conjunto inicial de datos de entrenamiento	<p>Estas son las imágenes que las y los estudiantes deberán usar para “enseñar” a su modelo de aprendizaje automatizado qué fotos son un gato o un perro.</p> <p>Note que hay muchos más gatos y que los gatos son más diversos en apariencia que los perros. Esto significa que el clasificador tendrá más información e identificará con mayor precisión a los gatos que a los perros.</p>
Conjunto de datos de prueba	<p>Estas son las imágenes que las y los estudiantes deberán usar para testear su clasificador después del entrenamiento. Deberán mostrar estas imágenes a su modelo y registrar si su clasificador predice si la imagen es de un perro o de un gato.</p> <p>Nota: Las y los estudiantes <u>no deben usar</u> estas imágenes para entrenar a su clasificador, produciendo un error metodológico esencial. Si usan una imagen para entrenar a un clasificador, la máquina ya habrá grabado la etiqueta correspondiente a esa imagen en particular, por lo que sólo</p>

	<p>estará reconociéndola y no extrapolando lo que aprendió. En otras palabras, no medirá qué tan bien generaliza el modelo.</p>
<p>Conjunto de datos redefinido</p>	<p>Corresponde las imágenes que los y las estudiantes usarán para reentrenar el modelo, una vez se den cuenta del sesgo del conjunto inicial. Este conjunto cuenta con menos imágenes de cada clase, pero con una mayor diversidad.</p>

El conjunto de datos de prueba deberá usarse dos veces, una para testear el clasificador inicial y otra vez para evaluar el entrenamiento del modelo con el conjunto de datos redefinido.

Las imágenes a continuación están organizadas para facilitar su impresión y corte (con dos imágenes por página, organizadas por conjunto de datos). Imprimir las imágenes permitirá a las y los estudiantes sostenerlas frente a la cámara web para el entrenamiento. Si va a utilizar las fotos en formato digital, ignore este párrafo.

Conjunto inicial de datos de entrenamiento

Perros









Gatos















©Warren Photographic



©Warren Photographic



Conjunto de datos de prueba







©Warren Photographic





Conjunto de datos redefinido

Perros









Gatos







ACTIVIDAD 5

CASO I: “FALSIFICACIÓN DE INFORMACIÓN” (PARTE 1)

Oportunidades curriculares

Tecnología

- **I° Medio - OA 5:** Analizar las formas en que los productos tecnológicos y los entornos evolucionan, caracterizando los diversos factores que influyen en ese cambio.
- **II° Medio - OA5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.
- **II° Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **II° Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.

Inglés

- **I° Medio - OA 1:** Demostrar comprensión de ideas generales e información explícita en textos orales adaptados y auténticos simples, literarios y no literarios, en diversos formatos audiovisuales (como exposiciones orales, conversaciones, descripciones, instrucciones y procedimientos, avisos publicitarios, entrevistas, noticieros, anécdotas, narraciones, canciones), acerca de temas variados (experiencias personales, temas de otras asignaturas, del contexto inmediato, de

actualidad e interés global o de otras culturas) y que contienen las funciones del año.

- **Iº Medio - OA 9:** Demostrar comprensión de ideas generales e información explícita en textos adaptados y auténticos simples, en formato impreso o digital, acerca de temas variados (como temas de interés de los y las estudiantes, temas de otras asignaturas, del contexto inmediato, de actualidad e interés global, de otras culturas y algunos temas menos conocidos) y que contienen las funciones del año.
- **IIº Medio - Unidad 1:** Globalization and communication

Visión general

En esta actividad, las y los estudiantes conocen y reflexionan en torno a sistemas de inteligencia artificial que pueden imitar la realidad o simular actividades humanas con un alto nivel de verosimilitud, comprendiendo sus implicancias en la generación de información falsa (por ejemplo, mediante la suplantación de identidad o la generación de textos noticiosos). Se presentan sistemas como Flow Machines de Sony, Deepfake y GPT para representar los resultados de este tipo de tecnologías.

Objetivos

- Experimentar con herramientas que permiten crear contenido nuevo a partir de datos preexistentes.
- Reflexionar sobre las implicancias de sistemas avanzados de IA que permiten imitar la realidad con un alto nivel de verosimilitud.

Agenda (75 minutos)

- Introducción (15 min)
- Desarrollo: presentación teórica y demostraciones (45 min)
- Cierre (15 min)

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas + acceso a Internet
- Hoja del estudiante (una por grupo)

Palabras clave

Sistemas de IA, IA creativa, algoritmos generativos, deepfake, noticias falsas, realidad.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1_Ihds-3mNn1KscblqAJuliAFQwZ9m09neZdS_6uMs6k/edit?usp=sharing

GUÍA DOCENTE

Introducción (15 minutos)

Esta es la primera actividad del programa relacionada con análisis de casos. En esta oportunidad, las y los estudiantes conocerán cómo la realidad, mediante el uso de algunos sistemas de IA, puede ser imitada y/o falsificada.

Para acompañar la explicación, la actividad cuenta con una presentación de diapositivas de apoyo, donde encontrará todos los elementos necesarios. Puede utilizar los textos a continuación para contextualizar el caso o leerlos en voz alta.

Motivación

—Comenzaremos con un ejemplo histórico, el primer caso de una fotografía falsa.

En enero de 1839, se celebraba en París la invención de la fotografía por parte de Louis Daguerre (durante años, las fotografías se conocían como daguerrotipos). Sin embargo, Daguerre no era el único que investigaba este proceso. Hippolyte Bayard, uno de sus competidores, clamaba haber inventado un proceso fotográfico distinto al de Daguerre, con anterioridad al que éste había presentado en 1839. Un año después, en 1840, se publicaba la siguiente imagen titulada “Autorretrato de un hombre ahogado” acompañado por el siguiente texto:

“El cuerpo que usted aquí ve le pertenece al Sr. Bayard, inventor del proceso que se despliega ante usted (la fotografía). El gobierno, que ha sido muy generoso con el Sr. Daguerre, ha dicho que no podía hacer nada por el Sr. Bayard, y el pobre se ha suicidado. ¡Oh... la frivolidad humana! Su cuerpo ha estado abandonado en la morgue por varios días y nadie lo ha reconocido ni reclamado. Señoras y Señores, mejor que ni se acerquen si no quieren exponerse a tan horrible olor, pues como pueden ver, sus manos y su cara ya han empezado a descomponerse”. [1]



Imagen: "Self-Portrait as Drowned Man"¹²

(Autorretrato como
hombre ahogado)

La fotografía, sin embargo, fue tomada por el mismo Bayard, quien, decepcionado de que la academia científica francesa hubiese invertido en el proceso de Daguerre y no en el suyo, protestaba con lo que hoy es considerada la primera fotografía falsa, inventada solo un año después de la fotografía y más de 150 años antes de los programas de retoque de imágenes. Efectivamente, Bayard alteró la imagen para que pareciera que tanto sus manos como su cara se encontraban en proceso de descomposición.

Reflexión con el curso

—¿Nos podemos imaginar el impacto que esta imagen debe haber tenido sobre aquellos que la veían a mediados del Siglo XIX?

Las fotografías hoy son parte de nuestro diario vivir, sin embargo, quienes se enfrentaron a la imagen anterior en 1840 debían haber visto en su vida solo algunas pocas que, hasta ese entonces, representaban fielmente la realidad. Aquellos que se enfrentaron a la figura de Bayard con las manos y la cara pintadas habrán creído sin dudas que se trataba efectivamente de la imagen de un cadáver.

—¿Son importantes las imágenes para configurar nuestra forma de mirar el mundo?
¿Podrían dar algún ejemplo?

¹² Fuente: <https://arthistoryproject.com>

(Si no surgen ejemplos rápidamente, puede hablar de imágenes de guerras, atentados, protestas, memes, etc.).

—Un impacto similar ocurre con quienes se enfrentan por primera vez al “deepfake”, término con el que nos referimos a contenidos manipulados con técnicas de inteligencia artificial, cuyo uso más conocido es el de “trasplantar” rostros y expresiones de una persona a otra (profundizaremos en esta técnica a lo largo de la actividad).

Desarrollo: presentación teórica y demostraciones (45 minutos)

—Los algoritmos de inteligencia artificial pueden ser **predictivos** (como vimos en las actividades 1 y 4) o **generativos**. En el segundo caso, utilizan como entrada un conjunto de ejemplos desde los cuales aprenden a generar nuevas muestras que no pertenecen al conjunto de ejemplos. Esta tecnología ha sido llamada IA creativa, pues tiene aplicaciones como el desarrollo de música, texto, imágenes e incluso videos.

A continuación, se muestran algunos ejemplos para revisar. Apóyese en las láminas de la presentación y diríjase a los links que se detallan a pie de página:

- A. Algoritmos generativos para música: **Flow Machines de Sony**
 - a. Daddy’s Car: canción creada por inteligencia artificial en el estilo de los Beatles¹³
 - b. Himno de la Alegría en distintos estilos generado por IA¹⁴
- B. Algoritmos generativos para Imágenes: **Karras et. al. y Nvidia**
 - a. This person does not exist: imágenes de personas que no existen, generadas a partir del aprendizaje de rasgos de diferentes fotos¹⁵
- C. Algoritmos generativos para Texto: **GPT-2**
 - a. Talk to a Transformer - Inferkit¹⁶ (disponible sólo en inglés)

Los dos primeros ejemplos (Flow Machines), generan música nueva en base a composiciones preexistentes; en el primer caso, a partir de canciones de The Beatles y en el segundo aplicando algoritmos de síntesis sobre diferentes versiones del Himno de la Alegría, en lo que se conoce como “transferencia de estilo”.

¹³ https://www.youtube.com/watch?v=LSHZ_b05W7o

¹⁴ <https://www.youtube.com/watch?v=buXqNqBFd6E>

¹⁵ <https://thispersondoesnotexist.com/>

¹⁶ <https://app.inferkit.com/demo>

“This person does not exist” es un sitio que contiene imágenes de personas generadas por IA a partir de una base amplia de fotos reales, gracias a una aplicación de Redes Generativas Adversarias (RGA o GAN). Estas redes antagónicas “se entrenan” mutuamente, una especializada en generar contenido realista y la otra en discriminar su calidad, generando, como resultado, contenidos muy difíciles de distinguir como falsos. **Esta tecnología es clave para entender el fenómeno del deepfake.**

“Talk to transformer” es un sistema que genera texto y conversaciones a partir de palabras ingresadas por el usuario, utilizando el sistema GPT-2, de Open AI, que trabaja en un campo llamado Procesamiento de Lenguaje Natural (Natural Language Processing o NLP). Esta tecnología funciona con un modelo entrenado a partir de textos reales, que es capaz de generar textos realistas en base a análisis estadísticos.

—Hagamos una prueba con esta última herramienta.

Ingrese a <https://app.inferkit.com/demo> y muestre cómo funciona. Seleccione uno de los temas preestablecidos, como “Sci-fi” (ciencia ficción) o “Unicorns that speak English” (unicornios que hablan inglés) y muestre cómo se genera texto de manera automática, a partir de un conjunto de palabras inicial dado.

El modelo está entrenado para interactuar en inglés, por lo que se recomienda tener abierto en paralelo un traductor en línea, como el de Google, Bing o DeepL, o realizar esta parte de la actividad en coordinación con la asignatura de inglés. Es ideal contar con una cuenta creada con anticipación, para que el sistema no limite el número de interacciones con la plataforma.

—Probemos ingresando los siguientes textos inconclusos y veamos lo que nos arroja el sistema. Tengamos siempre en mente las posibles ventajas y perjuicios en cada caso.

Seleccione “custom prompt” (texto personalizado) y haga pruebas con los textos que se sugieren en la tabla (use los textos en inglés). Definimos la primera interacción y generamos el resto de manera automática, presionando “Complete text”. Si el primer intento arroja algo inconexo o extraño, haga clic sobre el botón “Generate another” para obtener un texto diferente.

Para generar una conversación:

Pedro: Is everything OK? *

Antonia:

* [¿Está todo bien?]

Para generar un texto argumentativo o expositivo:

- Chile is a good place to live because
[Chile es un buen lugar para vivir porque...]
- Water has three basic states
[El agua tiene tres estados básicos...]

Para generar una historia de ficción:

- The little bear was
[El pequeño oso estaba...]
- There was a huge spaceship under the
[Había una nave espacial enorme bajo la...]
- Once upon a time
[Había una vez...]

Para generar textos sobre noticias plausibles:

- The congressman was found guilty of
[El diputado fue hallado culpable de...]
- Scientists discover
[Científicos descubren...]

Otra aplicación de IA creativa consiste en transferir los gestos y movimiento de una cara a otra. Dado que el algoritmo de aprendizaje utilizado como base para esta técnica se conoce como “redes neuronales profundas” o Deep Neural Networks, los resultados de estos algoritmos han sido denominados **deepfake**, una mezcla de los términos deep (profundo) y fake (falso). Una de las herramientas más utilizadas para la generación de deepfakes es Deepfacelab¹⁷, un sistema gratuito y abierto que presenta resultados muy convincentes y

¹⁷ <https://github.com/iperov/DeepFaceLab>

cuya aplicación más conocida es la de “trasplantar” el rostro de personajes famosos en otros.

(Si cuenta con tiempo, sugerimos visitar el sitio del proyecto, donde podrá encontrar aplicaciones tales como reemplazo de cara, cambio de cabeza manteniendo el mismo cuerpo, rejuvenecimiento facial y manipulación de discursos públicos).

Actualmente, la principal aplicación de los *deepfake* es en el área del entretenimiento. La industria del cine sigue utilizando tecnologías de captura de movimiento, maquillaje y animación digital, pero es probable que en el futuro cercano la tecnología que permite la creación de *deepfakes* sea lo suficientemente convincente como para que su uso tenga aplicaciones comerciales. A pesar de lo anterior, la popularización de esta tecnología ha levantado una intensa discusión ética sobre el impacto que podría ocasionar en nuestra sociedad.

Cierre (15 minutos)

—De la misma manera que la fotografía de Bayard pudo engañar a algunas personas, dada la novedad de la fotografía en el siglo XIX, los *deepfake* pueden tener el mismo impacto si su uso es inescrupuloso. Hoy, cuando la tecnología tiene el potencial real de imitar la realidad, nos preguntamos sobre las implicancias que esto puede tener en nuestra sociedad.

Los y las estudiantes reflexionan y responden sus guías en grupos de cuatro personas. Las preguntas tienen que ver con el impacto de este tipo de tecnologías en la vida real.

Termine la clase explicando que la actividad tiene una segunda parte, en la que las y los estudiantes podrán explorar nuevas herramientas de IA y someterse al llamado “Test de Turing” (puede incentivarlos, invitándoles a investigar en sus casas de qué se trata esta prueba).

Bibliografía:

[1] <https://arthistoryproject.com/artists/hippolyte-bayard/self-portrait-as-drowned-man/>

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana

Caso I: Falsificación de información (partes 1 y 2)

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales/ Inglés

Fecha: _____ Curso: _____

Objetivos:

- Experimentar con herramientas que permiten crear contenido nuevo a partir de datos preexistentes.
- Reflexionar sobre las implicancias de sistemas avanzados de IA que permiten imitar la realidad con un alto nivel de verosimilitud.

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____
4. _____

Primera parte (actividad 5)

Ahora que entiendes los conceptos básicos y conoces las implicancias de los sistemas de inteligencia artificial expuestos por el profesor o profesora, te invitamos a responder las siguientes preguntas (si quedaste con alguna duda, ¡es el momento para preguntarla!):

1. ¿Cuál es el impacto pueden tener los mensajes televisivos de nuestras autoridades en la confianza de la ciudadanía? ¿Y si los mensajes son falsos? ¿Cómo contribuyen a la propagación de noticias falsas?

2. Actualmente, muchas estafas telefónicas y por redes sociales consisten en la suplantación de familiares que, desesperados en una supuesta emergencia, nos piden dinero. ¿Qué impacto podría tener el tipo de tecnologías vistos en esta clase de estafas?

3. ¿Qué beneficios y perjuicios podrían traer este tipo de aplicaciones de la inteligencia artificial (pensando en tecnologías tales como generación de piezas artísticas, creación de rostros humanos, transferencia de caras y expresiones y generación de diferentes tipos de textos falsos)?

Segunda parte (actividad 6)

Ingresen al sitio <https://detectfakes.media.mit.edu> y sigan las instrucciones de su profesor o profesora.

1. Completen la siguiente tabla, indicando sus respuestas y aciertos.

Video	Respuesta del grupo	¿Acertaron?
	Por ejemplo: 100% deepfake	Sí/No
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

2. ¿Qué se siente ser “engañado/a” por una máquina? ¿Qué pasa cuando este engaño se usa para robar o generar un testimonio o evidencia falsos?

3. Observen sus respuestas de la actividad anterior. ¿Qué cambió? ¿Qué efectos creen que podría tener el uso malicioso de estas tecnologías, ahora que han experimentado el fenómeno del “engaño”?

ACTIVIDAD 6

CASO I: “FALSIFICACIÓN DE INFORMACIÓN” (PARTE 2)

Oportunidades curriculares

Tecnología

- **I° Medio - OA 5:** Analizar las formas en que los productos tecnológicos y los entornos evolucionan, caracterizando los diversos factores que influyen en ese cambio.
- **II° Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.
- **II° Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia

- **II° Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.

Visión general

En esta actividad, las y los estudiantes experimentan una adaptación del test de Turing, teniendo que diferenciar contenidos reales de los elaborados por sistemas de inteligencia artificial. Después de revisar un caso cotidiano aplicado a redes sociales, utilizan la plataforma *Detect Fakes* para identificar videos que han sido manipulados a través de IA para distorsionar caras o intercambiar

rasgos faciales. Finalmente, conocen y reflexionan en torno a un caso real de fraude, para volver a responder las preguntas de la guía de la actividad anterior.

Objetivos

- Experimentar con herramientas que permiten crear contenido nuevo a partir de datos preexistentes.
- Reflexionar sobre las implicancias de sistemas avanzados de IA que permiten imitar la realidad con un alto nivel de verosimilitud.

Agenda (60 minutos)

- Introducción (15 min):
 - Conocimientos previos: filtro de Snapchat
 - El test de Turing
- Desarrollo (30 min):
 - Experimentación con plataforma Detect Fakes
 - Caso real: engaño con deepfake
- Cierre (15 min):
 - Reflexión sobre el fenómeno del engaño
 - Comparación de respuestas de la actividad anterior

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Celular con Internet con Snapchat y sesión iniciada
- Computadores con Internet
- Hoja del estudiante (misma que para parte I)

Palabras clave

Test de Turing, IA creativa, deepfake, contenido falso, realidad.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1JRiACuJSVsiDyWZIXMF4zP_a4B6z2rufgnom1lI3lo/edit?usp=sharing

GUÍA DOCENTE

En la primera parte de esta actividad, las y los estudiantes reflexionaron acerca de las implicancias éticas de estos sistemas de IA. El objetivo de esta segunda parte será que puedan conocer nuevos elementos y experimentar el fenómeno del engaño por sus propios medios.

Introducción (15 minutos)

Esta parte introductoria tiene el objetivo de activar los conocimientos previos de las y los estudiantes mediante la revisión de aplicaciones cotidianas de la IA, llamando de paso la atención en torno a lo vinculados/as que están, quieran o no, con estos sistemas.

—¿Cuántos/as de ustedes tienen redes sociales como Snapchat, TikTok o Instagram?

Seguramente una amplia mayoría contesta afirmativamente.

—¿Alguien ha utilizado el filtro de intercambio de caras?

Pídale a un par de parejas que quieran hacer la prueba. Para esto, indíqueles que vayan al buscador de filtros de una de estas aplicaciones y pongan “FaceSwap”. Puede reforzar las instrucciones con fuentes externas¹⁸. Verá cómo sus estudiantes ríen y comentan los resultados del intercambio facial.

—¿Sabían que este filtro usa inteligencia artificial?

Comente a las y los estudiantes que éste es otro ejemplo de cómo este tipo de sistemas de IA generativa se están utilizando con fines recreativos, enfatizando que tecnologías muy similares también pueden usarse con objetivos delictivos. Esta reflexión inicial es importante para el análisis que viene a continuación.

IA Creativa e Inteligencia

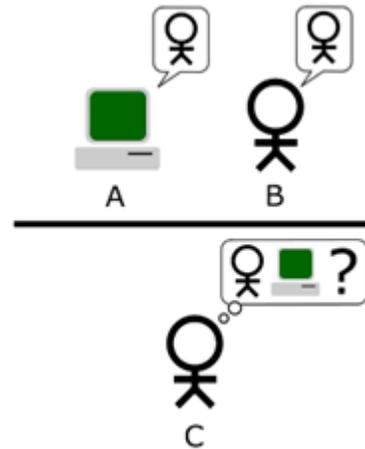
Como mencionamos en la actividad anterior, las tecnologías basadas en **IA generativa** son a veces llamadas **IA creativa**, pues su uso permite la creación de dibujos, pinturas, imágenes y canciones. Hasta hace solo unos años, el adjetivo “creativo” era exclusivamente

¹⁸<https://www.infobae.com/america/tecno/2020/04/23/face-swap-como-usar-el-filtro-de-snapchat-para-intercambiar-rostros>

utilizado para describir a los seres humanos. Más aún, **la creatividad artística es aún considerada una de las principales características de la inteligencia humana**. Sin embargo, los resultados de la IA creativa nos hacen preguntarnos si la creatividad es aún una cualidad exclusiva de los seres humanos. Más aún, ¿son los sistemas basados en IA “inteligentes”?

Esta última pregunta es tan antigua como la informática. **Alan Turing**, padre de la computación moderna, se pregunta en su artículo de 1950 “Máquinas computacionales e inteligencia” la pregunta: **¿pueden las máquinas pensar?** Al iniciar su artículo, Turing indica que las definiciones de “inteligencia” y “capacidad de pensamiento” son difíciles y polémicas. En vez de intentar definir estos términos, Turing propone una prueba llamada “El juego de la imitación” (The imitation game), en el cual un interrogador aislado en una sala se comunica a través de mensajes escritos con personas y con máquinas sobre las cuales se espera evaluar inteligencia. El interrogador, que no sabe quiénes de sus interlocutores son máquinas o personas, puede hacer todas las preguntas que estime necesarias. Al final del juego, el interrogador debe identificar quiénes de sus interlocutores son máquinas o personas. Si en varias rondas del juego, con distintos interrogadores, la máquina y las personas interrogadas no logran diferencias sustantivas en ser identificadas, entonces la máquina es considerada “inteligente”, o lo suficientemente “inteligente” como para suplantar a una persona¹⁹.

¹⁹ Puede apoyarse con este video de TED-Ed, activando los subtítulos en español:
<https://www.youtube.com/watch?v=3wLqsRLvV-c>

Alan Turing a los 16 años²⁰.

Prueba de Turing²¹. El interrogador (C) desconoce quién es el ser humano (B) y quién la máquina (A).

Esta “prueba de inteligencia” es actualmente conocida como el “test de Turing”.

La IA creativa ha tenido éxito en *imitar* la inteligencia humana en el sentido de la creatividad artística. Las “pruebas de Turing creativas” han sido utilizadas para evaluar la “creatividad” de las máquinas, pidiéndoles a personas intentar distinguir entre imágenes generadas por IA creativa o por seres humanos. Los resultados de estos tests son mixtos, aunque en algunas aplicaciones son muy positivos en favor de los algoritmos de inteligencia artificial [1].

—¿Les gustaría probar una adaptación del test de Turing? ¿Podrán identificar qué es real y qué es elaborado por máquinas?

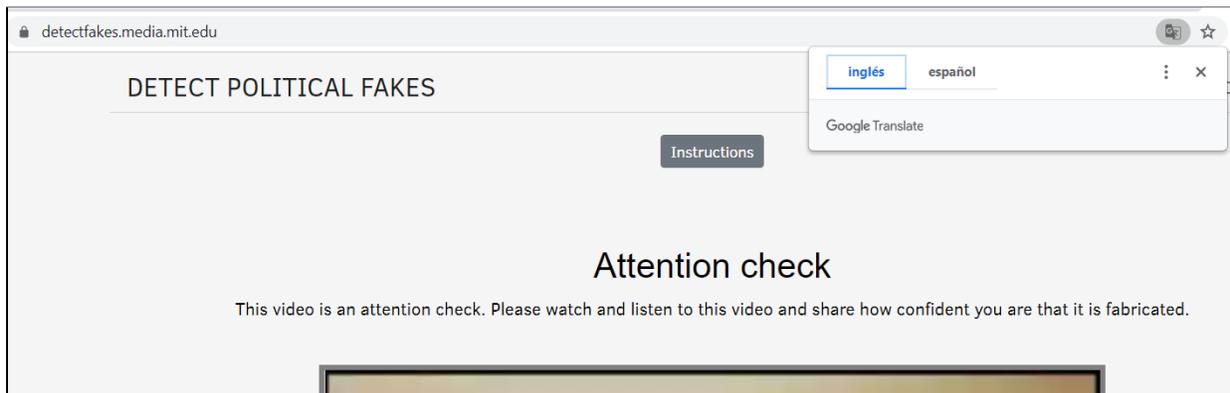
Desarrollo (30 minutos)

Para esta actividad, diríjase a la sala de computación o Enlaces para que sus estudiantes tengan acceso a un computador con Internet. Una vez allí, pídale ingresar al sitio con el navegador Google Chrome: <https://detectfakes.media.mit.edu>. Indíqueles que activen la traducción automática desde la esquina superior derecha de la pantalla (como se indica en

²⁰ Imagen de dominio público vía Wikimedia Commons

²¹ Imagen de dominio público vía Wikimedia Commons. Autor: [Bilby](#)

la imagen). Note que la plataforma solicita el consentimiento informado de los padres o apoderados, dado que recolecta los datos de las respuestas con fines investigativos²². Considere un tiempo para informar esta situación antes de realizar la actividad.



Una vez sus estudiantes hayan ingresado a la página, explíqueles que la actividad se trata de identificar si los videos que se muestran son reales o falsos, es decir, si están alterados o no por un algoritmo de IA (contenido falso o *fake*). El sitio define un *deepfake* como aquel contenido en el que “una cara ha sido manipulada por IA para distorsionarla parcialmente o intercambiar rasgos faciales”. Para hacer la prueba, tienen que ver los videos e indicar en cada caso qué tan seguros/as están de que se trata de un *deepfake* y presionar “enviar”.

Una vez enviada la respuesta, el sistema retroalimentará con la predicción de un sistema automatizado y, posteriormente, con una respuesta taxativa. Para ver otro video, es necesario pulsar “próximo”. La plataforma irá entregando indicaciones y explicaciones mientras se usa, como el número de preguntas que se ha respondido y la advertencia sobre el uso de distractores.

Deje que sus estudiantes revisen entre 5 y 10 videos, de modo que puedan experimentar el fenómeno del engaño. Para ello, deben usar la hoja del estudiante de la clase anterior y completar el cuadro con sus respuestas. En lo posible, mantenga los mismos grupos.

—¿Cómo les fue con el test? ¿Cuántas veces fallaron en identificar el contenido falso?

²² <https://detectfakes.media.mit.edu/informed-consent>

Una vez que hayan comentado distintas experiencias, presente el siguiente caso:

—¿Recuerdan que dijimos que estos sistemas estaban relacionados a la entretención, como en el caso de Instagram, TikTok o Snapchat? Pues bien, la IA creativa también puede ser usada para usos maliciosos.

Presente el titular de esta noticia “Los deepfakes de audio se usan para realizar estafas telefónicas”. Ingrese al link a continuación. Haga una lectura general y apóyese con la presentación de diapositivas.

<https://www.lavanguardia.com/tecnologia/20190910/47263675640/deepfake-audio-voz-suplantacion-estafa-telefonica-llamada-inteligencia-artificial-seguridad.html>

Si cuentan con tiempo, recomendamos revisar también los titulares y el contenido principal de los siguientes contenidos noticiosos:

- “Los videos virales del falso Tom Cruise en TikTok con tecnología deepfake: ¿cuál es el peligro?” El objetivo de este artículo es reflexionar sobre la proliferación y los riesgos de generar perfiles falsos de figuras públicas.

<https://www.infobae.com/america/tecno/2021/03/01/los-videos-virales-del-falso-tom-cruise-en-tiktok-con-tecnologia-deepfake-cual-es-el-peligro/>

- “ZAO, la app móvil china que a través de deepfake te convierte en DiCaprio en segundos”. El objetivo de este video es graficar la facilidad con que se pueden desarrollar hoy videos con suplantación de identidad, incluso sin la necesidad de software o conocimiento especializado.

https://www.youtube.com/watch?v=VnUvGeLK_8w

Cierre (15 minutos)

Para finalizar esta actividad de dos partes, realice la siguiente dinámica de cierre y consolidación.

—¿Qué se siente ser engañado/a por una máquina? ¿Qué pasa cuando este engaño ya no es un simple juego o entretenimiento, sino para robar o generar un testimonio o evidencia falsos?

Entréguales un espacio para contestar estas preguntas.

—Observen nuevamente sus respuestas de la actividad anterior. En ellas podrán encontrar las implicancias que ustedes indicaron que podía tener el uso de este tipo de sistemas de IA. ¿Qué cambió? ¿Qué efectos creen que podría tener el uso malicioso de estas tecnologías, ahora que han experimentado el fenómeno del “engaño”? Agreguen la respuesta a esta pregunta en la última parte de la guía anterior.

Bibliografía

[1] <http://bregman.dartmouth.edu/turingtests/node/72>

ACTIVIDAD 7

CASO II: ¿CÓMO ENSEÑAREMOS A LA IA A MANEJAR?

Oportunidades curriculares

Tecnología

- **I° Medio - OA 5:** Analizar las formas en que los productos tecnológicos y los entornos evolucionan, caracterizando los diversos factores que influyen en ese cambio.
- **I° Medio - OA 6:** Inferir, basándose en la evolución de los productos tecnológicos y los entornos, los efectos positivos o negativos que estos han tenido en la sociedad.
- **II° Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.
- **II° Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **II° Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.
- **II° Medio - OA 24:** Analizar y debatir sobre los desafíos pendientes para el país, por ejemplo, reducir la pobreza y la desigualdad, garantizar los derechos de los grupos discriminados, lograr un desarrollo sustentable, perfeccionar el sistema político y fortalecer la relación con los países vecinos, y reconocer los deberes del Estado y la responsabilidad de todos los miembros de la sociedad para avanzar en ellos.

Visión general

Luego de conocer y experimentar diferentes aplicaciones del aprendizaje supervisado (actividades 4, 5 y 6), las y los estudiantes se enfrentan al dilema ético de la inteligencia artificial aplicada a la conducción de automóviles autónomos. El curso reflexiona sobre las implicancias de diferentes criterios de respuesta frente a accidentes de tránsito, identificando las partes interesadas (*stakeholders*) en el diseño e implementación de los algoritmos que los sustentan y comprendiendo el fenómeno como un problema sociotécnico en el que debe verse reflejada la visión de la sociedad en su conjunto.

Objetivos

- Reflexionar sobre las consecuencias e implicancias éticas de sistemas de inteligencia artificial que se encuentran en fase de desarrollo.

Agenda (90 minutos)

- Introducción (15 min)
 - Preguntas iniciales y nueva tecnología disponible
- Desarrollo (60 min)
 - Video
 - Desarrollo de guía del estudiante
- Cierre (15 min)
 - Puesta en común

Materiales para esta actividad

- Presentación de diapositivas
- Hoja del estudiante
- Proyector o laboratorio de computación/ Enlaces

Palabras clave

Ética, partes interesadas, sistemas de IA, sistemas sociotécnicos, vehículos autónomos.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/19YppyYpL3wouw35GyZOe1MbgAW_pxD3LPGCWaRpSEFo/edit?usp=sharing

GUÍA DOCENTE

En esta actividad, revisaremos algunos temas éticos asociados al desarrollo de algoritmos de inteligencia artificial para la conducción de vehículos autónomos (automóviles, camiones, buses, etcétera). Al tratarse de que una tecnología que permite definir deliberadamente el repertorio de respuestas de los vehículos frente a accidentes de tránsito, resulta imprescindible discutir sobre los principios que orientarán la decisión de los algoritmos y a quién corresponde establecerlos, entrando en juego consideraciones éticas tales como la discriminación y la minimización de daños.

En la primera parte de la actividad, revisaremos conceptos y ejemplos asociados a la conducción autónoma de vehículos, con el objetivo de introducir un lenguaje y conocimiento comunes. En la segunda, veremos un video de TED-Ed que problematiza la respuesta de estos vehículos frente a accidentes de tránsito, en las palabras de Patrick Lin (director del grupo de ética y ciencias emergentes de la Universidad Politécnica de California). Para finalizar, las y los estudiantes reflexionan en grupo acerca de las implicancias de diferentes criterios de respuesta a adoptar, así como las partes interesadas en su diseño e implementación.

Introducción (15 minutos)

—¿Alguien alguna vez ha visto un automóvil que se maneje solo? (Probablemente le respondan que en alguna película o en las noticias)

Exponga lo siguiente, apoyándose en la presentación de diapositivas:

A pesar de que esta aplicación puede sonar a ciencia ficción, está lejos de serlo. En 2020, el estado de California, en Estados Unidos, otorgó permiso a 5 compañías distintas para probar automóviles completamente autónomos y a más de 50 para hacerlo con vehículos semiautónomos (donde la IA puede devolver el control a un conductor humano, si es necesario). Entre estas compañías se incluyen fabricantes de automóviles (BMW, Ford, Volkswagen, entre otras), compañías de servicio de transporte de pasajeros (Uber, Didi, Lyft, entre otras), fabricantes de equipos computacionales (Apple, Intel, Nvidia, entre otras) y compañías completamente dedicadas al desarrollo de automóviles autónomos como Waymo (ex-Google Car), Phantom AI e EasyMile [1].

Precisamente, esta última (EasyMile) recientemente realizó una prueba de uno de sus buses autónomos en Chile [2] con muy buenos resultados [3]. El bus, que recorrió más de 50 kilómetros diarios transportando a más de 200 pasajeros al día, no reportó incidentes en su corto recorrido de 1 kilómetro [4].

Como podemos observar, la tecnología se encuentra en una etapa de pruebas en condiciones reales. Por lo mismo, los países están discutiendo sobre el tipo de legislación que se debe crear para regular las condiciones de operación de estos vehículos.

En lo que sigue, veremos por qué esta discusión es difícil y es necesario desarrollar un punto de vista propio sobre este tema.

Desarrollo (60 minutos)

Luego de la introducción de la temática y los conceptos básicos, el curso verá el video de Patrick Lin en TED-Ed sobre los dilemas éticos asociados a la programación de sistemas de IA para la conducción autónoma (<https://youtu.be/ixloDYVfKA0>). El video se encuentra narrado en inglés, pero se pueden activar los subtítulos en español, seleccionando la rueda de configuración en la parte inferior derecha del video (como se observa en las siguientes imágenes). Se recomienda que el/la docente lo vea con anterioridad para manejar el tema y anticipar las observaciones y preguntas que surjan.





En la siguiente tabla, se presenta un breve resumen con las principales ideas del video.

El material inicia con un “experimento mental” similar al dilema del tranvía²³. En este experimento, se nos presenta un escenario donde somos pasajeros de un vehículo autónomo conducido por una IA que se enfrenta a la caída de carga desde el camión que va adelante. El algoritmo, entonces, debe decidir entre chocar de frente el contenido del camión, poniendo un riesgo alto sobre nuestras vidas; chocar una moto a la derecha, poniendo un riesgo muy bajo sobre nuestras vidas, pero un riesgo muy alto sobre el conductor de la moto; y finalmente chocar una camioneta 4x4 a la izquierda, poniendo un riesgo medio sobre nuestras vidas y las de quienes viajan en ese vehículo (**minimización de daños**).

Luego, hace la distinción ética entre la reacción instintiva que puede atribuirse a una persona al volante y una **decisión deliberada**, programada con anterioridad en un vehículo autónomo, en lo que denomina “**discriminación sistemática**”. Finalmente, se complejiza el problema poniendo el caso de que el vehículo deba elegir entre atropellar a un motociclista con casco o a otro que no lo usa. En el primer caso se está aplicando el criterio de minimización de daños, mientras que en el segundo se está decidiendo activamente proteger la vida del conductor más responsable.

Para un computador, unos pocos segundos es suficiente para realizar miles de millones de cálculos, por lo que, en el escenario anterior, una IA bien podría decidir qué hacer,

²³ https://es.wikipedia.org/wiki/Dilema_del_tranvía

considerando las condiciones del accidente y proyectando las consecuencias mucho mejor de lo que cualquier persona podría hacerlo.

—**¿Qué significa que una IA decida en este caso? (Enfatizar y basarse en el contenido de la presentación de diapositivas)**

Como hemos visto, las IA son diseñadas por personas. Más aún, en este diseño **influyen los intereses de las partes** que se pueden ver beneficiadas o perjudicadas por estos algoritmos. En otras palabras, **no es finalmente la IA la que toma decisiones** de forma autónoma, sino quienes crean estos algoritmos, mientras que los afectados somos todos como sociedad. Como hemos visto en actividades anteriores, cada uno de estos grupos es denominado “partes interesadas” (stakeholders, en inglés y en lenguaje profesional).

—**Consideremos en el escenario anterior que la IA fue diseñada para minimizar el daño que sufrimos nosotros, sus pasajeros.** Si, dado este diseño, el vehículo decide chocar al motorista y por esa determinación muere, ¿se trata de un homicidio premeditado? ¿Qué pasaría si la IA pudiese proyectar la muerte del motorista y, aun así, determina que lo más conveniente es chocarlo?

Espera a que las/los estudiantes se expresen sobre la complejidad del dilema, que implica tomar decisiones sobre a quién(es) proteger y a quiénes desatender, con consecuencias fatales.

Es importante en este punto destacar los beneficios que los vehículos autónomos pueden traer a la sociedad. Es sabido que la mayor parte de los accidentes vehiculares son producidos por errores humanos [5]. Es más, si únicamente redujéramos los accidentes producidos por conductores que manejan bajo los efectos del alcohol, tendríamos un impacto sustancial en la disminución de la accidentabilidad. Sin embargo, es esperable que accidentes por fallas mecánicas, de carga o de la vía sigan ocurriendo y es necesario determinar cómo los algoritmos de IA actuarán en estos casos antes de que los vehículos bajo su control se encuentren en las calles.

—Consideremos ahora que la IA se diseña tanto para minimizar el riesgo de los pasajeros como para minimizar el daño total provocado a todos los involucrados en el accidente. ¿Cómo se aplica este criterio a los casos vistos en el video?

En el escenario anterior, la IA debería chocar al vehículo de la izquierda pues produce un riesgo medio tanto a pasajeros como al resto de los involucrados. Sin embargo, el problema es más complejo en el caso donde se nos presenta un motorista con casco a la izquierda y uno sin casco a la derecha. Bajo los principios de diseño anterior, la IA debería chocar al motorista con casco, pues esto disminuye el riesgo de fatalidad con respecto al motorista que no lo usa.

—En tal caso, ¿no se encuentra el sistema castigando a quien sí tomó medidas de seguridad? ¿Por qué sufre las consecuencias el motorista responsable? ¿No se crea entonces una motivación para que los motoristas disminuyan sus medidas de protección personal y sean menos responsables? Consideremos también que si la IA decidiese chocar al motorista sin casco, estaríamos rompiendo el principio de diseño de minimización de daño y prácticamente introduciendo criterios de penalización de conductas irresponsables.

Finalmente, reflexione sobre en voz alta sobre el futuro del mercado de los vehículos autónomos, donde encontraríamos básicamente dos modelos de reacción programada: el primero es un algoritmo de IA diseñado para minimizar el riesgo de su dueño, es decir, de nosotros mismos; el segundo, para minimizar el riesgo de todos los involucrados. Es otras palabras, en una situación determinada, podría decidir ponernos en un alto riesgo si eso implica minimizar el daño total de todos los involucrados.

Incluso si decidiéramos adquirir el segundo vehículo, la evidencia indica que la mayor parte de la población elegiría un automóvil que minimice el daño de sus dueños. Más aún, precisamente porque la población es más proclive a comprar este tipo de autos, los fabricantes estarán motivados a crear algoritmos de IA que minimicen el daño de sus clientes para competir en el mercado y vender más unidades.

Divida la clase en grupos de hasta cuatro personas y entréguales la hoja de ejercicios, que contiene las siguientes preguntas:

- ¿Quiénes son las partes interesadas o stakeholders en la problemática del diseño de algoritmos en vehículos autónomos? ¿Qué intereses distintos pueden tener? (espere respuestas tales como la ciudadanía, el parlamento, las aseguradoras, los fabricantes de vehículos, entre otros)
- ¿Cuál es el problema principal del diseño de estos algoritmos?
- ¿Cuál es su posición con respecto al diseño de estos algoritmos? ¿Qué piensa usted que debería hacer el algoritmo en el primer y segundo escenario?
- En el tercer escenario, ¿qué vehículo compraría usted?
- Considere la siguiente situación: en el vehículo autónomo que ha perdido el control, se encuentran sólo pasajeros de edad avanzada (sobre los 80 años), mientras que a la izquierda y a la derecha se encuentran vehículos con niños a bordo. ¿Sería aceptable que el algoritmo decidiera poner en riesgo a los adultos mayores en vez de a los niños? ¿Por qué? ¿Es aceptable que el algoritmo evalúe la “conveniencia” de salvar a ciertas personas por sobre otras?

Cierre (15 minutos)

Para el cierre de la actividad, considere un espacio para compartir las respuestas. Ofrezca la palabra a dos o tres grupos y comenten. Concluya con que este tipo de problemas éticos o “sociotécnicos” tienen muchas veces más de una interpretación y que los intereses de los involucrados son clave en el diseño e implementación de tecnologías potenciadas por inteligencia artificial.

Alternativa sin conexión a Internet

En caso de no contar con dispositivos o una conexión que permita reproducir el video principal de la actividad, acceda a la transcripción al español desde la carpeta en línea del programa o a través del siguiente enlace:

https://docs.google.com/document/d/1jMBHbgiaT7VPCu5s6XO7n7UnbREf0zwxZL1_g7beU20/edit?usp=sharing

Referencias

[1] DMV California -

<https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-testing-permit-holders>

[2]

<https://www.latercera.com/nacional/noticia/probaran-chile-primer-vehiculo-autonomo-pasajeros/732081>

[3] <https://vimeo.com/401477096>

[4] <https://www.iadb.org/es/paises/piloto-vehiculo-autonomo>

[5] <https://www.conaset.cl/programa/observatorio-datos-estadistica/biblioteca-observatorio/estadisticas-generales>

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana

Caso II: ¿Cómo enseñaremos a la IA a manejar?

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales

Fecha: _____ **Curso:** _____

Objetivos:

- Reflexionar sobre las consecuencias e implicancias éticas de sistemas de inteligencia artificial que se encuentran en fase de desarrollo

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____
4. _____

Instrucciones

En la siguiente actividad, revisaremos algunos temas éticos asociados a la programación de vehículos autónomos (o vehículos sin conductor humano). Luego de escuchar la presentación del/la profesor/a y visto el video de TED-Ed, responde en grupo las siguientes preguntas:

1. ¿Quiénes son las partes interesadas o *stakeholders* en la problemática del diseño de algoritmos en vehículos autónomos? ¿Qué intereses distintos pueden tener?

ACTIVIDAD 8

BÚSQUEDA DEL TESORO EN YOUTUBE

Oportunidades curriculares

Tecnología

- **I°Medio - OA 3:** Evaluar el servicio desarrollado considerando criterios propios, técnicos y valóricos, y proponer mejoras asociadas tanto a los procesos como al producto final.
- **II°Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 22:** Evaluar situaciones de consumo informado y responsable, considerando los derechos del consumidor, los compromisos financieros, el sentido del ahorro y del endeudamiento, entre otros.

Visión general

En esta actividad, las y los estudiantes profundizan los conocimientos adquiridos en las primeras cuatro lecciones (conjuntos de datos, predicción y optimización), a través de un ejercicio en equipo en el que deben identificar y describir los diferentes algoritmos que usa YouTube (despliegue de publicidad, recomendación de videos, clasificador de comentarios y reproducción automática, entre otros).

Objetivos

- Identificar y caracterizar algoritmos de aprendizaje automático que operan en plataformas sociales de uso cotidiano.
- Reforzar los aprendizajes adquiridos en lecciones anteriores.

Agenda (60 minutos)

- Introducción (20 min)
 - Formación de grupos
 - Preguntas frente al computador o dispositivo móvil
- Desarrollo (30 min)
 - Aplicación en hoja de trabajo
- Cierre (10 min)
 - Preguntas de síntesis

Materiales para esta actividad

- Guía docente
- Computadores o dispositivos móviles con acceso a YouTube
- Hoja del estudiante (una por dupla o grupo)

Palabras clave

Algoritmos, optimización, aprendizaje y predicción; YouTube.

GUÍA DOCENTE

La siguiente actividad tiene como propósito que las y los estudiante identifiquen y analicen los algoritmos que operan en YouTube.

En la actividad, “Búsqueda del tesoro en YouTube”, el curso tendrá que buscar con qué tipos de algoritmos interactuamos cuando usamos la plataforma de videos e indagar acerca de los datos que utilizan, las predicciones que buscan establecer y los objetivos que persiguen.

Introducción (20 minutos)²⁴

—¿Recuerdan qué es un algoritmo? ¿Quién puede mencionar algún sistema con el que interactúen en su vida cotidiana?

Intente obtener respuestas centradas en IA, tales como reconocimiento de caras, filtros en tiempo real para redes sociales, traducción simultánea, asistentes para la optimización de la captura de fotografías en teléfonos celulares, recomendaciones de contenidos o productos, predicción de búsquedas, etcétera.

—**En esta clase vamos a enfocarnos en YouTube. ¿Se han preguntado cuántos algoritmos de inteligencia artificial diferentes conviven en la plataforma? Eso es lo que vamos a ver ahora.** Divídanse en grupos de 2 a 3 personas y completen la hoja de trabajo. Vamos a ver las formas en que YouTube usa la IA. Para cada característica distinta, 1) describan los datos que se utilizan para crear el algoritmo, 2) lo que el algoritmo está tratando de aprender o predecir y 3) la “opinión” general del algoritmo (para qué se está optimizando).

Traslade al curso a la sala de computación o sala Enlaces, pídale entrar a YouTube y entregue las hojas de trabajo.

—**Escriban las primeras tres letras de su raza de perro favorita en el buscador. ¿Qué pasa? (Ejemplo: lab [labrador]; Fox [Terrier]; San [Bernardo], etcétera).**

²⁴ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

Los/las estudiantes deberían indicar que YouTube completa el resto de su búsqueda.

—Así es, eso se llama **búsqueda sugerida**. ¿Qué notaron sobre las sugerencias que les dio YouTube?

Las/los estudiantes comparten sus observaciones.

—Intenten escribir algunas letras más. ¿Qué notan?

Las/los estudiantes comparten sus observaciones.

—¿Qué creen que el algoritmo está tratando de aprender en este caso?

Las/los estudiantes deberían responder que está tratando de adivinar lo que se busca.

—Sí. Está tratando de entender lo que están buscando antes de que terminen de escribirlo. ¿Qué datos creen que está utilizando el algoritmo para hacer eso?

Los estudiantes deberían decir que sus búsquedas anteriores, así como las de otras personas.

—Sí, correcto. YouTube está utilizando datos de sus búsquedas anteriores, así como de otras personas como ustedes. ¿Para qué creen que se están optimizando?

Las/los estudiantes deberían decir que para que vean más videos y así haya más ingresos por publicidad. Si esto no sucede, pregunte si hay algo más que aparezca aparte de los videos.

Desarrollo (30 minutos)

A partir de esta parte, el curso estará poniendo en práctica los contenidos aprendidos en las cuatro primeras actividades del programa, que son las que entregan las bases teóricas para entender los sistemas de inteligencia artificial.

—Ahora trabajen con sus grupos para investigar qué otros algoritmos utiliza la plataforma. Piensen en lo siguiente:

1. Lo que el algoritmo está tratando de aprender
2. El conjunto de datos está utilizando para aprender eso y
3. Para qué podría estar optimizándose.

—Hagan su mejor esfuerzo y levanten la mano si tienen alguna pregunta.

Las/los estudiantes trabajan por alrededor de 30 minutos con su hoja de trabajo. Una vez pasado el tiempo, replique la siguiente matriz en la pizarra (sólo el encabezado) y complete con los aportes del curso. Apóyese en el siguiente material:

Nombre	Datos	¿Qué está tratando de “aprender”?	¿Para qué se está optimizando?
Búsqueda sugerida (autocompletar)	Historial de búsqueda anterior realizado por ustedes y otras personas	Cuáles son las búsquedas más comunes, para predecir lo que ustedes desean ver	Más videos vistos
Recomendaciones	Historial de visualización del usuario/a, historial de visualización de otros/as usuarios/as	Qué videos les gustaría ver según lo que ya hayan visto	Clics en videos atractivos para el/la usuario/a (y, por lo tanto, más anuncios vistos)
Sección de comentarios	Cantidad de me gusta (<i>likes</i>) que tiene cada comentario, popularidad de quienes comentaron	Qué comentarios son los más relevantes	Comentarios que mantendrán al/la usuario/a en la plataforma por más tiempo
Resultados de búsqueda	Búsquedas anteriores por el/la usuario/a y personas con perfiles similares (en qué videos se hace clic, en última instancia)	Qué videos esperan ver los/as usuarios/as cuando realizan una búsqueda	Videos en los que se hace clic (y, por lo tanto, anuncios...)

Anuncios	Historial de visualización del/la usuario/a	Qué productos y/o servicios pueden atraer a un/a usuario/a	Interacciones con los anuncios
Reproducción automática	Historial de visualización del/la usuario/a, historial de visualización de otros/as usuarios/as	Qué video querrá ver un/a usuario/a a continuación	Videos vistos; mantener a los/las usuarios/as en el sitio web

Cierre (10 minutos)

—¿Cómo supieron qué datos se usaron para influir en el algoritmo? ¿Cómo averiguaron para qué estaba optimizado?

Recuerde que todos los algoritmos persiguen un fin, que a su vez obedece al interés de una persona, grupo o institución. En el caso del completo de la actividad 2, algunos pueden estar interesados en su sabor o su valor nutricional (esto es a lo que llamamos la “opinión”). Las redes sociales y las plataformas de contenido gratuito, por su parte, suelen recurrir a mecanismos de monetización por publicidad, que afectan la optimización y el fin último de la plataforma.

—¿Esta tarea fue fácil o difícil? ¿Hay algo que les haya sorprendido?

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Búsqueda del Tesoro de YouTube**

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales

Fecha: _____ **Curso:** _____

Objetivos:

- Identificar y caracterizar algoritmos de aprendizaje automático que operan en plataformas de uso cotidiano.
- Reforzar los aprendizajes adquiridos en lecciones anteriores

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____

Instrucciones

Escuchen atentamente al/la profesor/a y trabajen en equipo para descubrir las formas en que YouTube usa la IA. Para cada mecanismo, 1) describan los datos que se utilizan para crear el algoritmo, 2) lo que el algoritmo está tratando de aprender o predecir y 3) la “opinión” general del algoritmo.

Nombre	Datos	¿Qué está tratando de "aprender"?	¿Para qué se está optimizando?

ACTIVIDAD 9

REDISEÑO DE YOUTUBE

Oportunidades curriculares

Tecnología

- **I°Medio - OA 2:** Desarrollar un servicio que implique la utilización de recursos digitales u otros medios, considerando aspectos éticos, sus potenciales impactos y normas de cuidado y seguridad.
- **I°Medio - OA 3:** Evaluar el servicio desarrollado considerando criterios propios, técnicos y valóricos, y proponer mejoras asociadas tanto a los procesos como al producto final.
- **II°Medio - OA 3:** Evaluar las propuestas de soluciones que apunten a resolver necesidades de reducción de efectos perjudiciales relacionados con el uso de recursos energéticos y materiales, considerando aspectos o dilemas éticos, legales, económicos, ambientales y sociales.
- **II°Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 23:** Explicar que los problemas de una sociedad generan distintas respuestas políticas, ejemplificando mediante las posturas que surgieron frente a la “cuestión social” (por ejemplo, liberalismo, socialismo, anarquismo, comunismo y socialcristianismo) y de otras situaciones conflictivas de la actualidad.
- **II°Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel

nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.

Visión general

Las y los estudiantes continúan el trabajo realizado en la actividad N°8, donde identificaron y analizaron algunos de los algoritmos que usa YouTube para orientar el consumo de contenidos en la plataforma. En grupos, construyen una matriz ética sobre el sistema de recomendaciones de YouTube, identificando las partes interesadas y estableciendo un nuevo objetivo para el algoritmo. Finalmente, crean un prototipo en papel de su nueva versión de la plataforma, incorporando características que respondan a los intereses de las partes.

Objetivos

- Identificar las partes interesadas en el sistema de recomendaciones de una plataforma tecnológica real y los beneficios que cada una busca en sus resultados
- Reformular el algoritmo de recomendaciones en base a criterios éticos

Agenda (90 minutos)

- Introducción (15 min)
 - Repaso de los aprendizajes de la actividad anterior
 - Explicación de la actividad
- Desarrollo (60 min)
 - Resolución de la guía del estudiante (pasos 1 al 7)
- Cierre (15 min)
 - Puesta en común de prototipos de rediseño de YouTube

Materiales para esta actividad

- Guía docente
- Computadores o dispositivos móviles con acceso a YouTube
- Hoja del estudiante (una por dupla o grupo)
- Optativo: material para prototipar, como cartulinas, papel Kraft, cinta adhesiva y plumones

Palabras clave

Matriz ética, partes interesadas, optimización, prototipo, YouTube.

GUÍA DOCENTE

En esta actividad, el curso aplicará los contenidos aprendidos a lo largo del programa, por lo que su fuerte está en el trabajo grupal y en la orientación que pueda entregar el/la docente durante el proceso.

A partir de los resultados de la actividad anterior (N°8), las y los estudiantes realizarán un levantamiento de las partes que pueden estar interesadas o verse afectadas por los resultados del algoritmo, en base a lo cual rediseñarán la experiencia que ofrece la plataforma.

Para el prototipado, la guía del estudiante cuenta con un campo en blanco, en el que tendrán que dibujar un boceto. Alternativamente, se puede usar cartulinas, papel Kraft y plumones para crear modelos que se puedan exponer al curso, además de cinta adhesiva para pegar los trabajos.

Si no utiliza con frecuencia la plataforma de YouTube, recomendamos crear una cuenta y familiarizarse antes con su funcionamiento. Guíese con la siguiente tabla:

- ✓ Realizar búsquedas de contenidos
- ✓ Observar cómo se despliegan las recomendaciones de otros videos
- ✓ Observar en qué momentos y dónde se despliega la publicidad
- ✓ Conocer el mecanismo de reproducción automática (*autoplay*) al terminar un video
- ✓ Analizar cómo se van refinando la publicidad y las recomendaciones a partir de la visualización de videos
- ✓ Conocer los sistemas de comentarios y suscripción a canales

Al igual que en la actividad anterior, si no cuenta con computadores, puede trabajar con dispositivos móviles (celular o *tablet*) conectados a Internet. En cualquier caso, intente que el tipo de equipo a utilizar sea el mismo que en la clase anterior, ya que la experiencia es ligeramente diferente en ambos ambientes.

Introducción (15 minutos)²⁵

Ahora continuaremos el trabajo realizado en la actividad N°8, donde las y los estudiantes completaron una matriz en la que identificaron y caracterizaron algunos de los algoritmos que usa YouTube para orientar el consumo de contenidos. Tómese un tiempo para repasar las respuestas que se compartieron o grafique con la matriz de ejemplo que venía incluida en la guía docente.

Recuerde a la clase que todo algoritmo conlleva una “opinión” o, dicho en otras palabras, se encuentra optimizado para intereses determinados, sean éstos de personas, agrupaciones, empresas u otras organizaciones. Recuerde también que las redes sociales y las plataformas “gratuitas” para el usuario suelen recurrir a mecanismos de monetización por publicidad, por lo que se encuentran optimizadas para retenernos la mayor cantidad de tiempo. Es por ello que se dice que el “producto” que se vende es nuestro tiempo de atención como consumidores (exposición a la publicidad), nuestros datos —aquéllos que permiten identificarnos como potenciales consumidores de determinados productos— y las visitas o compras que podamos hacer en los sitios de los anunciantes.

—Bien, este repaso nos va a servir para retomar lo que hicimos la clase anterior. En esta ocasión, nuestro desafío va a ser reformular el sistema de recomendaciones de YouTube (o el “algoritmo de recomendación”, en el sentido amplio). ¿Recuerdan lo que es una matriz ética?

Espere respuestas relacionadas con la matriz de intereses o valores que diferentes personas, agrupaciones o instituciones sobre los resultados de un algoritmo. Puede reforzar con el ejemplo del completo de la actividad N°2, donde los apoderados se preocupaban por el valor nutricional y el sabor del producto; los profesores y médicos por el contenido alimenticio; las y los estudiantes por el sabor y las empresas por rebajar costos y aumentar las ventas (puede que hayan llegado a una matriz diferente).

—Vamos a aplicar todo lo aprendido sobre matrices éticas al algoritmo de recomendación de YouTube. Para eso, les voy a pasar una hoja y vamos a trabajar en grupo. Van a encontrar siete preguntas: entre la uno y la seis, tienen que definir las partes interesadas, construir una matriz ética, identificar un objetivo para su algoritmo, determinar los datos que usará para operar y las características que deberá tener la plataforma para cumplir el propósito establecido. En la pregunta siete, tendrán que crear un prototipo o maqueta gráfica de su nueva versión de YouTube.

²⁵ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

En caso de utilizar materiales como cartulinas y plumones, recomendamos usar este momento para prepararlos. Note que la pregunta siete solicita el prototipado de al menos tres funciones de la plataforma: el sistema de recomendación (videos sugeridos en el inicio o al visualizar otro contenido), la sección de comentarios y los resultados de búsqueda. Pueden estar contenidas en un mismo boceto o por separado. El trabajo puede incluir, además, el funcionamiento de las búsquedas sugeridas, la reproducción automática y las configuraciones de la publicidad.

Desarrollo (60 minutos)

Forme parejas o grupos de tres personas (puede reunir los mismos equipos de la actividad anterior) y entregue las hojas de trabajo. Oriente el proceso de llenado y evite entregar ejemplos demasiado concretos sobre la plataforma, que puedan predisponer las respuestas.

En el ejercicio de la matriz ética (tres primeros pasos), espere respuestas tales como que los padres se interesan por la cultura y el aprendizaje de hábitos y buenas costumbres, mientras que las empresas por vender productos y servicios y las y los jóvenes por entretenerse y estudiar. Otras partes interesadas pueden ser los/las docentes de asignatura, los/las profesores/as jefe, teatros y centros culturales, instituciones benéficas, el Ministerio de Educación, los “youtubers” o generadores de contenido y el propio YouTube.

En el paso 4, espere respuestas concretas, en las que se note que el algoritmo responde a los objetivos de una o más partes interesadas priorizadas.

Para el paso 5, dependiendo del prototipo que estén elaborando los grupos, puede que necesiten fuentes de datos especiales, tales como listas de videos bien rankeados por docentes y otros expertos, reputación de los generadores de contenido, etcétera, que no necesariamente use hoy la plataforma.

Deténgase un poco más el paso 7 para asegurarse de que haya coherencia entre lo propuesto en las preguntas precedentes y el prototipo que se está desarrollando. Si está

trabajando con prototipos en hojas de papel, se recomienda tener listos más folios en blanco, en caso de que los grupos lo necesiten.

Cierre (15 minutos)

Para finalizar, pida a los grupos que compartan su trabajo con el curso, explicando brevemente cómo funciona y por qué lo hicieron de esa manera. La puesta en común les ayudará a ampliar su mirada y corregir errores en los que puedan haber caído. Puede pegar los trabajos en la pared de la sala de clases por unos días para favorecer la consolidación de los aprendizajes.

Antes de concluir la clase, repase los conceptos básicos —tales como conjuntos de datos, partes interesadas, predicción y aprendizaje; objetivos, opinión y optimización— y las principales conclusiones de los grupos que presentaron.

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Rediseño de YouTube**

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales

Fecha: _____ **Curso:** _____

Objetivos:

- Identificar las partes interesadas en el sistema de recomendaciones de una plataforma tecnológica real y los beneficios que cada una busca en sus resultados
- Reformular el algoritmo de recomendaciones en base a criterios éticos

Nombres de las/los integrantes del grupo

1. _____
2. _____
3. _____

Instrucciones

El desafío de hoy es rediseñar el sistema o algoritmo de recomendación de YouTube, siguiendo los pasos que se presentan a continuación. Empezaremos identificando las partes interesadas y construyendo una matriz ética, para más tarde realizar nuestro propio boceto de la plataforma.

Paso 1: Identificar a las partes interesadas. Hagan una lluvia de ideas de al menos 10 partes interesadas que se preocupan por el algoritmo de YouTube.

1. _____
2. _____
3. _____

4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

Paso 2: Identifiquen algunos intereses o valores que sus partes interesadas podrían querer incluir en el algoritmo. Para cada una de las partes interesadas mencionadas anteriormente, identifiquen al menos dos formas en que se preocupan por los resultados del sistema de recomendaciones de YouTube.

Parte interesada	Valor 1	Valor 2

Paso 3: completen la siguiente matriz ética a partir de sus respuestas anteriores:

YouTube (la empresa)				

Paso 4: Identifiquen el objetivo de su algoritmo, a partir de la matriz ética anterior:

¿Por qué decidieron este objetivo? Explíqueno brevemente:

Paso 5: Describan qué datos necesitan para enseñar a su algoritmo las acciones correctas en base al objetivo definido anteriormente:

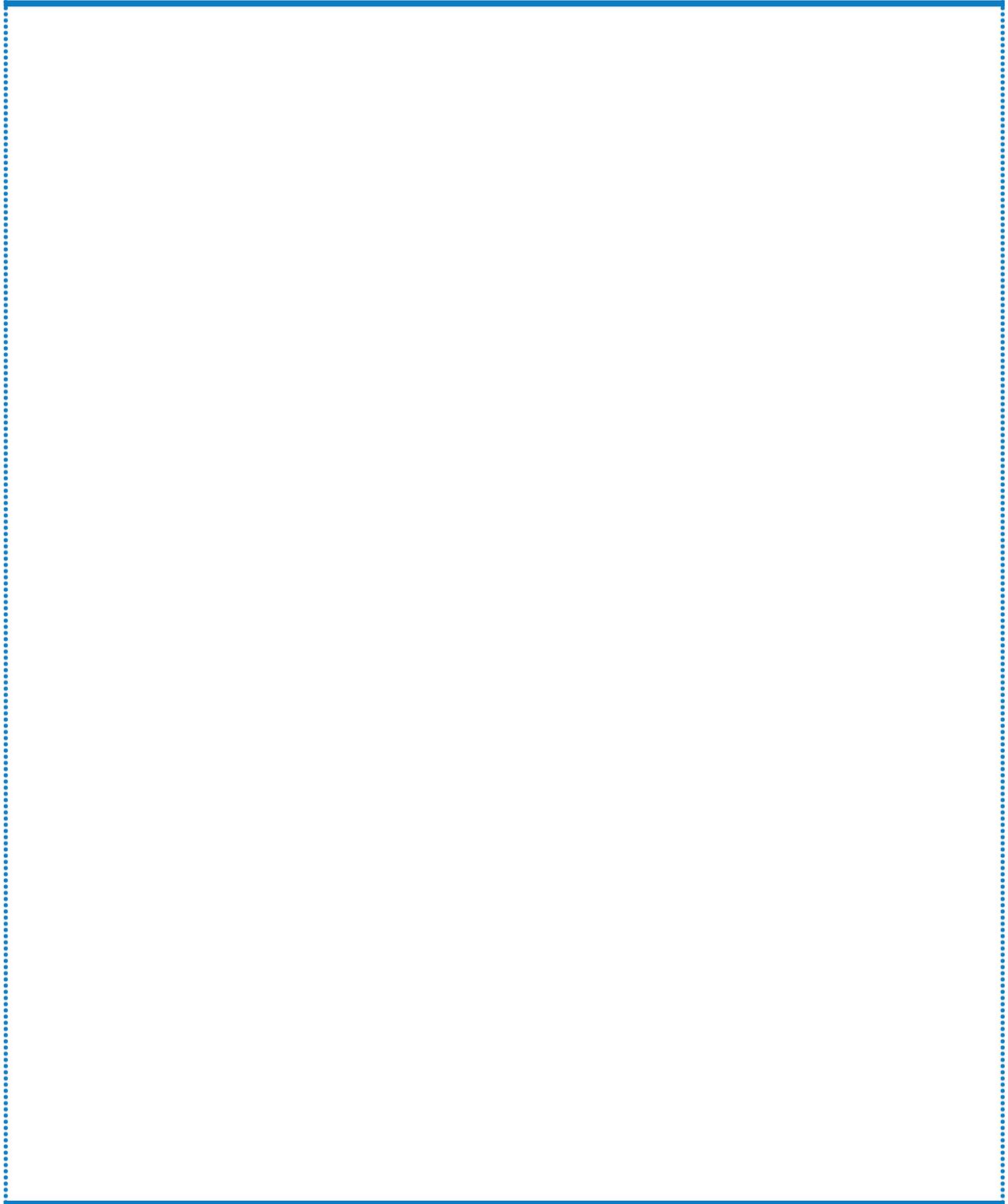
Paso 6: ¿Qué características tendrá su plataforma de YouTube rediseñada para lograr este objetivo?

Paso 7: Usando los materiales que tengan a mano, diseñen un boceto (dibujo) de la experiencia de usuario. Su modelo debe mostrar un prototipo de:

- **Sistema de recomendación**
- **Sección de comentarios**
- **Resultados de la búsqueda**

Opcionalmente, pueden incluir bocetos de:

- Búsquedas sugeridas (autocompletar)
- Reproducción automática (*autoplay*)
- Configuraciones de publicidad



ACTIVIDAD 10

SEMINARIO SOCRÁTICO DE YOUTUBE

Oportunidades curriculares

Tecnología

- **I°Medio - OA 4:** Comunicar el diseño, la planificación u otros procesos del desarrollo de un servicio, utilizando herramientas TIC, considerando diferentes tipos de objetivos y audiencias y teniendo en cuenta aspectos éticos.
- **II°Medio - OA 3:** Evaluar las propuestas de soluciones que apunten a resolver necesidades de reducción de efectos perjudiciales relacionados con el uso de recursos energéticos y materiales, considerando aspectos o dilemas éticos, legales, económicos, ambientales y sociales.
- **II°Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 23:** Explicar que los problemas de una sociedad generan distintas respuestas políticas, ejemplificando mediante las posturas que surgieron frente a la “cuestión social” (por ejemplo, liberalismo, socialismo, anarquismo, comunismo y socialcristianismo) y de otras situaciones conflictivas de la actualidad.
- **II°Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.

Lengua y Literatura

- **I° y II° Medio - OA 22:** Dialogar constructivamente para debatir o explorar ideas.

Visión general

Para concluir el ciclo de actividades basadas en el algoritmo de YouTube, los y las estudiantes leen el artículo titulado “YouTube evalúa traspasar los vídeos infantiles a su aplicación para niños”, basado en una publicación de The Wall Street Journal. Con esa información, participan en un seminario socrático para dialogar sobre los aspectos éticos en conflicto, el rol de las partes interesadas y la propuesta de la empresa tecnológica.

Objetivos

- Reflexionar sobre los riesgos asociados a los algoritmos de recomendación basados en inteligencia artificial
- Reconocer el rol que juegan los conflictos entre empresas tecnológicas, ciudadanía y grupos de interés en la discusión de estándares éticos y en el desarrollo de los servicios

Agenda (60 a 80 minutos)

- Introducción (10 min)
 - Explicación de la actividad
 - Preparación del lugar
- Desarrollo (40 a 60 min)
 - Lectura del texto
 - Desarrollo del seminario socrático
- Cierre (10 min)
 - Síntesis del/la docente

Materiales para esta actividad

- Guía docente
- Copias del texto a trabajar
- Hoja de preguntas para estudiantes (una por persona)

Palabras clave

Seminario socrático, partes interesadas, reproducción automática (*autoplay*), contenido inapropiado, optimización, YouTube, YouTube Kids.

GUÍA DOCENTE

La clase reflexionará en torno al significado y las implicancias de una serie de cambios evaluados por YouTube en 2019, según una publicación del diario estadounidense The Wall Street Journal, a raíz de una polémica asociada a la recomendación de videos inapropiados para niños, tras la visualización de contenidos de corte infantil. El curso leerá una adaptación del texto informativo y discutirá en torno a una pauta de preguntas orientadas por el/la docente. El caso permitirá evaluar la capacidad de síntesis y comprensión de textos de no ficción, poner en práctica conceptos técnicos como optimización y partes interesadas y analizar los aspectos éticos de una situación real.

Se recomienda organizar la sala en un círculo —o como el/la docente estime más práctico— y leer el texto en conjunto, buscando que la mayoría del curso participe, para lograr un mayor involucramiento en la actividad. La pauta de preguntas es referencial, por lo que el/la docente puede modificarla o agregar nuevas, en función de las características y la experiencia del grupo curso.

Introducción (10 minutos)²⁶

Esta actividad cierra la tríada de clases que analizan el funcionamiento y las implicancias del algoritmo de recomendación de YouTube para diferentes grupos de interés, aplicado a una polémica real ocurrida en 2019.

—Hoy vamos a trabajar como grupo curso, por lo que necesitamos que participen todos/as. ¿Alguien sabe lo que es el control parental?

Si no hay respuesta, insista **—¿Sabían que existen servicios para bloquear contenidos de Internet y aplicaciones en computadores y dispositivos móviles, que permiten crear cuentas de usuario “seguras” para niños (libres de contenido inapropiado)²⁷?**

Espere escuchar experiencias personales de cuando las/los estudiantes eran menores o casos que conozcan por familiares, vecinos o amigos.

²⁶ Actividad basada en el material *Un programa de Ética de la Inteligencia Artificial para estudiantes de 1° y 2° medio* (Blakeley H. Payne, MIT Media Lab). Traducido por Educarchile y Plan Nacional de Lenguajes Digitales del Ministerio de Educación de Chile.

²⁷ Use este enlace si desea profundizar:

<https://www.latercera.com/noticia/guia-para-configurar-control-parental-en-celulares-tabletas-y-computadores>

—¿Por qué los adultos algunas veces bloquean contenidos para que no los puedan ver los niños y niñas?

Espere respuestas relacionadas con el contenido y una edad no adecuada para verlo.

Presente la actividad a los/las estudiantes, comunicándoles que tendrán que analizar como clase un caso en el que se vio involucrado el sistema de recomendación de YouTube. Recuerde que el contenido de la plataforma es gratis y, en su mayoría, abierto a todo público, por lo que existe el riesgo de que niños y niñas accedan a contenido inapropiado para su edad. Esto, a pesar de la existencia de la plataforma YouTube Kids, que presenta únicamente contenidos de corte infantil.

—Ahora vamos a organizarnos en círculo para leer y comentar juntos/as un artículo de prensa.

Use el texto de la agencia EFE **“YouTube evalúa traspasar los videos infantiles a su aplicación para niños”**, una adaptación del artículo publicado por el diario estadounidense The Wall Street Journal. Imprímalo o proyéctelo para compartir con la clase.

<https://www.efe.com/efe/usa/sociedad/youtube-evalua-traspasar-los-videos-infantiles-a-su-aplicacion-para-ninos/50000101-4004615>

Desarrollo (40 a 60 minutos)

Una vez organizada la sala, asigne turnos para leer el texto entregado (puede ser un párrafo o un punto seguido por persona). Si no cuenta con proyección, entregue una hoja a cada estudiante.

Para corroborar si comprendieron, puede pedir a alguien al azar que sintetice la noticia. Una vez finalizado este ejercicio, dé inicio al seminario socrático.

Preguntas para el seminario socrático:

En esta instancia, se espera que participe la mayoría de las y los estudiantes, pudiendo intervenir con opiniones, comentarios o análisis diversos. La idea es que la/el docente incentive el diálogo, poniendo sobre la mesa algunas perspectivas que las/los estudiantes no hayan podido visualizar. Use la siguiente pauta de preguntas como guía:

1. ¿Cuál es el objetivo de este rediseño? ¿Para qué se está optimizando esta plataforma?
2. ¿Alguien puede nombrar las partes interesadas abordadas en este artículo?
 - a. ¿Cuál es la parte interesada más importante?
 - b. ¿Qué parte interesada está haciendo el mayor cambio o tiene más poder?
3. ¿Creen que el contenido infantil debería estar sólo en YouTube Kids? ¿Por qué? ¿Por qué no?
4. ¿Alguna vez han visto un contenido inapropiado en YouTube? ¿Qué hicieron?
5. ¿Usarían la aplicación YouTube Kids con hermanos o primos menores? ¿Creen que les gustaría a ellos/as?
6. ¿Creen que esta propuesta sería popular?
7. ¿Cómo creen que se sienten los anunciantes al respecto?
8. ¿Creen que YouTube bajaría sus ganancias? ¿Está bien si pierden ganancias?
9. ¿Qué sucede si hay más o menos contenido inapropiado?
10. ¿Por qué existe la reproducción automática (*autoplay*)? ¿Quién se beneficia de ella? ¿Debería existir?

Cierre (10 minutos)

Para finalizar la actividad, se sugiere hacer énfasis en aquellos aspectos en los que hubo mayor desacuerdo, destacando la capacidad que debemos tener, por un lado, de respetar los argumentos de otro/as y, por el otro, de ser conscientes de los riesgos asociados a los sistemas de recomendación basados en algoritmos automatizados de inteligencia artificial.

Recuerde que YouTube y las redes sociales cuentan con muchos contenidos que, sin ser abiertamente violentos o sensibles, pueden representar algún tipo de riesgo, como sucede con las campañas antivacunas y otras acciones de desinformación, las teorías conspirativas y los contenidos que requieren una madurez mínima para interpretarse correctamente. A modo de conclusión, abra la siguiente pregunta: **¿es correcto que plataformas como YouTube y las redes sociales nos sugieran contenidos sólo en base al número de reproducciones e interacciones de otros usuarios?**

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana **Seminario socrático de YouTube**

Asignatura: Tecnología/ Historia, Geografía y ciencias Sociales/ Lengua y Literatura

Fecha: _____ **Curso:** _____

Objetivos:

- Reflexionar sobre los riesgos asociados a los algoritmos de recomendación basados en inteligencia artificial
- Reconocer el rol que juegan los conflictos entre empresas tecnológicas, ciudadanía y grupos de interés en la discusión de estándares éticos y en el desarrollo de los servicios

Nombre:

Instrucciones

Lee el texto “YouTube evalúa traspasar los vídeos infantiles a su aplicación para niños” en conjunto con el curso (esperen a que el profesor o profesora lo comparta) y entrega tu visión sobre las preguntas que están más abajo. Intenta aplicar lo que has aprendido hasta el momento, como las partes que influyen en los objetivos de un algoritmo, los modelos de financiamiento basados en publicidad y tu experiencia usando YouTube.

Las preguntas para el seminario socrático son:

1. ¿Cuál es el objetivo de este rediseño? ¿Para qué se está optimizando esta plataforma?
2. ¿Puedes nombrar las partes interesadas abordadas en este artículo?

- a. ¿Cuál es la parte interesada más importante?
- b. ¿Qué parte interesada está haciendo el mayor cambio o tiene más poder?
3. ¿Crees que el contenido infantil debería estar sólo en YouTube Kids? ¿Por qué? ¿Por qué no?
4. ¿Alguna vez has visto un contenido inapropiado en YouTube? ¿Qué hiciste?
5. ¿Usarías la aplicación YouTube Kids con hermanos/as o primos/as menores? ¿Crees que les gustaría a ellos/as?
6. ¿Crees que esta propuesta sería popular?
7. ¿Cómo crees que se sienten los anunciantes al respecto?
8. ¿Crees que YouTube bajaría sus ganancias? ¿Está bien si pierden ganancias?
9. ¿Qué sucede si hay más o menos contenido inapropiado?
10. ¿Por qué existe la reproducción automática (*autoplay*)? ¿Quién se beneficia de ella? ¿Debería existir?

ACTIVIDAD 11

CASO III: ¿CÓMO NOS PREPARAMOS PARA LA SIGUIENTE PANDEMIA?

Oportunidades curriculares

Tecnología

- **I°Medio - OA 2:** Desarrollar un servicio que implique la utilización de recursos digitales u otros medios, considerando aspectos éticos, sus potenciales impactos y normas de cuidado y seguridad.
- **I°Medio - OA 4:** Comunicar el diseño, la planificación u otros procesos del desarrollo de un servicio, utilizando herramientas TIC, considerando diferentes tipos de objetivos y audiencias y teniendo en cuenta aspectos éticos.
- **I°Medio - OA 5:** Analizar las formas en que los productos tecnológicos y los entornos evolucionan, caracterizando los diversos factores que influyen en ese cambio.
- **I°Medio - OA 6:** Inferir, basándose en la evolución de los productos tecnológicos y los entornos, los efectos positivos o negativos que estos han tenido en la sociedad.
- **II°Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.
- **II°Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 22:** Evaluar situaciones de consumo informado y responsable, considerando los derechos del consumidor, los compromisos financieros, el sentido del ahorro y del endeudamiento, entre otros.

- **I°Medio - OA 25:** Analizar el impacto del proceso de industrialización en el medioambiente y su proyección en el presente, y relacionarlo con el debate actual en torno a la necesidad de lograr un desarrollo sostenible.
- **II°Medio - OA 22:** Analizar el concepto de Derechos Humanos, considerando características como su universalidad, indivisibilidad, interdependencia e imprescriptibilidad y la institucionalidad creada para resguardarlos a nivel nacional e internacional, reconociendo, en este contexto, los derechos vinculados a los grupos de especial protección.
- **II°Medio - OA 23:** Explicar los elementos que constituyen un Estado de derecho, como la Constitución, la división de los poderes del Estado, el respeto a los derechos humanos, entre otros; analizar su importancia para la gobernabilidad, la representación, la participación ciudadana y la convivencia pacífica; y debatir sobre su perfeccionamiento, defensa y riesgos a los que está expuesto.

Lengua y Literatura

- **I° y II°Medio - OA 10:** Analizar y evaluar textos de los medios de comunicación, como noticias, reportajes, cartas al director, propaganda o crónicas.
- **I° y II°Medio - OA 21:** Dialogar constructivamente para debatir o explorar ideas.

Visión general

Las y los estudiantes debaten en torno al dilema entre la protección de datos personales y su uso para fines investigativos y de salud pública, a partir de la contraposición de las estrategias de Corea del Sur y Chile en el abordaje del primer año de la crisis sanitaria del COVID-19. Se entregan nociones para entender el rol que juega la ciencia de datos —una aplicación de la inteligencia artificial que permite encontrar patrones útiles en datos masivos— en la toma de decisiones de relevancia pública y el dilema que genera con respecto a la privacidad de la información personal.

Objetivos

- Comprender el dilema entre la protección de datos y los beneficios del acceso a fuentes de información que permiten la toma de decisiones de relevancia pública
- Discutir sobre los riesgos asociados a la vulneración de la privacidad

Agenda (90 minutos)

- Introducción (15 min)
 - Conceptos: *Data Science* y *Big Data*
- Desarrollo (60 min)
 - Ciencia de datos y COVID-19
 - Problemas de privacidad
 - Debate
- Cierre (15 min)
 - Síntesis del debate

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Hoja del estudiante

Palabras clave

Ciencia de datos (*Data Science*), datos masivos (*Big Data*), COVID-19, privacidad de datos, uso de datos; debate.

Acceso a presentación de diapositivas

<https://docs.google.com/presentation/d/1I9hpCMD8qHd26sGWWSk9AchuQunr8E7X3p9t0Fa2o/edit?usp=sharing>

GUÍA DOCENTE

Esta actividad busca que las y los estudiantes reconozcan la ciencia de datos como una aplicación de la inteligencia artificial que permite tomar decisiones relevantes de política pública, así como las implicancias éticas que supone el uso de información privada. Por su amplio alcance y variedad de distinciones, se trabajará esta guía de forma expositiva, agregando ejemplos y orientando la lectura por temas. Si lo considera necesario, puede sintetizar algunos pasajes y/o resaltar los casos que le parezcan más adecuados para su grupo curso, cuidando mantener los aspectos técnicos que resulten esenciales para entender los fenómenos expuestos. Apóyese en la presentación de diapositivas incluida y considere leer las fuentes citadas.

La actividad continúa con un debate, para el cual se dividirá el curso en dos posiciones: a favor y en contra de la creación de una base de datos para la trazabilidad de casos COVID-19 a partir del análisis de datos de movilidad y compra en comercios. La guía incluye una pauta referencial de preguntas para orientar la discusión.

Introducción (15 minutos)

Utilice el texto a continuación como referencia para explicar al curso los conceptos de *Data Science* y *Big Data*, que se plantean de forma sintética en la presentación de diapositivas. Vaya explicando y reflexionando en voz alta con la clase.

La **ciencia de datos** o DS, por sus siglas en inglés (*Data Science*), es una aplicación de la inteligencia artificial. DS es un campo de estudio que busca patrones en grandes conjuntos de datos (*Big Data*) que pueden ser de utilidad para alguna tarea en particular. Las aplicaciones de DS son múltiples y variadas, aunque su elemento central son siempre los datos y cómo se puede encontrar valor en ellos.

Un ejemplo de una aplicación de DS lo presenta la farmacéutica Bayer, la cual ha reportado el uso exitoso de DS en la agricultura o lo que ellos llaman la **“agricultura digital”** (*digital farming*) [1]. Este enfoque se basa en el uso exhaustivo de distintos sensores para el control de una diversidad de variables (humedad de la tierra, temperatura, ventilación, etc.),

además de otros tipos de datos (GPS, imágenes digitales, entre otros). Bayer indica que “Las herramientas de inteligencia artificial analizan rápidamente estos datos, creando reportes para que los agricultores puedan tomar decisiones críticas, a tiempo y en terreno”.

La ciencia de datos también ha demostrado utilidad en su aplicación en la **prevención del crimen**. Por ejemplo, el Instituto Nacional de Justicia de Estados Unidos reporta su uso en distintas tareas de predicción: Predicción del lugar del Delito, Predicción de Potenciales Delinquentes, entre otros [2]. El mismo instituto liberó en 2017 un concurso que permitía a distintos investigadores acceder a datos sobre miles de reportes policiales que les permitieran presentar sus propios algoritmos predictivos [3]. De manera similar, pero en otro ámbito, Netflix realizó un concurso entre los años 2007 y 2009 llamado Netflix Prize, en los que prometía un millón de dólares al equipo que pudiera mejorar sus **sistemas de recomendación**. Para ello, presentaba a los grupos de investigadores participantes un conjunto de datos reales de casi medio millón de usuarios de Netflix y las películas que vieron.

Como hemos mencionado, DS es una aplicación de los algoritmos de inteligencia artificial con algunas particularidades especiales. La más importante es que un algoritmo de IA utilizado para DS recibe como entrada un conjunto de datos que suele ser muy grande. Normalmente, el conjunto de datos es tan grande que su simple almacenamiento suele ser un problema, lo que requiere de técnicas especiales para su gestión y procesamiento. Este conjunto de técnicas para gestionar grandes volúmenes de datos asegurando velocidad de acceso y veracidad de la información es lo que formalmente se conoce como **Big Data**, aunque muchas veces el término se ocupa para denominar también el mismo análisis de grandes volúmenes de datos.

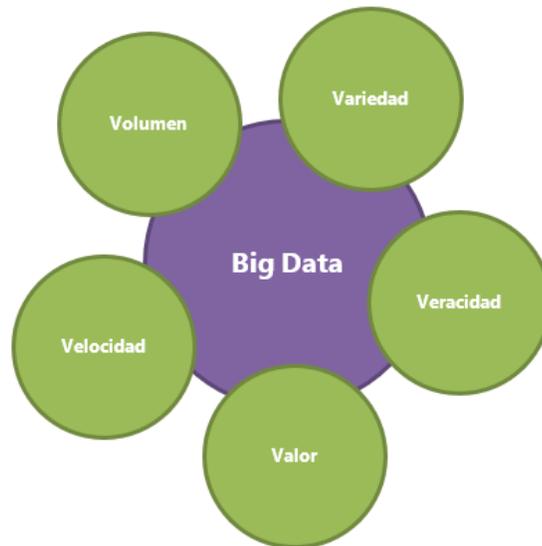


Diagrama: las 5 "V" del Big Data

Desarrollo (60 minutos)

En esta parte de la actividad, usaremos el contraste como herramienta para analizar el dilema ético que implica el trabajo con datos, lo que permitirá preparar puntos de vista antagónicos para el debate. Nos vamos a enfocar en las estrategias seguidas por Corea del Sur y Chile durante el primer año de la crisis sanitaria del COVID-19, evaluando sus implicancias en términos de los beneficios que reporta en materia de salud pública y los posibles riesgos ante la vulneración de la privacidad de los datos.

Ciencia de datos y COVID-19

Durante la pandemia del COVID-19, la recolección y análisis de datos han sido una de nuestras principales herramientas para encontrar mecanismos que nos ayuden a caracterizar y prevenir focos de contagio. Por ejemplo, en nuestro país, la Universidad del Desarrollo en conjunto con la compañía Telefónica, analizaron los datos de movilidad en todas las comunas del país [4] para determinar si las medidas de cuarentena habían tenido

éxito disminuyendo la circulación de las personas. Para ello, utilizaron como conjunto de datos la información de movilidad de teléfonos móviles, los cuales, dependiendo de su ubicación, seleccionan la antena más cercana para conectarse.

Los gobiernos de todo el mundo construyeron grandes bases de datos de trazabilidad, las cuales contienen la información de los contagios registrados en sus poblaciones (quién contagió a quién). Analizando estas bases de datos con algoritmos de DS e IA, los investigadores en DS y epidemiología intentan predecir el riesgo de un potencial nuevo foco de contagio y así tomar las medidas necesarias para prevenirlo. Como ejemplo, las principales Universidades de Estados Unidos desarrollaron modelos predictivos (muchos de ellos basados en algoritmos de IA) para determinar la curva de contagio de la enfermedad [5]. De la misma manera, muchas de las medidas tomadas por el gobierno de Corea del Sur se basaron en la aplicación de DS sobre bases de datos de interacciones ciudadanas para determinar la trazabilidad de contactos pasados y futuros de los infectados por COVID-19. Cabe destacar que, a mediados de noviembre de 2020, Corea del Sur reportaba una décima parte de los infectados en Chile y una centésima parte de los fallecidos por la enfermedad [6, 7], esto con una población casi 4 veces más grande.

En nuestro país, el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación creó un banco de datos con distintas fuentes que se fueron consolidando desde abril de 2020 [8]. Estas fuentes incluyeron datos de movilidad e información agregada de infectados a nivel comunal.

Problemas de privacidad

Desafortunadamente, las bases de datos útiles para análisis requieren información muy sensible sobre los pacientes. Esta información puede incluir su residencia, sus contactos personales y hábitos (lugares que frecuenta en la ciudad). En Corea del Sur, uno de los países que más éxito ha tenido en el control de la pandemia, las autoridades unificaron bases de datos que incluían información de compras con tarjetas de débito y crédito de todos los ciudadanos, así como la información de movilidad entre las antenas de telefonía celular [9]. De tal manera, las autoridades podían trazar de forma muy precisa los lugares

por los cuales una persona que había sido infectada se había desplazado, con el propósito de identificar y aislar a quienes se podían haber visto expuestos a la infección.

Contrastamos esto con el caso de Chile, donde la información disponible para análisis se encontraba anonimizada para proteger la privacidad de los ciudadanos. Más aún, la mayor parte de la información publicada por el Ministerio de Ciencia ni siquiera se encontraba individualizada, sino agregada (en vez de publicar la información de individuos, se publican estadísticas de grupos de individuos) con este mismo propósito.

¿Por qué esta diferencia? En general, reconocemos que la existencia de tales bases de datos es un arma de doble filo. Por un lado, pueden ser de alta utilidad para la traza de contactos de personas infectadas. Sin embargo, al mismo tiempo esta información puede ser utilizada con fines menos benevolentes. Por ejemplo, en una situación de estados autoritarios o dictaduras, estas bases de datos podrían ser utilizadas para perseguir y monitorear las actividades de opositores políticos. Incluso si estas bases de datos son creadas en democracia, la permanencia de su existencia representa un riesgo también a futuro.

Existen otros riesgos asociados a la simple existencia de estas bases de datos. En nuestro país, se reportaron ataques con piedras a residencias sanitarias donde se hospedaban personas contagiadas por la enfermedad [10]. Quienes se oponían a la presencia de estas residencias sanitarias temían contagiarse por aquellos que precisamente habían sido aislados para proteger a la población general. No es difícil imaginar que una reacción similar de estigmatización e incluso violencia se podría desatar contra las residencias de personas infectadas, si las bases de datos creadas por los gobiernos llegaran a ser filtradas al público. De la misma manera, empresas inescrupulosas podrían utilizar esta información para subir los planes de salud de aquellos que fueron víctimas de la enfermedad.

Más importante aún, las personas tienen derecho a la privacidad de su información clínica, según la ley vigente en nuestro país [11]. Esto quiere decir que el Estado, al crear esta base de datos, debe también ser garante de su seguridad e integridad, es decir, deben asegurarse de que esta información no caiga en "malas manos".

La privacidad y la utilidad de la información siempre están en tensión. Por un lado, queremos evitar a toda costa que la información privada de las personas llegue a malas manos. Por otro lado, queremos que los científicos que puedan proponer invenciones y desarrollos que nos ayuden a mantenernos sanos y a superar esta pandemia lo más rápido posible, tengan acceso a información relevante y útil.

El caso de Corea del Sur ha expuesto de manera clara cómo un enfoque de integración de bases de datos y herramientas de DS puede ser altamente efectivo para combatir los efectos de la pandemia. Sin embargo, es importante entender que nosotros valoramos nuestra privacidad de manera distinta a los ciudadanos de Corea del Sur y esto se refleja en nuestras instituciones y en las decisiones que ellas toman. En un reporte de la BBC [9], se indica que los ciudadanos de Corea del Sur *“han presentado poca oposición al nivel de vigilancia”, lo cual, explican, se podría deber a temas culturales.*

En nuestro país ocurrió algo completamente distinto. Solo un mes después de la primera publicación de datos, se produjo un quiebre en la mesa conformada por el Ministerio de Ciencias en conjunto con otras instituciones científicas. Particularmente, el Instituto Milenio Fundamentos de los Datos (IMFD) retiró su participación, acusando falta de disponibilidad de datos relevantes necesarios para la investigación científica [12]. En su comunicado, el IMFD indica que *“es válido cuestionar si la privacidad de los datos debe limitar el uso que puede hacerse de ellos cuando la salud del país está en juego”* [12]. Sin embargo, reconoce el riesgo de que estas bases de datos caigan en malas manos indicando que *“la liberación de datos sobre pacientes Covid-19 puede llevar a formas gravísimas de discriminación, tales como el apedreamiento de viviendas de pacientes COVID-19 y las amenazas de muerte y eventuales asesinatos de pacientes y sospechosos de estar contagiados”* [12].

Como podemos observar, la creación de estas bases de datos representa un dilema entre su utilidad como herramienta de prevención y el riesgo que representan. Como cualquier dilema, no hay una “mejor solución”. Más bien, como sociedad debemos decidir sobre si su creación es algo que queremos, sabiendo de antemano cuáles son los riesgos y tomando las medidas prácticas necesarias para mitigar los daños que éstas puedan generar.

Debate

Separe el curso en dos grupos y entrégueles unos minutos para exponer su posición. Solicite a ambos bandos que consideren abordar algunas de las preguntas durante su intervención. Una vez finalizada la exposición de ambos equipos, genere un intercambio cruzado de ideas, utilizando las mismas preguntas como guía. La idea es que se produzca una discusión libre en torno a la visión que ambas partes tienen sobre los medios, los riesgos y los beneficios de esta tecnología. En este sentido, el rol de la/el docente es el de moderar, orientar la discusión y solicitar precisión en las posturas.

Las posiciones son las siguientes:

1. Crear una base de datos de trazabilidad detallada con los movimientos en la ciudad de todos los ciudadanos (utilizando información de telefonía móvil), integrada con datos de compras con tarjetas de débito y crédito, de manera tal de identificar los movimientos de las personas que resulten infectadas.
2. Evitar la creación de la base de datos por considerarla peligrosa.

Las preguntas sugeridas para estructurar el debate son las siguientes:

- ¿Qué información es relevante para esta base de datos? ¿Qué otra información se podría incorporar para aumentar el éxito de la trazabilidad de las actividades que realizan las personas infectadas?
- ¿Qué otras razones hay para estar a favor o en contra de la creación de esta base de datos?
- ¿Qué malos usos se podrían dar a la información?
- ¿Cómo podemos asegurar que la información se encuentre disponible para los científicos que tienen las herramientas para disminuir la tasa de contagios y no para otras personas?

Cierre (15 minutos)

Para finalizar la actividad, haga una síntesis general de los resultados del debate. No es necesario dar por ganador o perdedor a ninguno de los grupos. Lo importante es resaltar

sus argumentos principales, de modo que el curso comprenda que se trata de un problema complejo y debatible y que, por lo tanto, no puede haber visiones arbitrarias ni unívocas.

Recuerde que este tipo de problemas que nos atañen a toda/os no tienen una sola verdad y que requieren un debate ciudadano amplio que permita llegar a consensos que tengan sentido de realidad en el contexto donde se apliquen. De eso se trata el debate ciudadano, de generar consensos ante los problemas y dificultades.

Referencias

- [1] <https://www.bayer.com/en/crop-science/crop-science-innovations-data-science>
- [2] <https://nij.ojp.gov/topics/articles/crime-mapping-crime-forecasting-evolution-place-based-policing#note30>
- [3] <https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge?amp%3Bamp%3Bamp%3B=#noteone>
- [4] https://www.udd.cl/wp-content/uploads/2020/05/indice-movilidad_ingenieriaudd.pdf
- [5] https://github.com/cdcepi/COVID-19-Forecasts/blob/master/COVID-19_Forecast_Model_Descriptions.md
- [6] Revisado el 17 de noviembre de 2020 - <https://covid19.who.int/region/wpro/country/kr>
- [7] Revisados el 17 de noviembre de 2020 - <https://covid19.who.int/region/amro/country/cl>
- [8] <https://www.minciencia.gob.cl/covid19>
- [9] <https://www.bbc.com/news/technology-52681464>
- [10] https://www.cnnchile.com/coronavirus/es-lo-mismo-que-lanzarle-una-piedra-a-un-hospital-gobierno-critica-ataques-a-residencias-sanitarias_20200424/
- [11] <https://www.bcn.cl/leychile/navegar?idNorma=141599>
- [12] <https://www.ciperchile.cl/2020/05/01/instituto-milenio-fundamentos-de-los-datos-congela-participacion-en-la-mesa-de-datos-del-covid-19/>

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana

Caso III: ¿Cómo nos preparamos para la siguiente pandemia?

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales/ Lengua y Literatura

Fecha: _____ **Curso:** _____

Objetivos:

- Comprender el dilema entre la protección de datos y los beneficios del acceso a fuentes de información que permiten la toma de decisiones de relevancia pública
- Discutir sobre los riesgos asociados a la vulneración de la privacidad

Nombre:

Instrucciones

En esta actividad debatirás sobre el uso sistemas de ciencia de datos e inteligencia artificial para el control, prevención y manejo de pandemias como la del COVID-19. Escucha atentamente las explicaciones de tu profesor/a y luego reúnete con tu grupo (la mitad del curso que te hayan asignado) para discutir los argumentos que presentarán al inicio del debate. Durante su intervención, intenten abordar algunas de las preguntas que se presentan más abajo. En la segunda parte del debate, el/la docente repasará las preguntas en voz alta y solicitará la intervención de ambos bandos para resolverlas a partir de las perspectivas que defienden.

Posiciones en debate:

1. Crear una base de datos de trazabilidad detallada con los movimientos en la ciudad de todos los ciudadanos (utilizando información de telefonía móvil), integrada con datos de compras con tarjetas de débito y crédito, de manera tal de identificar los movimientos de las personas que resulten infectadas.
2. Evitar la creación de la base de datos por considerarla peligrosa.

Las preguntas sugeridas para estructurar el debate son las siguientes:

- ¿Qué información es relevante para esta base de datos? ¿Qué otra información se podría incorporar para aumentar el éxito de la trazabilidad de las actividades que realizan las personas infectadas?
- ¿Qué otras razones hay para estar a favor o en contra de la creación de esta base de datos?
- ¿Qué otros malos usos se podrían dar a la información?
- ¿Cómo podemos asegurar que la información se encuentre disponible para los científicos que tienen las herramientas para disminuir la tasa de contagios y no para otras personas?

ACTIVIDAD 12

CASO IV: ¿DEBEMOS ADOPTAR LAS TECNOLOGÍAS BASADAS EN IA?

Oportunidades curriculares

Tecnología

- **I°Medio - OA 5:** Analizar las formas en que los productos tecnológicos y los entornos evolucionan, caracterizando los diversos factores que influyen en ese cambio.
- **I°Medio - OA 6:** Inferir, basándose en la evolución de los productos tecnológicos y los entornos, los efectos positivos o negativos que estos han tenido en la sociedad.
- **II°Medio - OA 3:** Evaluar las propuestas de soluciones que apunten a resolver necesidades de reducción de efectos perjudiciales relacionados con el uso de recursos energéticos y materiales, considerando aspectos o dilemas éticos, legales, económicos, ambientales y sociales.
- **II°Medio - OA 5:** Evaluar críticamente cómo las innovaciones tecnológicas actuales afectan a la sociedad y el ambiente, considerando criterios éticos, económicos, ambientales y sociales.
- **II°Medio - OA 6:** Proyectar escenarios de posibles impactos positivos o negativos de las innovaciones tecnológicas actuales en ámbitos personales, sociales, ambientales, legales, económicos u otros.

Historia, Geografía y Ciencias Sociales

- **I°Medio - OA 25:** Analizar el impacto del proceso de industrialización en el medioambiente y su proyección en el presente, y relacionarlo con el debate actual en torno a la necesidad de lograr un desarrollo sostenible.
- **II°Medio - OA 23:** Explicar los elementos que constituyen un Estado de derecho, como la Constitución, la división de los poderes del Estado, el respeto a los derechos humanos, entre otros; analizar su importancia para la gobernabilidad,

la representación, la participación ciudadana y la convivencia pacífica; y debatir sobre su perfeccionamiento, defensa y riesgos a los que está expuesto.

Lengua y Literatura

- **I° y II° Medio - OA 10:** Analizar y evaluar textos de los medios de comunicación, como noticias, reportajes, cartas al director, propaganda o crónicas.
- **I° y II° Medio - OA 21:** Dialogar constructivamente para debatir o explorar ideas.

Visión general

Retomando el caso de los automóviles sin chofer (actividad N°7), las y los estudiantes son confrontados/as a las consecuencias sociales de la automatización y, en particular, a la pérdida masiva de puestos laborales. Utilizando como contexto el modelo de adopción tecnológica de la comunidad Amish en Estados Unidos —quienes eligen cuidadosamente los medios que reportan beneficios a su estilo de vida— el curso problematiza la conveniencia y las desventajas de adoptar sistemas basados en inteligencia artificial.

La segunda parte de la actividad se realiza en Twitter, de manera asíncrona y distribuida en tres días (se proponen alternativas sin conexión, en caso de que no se cuente con los medios). Esto permite simular una discusión pública, donde todas y todos las y los estudiantes puedan manifestar sus puntos de vista, apoyar y complementar los de sus compañeros y compañeras. A diferencia de los ejercicios anteriores, esta dinámica busca la deliberación y desarrollo de puntos de vista individuales.

Objetivos

- Comprender y evaluar las consecuencias de la adopción de tecnologías basadas en inteligencia artificial
- Poner en perspectiva el actual modelo de desarrollo económico y social

Agenda (90 minutos)

- Introducción (30 min)
 - Explicación inicial
 - La comunidad Amish y la tecnología
- Desarrollo (30 min)

- Preguntas en torno a la IA y su uso
- Debate en Twitter (alternativas sin conexión)
- Cierre (30 min)
 - Análisis de las respuestas en el aula
 - Síntesis de la actividad

Materiales para esta actividad

- Guía docente
- Presentación de diapositivas
- Hoja del estudiante

Deseable:

- Cuentas en Twitter
- Computadores o celulares

Palabras clave

Adopción de tecnologías, modelo de desarrollo, progreso, transformación laboral, inteligencia artificial, Twitter.

Acceso a presentación de diapositivas

https://docs.google.com/presentation/d/1VKqHWkX4VBCL552oOO_BI_xKeU29_n8jC7vqkTh4SjE/edit?usp=sharing

GUÍA DOCENTE

El propósito de esta actividad es cuestionar nuestra relación con la tecnología, especialmente con respecto a la adopción de sistemas que generan cambios disruptivos, como aquéllos basados en la inteligencia artificial. A partir de la exposición del caso de la comunidad estadounidense Amish —que evalúa la incorporación de tecnologías en función de los valores y necesidades de la comunidad y la familia—, el objetivo es que las y los estudiantes discutan reflexivamente sobre las características de nuestro actual modelo de desarrollo (binomio tecnología-economía), en particular, en torno a las consecuencias para el mercado laboral de incorporar la automatización de la conducción, los valores que como sociedad debemos proteger y la posibilidad de frenar la adopción de estos avances o establecer medidas para mitigar el potencial daño que acarreen.

Esta actividad requiere el uso de computadores o equipos móviles y considera un tiempo para que cada estudiante cree una cuenta en Twitter. El uso de esta red social permitirá emular una discusión real en esta plataforma y evaluar los beneficios y vicios a los que está sujeta. No obstante lo anterior, se ofrecen variantes de la actividad sin necesidad de conexión a Internet, en caso de no contar con acceso o ancho de banda suficientes.

Para implementar esta actividad de forma completa, es necesario conocer las dinámicas básicas de Twitter y contar con una cuenta en la plataforma. Guíese con la siguiente tabla:

- ✓ Cree una cuenta en <https://twitter.com/i/flow/signup>
- ✓ Familiarícese con los conceptos *tweet*, *follow*, *follower*, *retweet*, menciones y *hashtag*.

Utilice este glosario como apoyo:

<https://www.cosmocudadano.mx/glosario-de-terminos-para-twitter>

- ✓ Familiarícese con los controles básicos de la plataforma, como se muestra a continuación:



Imagen 1. Encontrará estos íconos al pie de cualquier publicación (*tuit* o *tweet*). El primero permite responder (*reply*) y el segundo replicar y compartir la publicación (*retuitear* o *retweet*) con sus seguidores (*followers*).



Imagen 2 (izquierda). Este ícono permite crear un *tweet*

What's happening?



Tweet

Imagen 3. Caja para escribir un *tuit*. Escriba su mensaje (si desea, agregue imágenes, íconos u otros contenidos con los controles de abajo) y presione "Tweet".

Introducción (30 minutos)

Utilice la presentación de diapositivas para exponer el caso de la comunidad Amish y preparar la discusión. Explique y reflexione en voz alta con la clase, a partir del siguiente texto.

La comunidad Amish en Estados Unidos es un grupo religioso cristiano, popularmente conocido por un estilo de vida austero que, mirado por un o una visitante, pareciera haberse quedado congelado en el siglo XVII. Al observar su comunidad, resalta el uso de carretas tiradas por caballos en vez de automóviles, la falta de teléfonos celulares e incluso teléfonos fijos, además de la notoria escasez de enchufes eléctricos en sus hogares.



Miembros de la Comunidad Amish en Carruaje

Fuente: https://es.wikipedia.org/wiki/Amish#/media/Archivo:Amish_people.jpg

A pesar de que este estilo de vida nos puede parecer extraño, al mismo tiempo esconde una lección de la cual podemos aprender. Los Amish no rechazan la tecnología moderna por considerarla "malvada". En cambio, deciden adoptar sólo avances tecnológicos que consideren **un aporte** a su estilo de vida basado en la familia, la comunidad, la autosuficiencia y en su fe. Otras tecnologías que puedan interferir con estos valores o deteriorar este estilo de vida no son adoptadas por la comunidad.

Un ejemplo es el **automóvil**. Es indudable que esta tecnología es útil, pero su adopción masiva tiene consecuencias negativas claras en la contaminación que generan y en cómo impactan a la ciudad. Los Amish consideran que los carruajes imponen una restricción sobre la cercanía geográfica con la que los miembros de una familia y comunidad pueden elegir vivir (si debes ir en carreta a todas partes, es mejor que tu colegio, familia y amigos se encuentren cerca cuando debes ir a visitarlos). Esta evaluación parece ser acertada, pues efectivamente las ciudades tienden a crecer fragmentando a las familias y creando comunidades donde los vecinos no se conocen entre ellos. Por tanto, las familias Amish eligen no poseer automóviles y, en cambio, desplazarse con carruajes tirados por caballos [1].

Por otro lado, las familias Amish sí eligen tecnologías modernas si éstas aportan a su estilo de vida. Por ejemplo, una familia Amish puede poseer una moderna parrilla a gas si consideran que esta es útil para juntar a la familia y a la comunidad alrededor de un asado [2].

Las sociedades modernas, incluyendo a la sociedad chilena, no hacen este tipo de evaluaciones de manera explícita. Con contadas excepciones, **en general adoptamos las tecnologías cuando estas se encuentran disponibles y nuestras aprensiones con respecto a su uso son más bien reactivas que proactivas**. Es decir, regulamos el uso y accesibilidad a estas tecnologías cuando han causado algún problema o percibimos un peligro inminente en su uso. Por ejemplo, restringimos la calefacción a leña cuando los índices de contaminación son demasiado elevados o regulamos el uso de drones cuando su uso se ha vuelto una amenaza a la seguridad de las personas [3], pero no evaluamos sus efectos a priori.

El ejemplo de la comunidad Amish es un caso concreto de una población que ha deliberado en torno al uso y las consecuencias que tienen ciertos avances tecnológicos. Sin embargo, debemos considerar que en nuestra realidad actual resultaría complejo tomar decisiones tajantes en relación con el uso de nuevas tecnologías, dado el modelo de desarrollo que está intrínsecamente ligado con nuestra actividad económica.

En esta actividad, intentaremos apropiarnos un **enfoque proactivo** para la evaluación de la adopción de tecnologías de inteligencia artificial. Hasta ahora, hemos descubierto diferentes aplicaciones basadas en esta tecnología y hemos visto sus sorprendentes resultados. Sin embargo, al momento de implementarlas, no nos hemos cuestionado seriamente sobre sus potenciales impactos en la sociedad. Si aplicamos la filosofía de la comunidad Amish: **¿Debemos adoptar las tecnologías basadas en inteligencia artificial?** [Deje esta pregunta planteada, sin esperar respuesta, para preparar la discusión que viene a continuación].

Desarrollo (30 minutos)

En la actividad 7 discutimos sobre las capacidades de la IA para manejar de manera autónoma vehículos de transporte colectivo como microbuses y taxis. Además, enfocamos nuestra atención en el tipo de decisiones que debería tomar la IA frente a un accidente de tránsito, sin preocuparnos aún por otro impacto importante de esta tecnología, que es su capacidad de reemplazar empleos de manera masiva. Use el texto a continuación para preparar su exposición.

Inteligencia artificial, desempleo y desigualdad

En Chile, el Servicio de Impuestos Internos (SII) informa la existencia de 500 mil empleos asociados al rubro de Transporte y Almacenamiento [4], mientras que sólo Uber tiene registrados 70 mil conductores [5]. Todos estos empleos pueden volverse obsoletos con la introducción de tecnologías de inteligencia artificial para la conducción autónoma.

Ante este escenario, vale la pena preguntarnos qué es lo que valoramos como sociedad y cómo esta tecnología puede impactar esos valores. La posibilidad de un transporte seguro y eficiente es la promesa de la inteligencia artificial, pero el potencial costo es la pérdida de cientos de miles de puestos de trabajo remunerado.

Por otro lado, es importante considerar que las eficiencias y ahorros que permite la introducción de estas tecnologías benefician más a los que pueden pagarlas, redundando en una distribución de riqueza desigual. En otras palabras, quienes actualmente contratan a choferes para manejar un microbús deben invertir en la adquisición y despliegue de la tecnología, esperando un beneficio a largo plazo en el ahorro derivado de no necesitar pagar esos sueldos. Mientras tanto, los choferes desplazados dejan de percibir sueldo y deben invertir en capacitarse para adquirir un nuevo empleo que puede tener una remuneración menor, dada el alza en la demanda por empleos.

Luego de la lectura del texto, motive al curso a responder las siguientes preguntas, para preparar la discusión:

—En base a este escenario, ¿cuáles creen que son nuestros valores como sociedad?

Se espera que las y los estudiantes puedan identificar aquellos principios que como sociedad consideramos importantes, tales como oportunidades de desarrollo, acceso a la educación, acceso a la salud, libertad de expresión, desarrollo económico, seguridad, etc. Puede complementar esta lista, relevando aquéllos que sean adecuados al escenario (por ejemplo, la libertad de expresión no es un valor que se aplique de manera directa a este escenario). Puede hacer una lluvia de ideas.

—¿El desplazamiento masivo de empleos y la distribución desigual de la riqueza son costos adecuados para la adquisición de esta tecnología y sus beneficios? ¿Por qué? ¿Por qué no?

Escuche los puntos de vista del curso y guíe la discusión, relevando el dilema de elegir entre dos opciones no comparables. Por ejemplo, resulta complejo confrontar la seguridad vial y las vidas que se pueden salvar con la pérdida de empleo y el daño económico a miles de familias. El objetivo de esta pregunta es precisamente que las y los estudiantes se den cuenta de este dilema y puedan comprender el peso de la decisión. De ser posible, apóyese en quienes tengan parientes o amigos cuyas familias trabajen en transporte y logística, para generar una mayor conexión con el tema.

—¿Es posible pensar en no adoptar esta tecnología en nuestra sociedad? ¿Podría ocurrir realmente?

Normalmente, la adopción de este tipo de tecnologías es inevitable en economías abiertas como la de nuestro país. Se puede preguntar a las y los estudiantes qué elegirían si tuvieran que hacer un viaje de 12 horas a Arica o a Puerto Montt en bus, si entre todas las empresas que operan en el terminal hay una que ofrece un servicio cuya tasa de accidentabilidad es 100 veces menor a la de la competencia y a la mitad del precio. El objetivo de esta pregunta es permitir a los alumnos entender las dinámicas económicas que hacen casi inevitable la adopción de tecnologías de automatización.

—Si la adopción de esta tecnología es inevitable, ¿qué podemos hacer para mitigar sus impactos?

Motive la creatividad de las y los estudiantes con algunos ejemplos. Posibles medidas de mitigación incluyen la imposición de impuestos a estas tecnologías, de forma que su adopción sea paulatina. Esos mismos impuestos se pueden utilizar en programas de capacitación de choferes en otras áreas de trabajo.

Actividad en Twitter

Una vez preparado el tema, y con las reflexiones de la primera parte de la actividad, el curso estará en condiciones de sostener una dinámica de discusión donde todas y todos puedan aportar sus puntos de vista y apoyar y responder a los del resto de la clase. Antes de implementar la actividad, recomendamos leer el material desarrollado por la UNESCO para el uso de Twitter como herramienta educativa [6].

Diríjase a la sala de computación (también funciona con dispositivos móviles) y pida a sus estudiantes que creen una cuenta en Twitter, en cuyo perfil pueda distinguirse el nombre y la primera letra del apellido, de modo que puedan identificarse. Para evitar el uso de fotos personales, puede solicitarles identificarse con algún personaje de su gusto.

Use su cuenta para monitorear la conversación y realizar las preguntas que se detallan más abajo. Comparta su nombre de usuario con la clase, para que puedan seguirlo/a (*follow*). Las preguntas están planteadas para publicar una por día durante tres jornadas, de manera que las y los estudiantes cuenten con tiempo para contestar tranquilamente antes de evaluar.

Para cada pregunta, las y los estudiantes deben proveer una respuesta y *retuitear* la respuesta de un compañero o compañera. Para ello, deben elegir la respuesta con la que estén más de acuerdo. Antes de publicar la siguiente pregunta, la/el profesor debe elegir las tres respuestas más retuiteadas y compartirlas a la vez desde su cuenta. Se recomienda utilizar un *hashtag* o etiqueta en todas las publicaciones (por ejemplo, #actividad_IA) para que puedan reconocer la conversación en dicha red social.

Las preguntas de referencia son las siguientes:

Día 1: ¿Qué partes podrían estar interesadas en esta tecnología? ¿Quién podría verse más afectado/a por esta tecnología?

Día 2: Si esta tecnología se usara para hacer daño a las personas, ¿cómo podría hacerse?

Día 2: Si esta tecnología se usara para ayudar a las personas, ¿a quién y cómo podría ayudar?

Día 3: En 50 años, ¿Cómo piensa usted que la IA en automóviles ayudará a las personas?

Si las/los estudiantes tienen problemas respondiendo esta pregunta, puede motivar la discusión pidiéndoles que piensen en aplicaciones diferentes al transporte. Por ejemplo, puede mostrarles el video de Amazon Kiva, los robots de carga que usa Amazon en sus bodegas [7].

Día 3: En 50 años, ¿cómo piensa usted que la IA en automóviles podría causar daño a las personas?

De la misma forma, puede motivar la discusión preguntando cómo podría ser utilizada en conflictos bélicos.

Cierre (30 minutos)

Deje los primeros minutos de una próxima clase para presentar un resumen de la actividad y discutir con el curso las respuestas entregadas. Si lo considera adecuado, contraste las respuestas más populares (las más *retuiteadas*) con aquéllas que considere las mejores. Destaque que las respuestas más populares no son necesariamente las mejores y que esto es un defecto intrínseco de las redes sociales. Finalmente, recuerde que las opiniones compartidas en Twitter no se pueden homologar a la esfera pública —un error que cometemos con frecuencia cuando hablamos de la opinión de la ciudadanía— sino como una parte de ella, que se complementa con instancias de participación tales como juntas de vecinos, movimientos ciudadanos, agrupaciones de representación profesional y partidos políticos, entre otras.

Alternativa sin conexión a Internet

En caso de no contar con dispositivos o una conexión que permita usar Twitter, sugerimos dos alternativas:

- A. Cree un debate seleccionando tres de las cinco preguntas. Entregue 10 minutos para reflexionar de manera individual y pida a tres estudiantes al azar contestar la primera pregunta. Ofrezca la palabra a quienes quieran complementar o disentir. Repita lo mismo con las otras dos preguntas escogidas.
- B. Entregue una hoja o dicte las preguntas y asígnelas como tarea. Solicite al curso compartir sus respuestas al comienzo de la clase siguiente. Entregue un espacio para que las y los estudiantes puedan contrastar sus opiniones y votar por las que más les representen como curso.

Referencias

- [1] <https://www.amish365.com/dont-amish-drive-cars/>
- [2] Michael J. Quinn. 2014. Ethics for the Information Age (6th. ed.). Addison-Wesley Professional.
- [3] <https://www.dgac.gob.cl/como-operar-un-dron-en-chile/>
- [4] http://www.sii.cl/sobre_el_sii/estadisticas_de_empresas.html
- [5] <https://www.t13.cl/videos/nacional/video-cifras-uber-chile>
- [6] <https://about.twitter.com/content/dam/about-twitter/company/twitter-for-good/es/teaching-learning-with-twitter-unesco.pdf>
- [7] <https://www.youtube.com/watch?v=4D9k3tO4LDA>

GUÍA DEL Y LA ESTUDIANTE

Programa de inteligencia artificial y educación ciudadana

Caso IV: ¿Debemos adoptar las tecnologías basadas en IA?

Discusión en Twitter

Asignatura: Tecnología/ Historia, Geografía y Ciencias Sociales/ Lengua y Literatura

Fecha: _____ **Curso:** _____

Objetivos:

- Comprender y evaluar las consecuencias de la adopción de tecnologías basadas en inteligencia artificial
- Poner en perspectiva el actual modelo de desarrollo económico y social

Nombre:

Instrucciones

Esta actividad tiene como propósito reflexionar sobre el uso que como sociedad damos a los avances tecnológicos y, en este caso, a los sistemas de inteligencia artificial, que están transformando drásticamente el mundo laboral.

Sigue activamente la clase, hasta que tu profesor o profesora presente las siguientes preguntas (participa lo más que puedas y toma apuntes, ya que esto te ayudará a formar un punto de vista para la interacción en Twitter):

1. ¿Cuáles son nuestros valores como sociedad?
2. ¿El desplazamiento masivo de empleos y la distribución desigual de riqueza son costos adecuados para la adquisición de esta tecnología y sus beneficios? ¿Por qué? ¿Por qué no?
3. ¿Es posible pensar en no adoptar esta tecnología en nuestra sociedad? ¿Podría ocurrir realmente?

4. Si la adopción de esta tecnología es inevitable, ¿qué podemos hacer para mitigar sus impactos?

Ahora viene la parte más entretenida...

¿Alguna vez has participado en un debate por Twitter?

Esta puede ser una oportunidad genial para experimentar con la plataforma y conocer cómo son las dinámicas en esa red social.

Las instrucciones son simples:

1. **Crea una cuenta**, siguiendo las instrucciones de tu profesor o profesora
2. Espera los *tuits* (publicaciones) de tu profesor o profesora.
3. **Responde las preguntas** en forma de *tuits*. Recuerda que tienes que ser sintético/a, porque los caracteres son limitados (máximo 280 caracteres). Utiliza el *hashtag* (etiqueta) entregado por tu profesor o profesora, para mantener el hilo de la conversación.
4. **Retuitea la respuesta que más te haya gustado** de las que entregaron tus compañeros o compañeras.
5. **Repite este ciclo por cada una de las preguntas** que *tuitee* tu profesor o profesora.

Espera la próxima clase. Se seleccionarán las respuestas más *retuiteadas* y se reflexionará en torno a las controversias que hayan surgido.

AGRADECIMIENTOS

Agradecemos especialmente a las y los docentes y directivos escolares y comunales que apoyaron el proceso de testeo de los materiales, aportando ideas para mejorar la pertinencia del programa en el contexto escolar. Muchos de los cambios en la primera versión lanzada públicamente se deben a su generosidad.

Marco Aguilera
Ana María Benavente
Marta Ramírez
Valeria Zúñiga
Isabel Ventura
Mathias Mellado
Ricardo Ortega
Fabiola Quezada
Fabián Andrade
Camila Palacios
Daniel Soto
Patricio Acuña