

5. Comprender la IA desenchufados.

¿Cómo trabajarlo en el aula? Segundo y tercer ciclo

- [Consideraciones previas en segundo y tercer ciclo de primaria](#)
- [Actividad 1: El camino de decisiones: construimos un algoritmo \(tipo árbol\) y vemos el sesgo](#)
- [Actividad 2: ¡Crea una IA justa! El laboratorio de sesgos](#)

Consideraciones previas en segundo y tercer ciclo de primaria

En este capítulo, vamos a explorar cómo aprende una máquina y cómo podemos usar esa tecnología de manera creativa y responsable.

Esta pensado para alumnado de **segundo y tercer ciclo de Primaria**, que ya tiene capacidad de observar patrones, comprender relaciones de causa y efecto y hacerse preguntas sobre el mundo que les rodea debido a su **mayor desarrollo madurativo y capacidad de abstracción** con respecto al alumnado del anterior bloque.

En **este capítulo**, veremos propuestas de **actividades desenchufadas para segundo y tercer ciclo**, y en el capítulo siguiente aprenderemos de forma práctica cómo se entrenan los modelos de inteligencia artificial, de dónde vienen los datos que utilizan y cómo esos datos pueden influir en sus decisiones.

¿Sabías que los sistemas reCAPTCHA de google se utilizaron para entrenar inteligencias artificiales?

1. Digitalización de libros (Texto)

Los primeros reCAPTCHA pedían a los usuarios que descifrarán dos palabras distorsionadas. Una palabra era una prueba de control, cuya respuesta ya se conocía, para verificar que el usuario era humano.

La otra palabra provenía de textos que estaban siendo escaneados y digitalizados (como libros y periódicos antiguos). El sistema no estaba seguro de la transcripción correcta de esta segunda palabra, por lo que utilizaba la entrada humana para validarla y mejorar los algoritmos de Reconocimiento Óptico de Caracteres (OCR).

2. Reconocimiento de imágenes (Visión por ordenador)

Con el tiempo, los CAPTCHA evolucionaron a la selección de imágenes con objetos específicos (semáforos, pasos de peatones, coches, etc.). Al igual que con el texto, algunas imágenes tenían respuestas conocidas, mientras que otras no estaban completamente definidas. Las respuestas humanas masivas y consistentes se utilizaron para crear grandes conjuntos de datos etiquetados, que son fundamentales para entrenar sistemas de visión por ordenador, como los utilizados en los coches autónomos.

3. Análisis de comportamiento

Los sistemas de reCAPTCHA más modernos, como reCAPTCHA v3 y reCAPTCHA "No soy un robot" (la casilla de verificación), también analizan el comportamiento del usuario en la página web, como la velocidad de los clics, el movimiento del cursor y el historial de navegación, para determinar si es un humano o un bot. Estos datos de comportamiento también ayudan a refinar los algoritmos y la detección de patrones humanos.

En resumen, los CAPTCHA convirtieron una tarea de seguridad en una forma eficiente de obtener millones de etiquetas de datos gratuitos y de alta calidad para entrenar modelos de machine learning a gran escala.

Este enfoque, nos permitirá desmitificar la IA de cara a nuestro alumnado, mostrando que **no es magia**, sino una herramienta creada por personas, y entender que siempre **hay un punto de vista humano detrás de cada decisión de la máquina**.

Con ejemplos cercanos y actividades experimentales, este capítulo no sólo busca despertar la curiosidad por la tecnología, sino también ayudar a nuestro alumnado a **pensar de manera crítica, ética y reflexiva** sobre cómo usamos la IA en nuestra vida cotidiana.

Lejos de ser magia, la inteligencia artificial se basa en procesos que pueden explicarse de forma sencilla y experimentarse en el aula con herramientas como **Teachable Machine o LearningML**. En este punto resulta clave comprender cómo se entrenan los modelos y cómo los datos influyen directamente en sus decisiones, ya que es precisamente en ese proceso donde se originan muchos de los sesgos de la IA.

Actividad 1: El camino de decisiones: construimos un algoritmo (tipo árbol) y vemos el sesgo

Actividad 1. “El camino de decisiones: construimos un algoritmo tipo árbol.”

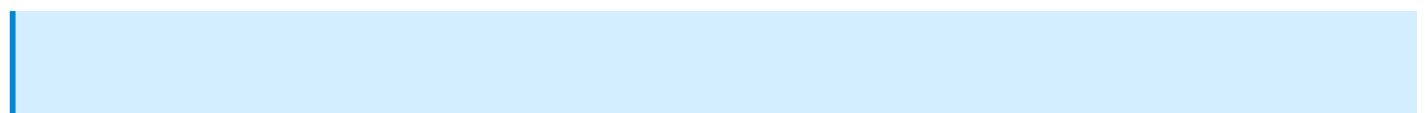
JUSTIFICACIÓN

El alumnado entiende cómo “decide” una IA mediante un algoritmo en forma de preguntas (árbol de decisiones), usando objetos reales o a través de tarjetas. Comprende que el modelo es el árbol final construido con datos (ejemplos) y que aparece sesgo si los ejemplos no representan bien la realidad. Se trata de una actividad muy manipulativa y favorece el razonamiento del alumnado..

Datos de la actividad

- **Curso:** 3º Primaria (8-9 años)
- **Tiempo:** 45-60 min (recomendable realizar en 2 sesiones)
- **Agrupación:** equipos de 4-5 y una puesta en común final.
- **Espacio:** aula (suelo o mesas grandes para montar el “camino”)
- **Formato:** manipulativa y sin pantalla.

VINCULACIÓN CURRICULAR



Esta actividad nos va a permitir realizar una vinculación no solo teniendo en cuenta su utilidad en el aprendizaje del funcionamiento de un modelo de IA, sino también en este caso con su utilidad para trabajar aspectos relacionados con el área de lengua como veremos a continuación.

AREA: Ciencias de la Naturaleza

- **CE.CN.3.** Resolver problemas a través de proyectos de diseño y de la aplicación del pensamiento computacional, para generar o reelaborar cooperativamente un producto creativo e innovador que responda a necesidades concretas. (relacionada con el marco de proyecto + pensamiento computacional: diseñar un algoritmo, probarlo, evaluarlo y mejorarlo, sin necesidad de tecnología)

AREA: Lengua Castellana y Literatura

- **CE.LCL.3.** Producir textos orales y multimodales con coherencia, claridad y registro adecuados para expresar ideas, sentimientos y conceptos; construir conocimiento; establecer vínculos personales; y participar con autonomía y una actitud cooperativa y empática en interacciones orales variadas.
- **CE.LCL.5.** Producir textos escritos y multimodales, con corrección gramatical y ortográfica básicas, secuenciando correctamente los contenidos y aplicando estrategias elementales de planificación, redacción, revisión y edición para construir conocimiento y dar respuesta a demandas comunicativas concretas.

La vinculación con el área de Lengua Castellana y Literatura se considera adecuada ya que el algoritmo se construye a través de formulación de preguntas, se valida con explicaciones a través del debate grupal (“falló porque...”) y se cierra con exposición oral y texto escrito basada en las conclusiones y trabajo durante el proceso de trabajo de la actividad.

OBJETIVO DIDÁCTICO

Construir un árbol de decisiones para clasificar objetos y comprender el proceso de aprendizaje seguido por la IA para crear un modelo; **datos, algoritmo, modelo y sesgo**, de esta manera entender como influyen en los aciertos y errores los datos aportados y las modificaciones que se realicen en el modelo.

DESARROLLO

1) **Pregunta detonante**

“¿Podemos construir un camino de preguntas SÍ/NO para que una “máquina” clasifique objetos?”

2) Organización y roles

Podemos utilizar los siguientes roles para el trabajo en equipo:

- **Ordenador/a:** coloca preguntas y flechas.
- **Comprobador/a:** pasa objetos por el camino y comprueba si funciona o no.
- **Detector de errores:** registra fallos y causas.
- **Portavoz:** explica el modelo al final.

3) Datos iniciales

El docente entrega al equipo un set de [TARJETAS DE OBJETOS \(enlace para descargar\)](#) ya “etiquetados” en una de las 3 categorías elegidas, se pueden utilizar objetos reales para hacerlo más realista.

Podemos usar las siguientes apartados:

- **ESCRIBIR:** lápiz, boli, rotulador, tiza
- **CORTAR:** tijeras, cúter de cartón (solo imagen)
- **SUJETAR:** clip, pinza, goma elástica, celo, grapadora(tarjeta)

Aquí se introduce: **DATOS = ejemplos con etiqueta (la respuesta correcta).**

4) Construcción del algoritmo

El equipo, en una cartulina va creando un **camino de preguntas SÍ/NO** para llegar a una categoría.

Ejemplos de preguntas útiles:

- “¿Sirve para hacer marcas o letras?”
- “¿Tiene filo o corta?” (si se usan tarjetas, no objetos peligrosos)
- “¿Sirve para sujetar/juntar cosas?”
- “¿Se usa con las manos para apretar?”

Reglas del juego (para que sea claro):

- Máximo **3 preguntas** antes de llegar a un destino
- Cada pregunta debe poder responderse **SÍ o NO**
- Cada objeto debe poder “viajar” por el camino

Aquí se nombra: **ALGORITMO = secuencia de preguntas para decidir.**

5) El modelo

Cuando el camino funciona con los 12 ejemplos, se pega un cartel como que el modelo ya esta entrenado y listo para comprobar:

MODELO = nuestro árbol final de preguntas (lo aprendido con los datos).

6) Reto de comprobación: objetos nuevos

El docente entrega **objetos nuevos** (no estaban en los 12 ejemplos) y el equipo los pasa por el modelo.

Registran si funciona el modelo o no.

Ejemplos de objetos “nuevos” que provocan pensamiento:

- regla (¿es para medir o para dibujar?)
- pegamento (¿juntar = sujetar?)
- estuche (¿sirve para...?)
- goma de borrar (¿es escribir o corregir?)
- botella (ninguna categoría: obliga a revisar)

7) Sesgo: faltaban ejemplos

Ahora el docente puede orientar sobre los posibles problemas con los objetos nuevos probados:

“Vuestros datos tenían muchos ejemplos de ESCRIBIR y muy pocos de SUJETAR (o faltaba ‘pegamento’, ‘goma’, etc.).”

Se repite una mini-prueba con algún objeto “conflictivos” de los utilizados para comprobar.

El equipo identifica posibles errores como:

- ¿Qué tipo de ejemplo faltó?
- ¿Qué pregunta está mal o no es clara?

Aquí se nombra: **SESGO** = cuando el conjunto de ejemplos no representa bien y el modelo aprende una idea incompleta.

8) Mejora del modelo

Cada equipo aplica 1 mejora:

- Añadir 2-3 ejemplos nuevos (datos) y rehacer una pregunta
- Cambiar el orden de preguntas
- Crear una "salida extra": **OTROS** donde se puedan clasificar los objetos más "conflictivos"

Este es un paso importante para entender como trabaja y mejora la IA cuando se encuentra con aspectos nuevos o corrige errores.

Vuelven a probar con los 4 objetos conflictivos.

9) Cierre y puesta en común

- "¿Qué ayudó más: cambiar preguntas o añadir ejemplos?"
- "¿Puede un modelo dar errores si solo ve algunos casos?"
- "¿Qué harías para que el modelo fuese más fiable?"

DUA (ADAPTACIONES)

- **Representación:** uso de pictos ([arasaac](#)) y colores por categorías lo que reduce la carga lectora y mejora la comprensión.
- **Andamiaje lingüístico:** "banco de preguntas" ("¿Sirve para...?, ¿Se usa para...?") o uso de frases guía ("Falló porque..., lo mejoramos cambiando..., faltaban ejemplos de..."). Elementos que facilitan la argumentación y elaboración de conclusiones en 2º Ciclo.
- **Función ejecutiva:** Secuenciar por mini-retos. Dividimos en pasos visibles: 1) 3 preguntas, 2) probar con 3 objetos, 3) ajustar, 4) prueba de 10, 5) mejora.

EVALUACIÓN

Para la evaluación de esta actividad podemos realizarla a través de una lista de cotejo establecida para 4 momentos específicos a lo largo de la actividad:

- Momento 1: Diseño del árbol (**CE.CN.3**)

- Momento 2: Prueba del modelo y registro de resultados **(CE.CN.3 + CE.LCL.5.)**
- Momento 3: Análisis y mejora del árbol **(CE.CN.3)**
- Momento 4: Comunicación final **(CE.LCL.3.)**

Se pueden establecer varios ítems por cada momento y establecer 3 categorías (sí, en parte, no) lo que nos permite evaluar la actividad en varios momentos distintos, así como el producto final. Este trabajo lo realizamos a través de un porcentaje que establezcamos de ítems con respuesta sí. De esta manera podemos observar cual es el momento o momentos que más problemas ha tenido el alumnado a la hora de llevarlo a cabo y poder actuar de forma más concreta.

Glosario rápido para explicarlo en clase

- **Datos:** ejemplos con etiqueta (respuesta correcta).
- **Algoritmo:** preguntas SÍ/NO que guían la decisión.
- **Modelo:** el producto final, en este caso el árbol final que usamos para clasificar.
- **Sesgo:** cuando el árbol falla más con ciertos objetos.
- **Prueba/validación:** comprobar el modelo con casos nuevos.

Actividad 2: ¡Crea una IA justa! El laboratorio de sesgos

¡Crea una IA justa! El laboratorio de sesgos

JUSTIFICACIÓN

Esta actividad es útil para que el alumnado conozca que “aprender” en IA significa **crear un modelo** a partir de **datos** y luego **probarlo** con casos nuevos. El alumnado detecta **sesgos** cuando ciertos grupos quedan poco representados y entiende por qué eso puede ser injusto. En esta actividad se integra pensamiento crítico, convivencia y conocer el lenguaje de IA sin usar pantallas.

Datos de la actividad

- **Curso:** 5º Primaria (10-11 años)
- **Tiempo:** 55-70 min ((recomendable realizar en 2 sesiones))
- **Agrupación:** equipos de 4-5 + puesta en común
- **Espacio:** aula (mesas)

RELACION CURRICULAR

La vinculación curricular de esta actividad, aunque trabajamos de forma desenchufada el funcionamiento de la IA, nos permite realizar la siguiente vinculación curricular:

AREA: Educación en valores Cívicos y Éticos

- **CE.EVCE.1:** Deliberar y argumentar sobre problemas de carácter ético referidos a sí mismo y su entorno, buscando y analizando información fiable y generando una actitud reflexiva al respecto, para promover el autoconocimiento y la autonomía moral. *(Deliberar y argumentar sobre problemas éticos; construir posición moral autónoma, incluyendo cuestiones sobre el uso responsable y crítico de medios/redes)*

- **CE.EVCE.2:** Actuar e interactuar de acuerdo con normas y valores cívicos y éticos, reconociendo su importancia para la vida individual y colectiva, y aplicándolos de manera efectiva y argumentada en distintos contextos, para promover una convivencia democrática, justa, respetuosa y pacífica. (*promover convivencia democrática y excluir arbitrariedad, injusticia y discriminación; analizar conflictos también en entornos virtuales y proponer medidas*).

AREA: **Ciencias de la Naturaleza** (dimensión digital y de datos)

- **CE.CN.1:** Utilizar dispositivos y recursos digitales de forma segura, responsable y eficiente, para buscar información, comunicarse, trabajar de manera individual, en equipo y en red y, para reelaborar y crear contenido digital de acuerdo a las necesidades digitales del contexto educativo.

Esta actividad curricularmente nos permite en tercer ciclo combinar: el análisis crítico y la deliberación ética en contextos cercanos (IA y decisiones automatizadas) y el uso responsable de información y datos en el apartado digital.

OBJETIVO DIDÁCTICO

Construir y evaluar un modelo sencillo para tomar decisiones, en este caso centrada en la toma de un solo tipo de decisión, usando datos de ejemplo, y comprobar cómo aparece el **sesgo** por representación y por “reglas mal diseñadas”.

DESARROLLO

1) Pregunta detonante

“Si quisiéramos una ‘máquina’ que nos ayudará a decidir nuestro **Robot ayudante del aula...** ¿cómo le enseñaríamos a decidir?”

2) Organización

- **Diseñador/a del modelo** (escribe reglas/puntos)
- **Entrenador/a de datos** (analiza ejemplos)
- **Probador/a** (aplica el modelo a nuevos casos)
- **Auditor/a de sesgo** (revisa a que casos perjudica y por qué)
- (Opcional) **Portavoz**

Materiales

- 30 “fichas de caso” con Robots y rasgos neutros del aula (sin datos sensibles) que queramos que analicen, por ejemplo: respeta turnos, interrumpe explicaciones, cuida el material, etc...

En este caso como ejemplo, los rasgos elegidos serían los siguientes:

POSITIVOS

- Respetar turnos
- Cuidar el material
- Trabajar bien en equipo
- Respetar la fila
- Ayudar a ordenar

NEGATIVOS

- Interrumpe
- No cuida el material
- Le cuesta trabajar en equipo
- Olvida los materiales
- Se enfada cuando no gana

A partir de estos datos, que en clase puede ser un **trabajo previo** donde el alumnado seleccione los rasgos que piensan, tanto positivos como negativos, que debería tener el robot ayudante de su clase. Se crean combinaciones para crear las fichas que vamos a utilizar para desarrollar esta actividad.

Como ejemplo tenemos estas fichas "[ROBOTS EJEMPLO](#)"

Podemos determinar el nivel de dificultad de la tarea añadiendo mayor o menor número de rasgos a cada robot

- 3 tipos de sobres:
 1. **Fichas de entrenamiento.** estarían ya etiquetadas por el docente como Sí es elegido o NO es elegido.
 2. **Fichas de prueba:** Estas fichas irían sin etiqueta para el alumnado. El docente si tendría la solución elegida para comparar los resultados obtenidos por el alumnado.
 3. **Fichas difíciles:** Estas tarjetas promueven que se origine debate, como plantea la actividad: se trata de crear casos mixtos como por ejemplos 2 rasgos positivos y uno negativo, 2 negativos y 1 positivo o incluir rasgos no contemplados anteriormente. Aquí es donde puede aparecer el sesgo del modelo.
- Plantilla "[Nuestro modelo](#)" (tabla de puntos o reglas)
- Hoja de registro: 12 predicciones señalando CORRECTA/INCORRECTA



- Post-its: DATOS / ALGORITMO / MODELO / SESGO / PRUEBA de esta manera pueden ir identificando en cada momento el momento del proceso en el que se encuentran.

Nota: se puede utilizar en vez de robots, cualquier tipo de ejemplo que resulte más motivante al grupo como animales, superhéroes, personajes de dibujos, plantas, deportistas, Pokemons,... Imaginación al poder.

3) Datos (entrenamiento)

Cada equipo recibe 10 tarjetas del sobre **ENTRENAMIENTO**, ya etiquetadas por el docente:

El equipo analiza:

- ¿Qué rasgos aparecen mucho en los "SÍ"?
- ¿Qué rasgos aparecen mucho en los "NO"?

Se nombra: **DATOS = ejemplos con respuesta correcta.**

4) Construcción del algoritmo (modelo por puntos)

El equipo crea un **algoritmo** tipo "puntos" (muy manipulativo y fácil de aplicar) como por ejemplo:

- +2 "respeta los turnos"
- +1 "ayuda a ordenar"
- -2 "interrumpe"
- -1 "olvida los materiales"

Puede ser una decisión de equipo o de clase determinar los valores de cada rasgo. En este caso sería recomendable que fuera una decisión a nivel de clase ya que favorece la comprobación de la tarea de forma igualitaria por todos los grupos.

Finalmente se establece una **Regla final** por ejemplo:

- 3 o más puntos → **SÍ**
- Menos de 3 → **NO**

Se nombra: **ALGORITMO = las reglas/pasos para sumar y decidir y MODELO = la tabla final de puntos (lo aprendido con los datos).**

5) Prueba (validación) con datos nuevos



Se entrega el sobre **PRUEBA** con 8 tarjetas nuevas (sin solución).
El equipo aplica su modelo, predice SÍ/NO y registra.

Después el docente revela la etiqueta real (o una hoja de soluciones común).

6) Mini-laboratorio de sesgos

Ahora entra en juego el sobre **DIFÍCILES** con 4 tarjetas pensadas para mostrar el sesgo.

Ejemplos:

- Casos con rasgos “mixtos” que no aparecían en entrenamiento (“no sigue las instrucciones” o “escucha con atención”)
- Casos con un rasgo que el modelo penaliza demasiado (“Interrumpe”)

El auditor responde en la hoja:

- ¿Qué tipo de caso recibe más “NO”?
- ¿Ese rasgo estaba poco representado en los datos?
- ¿La regla está exagerando una pista?

Se nombra: **SESGO = cuando el modelo se equivoca más con ciertos casos porque los datos eran pobres o las reglas no eran justas.**

7) Mejora del modelo

Cada equipo elige UNA mejora (obligatorio justificarla):

- **Mejora de datos:** añadir 2 tarjetas de entrenamiento que faltaban (más variedad de casos)
- **Mejora de reglas:** bajar rasgos de No ($-2 \rightarrow -1$), subir un criterio positivo, o añadir “NO SÉ” si está en zona gris
- **Mejora de prueba:** crear una regla de “revisión humana” para casos límite, en este caso la decisión final es tomada por el grupo.

Como ejemplo se puede crear el siguiente **modelo mejorado** teniendo en cuenta solo la puntuación:

- 3 o más puntos -> **SÍ**
- 1 o 2 puntos -> **REVISAR/ NO SÉ o DECISIÓN HUMANA**
- 0 o menos puntos -> **NO**

Se repiten las 4 tarjetas DIFÍCILES y se comparan resultados.

8) Debate final

- “¿Qué mejoró más: cambiar datos o cambiar reglas?”
- “¿Cuándo sería peligroso usar este modelo sin revisar?”
- “¿Qué significa que un modelo sea ‘justo’?”

DUA (adaptaciones)

- **Representación:** tarjetas con iconos/pictos ([arasaac](#)) y colores por categorías lo que reduce la carga lectora y mejora la comprensión.
- **Acción/expresión:** para construir el modelo por puntos, usad fichas: una ficha = 1 punto (o fichas dobles para +2). Así el alumnado “ve” la suma y puede demostrar el aprendizaje aunque escriba menos; después explican el resultado oralmente o con un mini-esquema “si suma $\geq 3 \rightarrow$ Sí”.

EVALUACIÓN

A modo de evaluación de la actividad proponemos como ejemplo la siguiente rúbrica.

Indicador (criterios)	1 · Inicial	2 · En proceso	3 · Adecuado	4 · Avanzado
1) Comprensión del funcionamiento de la IA (desenchufada) (CE CN 1)	Confunde datos/reglas/resultados; necesita guía constante.	Identifica partes del proceso con ayuda, con algunos errores.	Explica con claridad: datos → reglas/modelo → prueba → errores/sesgo .	Además, relaciona causas del sesgo y anticipa cómo prevenirlo.
2) Uso y organización crítica de datos/información (CE EVCE 1)	Usa ejemplos/datos sin criterio; no justifica decisiones.	Organiza de forma básica; justifica poco o de manera confusa.	Selecciona y organiza datos con criterio; justifica decisiones de forma clara.	Detecta desequilibrios/ausencias en los datos y propone cómo corregirlos con buena justificación.
3) Deliberación ética: justicia, igualdad y no discriminación (CE EVCE 1 y 2)	Opina sin razones; dificultad para escuchar y respetar turnos.	Aporta razones simples; escucha parcialmente; necesita recordatorios.	Argumenta con razones y ejemplos; escucha y dialoga con respeto.	Considera otros argumentos, ajusta su postura y ayuda a construir acuerdos justos.

4) Propuestas de mejora para una IA más justa y convivencia (CE EVCE 2)	Propone cambios vagos o irrelevantes; no concretas medidas.	Propone una mejora concreta, pero poco viable o poco relacionada con el sesgo.	Propone mejoras concretas y viables (mejorar datos, ajustar reglas, revisión humana) y las justifica.	Propone varias mejoras, prevé efectos y define cómo comprobar si el sistema es más justo.
--	---	--	--	---

La asociación de los criterios con la actividad sería la siguiente:

- **Registro de entrenamiento/prueba** (Indicadores 1 y 2)
- **Observación del debate** (Indicador 3)
- **Mini-informe “auditor/a de sesgo”**: qué sesgo, a quién afecta, mejora propuesta (Indicador 4)

Aprender sobre la IA no es "usar pantallas", sino comprender como decide un sistema basado en datos y que riesgos éticos puede tener.