



IA RESPONSABLE

para gestión ágil



Scrum Manager®

IA responsable para gestión ágil

Formación complementaria
Versión 1.0

Scrum Manager®

Skill tag:  #AI for Agile Projects Disponible en [Open Knowledge \(Scrum Manager\)](#).

Fecha de la versión: abril 2024.

Autora de la versión: Marta Palacio.

Diseño de portada: Ana Andrés Soria.

Agradecimientos: [Kathy Tian](#), *data science manager* en VISA Inc. y [Rubén Álvarez González](#), *agile coach y trainer* en Estrategco.

Uncovering Better Ways SLU es la editora y propietaria de los derechos de distribución, que libera en los términos de la licencia Creative Commons-by-nd-nc 4.0.

Derechos registrados en Safe Creative. N° de registro: [2404167662398](#).

Índice de contenidos

Introducción.....	4
Introducción a las inteligencias artificiales.....	5
Evolución histórica.....	5
Distintos tipos de IA.....	7
Sets de datos.....	9
La IA como herramienta.....	11
Ingeniería concurrente vs. agilidad.....	11
División de competencias.....	12
Consideraciones prácticas.....	13
Consideraciones éticas.....	14
Ejercicio. Supongamos un chatbot.....	17
Aplicaciones prácticas.....	19
Gestión y seguimiento del proyecto.....	19
Soluciones para equipos.....	23
Mejora de productos y servicios.....	25
Ejercicio. Planifica con un lienzo de IA.....	27
Invertir en talento.....	29
Perfiles asociados a la IA.....	29
Formación continua.....	31
Ejercicio. Planes de aprendizaje.....	32
Tendencias futuras.....	34
Inteligencia artificial general.....	34
Fuentes para mantenerse al día.....	34
Ejercicio. Legislación.....	35
Bibliografía.....	36
Apéndices.....	39
Cómo se entrena un modelo de ML supervisado.....	39

Introducción

Curso de introducción a la IA para proyectos ágiles. Recomendado para gestores y propietarios de producto.

Consejos prácticos

Exploramos cómo aprovechar mejor la inteligencia artificial en las tareas típicas de gestión de proyectos, en el trabajo de los equipos, y en los productos y servicios destinados al usuario final.

El contenido no se centra en sugerir herramientas de terceros concretas (aunque se mencionan algunas), o cómo aprovechar mejor éstas mediante *prompts*. Da una base sólida sobre inteligencia artificial en general, con una perspectiva que puede adaptarse a diferentes tipos de proyectos.

El apéndice incluye además aspectos básicos sobre el entrenamiento de modelos de machine learning.

Con un enfoque responsable y ágil

La IA se encuentra en un momento crítico de su evolución, y hay consideraciones prácticas y éticas fundamentales ligadas a su aplicación que no se pueden dejar aparte de la conversación. Además de dedicar un apartado concreto a estas consideraciones, el curso ayuda a entender cómo hacer un uso de la IA compatible con los valores de la agilidad, y a ser conscientes de los pros y contras reales del uso de máquinas predictivas.

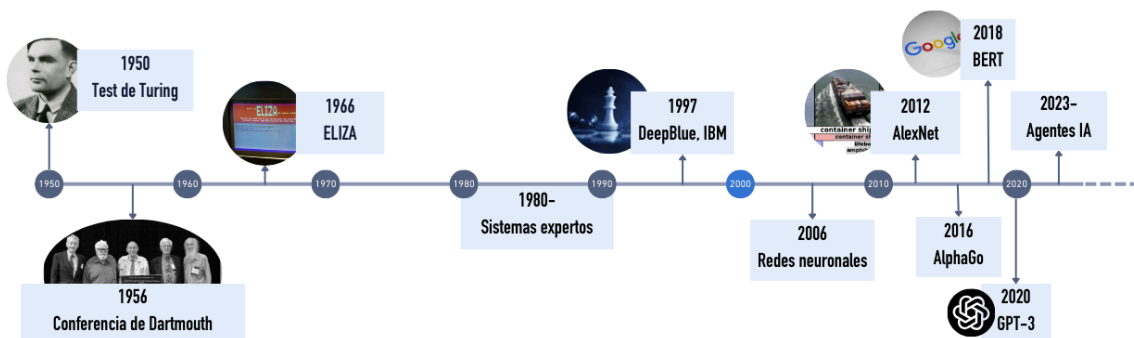
Introducción a las inteligencias artificiales

La inteligencia artificial (IA) se refiere a la simulación de procesos de inteligencia humana por parte de sistemas informáticos. Estos procesos incluyen aprendizaje, razonamiento, y auto-corrección; es decir: asimilar información y normas, seguir estas normas para alcanzar conclusiones, y aprender cuando estas conclusiones son erróneas para corregir y mejorar.

En el contexto de la gestión de proyectos la IA puede desempeñar un papel crucial en:

- La automatización de tareas repetitivas.
- La predicción de tendencias de mercado.
- La facilitación de la toma de decisiones basada en datos.

Evolución histórica



La IA comenzó a desarrollarse en la década de **1950**, con el objetivo inicial de realizar tareas que normalmente requerirían inteligencia humana. Es entonces cuando Alan Turing publicó su paper [Computing Machinery and Intelligence](#), donde aparece por primera vez el *imitation game*, conocido ahora como [Test de Turing](#). Una prueba para determinar si una máquina es capaz de hacerse pasar por un humano.

Desde entonces, la evolución de esta rama de las ciencias de la computación ha seguido adelante. Pese a algunos periodos de “estancamiento” en las primeras décadas, el ritmo de mejora actual es espectacular y está teniendo impactos en todos los ámbitos.

En **1956**, 6 años después del *paper* de Alan Turing, tuvo lugar la **conferencia de Dartmouth**. Se considera el nacimiento oficial de la inteligencia artificial como campo de estudio. El nombre lo eligió John McCarthy, uno de los expertos que organizaban la conferencia. Proponían un “estudio para proceder sobre la base de la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en

principio, ser descrito con tanta precisión que puede fabricarse una máquina para simularlo.” (McCarthy, 1955, [A proposal for the Dartmouth summer research project on artificial intelligence](#)).

La experimentación continuó durante las décadas de los 60 y 70, con éxitos moderados. En **1966**, Joseph Weizenbaum creó **ELIZA**, uno de los primeros programas de interacción en lenguaje natural con una computadora. Esta época también fue testigo de los fundamentos de la **retropropagación**, un algoritmo de aprendizaje supervisado para entrenar redes neuronales. Aunque su impacto no se reconocería hasta tiempo después.

El **efecto ELIZA** deriva su nombre de este primer chatbot. Se refiere a la tendencia a proyectar ciertas cualidades humanas en los textos producidos por estos programas informáticos. Creer que el programa piensa, es consciente de sí mismo, o siente empatía.

En los **80**, la IA experimentó un renacimiento con el desarrollo de los llamados “**sistemas expertos**”, diseñados para emular la toma de decisiones de humanos especializados en campos concretos. Con la evolución de estos sistemas llegó una inversión de capital significativa, para aplicaciones comerciales de inteligencia artificial.

Los resultados empezaron a hacerse poco a poco más evidentes. En **1997**, Deep Blue de IBM adquirió fama al convertirse en la primera computadora en ganar contra un campeón mundial de **ajedrez**, Garry Kasparov.

El comienzo de la **era moderna** del **aprendizaje profundo** (**deep learning**, **DL** por sus siglas en inglés) se da hacia **2006**, cuando Geoffrey Hinton y otros reintroducen las **redes neuronales** con técnicas innovadoras. En 2012, AlexNet, una red neuronal convolucional (*convolutional neural network*, CNN por sus siglas en inglés) de Alex Krizhevsky gana la competición ImageNet, superando a todas las entradas basadas en métodos tradicionales de aprendizaje automático o *machine learning* (ML). Esto impulsó el interés y la inversión en aprendizaje profundo.

Desde ese momento hasta hoy se amontonan los hitos a un ritmo cada vez más acelerado:

- **2016: AlphaGo**, de DeepMind, derrota a Lee Sedol, el campeón mundial de Go. Se trata de un juego considerado mucho más complejo que el ajedrez para la IA.
- **2018: Google** introduce **BERT** (*Bidirectional Encoder Representations from Transformers*), un nuevo método para preentrenar modelos de procesamiento de lenguaje natural, mejorando significativamente la comprensión del lenguaje y estableciendo un nuevo estándar en **procesamiento natural del lenguaje (NLP)** por sus siglas en inglés).
- **2020: OpenAI** lanza *Generative Pre-trained Transformer 3 (GPT-3)*, un modelo de lenguaje que marca un nuevo estándar en la capacidad de generar texto coherente y contextualmente relevante. GPT-3 demuestra una versatilidad impresionante en tareas de NLP, desde la creación de contenido hasta la generación de código.

- **2021:** la IA jugó un papel crucial en diversas áreas durante la pandemia de **COVID-19**, desde el desarrollo de vacunas hasta la gestión de la salud pública. Se usaron modelos predictivos para prevenir brotes, optimizar recursos y entender la propagación del virus, y en investigación médica la IA aceleró el proceso de identificación de candidatos para vacunas y tratamientos.
- **2023:** hay un creciente interés por los agentes inteligentes y los agentes inteligentes autónomos, capaces de percibir su entorno por sensores y actuar en base al estado actual de éste. Pueden ser tan sencillos como un termostato. Hoy en día, el foco está puesto sobre todo en sistemas multi-agente, que agrupan y coordinan varios agentes inteligentes para realizar tareas complejas de principio a fin.

También ha habido grandes progresos en robótica: los sistemas autónomos y drones son cada vez más capaces de realizar tareas complejas con mínima supervisión humana, con aplicaciones que van desde la logística y entrega de pedidos hasta operaciones de búsqueda y rescate.

Nos encontramos ahora en un punto de inflexión, en el que la IA está cada vez más presente en la vida diaria. Aunque aumentan las expectativas, también lo hace la preocupación. Han ido surgiendo a nivel global cada vez más iniciativas con las que establecer marcos éticos y legales que garanticen un desarrollo seguro y responsable de estas tecnologías.

Distintos tipos de IA

En el apartado anterior han aparecido términos con los que conviene familiarizarse. Dentro de la inteligencia artificial hay muchos subcampos: robótica, sistemas expertos, visión por computadora, inteligencia artificial general...

Para los objetivos del curso, los más interesantes de diferenciar son dos:

- Aprendizaje automático o *machine learning* (ML).
- Aprendizaje profundo o *deep learning* (DL).

Se trata de dos áreas muy amplias, que engloban muchas herramientas útiles para la gestión de proyectos. Pueden ayudar en tareas de planificación, personalización de experiencia de usuario, y a descubrir oportunidades de negocio, entre otras funciones que veremos a lo largo del curso.

Aprendizaje automático (ML)

Este subcampo de la IA permite a las máquinas aprender de los datos y mejorar su rendimiento sin ser explícitamente programadas para cada tarea. Utiliza algoritmos para analizar datos, aprender de ellos y hacer predicciones o tomar decisiones basadas en la información procesada.

Un sistema de ML puede usarse, por ejemplo, para analizar historiales de compra de clientes y predecir qué productos podrían interesarles en el futuro.

Los sistemas de ML se suelen dividir en dos:

- **Supervisados:** los más habituales. Se utilizan para predecir resultados en base a datos históricos. Es necesario comprobar que el algoritmo produce respuestas que coinciden con lo que se espera.
- **No-supervisados:** se usan cuando no se está buscando una respuesta concreta sino explorar los datos para encontrar estructuras ocultas o patrones. Esto permite, por ejemplo, agrupar datos de usuarios en grupos (*clusters*) con similitudes.

Aprendizaje profundo (DL)

Es una técnica de aprendizaje automático, el subcampo que acabamos de ver. Enseña a las computadoras a procesar entradas a través de capas, para reconocer patrones y características. El término "profundo" hace referencia al número de capas a través de las cuales pasa la información, permitiendo a la máquina aprender y realizar tareas de una complejidad sin precedentes, como el reconocimiento facial o de voz.

Las redes neuronales profundas, especialmente las redes neuronales recurrentes (RNN) y las redes neuronales convolucionales (CNN) procesan y aprenden usando grandes volúmenes de datos no estructurados, en formatos como imágenes, texto y audio.

El reconocimiento facial es uno de los usos típicos del aprendizaje profundo. Se usa en redes sociales, por ejemplo, para identificar y etiquetar a usuarios. Otro uso del DL son los modelos de lenguaje de gran tamaño (LLM por sus siglas en inglés), que pueden contener miles de millones de parámetros. El más famoso de ellos es ChatGPT, de OpenAI.

Diferencias entre ML y DL

El **aprendizaje automático** puede ser más efectivo en sistemas con **relaciones simples** o bien definidas entre los datos. El **aprendizaje profundo** sobresale en entornos donde hay **mayor volumen de datos**, éstos son complejos y las relaciones más sutiles.

Sin embargo, los sistemas de ML son (por ahora) más **trazables** y comprensibles que los de DL. Un algoritmo de DL, debido a su profundidad y complejidad, es muy difícil de interpretar. Por eso se suele decir que son "cajas negras".

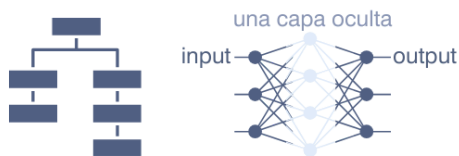
El campo de la "interpretabilidad de modelos" o "explicabilidad de la IA" está avanzando rápidamente. Hay un gran interés en desarrollar técnicas que permitan entender mejor cómo los modelos de DL toman sus decisiones. Existen enfoques como

LIME (*Local Interpretable Model-agnostic Explanations*) y SHAP (*SHapley Additive exPlanations*).

Los algoritmos de aprendizaje automático (ML), en cambio, sí pueden interpretarse. Si por motivos de investigación, regulación, éticos... queremos saber cómo el algoritmo llega a las decisiones que llega, es preferible **usar ML por su trazabilidad**.

Aprendizaje automático / Machine learning

Usa muchos modelos.
 Por ejemplo, árboles de decisiones y redes neuronales.



*ejemplos: un árbol de decisiones y redes neuronales prealimentadas, de *input* a *output*

Aprendizaje profundo / Deep learning

Una red neuronal de **más de 3 capas** se considera profunda.
 Los modelos de DL pueden contener **cientos o miles** de capas




Sets de datos

Dada la gran cantidad de datos que se necesitan para entrenar modelos, y que obtenerlos puede ser difícil, es habitual descargar sets de Internet. Esto significa que el uso extendido de un set puede causar daños a gran escala. Cuando los sets se aplican sin conocer lo que contienen, los modelos entrenados con ellos pueden dar **resultados erróneos o sesgados sin saber por qué**.

Un caso conocido es el de Joy Buolamwini, investigadora de ciencias de la computación, que demostró que los algoritmos de reconocimiento facial presentaban **sesgos significativos**: fallaban al identificar correctamente rostros de personas de piel oscura, sobre todo mujeres. Esto se debía a los conjuntos de datos de entrenamiento, con un porcentaje mucho más elevado de rostros de hombres blancos. Su trabajo subrayó la importancia de la diversidad y equidad en los datos utilizados para IA, y fue utilizado por IBM para avanzar en este sentido (Christian 2020).

Pero el trabajo de Buolamwini es sólo ilustrativo, y el problema es difícil de resolver. **La proporción de muestras de un tipo de contenido en el set afecta a la precisión del modelo cuando se aplica**. Si entrenamos a un modelo para que separe tipos de pelotas, y en el set de datos casi siempre presentamos las pelotas de fútbol como blancas y negras, una pelota de fútbol naranja puede acabar clasificada como una de baloncesto. Es porque el modelo no tiene suficiente información. Cuando un grupo es una minoría eso suele ir asociado a que, porcentualmente, genera menos datos.

Al igual que la trazabilidad de los modelos es importante, también lo es tener **acceso a los datos de entrenamiento, para poder auditar su calidad**. Muchas herramientas de IA no comparten información sobre sus conjuntos de datos, y según el

uso que vayamos a darles esto es un factor de riesgo a tener en cuenta. Se está empezando a regular en este sentido. Por ejemplo, la Ley de divulgación de derechos de autor de IA generativa en EEUU pide mayor transparencia, y obligaría a comunicar el uso de obras protegidas por derechos de autor en el entrenamiento de estos modelos.

La IA como herramienta

El rápido avance de la IA genera inquietud por muchas razones. Una de ellas es su potencial para reemplazar puestos de trabajo. En el contexto de la gestión de proyectos, se está viendo no obstante una evolución más que una sustitución. Los gestores y *product owners* que se adapten rápido y aprendan a utilizar estas herramientas efectivamente pueden asegurarse un lugar indispensable en el futuro del trabajo, y elevar el potencial de sus equipos y proyectos.

Pero la adopción debe hacerse de forma bien informada y consciente, para sacarles el mayor provecho.

Ingeniería concurrente vs. agilidad

El primer principio del *Manifiesto Ágil* dice:

***Valoramos a los individuos y sus interacciones
por encima de los procesos y las herramientas.***

Un proceso es un contenedor de conocimiento explicitado. Desde el punto de vista de la ingeniería concurrente, las personas supervisan que el proceso se haga bien y, en última instancia, la calidad del resultado depende de la calidad del proceso. Son entornos donde prima la predictibilidad y la estandarización, y que dificultan la innovación y reducen el empoderamiento de las personas. Desde el punto de vista de la agilidad, en cambio, priman los individuos y sus interacciones, y la calidad del resultado depende de éstos.

¿Dónde reside el conocimiento?

La distinción clave entre si se trabaja con ingeniería concurrente o con agilidad es dónde reside el conocimiento del proyecto.

- **Ingeniería concurrente:** se trabaja de forma iterativa y con fases solapadas, pero tiene prioridad el **conocimiento explícito** sobre el tácito. Los empleados siguen instrucciones y protocolos claros, con poco margen de maniobra.
- **Agilidad:** se trabaja de forma ágil y además tiene prioridad el **conocimiento tácito** de los equipos. Las personas del proyecto tienen el juicio, el talento y la confianza para experimentar y tomar decisiones.

En empresas donde el conocimiento es explícito, y los procesos guían y prescriben el trabajo, las personas ayudan al proceso. Pero si seguimos el primer principio del *Manifiesto Ágil*, la relación debe ser a la inversa: **el proceso ayuda a las personas a**

mejorar la calidad de su trabajo. Perder esto de vista es sacrificar el espíritu de la agilidad.

Los modelos actuales de IA encajan dentro de la categoría de “**procesos**”. Para trabajar con IA en un entorno ágil es útil entenderlas como **herramientas procedimentales** muy complejas. Contienen **conocimiento explicitado**: el algoritmo extrae conclusiones siguiendo unos criterios definidos e instrucciones fijas. Al emplear herramientas de IA para crear listas priorizadas de tareas o diagramas de Gantt lo que estamos aprendiendo son técnicas y herramientas que pueden servir de apoyo, pero no de guía.

División de competencias

Fortalezas de las máquinas predictivas

Ya hemos visto que hay diferentes tipos de inteligencia artificial. El nombre “inteligencia” viene del propósito de estas tecnologías: imitar capacidades propias de la inteligencia humana. No obstante, ninguno de los modelos actuales (ML, DL) son inteligentes, sino **máquinas predictivas**, como señalan Ajay Agrawal, Joshua Gans y Avi Goldfarb en su libro, así titulado. Al elaborar nuestras estrategias es útil recordar este concepto y entender así estos modelos.

La predicción no es inteligencia, ni tampoco una predicción equivale a una decisión. Pero es un componente de ambas. La capacidad de extraer patrones a partir de grandes cantidades de datos no estructurados y realizar predicciones tiene mucho potencial en nuestro ámbito.

Las áreas con mayor potencial para automatizar con IA pueden ser:

1. “Allí donde todo lo demás, excepto la predicción, ya se ha automatizado.
2. Áreas donde los beneficios de la velocidad de acción en respuesta a la predicción son elevados.
3. Áreas donde los beneficios de reducir el tiempo de espera de las predicciones son elevados.” (Ajay Agrawal et al. 2019:142)

No obstante, a la hora de decidir si automatizar o no y en qué medida, hay otros factores que considerar.

Fortalezas humanas y sinergias

Las máquinas de predicción no aportan **juicio crítico**, **conocimiento tácito** ni **talento**. Una organización ágil debería **invertir** precisamente en estas habilidades, pues son las que más valor van a adquirir en este nuevo paradigma.

La colaboración entre inteligencia humana y artificial promete nuevos niveles de innovación, eficiencia y personalización de experiencias de usuario. Pero para garantizar que estas sinergias sean compatibles con una cultura ágil, hay que tomar medidas que garanticen que prime el conocimiento tácito de las personas. Lo que a menudo es referido como “intuición”, “experiencia” o “saber hacer”.

En muchas de las sinergias que observamos, la IA aporta información, la ordena o la interpreta de forma útil. Es una fuente más de información a sopesar a la hora de tomar decisiones.

Otra función destacable es ahorrar tiempo, automatizando tareas repetitivas o al menos aumentando su velocidad de ejecución. Reduciendo la carga cognitiva de este tipo de trabajos, las personas pueden dedicar su atención a otros que requieren de más criterio o creatividad.

Consideraciones prácticas

La implementación de soluciones con inteligencia artificial conlleva sus propios desafíos, inversión de tiempo y recursos, y una cierta curva de aprendizaje. Independientemente de que desarrollemos soluciones propias o utilicemos servicios de terceros.

Conforme la tecnología avance, más casos de estudio aparecerán y más herramientas saldrán al mercado. Si nuestra capacidad para experimentar y nuestros datos son limitados, podemos:

- Seguir con atención los experimentos con IA de otras empresas en nuestro sector, para aprender de sus aciertos y sus errores.
- Identificar las áreas con un componente de predicción donde la IA puede sernos más útil y centrar ahí la implementación y la experimentación.

Hay que tener en cuenta los recursos disponibles y el posible impacto a corto, medio y largo plazo de las implementaciones. Por ejemplo:

- Un *chatbot* mal entrenado de cara a los usuarios puede ser más alienante que un correo de atención al cliente, aunque se tarde unas horas en recibir respuesta. Se necesita tiempo para mantener y entrenar estos modelos, sobre todo al principio.
- En un equipo pequeño, querer automatizar ciertos procesos puede ser un malgasto de recursos. Hay que considerar si aligeraría la carga de trabajo lo suficiente para dar más cancha a la experimentación y la innovación sin sacrificar seguridad o calidad.
- La calidad de los resultados depende de la cantidad y la calidad de los datos disponibles. Recopilar y pre-procesar estos datos es clave para implementar IA, pero ambas tareas traen consigo consideraciones de privacidad y sesgos, entre

otras. Hay un equilibrio delicado entre proporcionar suficiente contexto para que el algoritmo sea útil y sobrecargar el *prompt* con información innecesaria que confunda al modelo. Según nuestros propósitos, se requieren conocimientos especializados propios de las ciencias de datos, por lo que puede que tengamos que invertir en formación o personal. Más información sobre conocimientos técnicos en el capítulo 4.

Consideraciones éticas

Discriminación

Se ha observado que los sistemas de IA pueden estar sesgados en contra de ciertos colectivos, lo que puede tener consecuencias discriminatorias en áreas como el empleo, la vivienda o el acceso a la justicia. Estos sesgos no siempre son evidentes. Para evitarlos, como señalamos en el capítulo 1, es importante que los modelos sean trazables y comprensibles y conocer cuáles son los datos de entrenamiento.

Por ejemplo, cuando salió Alexnet, la IA etiquetaba a personas negras en los resultados de Google Imágenes como “gorilas”. El problema era más difícil de resolver de lo que parece.

Por un lado, tenía que ver con la calidad de las fotografías en sí y la capacidad de las cámaras para capturar con precisión el color de piel de personas no blancas. No fue hasta los 90 que Kodak Shirley introdujo una carta con múltiples tonos de piel.

Por otro lado, las minorías por el hecho de ser minorías aparecen menos representadas en los datos. Investigadoras como Joy Buolamwini y Julia Angwin, por ejemplo, han llamado la atención sobre este tema.

Joy Buolamwini descubrió que los sistemas de inteligencia artificial utilizados para tecnología de análisis facial estaban sesgados: la detectaba mejor cuando llevaba una máscara blanca. Muchas de estas tecnologías emplean el set de datos *Labeled Faces in the Wild*, en el que hay, por ejemplo, el doble de fotografías de George W. Bush (530 imágenes) que de mujeres negras.

Julia Alwin en 2016 dio la voz de alarma sobre los sistemas de predicción con IA que se estaban aplicando en Estados Unidos para informar de las decisiones de las juntas para libertad condicional. Su artículo se titula “*Machine bias: se está utilizando software en todo el país para predecir futuros delincuentes. Y está sesgado contra los negros*”.

En conclusión, se han ido acumulando casos que demuestran que los algoritmos no sólo hacen ciertos sesgos evidentes, sino que también los van perpetuando al tomar decisiones basadas en los sesgos y generar un bucle. No se recomienda utilizar ciertos modelos de lenguaje más allá de su uso previsto (Christian 2020).

Privacidad y seguridad

Tanto Internet como la IA plantean riesgos para la privacidad de los usuarios y la seguridad de sus datos personales. Si recopilamos datos de usuarios para utilizar en nuestros modelos, debemos asegurarnos de que éstos están seguros, como mínimo según obliguen las regulaciones vigentes.

La IA aún no está bien regulada, lo cual genera incertidumbre jurídica y riesgos para los usuarios. La recopilación y el uso masivo de datos por parte de empresas y gobiernos pueden ser una amenaza para la libertad individual, y el rápido avance de la IA significa que los problemas éticos y legales pueden surgir de forma más rápida e inesperada.

Este entorno requiere seguir con atención la evolución de las regulaciones sobre IA, y noticias relacionadas con su uso en los sectores de negocio a los que pertenezca el proyecto. En Europa, los requisitos mínimos de seguridad y transparencia para la recopilación de datos de usuarios vienen marcados por la GDPR desde 2018.

Responsabilidad

¿Cuánta responsabilidad debería asumir la IA? No debemos olvidar que **la IA es una herramienta procedimental**. Y tal y como se expresa en el *Manifiesto Ágil*: “Valoramos a los individuos y sus interacciones por encima de los procesos y las herramientas.” Es decir, en ningún caso, ni en este curso ni en ningún otro, recomendamos el uso de la IA en lugar del talento humano. Su uso siempre debería ir guiado y supervisado por el de personas expertas en las áreas pertinentes, que deben responsabilizarse de su buen funcionamiento.

Los sistemas de IA pueden tener un mayor grado de autonomía que los sistemas informáticos tradicionales, lo que plantea dudas sobre la responsabilidad en caso de errores o daños. Es necesario mantenerlos supervisados y actualizados.

Según el proyecto, hay que considerar quién se hará responsable de las decisiones tomadas por una inteligencia artificial: a quién puede acudir un usuario damnificado, qué cosas pueden salir mal y qué soluciones o medidas podemos ofrecer en cada caso.

También hay consideraciones de responsabilidad importantes en torno a los sets de datos, no sólo a los modelos. Es decir: ¿cómo se controla lo que contienen? Cuando el volumen de datos es demasiado grande, se tiende a usar filtros automáticos. Pero a mayor volumen de datos, menor control. Un ejemplo de las consecuencias de estos métodos es LAION-5B, un set de datos de imágenes de uso abierto para fines de investigación. Aunque ellos no recomiendan el uso de su set en productos comerciales listos para usarse, Midjourney y Stable Diffusion lo utilizan, entre otros. El set ya no puede descargarse; se deshabilitó en diciembre de 2023 por una investigación ya que

contenía imágenes de abuso sexual a menores. Las imágenes se habían obtenido mediante *web crawl*, y se había tratado de eliminar el contenido inapropiado con filtros automáticos.

Desinformación

Los modelos de lenguaje grandes (LLM) y otras IA generativas se están usando de forma masiva para generación de contenido, contribuyendo a la cacofonía de información de baja calidad en Internet.

Muchas de estas inteligencias artificiales dan información falsa como si fuera cierta. Cuando esto sucede, se dice que tienen “alucinaciones”. Compartir estos datos puede dañar la credibilidad del proyecto y tener consecuencias imprevisibles para los usuarios. Es crucial revisar el contenido, contrastar fuentes y consultar a expertos.

Textos e imágenes generadas por IA están apareciendo en publicaciones revisadas por expertos (Guo X, Dong L y Hao D 2024), investigaciones científicas, informes jurídicos (Russell 2023)... En una publicación reciente del *Yale Journal of Biology and Medicine*, se plantea si es viable usar ChatGPT para revisar artículos académicos (Biswas et al. 2023). Quizá la sobrecarga de trabajo lleva a considerar estas soluciones, pero el criterio, saber hacer y los conocimientos técnicos de expertos siguen siendo fundamentales. No conviene dejarse llevar por los primeros resultados de modelos de IA para hacer predicciones, como por ejemplo si un artículo es adecuado o no. Existen casos que demuestran que, pese a unos buenos resultados iniciales, las IA toman sus decisiones de formas que no son fiables a largo plazo, a veces desconcertantes. Lo observaron por ejemplo los autores de *Automated Classification of Skin Lesions* al analizar una IA desarrollada para clasificar lesiones en la piel. Se descubrió que uno de los factores determinantes era si había una regla en la fotografía. En los datos de entrenamiento, si un médico tenía sospechas de que podía tratarse de un tumor maligno, solía aparecer una regla en la foto, así que el modelo aprendió que las reglas eran indicativas (Novoa et al. 2018).

En proyectos ágiles, donde prima el conocimiento de las personas, su talento y experiencia se vuelven clave para identificar información falsa o contradictoria y evitar su difusión. La **responsabilidad** de comprobar la veracidad de la información es del operador de la IA, o de las personas del proyecto responsables de revisar el contenido antes de su uso o publicación.

Propiedad intelectual

Todavía no se ha regulado de forma clara sobre si las obras generadas por IA tienen derechos de autor o no. En general, lo que el usuario puede hacer o no con la obra generada viene estipulado en los términos y condiciones de uso de la herramienta. Pero más allá de estos términos existe debate. Si queremos tener seguridad jurídica y defender nuestros derechos de propiedad intelectual sobre diseños, fotografías, música y

otras obras creativas asociadas a nuestra marca, es mejor no utilizar obras generadas por IA directamente.

Existen muchos casos en curso por infracción de derechos de autor contra las grandes empresas de IA generativa. Estos modelos han sido entrenados con grandes bases de datos que contienen obras protegidas, y a veces producen resultados que infringen la propiedad intelectual. El *New York Times*, entre otros muchos, ha denunciado a Microsoft y OpenAI por infracción de *copyright*. Según un análisis de Patronus AI, a fecha de 2024, GPT4 es la IA que más derechos de autor infringe (TIPS SafeCreative, 2024).

Falacias de autoridad y empiristas

Se tiende a asumir que algo debe ser cierto simplemente porque está argumentado o respaldado por datos. Este fenómeno, aunque no tiene un término específico, está relacionado con la falacia de apelo a la autoridad y el concepto de falacia empirista. Se asume que la mera presentación de datos empíricos es suficiente para validar un argumento, ignorando otros factores como la interpretación de los datos, el contexto, las limitaciones del estudio, o cómo se obtuvieron, procesaron y analizaron los datos.

Muchas decisiones van a empezar a tomarse en base a conclusiones obtenidas por inteligencia artificial. Debemos ser escépticos y críticos con el uso de datos en argumentos, ya que éstos pueden ser manipulados, mal interpretados o presentados de manera que apoyen una conclusión específica, ignorando deliberadamente datos que la contradigan.

Los modelos de IA pueden entrenarse mal, sus predicciones empezar a fallar con el tiempo si no se mantienen, y pueden confundirse tanto si la información proporcionada es insuficiente como si es excesiva.

Para contrarrestar este efecto, además de espíritu crítico, conviene utilizar modelos comprensibles y estar abiertos a evaluar y mejorar tanto los resultados como los métodos.

Ejercicio. Supongamos un *chatbot*

Supongamos que vamos a empezar a utilizar un *chatbot* avanzado basado en IA para manejar preguntas frecuentes y problemas comunes de los clientes. Se propone utilizar análisis predictivo para identificar patrones en las consultas de los clientes, para resolver problemas recurrentes con más rapidez.

1. En el contexto del servicio de atención al cliente, compara: ¿cuáles serían las fortalezas humanas frente a las de la máquina?
2. Supongamos que la empresa es una corredora de seguros, y el *chatbot* es capaz de dar cotizaciones iniciales en base a los datos del usuario. ¿Qué

consideraciones éticas deberían tenerse en cuenta al entrenar el chatbot? ¿Cómo abordarías cada una para mitigar riesgos?

Aplicaciones prácticas

Este módulo ofrece sugerencias y consideraciones importantes para optimizar tareas y mejorar la experiencia de los usuarios con IA.

Para saber más sobre herramientas que facilitan estas aplicaciones, o aspectos técnicos sobre cómo entrenar modelos de *machine learning*, consulta los [Apéndices](#) del curso.

Vamos a ver cómo podría optimizarse cada parte de un flujo de trabajo iterativo con herramientas de IA **de forma responsable y compatible con los principios ágiles**.

Gestión y seguimiento del proyecto

Priorización de *backlogs*

A la hora de priorizar el *backlog*, el valor del criterio del *product owner* sigue siendo clave. La IA puede servir como una fuente de información e inspiración más, y como un asistente que agilice la extracción e interpretación de datos. Los ítems del *backlog* pueden provenir de fuentes diversas y ser de muchos tipos. Además, conforme el proyecto crece, cada ítem puede tener dependencias a considerar. Refinar y considerar toda la información necesaria a la hora de priorizar es una tarea difícil, y de hecho existen dudas sobre la efectividad de aplicar modelos de IA en este terreno dada la gran complejidad que puede entrañar (Salehi Farhang 2023).

No obstante, siendo conscientes de las limitaciones actuales, sigue siendo una aplicación con potencial a explorar. Un algoritmo puede analizar datos históricos de nuestro proyecto u otros similares y realizar predicciones tales como:

- Qué tareas pueden tener mayor impacto en el éxito del proyecto.
- Cuáles priorizar en base a su impacto, valor comercial, complejidad y urgencia.

El primer paso sería, como siempre que entrenemos un modelo, reunir datos históricos relevantes. Por ejemplo:

- Impacto medido de características previas en la retención y satisfacción del usuario.
- Complejidad estimada y tiempo real de implementación de tareas anteriores.

Utilizando estos datos, se entrena un modelo de aprendizaje automático (ML). Por suponer, podría ser un clasificador que prediga el "impacto en la satisfacción del usuario" basado en factores como la naturaleza de la tarea (nueva característica, corrección de errores, mejora de la usabilidad), la complejidad estimada y datos históricos de impacto.

Una vez entrenado, el modelo se utiliza para evaluar el *backlog* actual. El *product owner* revisa las recomendaciones del modelo, ajustando la priorización según su criterio y conocimiento del mercado y los usuarios.

Reevaluación de user personas

Mediante el uso de modelos de aprendizaje automático no supervisados, podemos localizar patrones y agrupar usuarios con características comunes que, para un humano, pueden no ser evidentes.

Crear estos grupos o *clusters* con similitudes puede ser útil para revisar los arquetipos de usuarios ya creados, o crear nuevos.

Gestión de avance

Los datos actualizados del ritmo de avance pueden ayudar a detectar cuellos de botella o retrasos de forma temprana. Estas alertas pueden comentarse en equipo para actualizar el *roadmap* si es necesario, o centrar esfuerzos donde sea más crítico.

No obstante, hay detalles e información que pueden perderse si se depende en exclusiva de herramientas de gestión. Los equipos deberían seguir gestionando su avance con comunicación directa diaria. En caso de seguir el marco *scrum*, a través de las *dailies*.

Alertas y mitigación de riesgos

En algunos sectores concretos, la IA no sólo automatiza tareas repetitivas sino que esto hace posible que ciertos servicios funcionen. Es el caso de los sistemas de detección de fraude en el uso de tarjetas de crédito. La primera línea de defensa es una IA que, según parámetros, decide si el nivel de riesgo de la transacción es tolerable o no. La intervención humana sólo es necesaria cuando el riesgo de fraude supera ese nivel.

Según el tipo de proyecto, podemos configurar notificaciones automáticas y alertas para ciertos eventos o riesgos. Pueden ayudar a gestionar ciertas acciones de los usuarios, como en el caso de las tarjetas de crédito, o ayudarnos a nivel interno. Por ejemplo:

- Verificar las dependencias de *software* del proyecto para mantenernos al día con las últimas versiones.
- Monitorizar automáticamente servidores y sistemas.
- Monitorizar automáticamente el uso de recursos, inventario disponible y presupuestos, proporcionando alertas tempranas si se proyecta, por ejemplo, una falta de productos para una época en la que se demandan.

Optimización de recursos

La IA puede analizar datos históricos de consumo, producción e inventario para identificar áreas donde optimizar. Por ejemplo, se puede predecir la demanda de productos y servicios concretos y ajustar la producción y el inventario de manera más eficiente en base a estas predicciones.

Herramientas como PowerBi (Microsoft) o Tableau, además de ser excelentes soluciones para visualización de datos, permiten ver en tiempo real el consumo de materiales en una planta de producción, para identificar puntos de fuga o ineficiencias. Otra herramienta de uso muy extendido es BlueYonder, líder en la gestión de inventario y cadenas de suministro, utilizada por empresas como Walmart o Coca-Cola.

Soluciones integrales

Existen herramientas muy completas para integrarse en el flujo de trabajo y ayudar a optimizarlo con servicios de análisis predictivo, automatización y soporte.

Por ejemplo, Jira Software, de Atlassian, que permite añadir apps con inteligencia artificial. A finales de 2023 anunciaron *Atlassian Intelligence*, un servicio premium que incorpora IA a toda clase de tareas dentro del entorno de trabajo.

Usos desaconsejados

Aunque las herramientas lo permitan, incorporar soluciones con inteligencia artificial en algunas áreas puede tener **impactos negativos en una cultura ágil**. En general, es poco aconsejable implementar este tipo de soluciones para:

- Estimación de tareas.
- Creación de historias de usuario.
- Asignación de trabajo.
- Análisis de rendimiento.

Estimación de tareas

Las dinámicas de estimación no sirven sólo para llevar un control sobre los plazos de entrega. Son el momento para plantear dudas, entender mejor la visión y los objetivos y, si existe mucha discrepancia entre estimaciones, entender por qué: si alguien está viendo un problema que otros ignoran, o una solución que a otros no se les ha ocurrido.

El uso de IA para estimación puede tener consecuencias tóxicas en el ambiente laboral. Los modelos de ML utilizan datos cuantitativos, y fomentan caer en modelos de productividad y eficiencia que tratan de cuantificar al detalle.

La microgestión y monitorización de los empleados generan un ambiente de desconfianza. Se ha observado que, con el tiempo, [el estrés anula la motivación y la creatividad](#).

Si las métricas son visibles para todos, pueden generar competitividad tóxica del mismo modo que sucede a veces con los puntos de historia. Empleados y equipos se resienten y compiten por querer producir más en menos tiempo que otros. Al final no sólo la salud de los equipos se desgasta, sino que también puede salir perjudicada la calidad de las entregas.

En agilidad **quienes deben estimar**, en caso de que se estime, **son quienes realizan la tarea**. Daría igual que las estimaciones las hiciera una IA o un departamento de gestores dedicados a ahorrarles este trabajo a los desarrolladores. Sólo que una estimación hecha por una IA tiene el peligro añadido de percibirse como una medida objetiva, debido a falacias de apelo a la autoridad y empiristas. Al estimar en equipo logramos que haya fertilización cruzada de ideas y una visión mejor alineada.

Cuando los productos y servicios que construimos tienen un alto componente creativo, como sucede en el caso del desarrollo de *software*, las estimaciones son inexactas por la naturaleza del trabajo. La cantidad de variables que influyen es infinita: los sistemas de código son complejos, los entornos cambian constantemente, salen versiones nuevas de las herramientas que requieren actualizaciones y lectura de documentación... Es habitual que una tarea que parece sencilla se convierta en un cuello de botella por algo imposible de predecir. Del mismo modo, a veces algo que parece complejo se soluciona en pocas horas.

Creación de historias de usuario

Tampoco es conveniente usar inteligencia artificial para automatizar la creación de historias de usuario y criterios de aceptación. Estamos viendo esta funcionalidad en grandes plataformas de gestión de proyectos, y parece venir de una continua falta de comprensión.

Las historias de usuario **son más que tarjetas con documentación escrita** en un formato reducido desde el punto de vista del usuario. El término "historia" se utilizó en origen para indicar que debían ser narradas, habladas. La nota que se escribe sirve para refrescar la memoria y recordar lo hablado. Es durante la conversación que nos aseguramos de estar entendiendo exactamente lo que quiere el cliente y alineamos la visión: podemos plantear preguntas, realizar esbozos, contrastar ideas...

Una IA puede ayudar dando sugerencias sobre cómo descomponer una tarea grande en otras más asumibles, o sugerir criterios de aceptación una vez acordada y entendida la historia. Pero **no debe reemplazar nunca la comunicación directa**, y todo cuanto la IA produzca debe revisarse para verificar su adecuación, realismo y calidad.

Asignación de trabajo

Hay herramientas de gestión de proyectos que ofrecen soluciones con inteligencia artificial para automatizar la asignación de tareas. Bien aprovechadas, estas herramientas podrían evitar la sobrecarga de trabajo y la subutilización de habilidades. Pero no es sencillo.

El uso de una herramienta así puede empujar en dirección contraria a la creación de equipos, y más de equipos multidisciplinares, donde las personas han de salir de su área de confort. Además, la efectividad dependerá, en gran medida, del tipo de datos con los que entrenemos el algoritmo y del nivel de aceptación que el sistema tenga entre el equipo.

No es una herramienta a implementar sin considerar pros y contras antes. En caso de optar por una IA para asignación de tareas, habrá que controlar y revisar sus decisiones y la satisfacción, para ajustar el modelo y confirmar si responde a nuestras necesidades. Si en lugar de aliviar la carga de trabajo excesiva la herramienta da como resultado un mayor estrés, por ejemplo, habría que reconsiderar su uso. O si desincentivara el desarrollo de los miembros del equipo, lo cual puede ser difícil de observar pero muy negativo.

Análisis de rendimiento

Los análisis de rendimiento automáticos a través de herramientas de *tracking* o monitorización de empleados conllevan riesgos similares a los de las estimaciones automatizadas. Utilizar los datos y capacidades de estas herramientas de forma irresponsable puede [erosionar la motivación de los equipos y generar dinámicas tóxicas, como competir en horas de trabajo o velocidad.](#)

Si necesitamos calcular costes de entrega u obtener datos que ayuden a la mejora continua, existen otros métodos.

El objetivo de cada iteración es producir el incremento acordado listo para usarse. Para medir la eficiencia del equipo y mejorar continuamente, podemos observar tendencias y si se están alcanzando los objetivos marcados de manera satisfactoria en las retrospectivas. Podemos obtener datos a través de sistemas de seguimiento de errores, encuestas de satisfacción al cliente, y otras fuentes que fomenten el aprendizaje y la colaboración. **En una cultura ágil se gestiona el trabajo, no a las personas.**

Soluciones para equipos

¿Sabes qué tareas roban más tiempo y energía a los desarrolladores? La automatización de tareas rutinarias puede ser una conversación interesante y un punto de partida. Aquí sólo vamos a dar algunas ideas, pero hay un amplio mercado de soluciones que se

pueden ajustar exactamente a cada proyecto y a cada persona. De hecho, es probable que muchos en el equipo ya estén utilizando herramientas con capacidades de IA.

Automatizar tareas repetitivas

- **Seguimiento de errores y solicitudes de funcionalidades:** una IA puede clasificar y priorizar *tickets* basándose en reglas predefinidas, ahorrando el trabajo de filtrado en atención al cliente.
- **Envío de correos comerciales y seguimiento con clientes potenciales:** herramientas como reply.io pueden aligerar la carga de trabajo del departamento de ventas. La IA ayuda a identificar clientes potenciales, redactar y mejorar los mensajes, e interpreta la intención de las respuestas para informar a simple vista sobre si el destinatario ha mostrado interés en lo que se le ofrece o no.

Recopilación y análisis de datos

- **Feedback de usuarios:** para extraer la reacción general hacia campañas de *marketing*, productos, o saber si alguna característica o mejora se está demandando con más insistencia que otras.
- **Transcripción y resúmenes de reuniones.** Las IA de este tipo empiezan a aparecer integradas en la mayoría de *software* para vídeo-conferencias. Pueden aprovecharse en equipos asíncronos, para miembros que no puedan asistir a reuniones, con el permiso de los asistentes.
- **Generación de reportes.**
- **Visualizar datos:** PowerBi (Microsoft) y Tableau son dos de las herramientas más populares. Tableau Pulse es la nueva versión con IA disponible a través de Tableau Cloud.

Onboarding y formación

- **Gestión de documentación:** usando la IA desde una perspectiva ágil, preferimos enfocar el esfuerzo en el resultado por encima de la documentación. En proyectos que, por sus circunstancias, requieren de documentación (mantenimiento de *software*, *onboarding*...), la IA encaja para generar la documentación necesaria sin perder el foco. Las herramientas pueden generar automáticamente explicaciones a partir del código, por ejemplo, y mantener actualizados documentos de proyecto.
- Hay muchos otros **términos específicos** de la empresa o el proyecto que pueden confundir a nuevas incorporaciones. Esta información, igual que la documentación del *software*, debería estar documentada y mantenerse actualizada. Algunas soluciones integradas de gestión de proyectos como Jira permiten centralizar esta información para buscarla rápidamente cuando un nuevo empleado lo necesita.
- **Un área en desarrollo es el de los entrenadores IA para nuevas incorporaciones.** Pueden personalizarse para simular situaciones reales, y que los

empleados que vayan a trabajar de cara al público puedan practicar en un entorno seguro. SecondNature es un ejemplo para ventas.

Mejora de productos y servicios

Estrategia de producto

La IA no sólo va a transformar la forma de analizar datos, sino también la forma de las experiencias de usuario. Cada vez se habla más de *customer-first AI*. Están apareciendo nuevas formas de ofrecer soporte a usuarios de forma automatizada, así como experiencias de compra personalizadas, guiadas por asistentes virtuales.

También hay nuevos productos y servicios que son posibles gracias a IA y pueden encajar dentro de nuestro producto, según el tipo de proyecto. Cabe reflexionar sobre el producto y qué áreas transformar. Podemos guiar los esfuerzos mediante analíticas para identificar áreas de fricción:

- ¿Qué áreas deben cambiar?
- ¿Qué áreas deben mantenerse?
- ¿Qué adiciones completamente nuevas son ahora posibles?

La IA permite entregar valor más rápido a los usuarios. Otras empresas ya están empezando a rediseñar sus *user journeys*, de forma que sea posible realizar preguntas sencillas en lenguaje natural para llegar a lo que se busca, sin necesidad de atravesar varios pasos y *clicks*, por ejemplo. Podemos plantearnos experiencias de usuario que antes no eran posibles (Pendo n.d.).

Ejemplos:

- Adobe o Canva han integrado en sus plataformas nuevas soluciones posibles gracias a los nuevos modelos de IA. En el caso de [Canva](#), las llaman herramientas “mágicas”: *Magic write*, *Magic edit*, *Magic design*... y sirven para agilizar y automatizar parte de la creación de contenido.
- [Stitchfix](#) integra IA en 3 áreas: análisis de *feedback* de usuarios, creación de textos descriptivos de productos, y creación de sugerencias de estilo personalizadas, un aspecto central de su negocio.
- [Duolingo Max](#) ofrece una nueva experiencia de aprendizaje de lenguaje que explica las respuestas al usuario y permite tener conversaciones más realistas.

Chatbots

Los *chatbots* y asistentes virtuales pueden proporcionar respuestas rápidas a consultas comunes basadas en una base de datos de conocimientos generales de la empresa o de los departamentos.

En caso de que el *chatbot* no sea capaz de responder la duda, eso quiere decir que se necesita la intervención de una persona. Así, la atención al cliente puede centrarse en las dudas que una FAQ no puede resolver.

Personalizar la experiencia del usuario

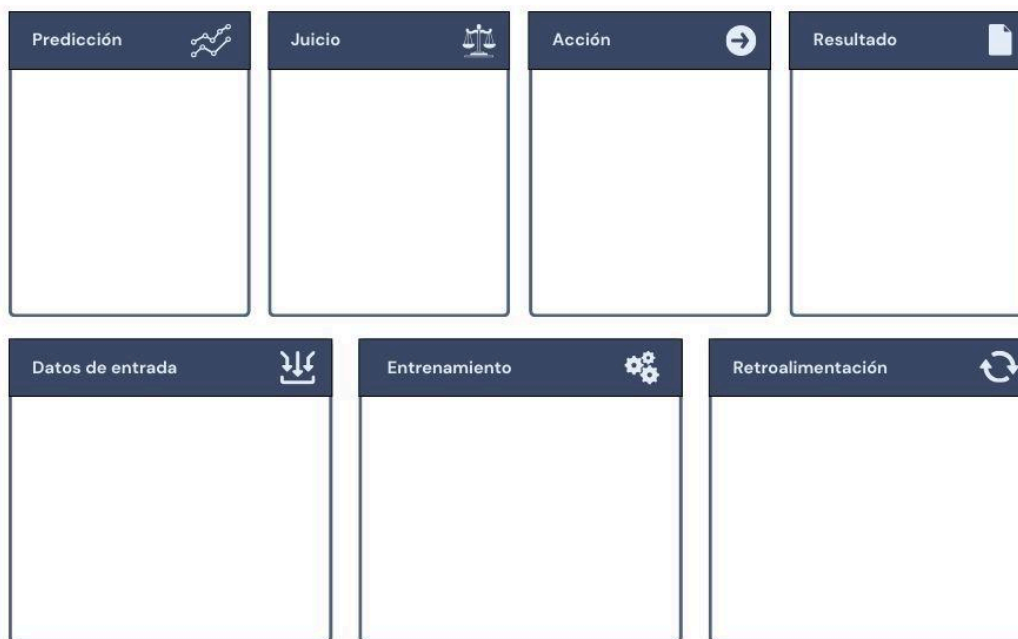
Los usuarios cada vez esperan más que los servicios y productos se adapten a sus necesidades y preferencias específicas. Algunas estrategias adicionales que considerar para la estrategia de producto incluyen:

- **Segmentación y recomendaciones personalizadas:** mediante ML no supervisado, los clientes pueden agruparse en *clusters* con similitudes relevantes pero que pueden no ser evidentes. Spotify, por ejemplo, utiliza esto para crear listas de reproducción como “descubrimiento semanal”. Estos *clusters* suelen aprovecharse, en general, para dar recomendaciones personalizadas a cada tipo de usuario.
- **Interfaces conversacionales y buscadores avanzados:** se pueden desarrollar asistentes virtuales más complejos que los *chatbots*, o entrenar buscadores con bases de datos para ayudar a los usuarios a encontrar lo que buscan cuando la oferta es muy amplia y depende de muchas variables. Por ejemplo, Newegg.com ofrece una herramienta llamada [PC Builder](#) optimizada con IA para ayudar a los clientes a componer su PC según sus necesidades y garantizando que todas las piezas sean compatibles.

Ejercicio. Planifica con un lienzo de IA

Vamos a aprender a utilizar un lienzo de IA, una herramienta propuesta en *Máquinas predictivas* (Agrawal A. 2019) para rediseñar procesos automatizando tareas. Para plantear, construir y evaluar herramientas de IA y si están cumpliendo con los objetivos deseados.

LIENZO DE IA



El lienzo de IA separa cada decisión en estos elementos:

- **Acción:** ¿qué está intentando hacer?
- **Predicción:** ¿qué necesita saber para tomar una decisión?
- **Juicio:** ¿cómo valora los diferentes resultados y errores?
- **Resultado:** ¿cuáles son sus parámetros del éxito de la tarea?
- **Datos de entrada:** ¿qué datos necesita para aplicar el algoritmo predictivo?
- **Entrenamiento:** ¿se usará el mismo tipo de datos que en los datos de entrada?
- **Retroalimentación:** ¿cómo podemos usar los resultados para mejorar el algoritmo?

Por ejemplo, supongamos que tenemos un negocio de venta o subasta en línea de ítems coleccionables. Pueden ser sellos. Un valor útil a mostrar en nuestro sitio web es el valor real estimado de cada ítem, para informar al usuario interesado sobre si lo que pide el vendedor está muy por encima o muy por debajo de ese precio, y también para guiar las ofertas de los vendedores. Estimar por cada sello a mano sería inviable, así que

decidimos entrenar un modelo. Apunta lo que crees que iría en cada sección del lienzo en este caso, y comprueba tus respuestas.

Invertir en talento

Ya hemos visto muchos ejemplos prácticos para implementar IA. También hemos señalado que, para una cultura ágil real, estas herramientas deben facilitar el trabajo de las personas del equipo. A continuación, hablaremos de qué nuevas habilidades se vuelven esenciales para crecer dentro del nuevo paradigma, cómo reclutar talento y mantenerlo.

Un consejo para captación y retención

Como ya vimos en el capítulo 3, ciertos usos de IA pueden ser desaconsejables e incluso incompatibles con una cultura ágil. Es posible que, en culturas no ágiles o debido al deseo de implementar estas soluciones y automatizar todo lo posible, estos usos se empiecen a extender.

Dado este contexto, no realizar este tipo de prácticas puede ser una gran ventaja competitiva en la captación y retención de talento.

Si queremos trabajar con personas hábiles y motivadas es clave que haya una confianza real en ellas y en sus capacidades. Al elegir entre varias opciones, gravitarán hacia proyectos donde se sientan realizadas, no controladas y evaluadas.

Perfiles asociados a la IA

Aunque este ámbito está evolucionando rápidamente, a día de hoy hay sobre todo dos tipos de perfiles profesionales especializados en IA:

- **Científico de datos:** el perfil clave que busca si necesitamos crear y entrenar modelos propios.
- **Ingeniero de aprendizaje automático (ML):** perfil especializado en despliegue y mantenimiento de los modelos en producción.

A continuación desglosamos **los conocimientos técnicos propios** de cada perfil, aunque puede haber excepciones o perfiles mixtos.

Científico de datos

Éstas son algunas de las habilidades esenciales para creación y entrenamiento de modelos de IA, y que encontramos en el perfil de científico de datos:

- **Python o R:** son los lenguajes de programación más relevantes para IA, junto con bibliotecas específicas como TensorFlow de Google y PyTorch de Meta (en el caso de *deep learning*) y Scikit-learn (en el caso de *machine learning*). R es un lenguaje algo más popular en ámbitos académicos que de negocio.

- **Análisis de datos:** para entender los *outputs* de los modelos de IA y entrenar los modelos se necesitan conocimientos de análisis estadístico y manipulación de datos. La manipulación incluye tareas como limpieza, transformación y visualización.
- **Interpretabilidad de modelos:** conocimientos sobre técnicas como LIME (*Local Interpretable Model-agnostic Explanations*) o SHAP (*SHapley Additive exPlanations*), para trazar y explicar cómo los modelos llegan a sus conclusiones.
- **Ingeniería de *features*:** habilidad para identificar, seleccionar y transformar las variables adecuadas para alimentar el modelo. Esto requiere no sólo conocimientos técnicos, sino también una comprensión profunda del dominio del problema e intuición.

Ingeniero de aprendizaje automático

Por otro lado, éstas son habilidades concretas para el despliegue y mantenimiento de sistemas de inteligencia artificial:

- **Python o R y bibliotecas asociadas:** este perfil también requiere conocimientos sobre estos lenguajes, para ajustes finos y adaptación de modelos a las necesidades específicas del proyecto. También se necesita poder trabajar con APIs de modelos de ML para integrarlos en aplicaciones existentes.
- **Docker y Kubernetes:** Docker se utiliza para encapsular los modelos de IA y sus dependencias, facilitando su despliegue y escalabilidad. Kubernetes ayuda a orquestar y gestionar estos contenedores en entornos de producción.
- **DevOps para integración y despliegue continuo (CI/CD):** capacidad para implementar prácticas de DevOps específicas para proyectos de ML, incluyendo la automatización de pruebas, la integración y el despliegue continuo de modelos de IA. Esto asegura actualizaciones fluidas y mantenimiento de los sistemas de ML.
- **Estadística y evaluación de modelos:** aunque menos enfocado que el científico de datos en la creación de modelos, el ingeniero de ML debe tener una sólida comprensión estadística para evaluar la eficacia y rendimiento de los modelos desplegados. Esto incluye comprender métricas de rendimiento, diagnóstico de problemas de sobreajuste o subajuste, y optimización de modelos para la producción.
- **Monitorización y mantenimiento de sistemas de ML:** herramientas y técnicas de monitorización del rendimiento de los modelos en producción, identificación de degradaciones o cambios en los patrones de datos que puedan requerir reentrenamiento. Esto es vital para asegurar la relevancia y precisión de los modelos a lo largo del tiempo.

Formación continua

La evolución del proyecto puede requerir ofrecer formación a los miembros actuales del equipo. Por ejemplo, si se usan herramientas de terceros conviene tener conocimientos sobre regulación, ética y seguridad, que son independientes del perfil técnico específico.

Comprensión de la IA y *Machine Learning* (ML)

No es necesario ser programador para conocer el funcionamiento de estos modelos. Existen muchos recursos, gratuitos y de pago, con formación básica de fundamentos de IA/ML y sobre cómo evaluar la efectividad y aplicabilidad de IA en escenarios de proyectos específicos. Si la IA va a formar parte de la estrategia del negocio, se debería ofrecer este tipo de formación a los equipos actuales que lo necesiten.

Pair programming y pruebas con IA

La IA está transformando el día a día de los programadores. Su uso se ha extendido rápidamente, y aunque diferentes herramientas tienen sus fortalezas y debilidades existe consenso en que son de gran ayuda tanto para escribir código nuevo como para testear y debuggear.

Ética y regulaciones de IA

Conocimiento de las regulaciones relevantes en IA como la GDPR, y comprensión de las implicaciones éticas y legales de desplegar este tipo de soluciones: sesgos, seguridad de datos, propiedad intelectual...

Seguridad específica para IA

Más allá de la seguridad de datos general, entender las vulnerabilidades específicas de los sistemas de IA (como ataques adversarios que manipulan modelos de ML) es una habilidad valiosa para proteger las soluciones y los datos.

Gestión de proyectos de datos

Habilidad para diseñar una estrategia de datos efectiva y establecer una gobernanza de datos que asegure la calidad, integridad y accesibilidad de los datos para proyectos de IA.

Automatización de procesos

Habilidad para automatizar tareas repetitivas mediante *scripts* o plataformas de automatización, como Zapier o Microsoft Power Automate.

Uso de plataformas AlaaS

Familiarizarse con plataformas de IA como servicio, que permiten implementar soluciones de IA sin construir modelos desde cero. Algunas de estas herramientas tienen su propia curva de aprendizaje, así como competidores entre los que elegir. Se necesita investigar para seleccionar, personalizar e integrar estos servicios dentro de los sistemas existentes.

Atención al cliente mediante *chatbots*

Algunos roles se están transformando significativamente con la IA. Si bien los *chatbots* y otras IA conversacionales reducen la necesidad de atender caso por caso a cada cliente, están apareciendo nuevas tareas relacionadas. Según el [informe de 2024 de Intercom](#), éstas incluyen:

- **Análisis de *chatbots*:** revisar las conversaciones que han tenido lugar para identificar oportunidades.
- **Diseño de conversaciones:** entrenar y corregir al *chatbot* para que dé respuestas correctas y en el tono deseado.
- **Mapeo de experiencia de usuario conversacional:** relacionado con lo anterior. Analizar las conversaciones para entrenar al *chatbot* de forma que las consultas se resuelvan de la manera más satisfactoria posible para el usuario.

Ejercicio. Planes de aprendizaje

Vamos a realizar un análisis de roles y competencias. Laura y Miguel son miembros de nuestro equipo, que va a empezar a desarrollar soluciones propias con IA. Laura, analista de datos, está interesada en evolucionar hacia un nuevo rol creando y entrenando modelos de IA para el proyecto. Miguel, ingeniero de *software*, está dispuesto a formarse para evolucionar hacia un nuevo rol como ingeniero de ML.

- **Perfil 1: Laura González. Analista de datos**
 - **Experiencia:** 3 años trabajando con grandes conjuntos de datos, análisis predictivo y visualización de datos en una empresa tecnológica. Ha liderado proyectos de análisis para optimizar procesos de negocio y tiene experiencia trabajando con equipos interdisciplinarios.
 - **Habilidades:** estadísticas, manejo de SQL y Python, especialmente con bibliotecas como Pandas y Matplotlib. Experiencia en la limpieza y preparación de datos, así como en la interpretación de tendencias y patrones.
 - **Formación:** Licenciatura en Matemáticas Aplicadas. Cursos complementarios en análisis de datos y ciencia de datos.

- **Competencias adicionales:** capacidad de trabajo en equipo, pensamiento crítico, y habilidades de comunicación para presentar hallazgos complejos de manera comprensible.
- **Perfil 2: Miguel Álvarez. Ingeniero de *software***
 - **Experiencia:** 5 años de experiencia en desarrollo de *software*, con un enfoque en el desarrollo de aplicaciones web y móviles. Ha colaborado en proyectos de *software* en todas las fases del ciclo de desarrollo, desde la concepción hasta el despliegue.
 - **Habilidades:** programación en Java, Python y JavaScript. Experiencia con *frameworks* de desarrollo y conocimientos básicos de contenedores y microservicios.
 - **Formación:** Grado en Ingeniería Informática. Cursos especializados en desarrollo de *software* y arquitecturas de aplicaciones.
 - **Competencias adicionales:** capacidad para resolver problemas complejos, trabajar en equipos multidisciplinares y adaptabilidad a nuevas tecnologías.

A continuación, **prepara la ruta de aprendizaje** para uno de ellos, que le facilite lograr sus objetivos en un plazo de tiempo razonable dadas sus habilidades actuales.

Tendencias futuras

Inteligencia artificial general

La inteligencia artificial general (AGI por sus siglas en inglés) se refiere a una máquina con la capacidad de entender, aprender y aplicar su inteligencia a cualquier problema intelectual, de manera similar a como lo haría un ser humano.

La mayoría de los modelos de inteligencia artificial que conocemos hoy en día son ejemplos de inteligencia artificial específica (IAE) o *narrow AI*: modelos diseñados y entrenados para realizar tareas muy específicas: reconocimiento de voz, traducción de idiomas, jugar ajedrez... Aunque son capaces de superar a humanos, sus capacidades están limitadas a un ámbito. Una IAG sería capaz de adaptarse a diferentes contextos y aprender de forma autónoma, llegando incluso quizá (se especula) a desarrollar niveles de autoconciencia o comprensión similares o superiores a los humanos.

Este curso habla de cómo aprovechar herramientas de IAE. En caso de que se llegue a desarrollar inteligencia artificial general, muchas de las consideraciones que hemos hecho no serán aplicables, y se abrirán nuevos desafíos e incógnitas.

No es posible predecir cuándo se alcanzará este objetivo, pero es un objetivo a largo plazo de la investigación en IA. En 2023 se publicó un *paper* que generó debate titulado “Destellos de AGI”: [Sparks of Artificial General Intelligence: Early experiments with GPT-4](#).

Ahora, con el interés de la industria puesto en modelos más avanzados y soluciones de *software* que actúan como agentes inteligentes, empieza a haber titulares que anuncian la llegada inminente de la AGI. La mayoría de expertos consideran que todavía estamos lejos.

Fuentes para mantenerse al día

Dada la popularidad del tema, no faltan fuentes a las que suscribirse: canales de divulgación como blogs y vídeo blogs, portales de noticias, canales especializados en redes sociales... Es importante tener en cuenta que la sobreinformación es casi tan dañina como la desinformación. En este momento existen unos tres tipos de fuentes a las que se puede acudir en busca de información respecto a la IA:

- La que proviene de investigadores independientes especializados en IA, que suelen publicar sus estudios en blogs propios, artículos, libros...
- La que proviene de las empresas desarrolladoras de IA.
- La que proviene de personas no especializadas, sino aficionadas o interesadas en el tema, a favor o en contra de estas tecnologías.

Ante la duda, el primer tipo de fuente suele ser el menos sesgado. En la bibliografía de este curso hemos incluido varios libros y artículos escritos por especialistas que no trabajan en empresas desarrolladoras de IA.

Teniendo esto en cuenta, hay varias fuentes que conviene conocer y que probablemente se mantengan relevantes a largo plazo:

- [Knowing Machines](#), un proyecto de investigación que analiza y explica cómo funcionan los sistemas de aprendizaje automático.
- **Google Scholar** puede ser una mejor fuente que el buscador estándar para búsquedas generales sobre IA aplicada a distintos ámbitos.
- El blog y los canales de divulgación de **IBM**.
- [arXiv.org \(sección de CS.AI\)](#): una plataforma para la publicación de *preprints* de investigaciones científicas que incluye los últimos trabajos en inteligencia artificial **antes** de que sean revisados.
- **MIT Technology Review**: ofrece artículos sobre últimos desarrollos en tecnología desde una perspectiva accesible pero informada. Su *newsletter* [“The Algorithm”](#) está especializada en IA.
- **Blogs** de las empresas cuyas soluciones utilizamos, así como de empresas líderes en IA como [DeepMind](#), [OpenAI](#) y [Google AI Blog](#). Si usamos sus herramientas, conviene seguir el blog y leer sus términos de uso cuando éstos se actualicen.
- [AI in Business de Emerj](#): se centra en aplicaciones prácticas de IA en diferentes industrias.

También hay proyectos abiertos con cursos sobre análisis de datos, estadística, programación aplicada a IA... Si nuestro proyecto va a usar o desarrollar este tipo de tecnologías, estar bien informados a nivel técnico es fundamental para tener conversaciones constructivas.

Ejercicio. Legislación

¿Sabrías decir qué leyes regulan el uso de la IA en tu país? ¿Sabes si hay iniciativas en curso o ya aprobadas que limiten o controlen el desarrollo y el uso de estas herramientas?

Busca información en línea si lo necesitas, y anota las leyes o proyectos de leyes para regulación IA aplicables en tu país. Considera si alguna de estas regulaciones es aplicable a tu proyecto. Si lo es, apunta 2 ó 3 ideas accionables para usar, desarrollar y mantener los sistemas de IA pertinentes cumpliendo con la normativa.

Ejemplo: en países de la Unión Europea, dos regulaciones relevantes son la GDPR (si vamos a recoger datos de usuarios) y el AI Act.

Bibliografía

Artículos y papers

Biswas, Som; Dobarra, Dushyant; L. Cohen, Harris (2023). "ChatGPT and the Future of Journal Reviews: A Feasibility Study", *Yale Journal of Biology and Medicine*:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10524821/>

Bubeck et al. (2023). *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. Recuperado de <https://arxiv.org/abs/2303.12712>.

Delgado, Paulette (2023). "El estrés y la ansiedad pueden provocar deterioro cognitivo", *Tec.mx Observatorio*. (n.d.). Recuperado de <https://observatorio.tec.mx/edu-news/el-estres-y-la-ansiedad-pueden-provocar-deterioro-cognitivo/>

Duolingo Team (2023) "Conoce Duolingo Max, una experiencia de aprendizaje potenciada con GPT-4", *Duolingo Blog*. Recuperado de <https://blog.duolingo.com/es/duolingo-max-experiencia/>

Guo X; Dong L; Hao D (2024). "Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway." *Front. Cell Dev. Biol*:
<https://cdn.arstechnica.net/wp-content/uploads/2024/02/fcell-11-1339390-1.pdf>

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. Recuperado de <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

Novoa et al. (2018). "Automated Classification of Skin Lesions: From Pixels to Practice", *Journal of Investigative Dermatology*, Volume 138, Issue 10, pp. 2108-2110.
<https://www.sciencedirect.com/science/article/pii/S0022202X18322930>

Polo, J. D. (2024). "ChatGPT para estructurar datos desordenados. IA para el uso práctico", *LinkedIn*. Recuperado de <https://www.linkedin.com/pulse/chatgpt-para-estructurar-datos-desordenados-juan-diego-polo-lpz8f/>

Russell, Josh (2023). "Sanctions ordered for lawyers who relied on ChatGPT artificial intelligence to prepare court brief", *Courthouse News Service*:
<https://www.courthousenews.com/sanctions-ordered-for-lawyers-who-relied-on-chatgpt-artificial-intelligence-to-prepare-court-brief/>

Safe Creative

- (2024). “El New York Times demanda a Microsoft y OpenAI”, *TIPS*: <https://www.safecreative.org/tips/es/el-new-york-times-demanda-a-microsoft-y-openai/>
- (2024). “GPT4, el modelo de IA que más infringe los derechos de autor”, *TIPS*: <https://www.safecreative.org/tips/es/gpt4-el-modelo-de-ia-que-mas-infringe-los-derechos-de-autor/>

Salehi, F., Aghaie, A., & Ghapanchi, A. H. (2023). “Role of artificial intelligence on agile planning and organizational performance in the ICT industry”. *ResearchGate*.

Recuperado de

https://www.researchgate.net/publication/368381945_Role_of_Artificial_Intelligence_on_Agile_Planning_and_Organizational_Performance_in_the_ICT_Industry

StitchFix (2023). “How We’re Revolutionizing Personal Styling with Generative AI”, *StitchFix*:

<https://newsroom.stitchfix.com/blog/how-were-revolutionizing-personal-styling-with-generative-ai/>

Libros

Agrawal, A., Gans, J., & Goldfarb, A. (2019). *Máquinas predictivas: La sencilla economía de la inteligencia artificial*. Editorial Reverté, S.A.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... & Thomas, D. (2001). *Manifiesto por el desarrollo ágil de software*. Recuperado de <http://agilemanifesto.org/iso/es/manifesto.html>

Christian, B. (2020). *The Alignment Problem: Machine Learning and Human Values*. W.W. Norton & Company.

Iansiti, M., & Lakhani, K. R. (2020). *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Harvard Business Review Press.

Lencioni, P. (2002). *Las 5 disfunciones de un equipo: Un inteligente modelo para formar un equipo*. Jossey-Bass.

Vídeos

Daugherty, Paul

- (2019). “Human + Machine: Reimagining Work in the Age of AI | Paul Daugherty”, Canal U-M Computer Science and Engineering: <https://www.youtube.com/watch?v=Qfpmggpsdtk>

- (2019). “Paul Daugherty: Human + Machine: Reimagining Work in the Age of AI”, Canal Northwestern Engineering:
<https://www.youtube.com/watch?v=qtEUHcz-Fo4>

Cursos y sitios web

Intercom. (2024). *The Intercom Customer Service Trends Report 2024*. Recuperado de <https://www.intercom.com/campaign/customer-service-trends-2024>

Pendo. (n.d.). *AI for Product Management*. Recuperado de <https://www.productledcertified.com/ai-for-product-management>

Wikipedia contributors. (n.d.). *Computing Machinery and Intelligence*. Recuperado de https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence

Wikipedia contributors. (n.d.). *Turing test*. Recuperado de https://en.wikipedia.org/wiki/Turing_test

Apéndices

Cómo se entrena un modelo de ML supervisado

Nuestro *data frame* debe tener bien definidos un objetivo a predecir y predictores que creemos influyen en éste. Al objetivo también se le llama “variable dependiente” o *target*, y a los predictores “variables independientes” o *features*.

Paso 1. Recopilar y preparar la información

Deberemos **reunir datos históricos relevantes**. En principio, cuanto más detallados y completos sean los datos, mejor será el modelo. Pero elegir las características adecuadas a incluir en el modelo es un asunto delicado. Una cosa es contexto y otra es ruido.

Una vez recopilados, es necesario **limpiar y preprocesar los datos**. Esto incluye manejar valores faltantes, eliminar *outliers* y normalizar o estandarizar.

Para que el modelo sea capaz de interpretar los datos y extraer patrones, las variables deben ser cuantitativas o convertirse a un valor numérico. Hay muchas técnicas para limpiar datos cualitativos o categóricos: *label encoding*, codificación *one-hot*...

Más información sobre limpieza de datos: [Limpieza de datos en tandas explicado con ejemplos \(FreeCodeCamp\)](#).

El conjunto de datos se suele dividir en dos partes: datos para el entrenamiento y datos para pruebas. Si el modelo de IA fuese un alumno que se prepara para realizar un examen, los datos de entrenamiento serían los exámenes de años pasados y los datos para pruebas un examen nuevo jamás visto. Al aplicar el modelo sobre los datos para pruebas, podemos evaluar si es capaz de aplicarse a nuevas situaciones.

Paso 2. Pruebas y entrenamiento

No existe un único sistema de ML, sino que hay varios modelos algorítmicos. La selección del modelo depende de la complejidad de los datos y el tipo de estimaciones que necesitemos realizar. Es habitual probar varios modelos y evaluar para ver cuál da mejores resultados.

Posibles modelos, del más sencillo a más complejo:

- **Regresión lineal:** cuando la relación entre el objetivo y los predictores parece ser lineal. Es un modelo simple, fácil de implementar e interpretar.
- **Regresión polinomial:** para relaciones más complejas. Se modela aumentando el grado de ecuación polinomial y puede conducir a sobreajuste si el polinomio es demasiado alto.

- **Árboles de decisión:** se suele usar para predecir las consecuencias de decisiones. Es útil en la predicción basada en decisiones categóricas o continuas. Resulta fácil de interpretar y permite capturar no linealidades, pero es propenso al sobreajuste cuando los árboles son muy profundos.
- **Bosques aleatorios:** es una mejora de los árboles de decisión, útil para mejorar la precisión de la predicción y controlar el sobreajuste. Consiste en ensamblar múltiples árboles, importante para datos complejos y de alta dimensionalidad. Estos modelos empiezan a ser más difíciles de interpretar, y suponen mayor carga computacional.
- **Redes neuronales:** permiten modelar relaciones extremadamente complejas y no lineales entre variables. Son muy escalables y excelentes para grandes volúmenes de datos, pero difíciles de interpretar y requieren de gran cantidad de datos para su entrenamiento. Al igual que los bosques aleatorios, la carga computacional es alta, y además son propensas al sobreajuste. Cuando una red neuronal tiene más de 3 capas, entramos en lo que se conoce como *deep learning*.

Las características elegidas, el número de iteraciones, cómo de profundo es el árbol de decisiones... todo esto afecta a los resultados y hay que experimentar. Es similar a ajustar los diales de una radio hasta encontrar una señal clara, y al proceso se le llama ajuste de hiperparámetros (*hyperparameter tuning*). Se realizan pruebas y ajustes sobre diferentes modelos y con *subsets* de datos hasta dar con resultados óptimos.

El proceso puede agilizarse con herramientas como **AutoML**. Es recomendable sobre todo para perfiles con conocimiento de estadística pero poca base de programación.

Paso 3. Implementación

Territorio de los ingenieros de ML. Con el tiempo, aparecen desviaciones o fallos en el modelo, a lo que suele referirse como **model drift**. Es un problema muy frecuente, por lo que parte del trabajo de los ingenieros es ver si el modelo se está desviando, porque entonces tus predicciones ya no sirven y hay que reentrenar el modelo.

