RED educativa digital escartes

COLECCIÓN
¿CÓMO FUNCIONA?

¿CÓMO FUNCIONA EL APRENDIZAJE AUTOMÁTICO?

Descubre los principios, algoritmos y aplicaciones del aprendizaje automático, la tecnología que permite a las máquinas aprender de los datos y tomar mejores decisiones cada día.

PRINCIPIOS CLAVE
Conceptos fundamentales explicados de forma clara.

ALGORITMOS ESENCIALES
Modelos y técnicas que impulsan el aprendizaje automático.

APLICACIONES REALES
Casos de uso que están transformando el mundo.

Juan Guillermo Rivera Berrío

¿Cómo funciona el aprendizaje automático?

Juan Guillermo Rivera Berrío

Fondo Editorial RED Descartes

Córdoba (España)
2026

¿Cómo funciona el aprendizaje automático?

Autor:
Juan Guillermo Rivera Berrío

Código JavaScript para el libro: Joel Espinosa Longi, IMATE, UNAM.
Recursos interactivos: DescartesJS, Herramientas de IA, ChatGPT y Gemini.
Fuentes: Lato y UbuntuMono
Imagen de portada: ilustración generada por GPT Image 2

Red Educativa Digital Descartes
Córdoba (España)
descartes@proyectodescartes.org
https://proyectodescartes.org

Proyecto iCartesiLibri
https://proyectodescartes.org/iCartesiLibri/index.htm

ISBN: 978-84-10368-64-4

Tabla de contenido

Prefacio

Vivimos en una época en la que la inteligencia artificial ha dejado de ser una tecnología reservada para laboratorios de investigación y grandes empresas para convertirse en una herramienta presente en la educación, la medicina, la industria, las finanzas, el entretenimiento y prácticamente todos los ámbitos de la vida cotidiana. Sin embargo, detrás de conceptos como aprendizaje automático, redes neuronales o modelos de inteligencia artificial existen principios científicos y matemáticos que con frecuencia parecen complejos o inaccesibles para quienes se acercan por primera vez a este fascinante campo.

Este libro nace precisamente con el propósito de acercar ese conocimiento a cualquier lector curioso.

¿Cómo funciona el aprendizaje automático? forma parte de la colección "**¿Cómo funciona?**", una serie de libros concebida para explicar, de manera clara, rigurosa y visual, el funcionamiento de tecnologías, fenómenos científicos, disciplinas y procesos que transforman nuestro mundo. Cada volumen de la colección parte de una pregunta sencilla —"**¿Cómo funciona...?**"— y la responde mediante explicaciones progresivas, ejemplos cotidianos, ilustraciones, recursos interactivos y un lenguaje accesible, sin renunciar al rigor científico.

La filosofía de esta colección consiste en ir más allá de la simple descripción de los conceptos. Nuestro objetivo es revelar los mecanismos internos que hacen posible cada tecnología, comprender por qué funciona y descubrir cómo sus diferentes componentes interactúan entre sí. Creemos que entender el funcionamiento de las cosas es el primer paso para utilizarlas de forma crítica, creativa y responsable.

En este libro, el lector recorrerá el camino que sigue una máquina para aprender a partir de los datos. Descubrirá cómo los algoritmos

identifican patrones, cómo se entrenan los modelos, cuáles son las diferencias entre el aprendizaje supervisado, no supervisado y por refuerzo, cómo funcionan las redes neuronales, de qué manera se evalúa el rendimiento de un modelo y cuáles son las aplicaciones que hoy están transformando múltiples sectores de la sociedad. Finalmente, explorará las tendencias que marcarán el futuro del aprendizaje automático y los desafíos científicos, tecnológicos y éticos que acompañarán su evolución.

Como todos los libros de la colección "**¿Cómo funciona?**", esta obra incorpora abundantes recursos visuales e interactivos que enriquecen la experiencia de aprendizaje. Las ilustraciones, diagramas, simulaciones y actividades permiten experimentar con los conceptos, facilitando una comprensión más profunda y favoreciendo un aprendizaje activo. El objetivo no es únicamente leer sobre el aprendizaje automático, sino también observarlo, explorarlo e interactuar con él.

Esperamos que este libro despierte nuevas preguntas, estimule la curiosidad científica y motive a seguir aprendiendo. Porque comprender **cómo funciona** una tecnología es también comprender mejor el mundo que estamos construyendo y prepararnos para participar activamente en su transformación.

Bienvenido a la colección "**¿Cómo funciona?**".

Esperamos que este sea solo el primero de muchos viajes para descubrir, comprender y disfrutar el extraordinario funcionamiento de la ciencia y la tecnología.

Juan Guillermo Rivera Berrío

2026

¿CÓMO FUNCIONA EL APRENDIZAJE AUTOMÁTICO?

Sistemas inteligentes que aprenden de los datos para tomar decisiones y mejorar con la experiencia.

PROPÓSITO:
DIVULGAR

Hacer comprensible la ciencia detrás del aprendizaje automático.

1 DATOS DE ENTRADA

- Se recopilan datos del mundo real.
- Pueden ser imágenes, textos, números, sonidos, etc.
- La calidad de los datos determina el aprendizaje.

2 PREPARACIÓN DE DATOS

- Los datos se limpian y organizan.
- Se transforman en formatos que el modelo puede entender.
- Se dividen en entrenamiento y prueba.

3 ENTRENAMIENTO DEL MODELO

- El algoritmo aprende patrones reconociendo relaciones.
- Ajusta parámetros para minimizar errores.
- Requiere muchos ejemplos.

4 EVALUACIÓN Y AJUSTE

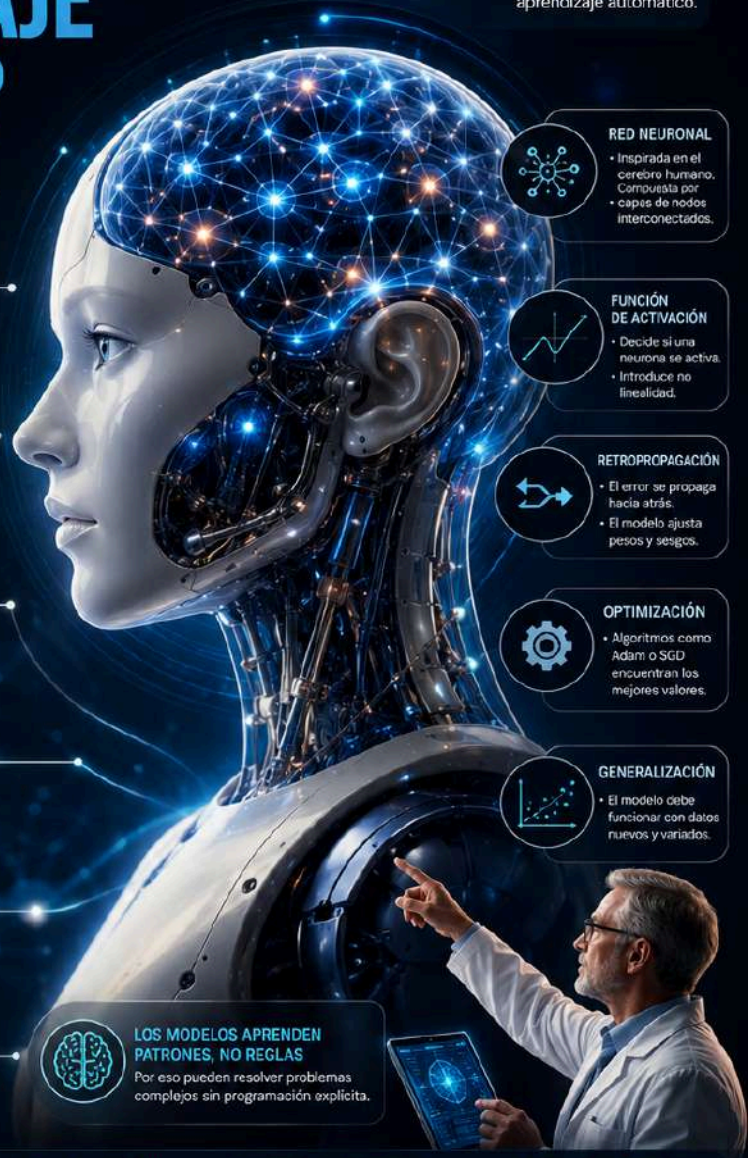
- Se prueba el modelo con datos nuevos (no vistos).
- Se mide su precisión.
- Se ajusta para mejorar resultados.

5 PREDICCIÓN E INFERENCIA

- El modelo usa lo aprendido para hacer predicciones.
- Puede clasificar, recomendar, detectar o predecir valores.

6 APRENDIZAJE CONTINUO

- El modelo mejora con nuevos datos.
- Se adapta a cambios y nuevos patrones.
- Aprender nunca termina.



RED NEURONAL

- Inspirada en el cerebro humano. Compuesta por
- capas de nodos interconectados.

FUNCIÓN DE ACTIVACIÓN

- Decide si una neurona se activa.
- Introduce no linealidad.

RETROPROPAGACIÓN

- El error se propaga hacia atrás.
- El modelo ajusta pesos y sesgos.

OPTIMIZACIÓN

- Algoritmos como Adam o SGD encuentran los mejores valores.

GENERALIZACIÓN

- El modelo debe funcionar con datos nuevos y variados.

LOS MODELOS APRENDEN PATRONES, NO REGLAS

Por eso pueden resolver problemas complejos sin programación explícita.

DATOS RÁPIDOS

60%+ del rendimiento depende de la calidad de datos.

10^3–10^9 ejemplos pueden ser necesarios para entrenar modelos.

TIPOS DE APRENDIZAJE

- Supervisado (con etiquetas)
- No supervisado (sin etiquetas)
- Por refuerzo (basado en recompensas)

APLICACIONES CLAVE

- Visión por computadora
- Procesamiento de lenguaje
- Recomendaciones
- Detección de fraudes
- Salud y diagnósticos

MODELO VS HUMANO

VS

Aprende más rápido en datos masivos.

Entiende contexto, emociones y valores.

MEJORAS CONTINUAS

El rendimiento mejora con más datos y mejor entrenamiento.

DATOS → ENTRENAMIENTO → EVALUACIÓN → PREDICCIÓN → MEJORA CONTINUA

APRENDER. ADAPTARSE. EVOLUCIONAR. EL PODER DE LOS DATOS, LA INTELIGENCIA DEL FUTURO.

Introducción

El **aprendizaje automático** (o Machine Learning) es una rama de la inteligencia artificial que transforma la manera en que las computadoras resuelven problemas complejos. A diferencia de la programación tradicional, donde un desarrollador escribe reglas lógicas explícitas paso a paso, el aprendizaje automático permite que los sistemas identifiquen patrones de manera autónoma a partir de datos históricos. En esencia, el **principio de funcionamiento** busca encontrar una función matemática desconocida f que mapee un conjunto de variables de entrada X hacia una salida deseada Y:

$$Y = f(X)$$

Figura 1. El futuro aprende de los datos.

Mediante esta formulación, el algoritmo adapta sus propios parámetros para aproximar la función f sin ser programado específicamente para cada escenario.

Para lograr esta adaptación, los **procesos internos** del aprendizaje automático se estructuran en torno a tres **componentes principales**: los datos de entrenamiento, un modelo predictivo y una función de costo o pérdida. Durante la fase de entrenamiento, el modelo toma las entradas y genera una estimación $\hat{y}$. A continuación, el sistema evalúa qué tan distante está su predicción del valor real y mediante la función de pérdida, como el error cuadrático medio:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

A través de un algoritmo de optimización (como el descenso de gradiente), el modelo ajusta recursivamente sus parámetros internos θ para minimizar el valor de $L(\theta)$, mejorando gradualmente su precisión a medida que procesa más información.

Este flujo de trabajo ha dado lugar a múltiples **aplicaciones prácticas** y **casos reales** que utilizamos a diario. Entre los ejemplos más destacados se encuentran:

Sistemas de recomendación: Plataformas como Netflix o Spotify analizan patrones de consumo masivos para predecir qué contenido ofrecer a cada usuario.

Diagnóstico médico: Algoritmos de visión por computadora procesan imágenes médicas (como resonancias o radiografías) para detectar patologías en etapas tempranas.

Procesamiento de lenguaje natural: Asistentes virtuales y traductores automáticos interpretan el contexto gramatical para generar respuestas coherentes.

Entre sus principales **beneficios y ventajas**, destaca la capacidad de automatizar la toma de decisiones en entornos altamente dinámicos y procesar grandes volúmenes de datos en múltiples dimensiones, algo imposible de realizar manualmente. Además, los modelos son capaces de adaptarse a nuevos datos de forma continua, mejorando su desempeño con el tiempo sin necesidad de reescribir la arquitectura base del software.

No obstante, esta tecnología presenta **desafíos y limitaciones** significativos. Los modelos son altamente dependientes de la calidad de los datos recibidos; si los datos de entrenamiento contienen sesgos o errores, el algoritmo replicará y amplificará esas deficiencias. Asimismo, existe el riesgo del **sobreajuste** (*overfitting*), un fenómeno donde el modelo memoriza perfectamente los datos de entrenamiento pero pierde la capacidad de generalizar ante datos nuevos, sumado a la falta de interpretabilidad en modelos complejos, comúnmente conocidos como "cajas negras".

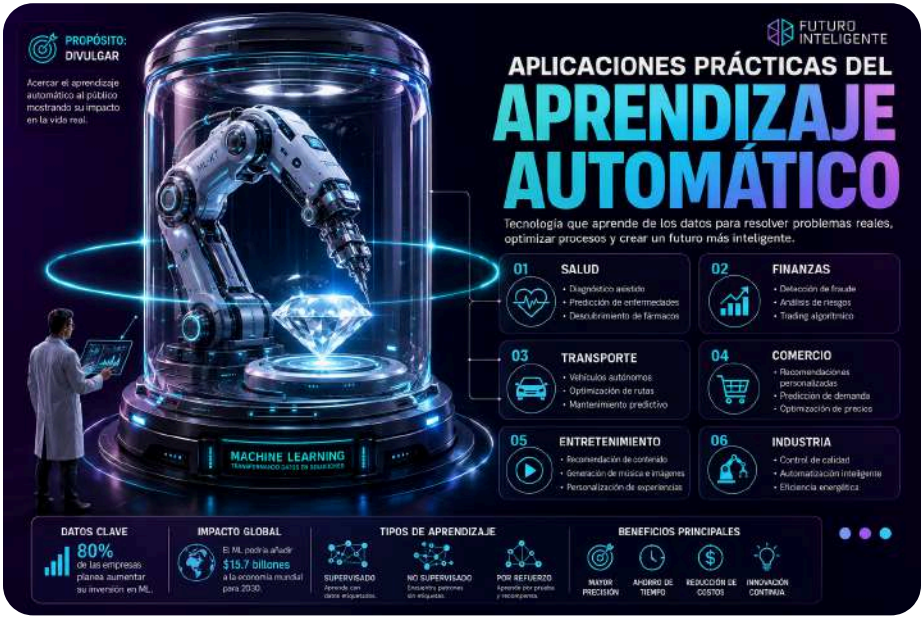


Figura 2. Aplicaciones prácticas del aprendizaje automático.

¿Cómo funciona el aprendizaje de las máquinas a partir de datos?

El aprendizaje de las máquinas a partir de datos se basa en el principio de que los sistemas informáticos pueden identificar patrones, extraer conocimiento y tomar decisiones de forma autónoma sin ser programados explícitamente para cada regla. En lugar de seguir instrucciones fijas del tipo "si ocurre A, haz B", el modelo analiza un conjunto de datos de entrada X para aproximar una función matemática desconocida f que conecta dichos datos con un resultado deseado y. Esta relación fundamental se representa conceptualmente como: $y = f(X) + \epsilon$, donde ϵ representa el error impredecible o ruido inherente a los datos reales.

Figura 3. Así funciona el aprendizaje de las máquinas a partir de datos.

Figura 4. Las máquinas aprenden identificando patrones en los datos.

A medida que el algoritmo procesa más ejemplos, ajusta paulatinamente sus parámetros internos para reducir este margen de error y capturar la verdadera estructura subyacente de la información. Para llevar a cabo esta transformación, el proceso interno de aprendizaje depende de tres componentes principales: los datos de entrenamiento (características X y etiquetas y), la **función de pérdida** ($\mathcal{L}$) y el **algoritmo de optimización**. La función de pérdida mide la discrepancia entre la predicción del modelo, denotada como $\hat{y}$, y el valor real y. Para corregir las desviaciones, se utiliza un método iterativo como el descenso del gradiente, el cual actualiza paulatinamente los parámetros del modelo (θ) mediante la siguiente expresión:

$$\theta_{t+1} = \theta_t - \alpha \nabla \mathcal{L}(\theta_t)$$

Donde α es la tasa de aprendizaje y $\nabla \mathcal{L}(\theta_t)$ es el gradiente de la pérdida. Mediante este ciclo continuo de predicción, evaluación de error y ajuste matemático, el sistema "aprende" progresivamente a minimizar el fallo global.

Este enfoque impulsado por datos ofrece valiosas **ventajas**, como la adaptabilidad a entornos cambiantes y la capacidad de descubrir patrones complejos no evidentes para el ojo humano. Entre sus **aplicaciones prácticas y casos reales** más destacados se encuentran:

Diagnóstico médico por imagen: Análisis de miles de tomografías para detectar anomalías o tumores con alta precisión.

Filtros de correo no deseado: Clasificación automática de mensajes analizando la frecuencia de palabras y la estructura del texto.

Predicción del valor inmobiliario: Estimación del precio de una vivienda en función de características como la ubicación, la superficie en metros cuadrados y el número de habitaciones.

A pesar de su enorme potencial, la extracción de conocimiento a partir de datos presenta **desafíos y limitaciones** críticos. El problema más recurrente es el sobreajuste (overfitting), que ocurre cuando el modelo memoriza en exceso el ruido del conjunto de entrenamiento y pierde la capacidad de generalizar ante datos nuevos. Asimismo, se cumple el principio fundamental de "garbage in, garbage out": si los datos de origen están sesgados, incompletos o mal etiquetados, el sistema aprenderá y perpetuará dichos errores de forma sistemática. Por ello, la calidad y el preprocesamiento de los datos resultan tan determinantes como la sofisticación del propio algoritmo.

Figura 5. Las máquinas aprenden a partir de datos.

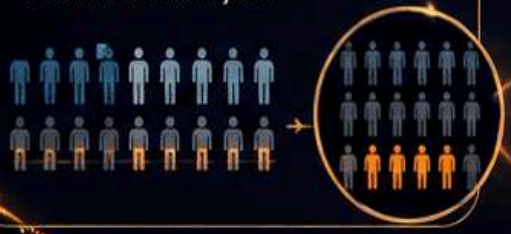
SI LOS DATOS DE
SESGADOS, IN
MAL ÉTIQ

EL SISTEMA APRENDE
DICHOS ERRORES DE F

FUNCIONES CLAVE • ADAPTACIONES •
IMPORTANCIA • CARAC

01 DATOS SESGADOS

- Reflejan prejuicios humanos
- Excluyen grupos o realidades
- Generan decisiones injustas



02 DATOS INCOMPLETOS

- Falta de contexto
- Huecos de información
- El modelo asume patrones incorrectos

03 DATOS MAL ETIQUETADOS

- Etiquetas incorrectas
- Confunden al modelo
- Aprende relaciones falsas

La calidad de la IA d
de la c.alidad de los
Datos limpios, decisio

DATOS RÁPIDOS

70% de los proyectos de IA fallan por datos deficientes.

1 de cada 3 sistemas de IA muestra sesgos significativos.

Hasta 80% más costoso corregir datos después del desarrollo.

COMPARATIVA DE IMPACTO

DATOS DE CALIDAD

- Decisiones justas
- Modelos confiables
- Impacto positivo

VS

DATOS

E ORIGEN ESTÁN COMPLETOS O QUETADOS,

ERÁ Y PERPETUARÁ FORMA SISTEMÁTICA.

ESTRUCTURA • COMPORTAMIENTO
CTERÍSTICAS ÚNICAS

epende
datos.
nes justas.

04 APRENDIZAJE EQUIVOCADO

- El modelo internaliza errores
- Ajusta patrones incorrectos
- Confirma sesgos existentes

DATOS DEFECTUOSOS → MODELO → PATRONES

05 ERRORES SISTEMÁTICOS

- Se repiten en cada predicción
- Impacto a gran escala
- Difíciles de detectar

ENTRADA → PREDICCIÓN → DECISIÓN → IMPACTO

06 PERPETUACIÓN DEL ERROR

- Retroalimentación negativa
- Refuerza los mismos fallos
- Normaliza la injusticia

ERRORES → RESULTADOS → NUEVOS DATOS → MÁS ERRORES

S DEFICIENTES
Decisiones sesgadas
Modelos inestables
mpacto negativo

TIPOS DE ERRORES COMUNES

Sesgo de selección

Datos faltantes

Etiquetas incorrectas

Medición errónea

Contexto ignorado

PROPÓSITO

DIVULGAR

el aprendizaje de las máquinas a partir de datos

Una aventura interactiva

Proyecto de Ciencias

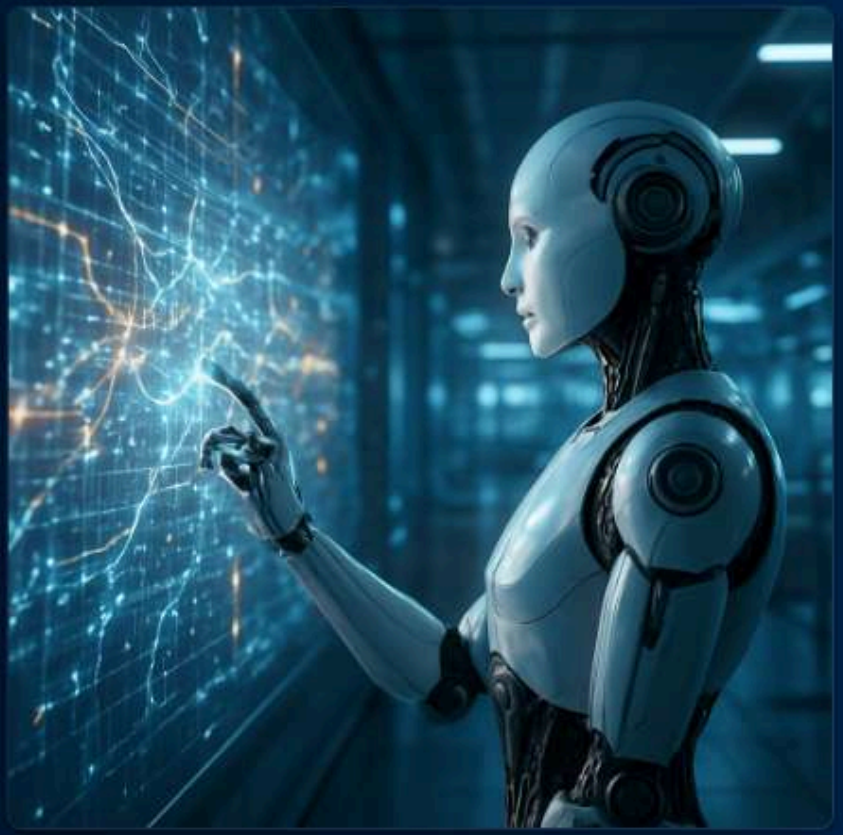
Ana:
¡No puedo creer cuánto texto tienen estos libros! Necesito una forma de aprender rápidamente todo esto.

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

IA que aprende



¿Cómo funciona el aprendizaje supervisado?

El **aprendizaje supervisado** es uno de los paradigmas fundamentales del aprendizaje automático, cuyo **concepto fundamental** se asemeja al proceso de enseñanza humana guiada por un tutor. En este enfoque, el algoritmo aprende a partir de un conjunto de datos previamente etiquetados, donde cada vector de entrada X está explícitamente vinculado a una respuesta o etiqueta objetivo y. El **principio de funcionamiento** radica en descubrir una función matemática implícita capaz de mapear estas entradas con sus correspondientes salidas, permitiendo realizar predicciones acertadas sobre datos futuros nunca antes vistos.

Figura 6. ¿Cómo funciona el aprendizaje supervisado?

Figura 7. El aprendizaje supervisado es un paradigma del aprendizaje automático.

Esta relación entre la variable dependiente y las independientes se expresa formalmente como:

$$y = f(X) + \epsilon$$

Donde $f(X)$ representa la función objetivo que el modelo intenta aproximar, X son las características de entrada, y es la salida real y ϵ simboliza el término de error no reducible o ruido aleatorio de la muestra.

Dentro de sus **componentes principales** se encuentran las variables de entrada o características (**features**), las etiquetas objetivo (**targets**), los parámetros optimizables (como los pesos w y el sesgo b) y la función de costo. Durante los **procesos internos** de aprendizaje, el algoritmo evalúa constantemente la discrepancia entre las predicciones realizadas ($\hat{y}$) y los valores verdaderos (y). En problemas de regresión numérica, un método habitual para medir este margen de error es el **Error Cuadrático Medio** (MSE):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

A través de métodos numéricos de optimización, como el **gradiente descendente**, el sistema ajusta iterativamente sus parámetros internos en la dirección opuesta al gradiente del error, logrando reducciones progresivas hasta converger en un punto de mínima pérdida.

Para garantizar la precisión de este proceso, el flujo de trabajo estándar en el aprendizaje supervisado se divide en las siguientes etapas clave:

Preparación y etiquetado: Recolección, limpieza y asignación rigurosa de las etiquetas (*ground truth*) a los datos de entrenamiento.

División del conjunto de datos: Separación de la muestra en conjuntos de entrenamiento, validación y prueba (frecuentemente en proporciones 80/20 o 70/30).

Fase de entrenamiento e inferencia: Optimización de los parámetros del modelo utilizando el conjunto de entrenamiento y posterior evaluación en el conjunto de prueba para medir la generalización.

En cuanto a sus **aplicaciones prácticas y casos reales**, el aprendizaje supervisado está detrás de tecnologías cotidianas como los filtros de spam en el correo electrónico, el diagnóstico de patologías mediante visón por computadora en radiografías, la estimación del precio de bienes raíces y los sistemas de reconocimiento de voz. Sus principales **beneficios y ventajas** incluyen un alto nivel de precisión en tareas bien delimitadas, la capacidad de cuantificar con certeza el margen de error del modelo y un comportamiento altamente predecible cuando se le alimenta con datos de alta calidad.

No obstante, este paradigma enfrenta **desafíos o limitaciones** considerables. El obstáculo más apremiante es la alta dependencia de grandes volúmenes de datos etiquetados, proceso que suele ser costoso, lento y sujeto a sesgos humanos. Asimismo, existe el riesgo constante del **sobreajuste** (*overfitting*), en el cual el modelo memoriza el ruido de los datos de entrenamiento perdiendo capacidad de abstracción, o del **subajuste** (*underfitting*), que ocurre cuando el algoritmo es demasiado simple para capturar la complejidad subyacente de la información.

APRENDIZAJE SUPERVISADO

LA INTELIGENCIA QUE APRENDE DE TUS EJEMPLOS

De spam a radiografías:
clasifica, predice y reconoce
usando datos etiquetados.

Aprendizaje supervisado

Una aventura interactiva

Encuentro con un Misterio

Ana:
¡Profe! ¿Cómo sabe la computadora si esta foto es de un perro o un gato?

¿Qué es el Aprendizaje Supervisado?


Generador de Iceberg Pro
Nivel 1: La Superficie
Nivel 1
IA con Tutor

¿Cómo funciona el aprendizaje no supervisado?

El **aprendizaje no supervisado** es una rama fundamental del aprendizaje automático donde el algoritmo se entrena utilizando conjuntos de datos que carecen de etiquetas o respuestas predefinidas. A diferencia del modelo supervisado, aquí el sistema no intenta predecir un resultado específico y, sino que explora la estructura intrínseca del conjunto de datos $X = \{x_1, x_2, \ldots, x_n\}$ para descubrir patrones ocultos, agrupaciones o relaciones subyacentes. Su **principio de funcionamiento** se apoya en la autoorganización de la información: el algoritmo evalúa continuamente métricas de similitud y distancia matemática entre los datos para construir representaciones estructuradas sin la intervención de un tutor humano.

Figura 8. ¿Cómo funciona el aprendizaje no supervisado?

Figura 9. El aprendizaje no supervisado es una rama del aprendizaje automático.

Entre sus **componentes principales** y **procesos internos**, el agrupamiento o clustering destaca como una técnica primordial, siendo el algoritmo K-Means uno de los más extendidos. Durante su ejecución, el modelo distribuye iterativamente cada observación x_i hacia el centroide μ_j más próximo, buscando minimizar la varianza total dentro de cada grupo. La función de costo o inercia que el sistema optimiza internamente se expresa mediante la suma de distancias euclidianas al cuadrado:

$$J = \sum_{j=1}^{k} \sum_{x_i \in S_j} \|x_i - \mu_j\|^2$$

Donde k es el número total de clusters, S_j representa el conjunto de datos asignados al cluster j, y $\|x_i - \mu_j\|$ cuantifica la distancia entre el dato x_i y su centroide μ_j.

Otro proceso esencial es la **reducción de dimensionalidad**, técnica diseñada para simplificar tablas de datos complejas eliminando la redundancia y conservando la máxima información posible. El Análisis de Componentes Principales (PCA) es el método matemático por excelencia para este fin; proyecta los datos originales a un nuevo sistema de coordenadas donde los ejes principales maximizan la varianza. La transformación lineal que mapea los datos a un espacio reducido se define como:

$$Z = XW$$

Donde X representa la matriz original de datos de dimensión $n \times p$, W es la matriz de proyección formada por los autovectores de la matriz de covarianza, y Z es la nueva representación simplificada en un espacio de menor dimensión.

Las **aplicaciones prácticas** y **casos reales** de esta metodología son indispensables en la industria moderna:

Segmentación de clientes: Identificación de perfiles de consumidores con hábitos de compra similares para optimizar campañas de marketing.

Detección de anomalías: Monitoreo de transacciones financieras para señalar comportamientos atípicos que sugieran fraude bancario.

Sistemas de recomendación: Descubrimiento de afinidades entre productos o contenidos audiovisuales en plataformas de *streaming*.

Esta aproximación ofrece destacados **beneficios y ventajas**, como la eliminación de la laboriosa tarea de etiquetar manualmente volúmenes masivos de datos y la capacidad para detectar conexiones imprevistas que el ojo humano pasaría por alto. No obstante, plantea significativos **desafíos y limitaciones**: la evaluación del desempeño suele ser subjetiva al no existir una "respuesta correcta" comprobable, y la interpretación de los grupos o dimensiones resultantes requiere un profundo conocimiento del dominio por parte del especialista.

Figura 10. La máquina encuentra patrones sin instrucciones previas.

CIENCIA DE DATOS
DIVULGACIÓN · EDUCACIÓN · INNOVACIÓN

COMPA

EL APRENDIZAJE SUPERVISADO VER

— DIFERENCIAS CLAVE, FUNCIONAMIE

APRENDIZAJE SUPERVISADO

El modelo aprende a partir de datos etiquetados, donde la salida deseada ya es conocida.

- GUIADO POR ETIQUETAS
- OBJETIVO DEFINIDO
- PREDICCIÓN PRECISA
- EVALUACIÓN EXACTA

1 DATOS UTILIZADOS
Datos etiquetados (con entradas y salidas conocidas).

2 OBJETIVO PRINCIPAL
Predecir o clasificar resultados futuros.

3 TIPOS DE ALGORITMOS
Regresión lineal, Árboles de decisión, SVM, Redes neuronales, KNN, etc.

4 EJEMPLOS DE USO
Clasificación de correos, diagnóstico médico, predicción de ventas.

5 EVALUACIÓN DEL MODELO
Se mide con métricas como precisión, recall, F1-score, RMSE, etc.

6 RESULTADO ESPERADO
Predicciones precisas y resultados medibles sobre datos nuevos.

EXPERTO EN DATOS

La elección del enfoque depende del objetivo del análisis y del tipo de datos disponibles. Ambos son complementarios y esenciales.

SUPERVISADO

Precisión	Datos	Control	Interpretabilidad
Alta	Etiquetados	Alto	Alta

RENDIMIENTO 90%

DOS ENFOQUES, UN MISMO OBJETIVO:

RATIVO

SUS APRENDIZAJE NO SUPERVISADO

NTO, APLICACIONES Y RESULTADOS

PROPÓSITO:
DIVULGAR

Comprender las diferencias clave entre dos enfoques fundamentales del aprendizaje automático y su aplicación en la ciencia de datos y la inteligencia artificial.

DATOS UTILIZADOS
Datos sin etiquetar, solo entradas sin salidas conocidas.

OBJETIVO PRINCIPAL
Descubrir patrones, grupos o estructuras ocultas.

TIPOS DE ALGORITMOS
K-Means, Clustering Jerárquico, PCA, DBSCAN, Autoencoders, etc.

EJEMPLOS DE USO
Segmentación de clientes, detección de anomalías, reducción de dimensiones.

EVALUACIÓN DEL MODELO
Se evalúa con métricas como silueta, varianza explicada, inercia, etc.

RESULTADO ESPERADO
Conocimiento profundo de la estructura interna de los datos.

APRENDIZAJE NO SUPERVISADO

El modelo encuentra patrones y estructuras ocultas en datos sin etiquetar ni respuestas previas.

- SIN ETIQUETAS
- DESCUBRIMIENTO DE PATRONES
- EXPLORACIÓN DE DATOS
- AGRUPACIÓN NATURAL

NO SUPERVISADO

Descubrimiento	Datos	Control	Interpretabilidad
Alto	Sin etiquetar	Bajo	Media

RENDIMIENTO 70%

EN POCAS PALABRAS

Supervisado: aprende con respuesta conocida.
No supervisado: descubre lo desconocido.

- SUPERVISADO
- NO SUPERVISADO
- COMPARATIVO

TRANSFORMAR DATOS EN CONOCIMIENTO

Aprendizaje no supervisado

Una aventura interactiva

El Misterio de las Galletas

Ana:
¡Mira todas estas galletas! ¡Hay tantas! ¿Cómo las separo?

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

Clústeres Sin Etiquetas -

¿Cómo funciona el aprendizaje por refuerzo?

El **aprendizaje por refuerzo** es una rama del aprendizaje automático inspirada en la psicología conductista, cuyo **concepto fundamental** se basa en el aprendizaje interactivo a través del ensayo y error. A diferencia del aprendizaje supervisado, donde un modelo se entrena con un conjunto de datos previamente etiquetados por humanos, en este paradigma una entidad autónoma denominada **agente** aprende a tomar decisiones secuenciales interactuando directamente con un **entorno**. El **principio de funcionamiento** radica en que el agente no es instruido sobre qué acciones tomar, sino que debe descubrirlas evaluando cuáles producen la mayor **recompensa** a largo plazo.

Figura 11. ¿Cómo funciona el aprendizaje por refuerzo?

Figura 12. Aprendizaje interactivo a través de ensayo y error.

Para estructurar este **proceso interno** de toma de decisiones de manera rigurosa, el problema se formaliza habitualmente mediante los Procesos de Decisión de Markov (MDP). Los **componentes principales** del sistema interactúan en un ciclo continuo: en cada paso temporal t, el agente evalúa el **estado** actual s_t, ejecuta una **acción** a_t y el entorno responde devolviendo una **recompensa** numérica r_{t+1} junto con un nuevo estado s_{t+1}. Para calcular el valor futuro acumulado de tomar una decisión concreta, el agente utiliza la ecuación de Bellman, expresada formalmente como:

$$Q(s, a) = r + \gamma \max_{a'} Q(s', a')$$

En esta formulación, $Q(s, a)$ representa el valor estimado de ejecutar la acción a estando en el estado s; r es la recompensa inmediata; s' es el nuevo estado al que se transiciona; a' representa la mejor acción posible en ese nuevo estado, y $\gamma \in [0, 1)$ es el **factor de descuento**, un parámetro crucial que pondera la importancia de las recompensas futuras con respecto a las inmediatas.

Entre las **ventajas y beneficios** más destacados de este método se encuentra su capacidad para operar en entornos inciertos y altamente dinámicos sin necesidad de conocimiento previo sobre las reglas internas del sistema, superando frecuentemente los límites de la intuición humana. Sin embargo, la técnica enfrenta importantes **desafíos y limitaciones**. Uno de los más complejos es el dilema entre exploración (probar nuevas acciones no evaluadas) y explotación (repetir acciones conocidas que ya ofrecen recompensas altas). Asimismo, el problema de la asignación de crédito complica el entrenamiento, pues una recompensa positiva o negativa puede ser el resultado diferido de una cadena de cientos de decisiones previas, requiriendo un alto costo computacional para converger a una solución óptima.

A pesar de estas complejidades, las **aplicaciones prácticas** e **historias de éxito reales** han demostrado un impacto transformador en diversas industrias avanzadas:

Sistemas de control en robótica: Permite a brazos mecánicos aprender a manipular objetos frágiles o a robots bípedos mantener el equilibrio al desplazarse por terrenos irregulares.

Dominio de juegos complejos: Ejemplares como *AlphaGo* de DeepMind o *OpenAI Five*, que derrotaron a campeones mundiales en el juego de Go y en videojuegos de estrategia en tiempo real, aprendiendo mediante simulaciones masivas contra sí mismos.

Navegación autónoma: Optimización en tiempo real de trayectorias, aceleración y maniobras de adelantamiento en vehículos autónomos.

Eficiencia en centros de datos: Ajuste dinámico de los sistemas de refrigeración industrial, logrando reducir drásticamente el consumo energético.

Figura 13. Descubre, explora y aprende por refuerzo.

CIENCIA DE DATOS
DIVULGACIÓN • EDUCACIÓN • INNOVACIÓN

COMPA

EL APRENDIZAJE SUPERVISADO VE

DOS ENFOQUES, UN MISMO OBJETIVO: APR

APRENDIZAJE SUPERVISADO

El modelo aprende a partir de datos etiquetados, donde el resultado correcto ya es conocido.

FLUJO DE APRENDIZAJE

DATOS ETIQUETADOS → MODELO → PREDICCIÓN CORRECTA

EJEMPLOS DE USO

- Clasificación de correos (spam/ham)
- Diagnóstico médico
- Predicción de precios
- Reconocimiento de imágenes

DATOS UTILIZADOS

Datos etiquetados (entrada + salida conocida).

1

OBJETIVO PRINCIPAL

Predecir o clasificar resultados futuros.

2

TIPO DE APRENDIZAJE

Aprende de ejemplos ya conocidos.

3

RETROALIMENTACIÓN

Se recibe el resultado correcto para cada ejemplo.

4

COMPLEJIDAD

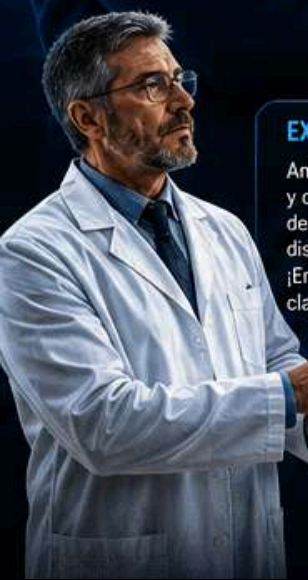
Menor complejidad de entrenamiento.

5

EVALUACIÓN

Precisión, Recall, F1-score, RMSE, etc.

6



EXPERTO EN DATOS

Ambos enfoques son poderosos y complementarios. La elección depende del problema, los datos disponibles y el objetivo final. ¡Entender sus diferencias es clave para innovar!

ENFRENTA

SUPERVISADO

Datos	Aprendizaje	Resultado	Estabilidad
Etiquetados	De ejemplos	Inmediato	Alta

NECESIDAD DE DATOS	TIEMPO DE ENTRENAMIENTO	FACILIDAD DE IMPLEM
Alta	Menor	Alta
80%		85%

APRENDER DE LOS DATOS O APRENDER DE LA EXPERIE

RATIVO

RSUS APRENDIZAJE POR REFUERZO

ENDER PARA TOMAR MEJORES DECISIONES

PROPÓSITO: DIVULGAR

Comprender y divulgar las diferencias clave entre dos enfoques fundamentales del aprendizaje automático para entender cómo aprenden las máquinas y cómo aplicarlas mejor.

APRENDIZAJE POR REFUERZO

El agente aprende mediante prueba y error, recibiendo recompensas o castigos según sus acciones en el entorno.

DATOS UTILIZADOS
Sin etiquetas. Solo recibe recompensas o castigos del entorno.

OBJETIVO PRINCIPAL
Maximizar la recompensa acumulada a lo largo del tiempo.

TIPO DE APRENDIZAJE
Aprende explorando y mejorando con la experiencia.

RETROALIMENTACIÓN
Recompensa positiva o castigo negativo después de cada acción.

COMPLEJIDAD
Mayor complejidad y tiempo de entrenamiento.

EVALUACIÓN
Recompensa acumulada, tasa de éxito, estabilidad, etc.

FLUJO DE APRENDIZAJE

AGENTE → ENTORNO → RECOMPENSA (REFUERZO)

EJEMPLOS DE USO

- Juegos (AlphaGo, Atari)
- Robótica y control
- Optimización de procesos
- Conducción autónoma

NMIENTO RÁPIDO

VS

POR REFUERZO

Datos	Aprendizaje	Resultado	Estabilidad
Sin etiquetas	Por experiencia	A largo plazo	Variable

ENTACIÓN	NECESIDAD DE DATOS	TIEMPO DE ENTRENAMIENTO	FACILIDAD DE IMPLEMENTACIÓN
%	Baja 30%	Mayor	Media 55%

NCIA. AMBOS CAMINOS NOS LLEVAN AL CONOCIMIENTO.

aprendizaje por refuerzo

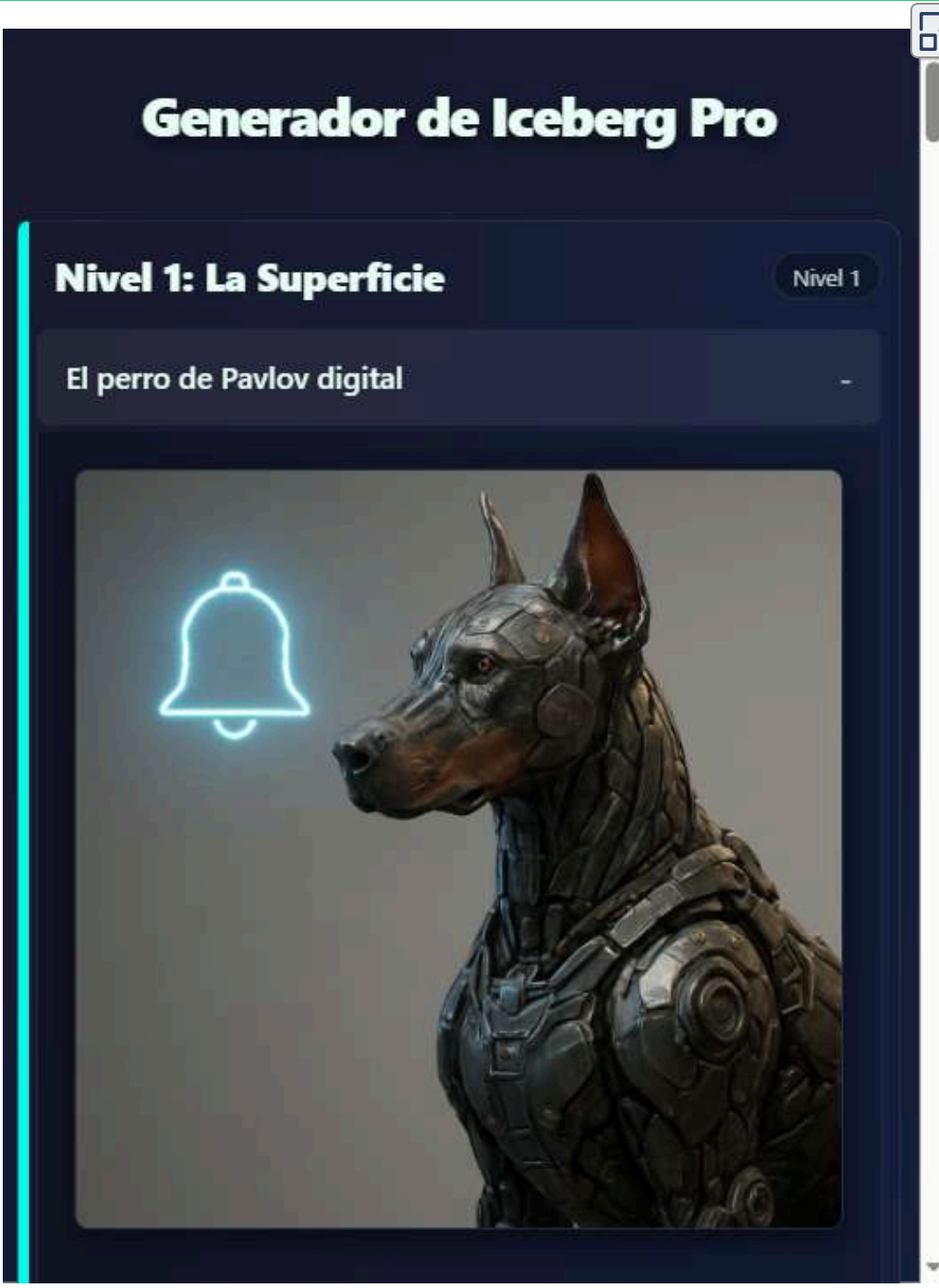
Una aventura interactiva

El Comienzo

Bolt:
¡Mentor! ¿Qué es este laberinto?
¡Parece un desafío!

¿Qué es el Aprendizaje por Refuerzo?

Es un tipo de aprendizaje automático donde un


Generador de Iceberg Pro
Nivel 1: La Superficie
Nivel 1
El perro de Pavlov digital

¿Cómo funcionan las redes neuronales?

Las **redes neuronales artificiales** son modelos matemáticos inspirados en la estructura y el comportamiento del cerebro humano, diseñados para reconocer patrones complejos en grandes volúmenes de datos. Su **concepto fundamental** se basa en interconectar unidades de procesamiento llamadas **neuronas artificiales** o nodos, organizadas en tres tipos de capas: la **capa de entrada**, que recibe los datos iniciales; las **capas ocultas**, que procesan la información en distintos niveles de abstracción; y la **capa de salida**, que entrega la predicción final.

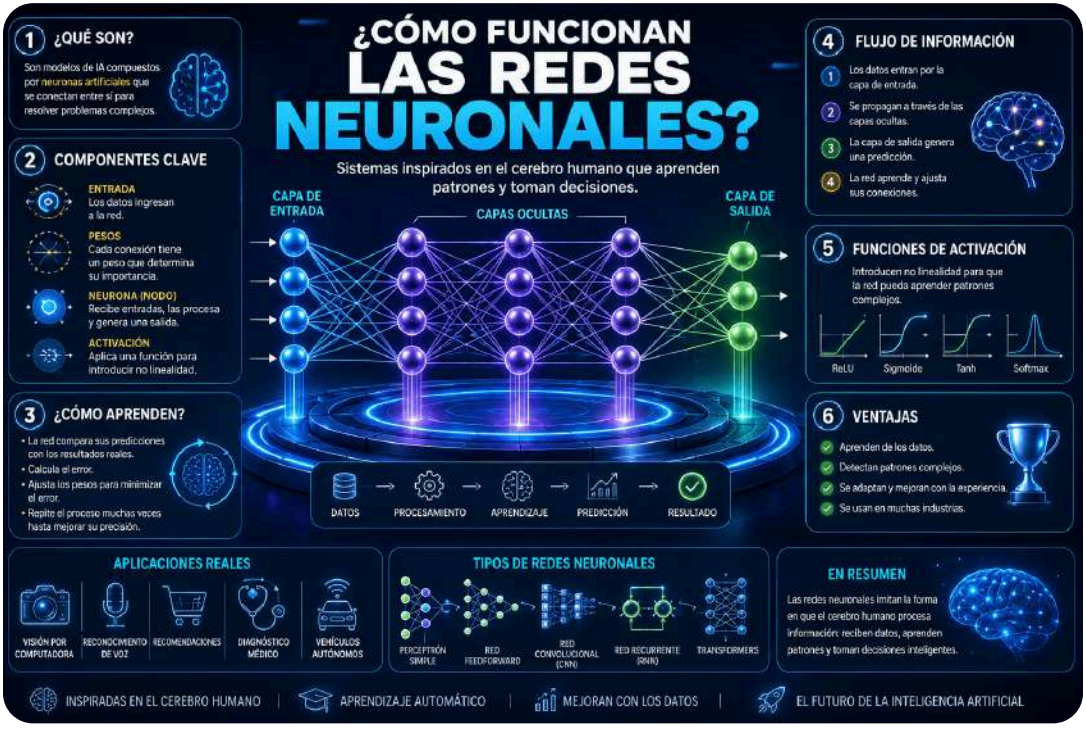


Figura 14. ¿Cómo funcionan las redes neuronales?

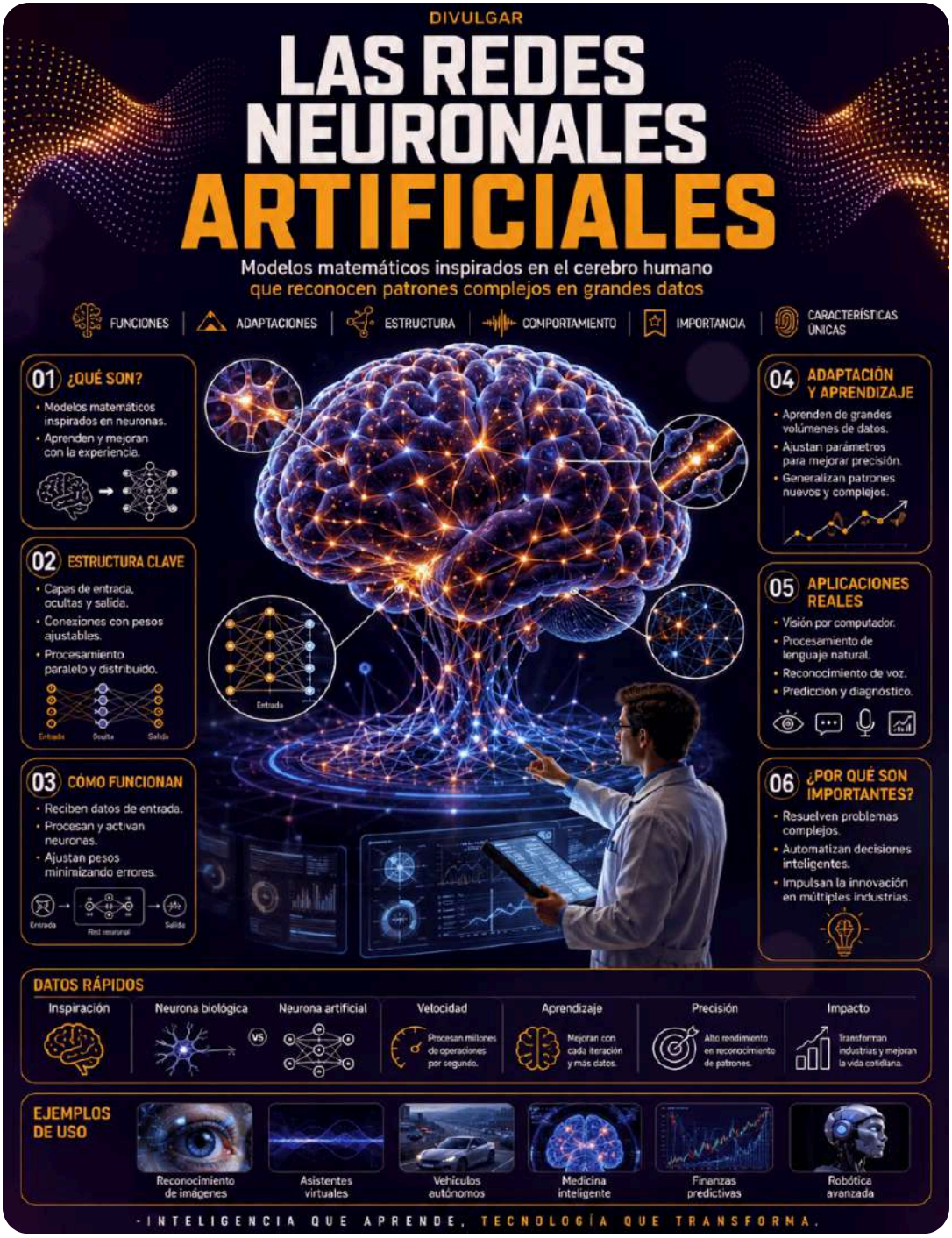


Figura 15. Modelos matemáticos inspirados en el cerebro humano.

Cada conexión entre neuronas posee un valor numérico denominado **peso** (w_i), el cual regula la importancia de esa señal, junto con un término de ajuste llamado **sesgo** (b). Matemáticamente, el valor de entrada combinado para una neurona se calcula como:

$$z = \sum_{i=1}^{n} w_i x_i + b$$

Donde x_i representa cada una de las entradas que recibe la neurona.

El **principio de funcionamiento** de esta estructura ocurre mediante un proceso conocido como **propagación hacia adelante** (*forward propagation*). Cuando el valor combinado z pasa por la neurona, no se transmite tal cual; en su lugar, se procesa a través de una **función de activación** (σ). Esta función introduce no-linealidad en la red, lo que le permite aprender relaciones sumamente complejas que van más allá de simples líneas rectas. Una de las funciones clásicas más utilizadas es la función sigmoide:

$$a = \sigma(z) = \frac{1}{1 + e^{-z}}$$

Donde a es la activación resultante que la neurona transmitirá como entrada hacia las neuronas de la siguiente capa, repitiendo el proceso hasta alcanzar la salida.

El verdadero aprendizaje ocurre en los **procesos internos** de corrección de errores mediante la **propagación hacia atrás** (*backpropagation*) y el **descenso del gradiente**. Cuando la red genera una predicción ($\hat{y}$), esta se compara con la respuesta real (y) utilizando una función de pérdida que mide qué tan alejada estuvo la red de la realidad. Por ejemplo, mediante el error cuadrático medio:

$$L = \frac{1}{2}(\hat{y} - y)^2$$

A partir de este error L, el algoritmo calcula la derivada parcial para saber cuánto contribuyó cada peso al fallo y actualiza los parámetros en dirección opuesta al error para mejorarlos gradualmente, ajustando cada peso según la regla:

$$w_{\text{nuevo}} = w_{\text{viejo}} - \alpha \frac{\partial L}{\partial w}$$

Donde α representa la **tasa de aprendizaje**, un hiperparámetro que controla qué tan grandes son los pasos en la corrección de los pesos.

En cuanto a sus **aplicaciones prácticas y casos reales**, las redes neuronales son la columna vertebral de tecnologías modernas revolucionarias. Destacan en campos como:

Visión por computadora: Diagnóstico médico mediante análisis de radiografías y detección de obstáculos en vehículos autónomos.

Procesamiento del lenguaje natural: Traducción automática instantánea, asistentes de voz e interactividad con modelos de lenguaje generativos.

Procesamiento de audio: Cancelación de ruido inteligente y reconocimiento de voz.

Los principales **beneficios y ventajas** de esta arquitectura residen en su extraordinaria capacidad para realizar extracción automática de características, lo que evita la necesidad de procesar o etiquetar manualmente cada detalle de datos no estructurados como imágenes, texto o audio.

A pesar de su enorme potencial, existen importantes **desafíos o limitaciones** que deben considerarse. El entrenamiento de redes neuronales profundas requiere un consumo masivo de recursos computacionales (GPUs/TPUs) y enormes cantidades de datos etiquetados para evitar el **sobreajuste** (*overfitting*), fenómeno en el que la red memoriza los datos de entrenamiento pero falla al generalizar ante datos nuevos. Asimismo, sufren del problema de la "caja negra": debido a los millones de parámetros matemáticos interactuando al mismo tiempo, interpretar exactamente **por qué** la red tomó una decisión específica sigue siendo un reto fundamental en el aprendizaje automático moderno.

Figura 16. Redes neuronales.

Simulación diseñada con Claude Sonnet 5 (ver en pantalla ampliada).

CIENCIA DE DATOS
DIVULGACIÓN • EDUCACIÓN • INNOVACIÓN

COMPAR

REDES NEURONALES HUMANAS VI

DOS INTELIGENCIAS, DIFERENTE ORIGEN. M

REDES NEURONALES HUMANAS

~86 MIL MILLONES DE NEURONAS

~100 BILLONES DE SINAPSIS

20 W CONSUMO APROXIMADO

EVOLUCIÓN MILLONES DE AÑOS

1 ESTRUCTURA
Neurona biológica con dendritas, axón y sinapsis electroquímicas complejas.

2 APRENDIZAJE
Aprendizaje continuo, flexible y adaptativo desde la experiencia y el contexto.

3 PLASTICIDAD
Alta plasticidad sináptica, se reorganiza con emociones, sueño y experiencias.

4 EFICIENCIA
Extremadamente eficiente, funciona con muy baja energía.

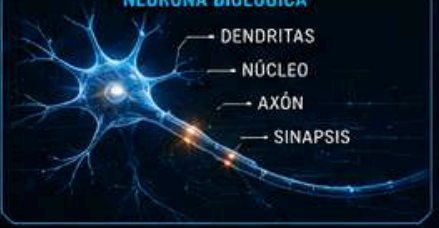
5 GENERALIZACIÓN
Generaliza con pocos ejemplos, entiende conceptos abstractos y transferencias.

6 ROBUSTEZ
Robusta al ruido, daño parcial y cambios inesperados.

EXPERTO EN IA Y NEUROCIENCIA

Ambos sistemas son extraordinarios a su manera. Comprender sus diferencias nos ayuda a construir mejores máquinas y a entendernos mejor como humanos.

NEURONA BIOLÓGICA



HUMANAS

NEURONAS	SINAPSIS	CONSUMO ENERGÉTICO	VELOCIDAD DE SEÑAL	FLEXIBILIDAD
~86 MIL MILLONES	~100 BILLONES	~20 VATIOS	0.5 - 120 m/s	MUY ALTA

LA BIOLOGÍA INSPIRA. LA TECNOLOGÍA POTENCIA

RATIVO

ERSUS REDES NEURONALES DE IA

ISMO OBJETIVO: APRENDER Y RESOLVER PROBLEMAS.

PROPÓSITO: DIVULGAR

Comprender y divulgar las diferencias clave entre las redes neuronales biológicas y las artificiales para entender cómo aprendemos nosotros y cómo aprenden las máquinas.

REDES NEURONALES DE IA

1. **ESTRUCTURA**
   Unidades artificiales (nodos) con conexiones numéricas (pesos y sesgos).
2. **APRENDIZAJE**
   Aprendizaje basado en datos masivos, entrenamiento por retropropagación.
3. **PLASTICIDAD**
   Ajuste de pesos durante el entrenamiento; depende del algoritmo y datos.
4. **EFICIENCIA**
   Alta capacidad de cálculo pero con mayor consumo energético.
5. **GENERALIZACIÓN**
   Puede requerir muchos datos para generalizar; depende del modelo.
6. **ROBUSTEZ**
   Sensible a datos adversarios y fallos fuera del entrenamiento.

MILLONES A MILES DE MILLONES
DE PARÁMETROS

CONECCIONES SIMULADAS
(MATEMÁTICAS)

100 W - 10 kW
CONSUMO TÍPICO
(DEPENDIENDO DEL MODELO)

DESARROLLO
DÉCADAS DE INVESTIGACIÓN

NEURONA ARTIFICIAL (NODO)

EJEMPLOS DE APLICACIÓN

HUMANAS	IA
Aprendizaje	Reconocimiento de imágenes
Memoria	Procesamiento de lenguaje
Creatividad	Predicción de datos
Intuición	Automatización
Emociones	Vehículos autónomos

IA

PARÁMETROS	CONECCIONES	CONSUMO ENERGÉTICO	VELOCIDAD DE CÓMPUTO	FLEXIBILIDAD
MILLONES A BILLONES	MILLONES A BILLONES	100 W - 10 kW (APROX.)	TFLOPS - PFLOPS (POR SEGUNDO)	ALTA (POR DISEÑO)

L EL FUTURO ES LA COLABORACIÓN ENTRE AMBAS.

redes neuronales

Una aventura interactiva

El Cerebro Artificial

Estudiante:
¡Profe! ¿Qué es una red neuronal?

Inspiración

Las redes neuronales artificiales se inspiran en

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

Cerebros digitales -

¿Cómo funciona el entrenamiento de un modelo de IA?

El entrenamiento de un modelo de inteligencia artificial es un proceso iterativo mediante el cual un algoritmo aprende a realizar tareas específicas ajustando sus parámetros internos a través de la exposición continua a datos. En términos sencillos, es análogo al aprendizaje humano: el modelo realiza una suposición, mide su nivel de error y ajusta su conocimiento previo para no repetir las equivocaciones. Los **componentes principales** de este ecosistema incluyen el conjunto de datos de entrenamiento (*dataset*), los **parámetros optimizables** (pesos w y sesgos b), una **función de pérdida** (o *loss function*) que cuantifica el margen de error, y un **algoritmo de optimización** encargado de aplicar las correcciones necesarias.

Figura 17. ¿Cómo funciona el entrenamiento de un modelo de IA?

Figura 18. El entrenamiento de un modelo de IA.

El **principio de funcionamiento** se basa en un ciclo iterativo. Durante la fase de propagación hacia adelante (*forward pass*), los datos de entrada x se combinan con los pesos w y el sesgo b para generar una predicción $\hat{y}$. Inmediatamente después, el modelo evalúa su precisión comparando la predicción con el valor real. En un problema de regresión, por ejemplo, una de las métricas de error más utilizadas es el **Error Cuadrático Medio** (MSE), expresado formalmente como:

$$L(w, b) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

Donde N representa el número total de muestras observadas, y_i indica el valor real esperado y $\hat{y}_i$ denota la estimación producida por el modelo. Una vez calculado el error, se inician los **procesos internos** más críticos del entrenamiento: la retropropagación (*backpropagation*) y la optimización. Mediante la regla de la cadena del cálculo diferencial, la retropropagación calcula la derivada parcial de la función de pérdida con respecto a cada peso, determinando cuánto contribuyó cada parámetro al error total. Con esta información, el algoritmo de **Descenso del Gradiente** actualiza los pesos en la dirección opuesta al gradiente para reducir la pérdida en la siguiente iteración, siguiendo la regla de actualización:

$$w_{nuevo} = w_{actual} - \eta \cdot \frac{\partial L}{\partial w}$$

En esta ecuación, η representa la **tasa de aprendizaje** (*learning rate*), un hiperparámetro clave que controla la magnitud del paso que da el modelo en cada iteración para alcanzar el mínimo de la función.

Este flujo de trabajo ofrece **beneficios y ventajas** extraordinarios, pues permite resolver problemas altamente complejos —como la

visión por computadora o el procesamiento del lenguaje natural— sin necesidad de programar reglas explícitas. En **casos reales** como el diagnóstico médico asistido por imagen, un modelo puede analizar miles de tomografías etiquetadas hasta reducir el valor de L a niveles óptimos, logrando identificar patologías con una precisión sobrehumana. Sin embargo, el entrenamiento también conlleva importantes **desafíos y limitaciones**:

Sobreajuste (*overfitting*): El modelo memoriza el ruido de los datos de entrenamiento y pierde la capacidad de generalizar ante datos nuevos e imprevistos.

Subajuste (*underfitting*): La arquitectura es demasiado simple o el entrenamiento insuficiente, impidiendo que el algoritmo capture la estructura subyacente de los datos.

Costo computacional elevado: Los modelos modernos requieren días o semanas de procesamiento en hardware especializado (GPU o TPU), implicando un consumo energético considerable.

Figura 19. El entrenamiento de un modelo de IA.

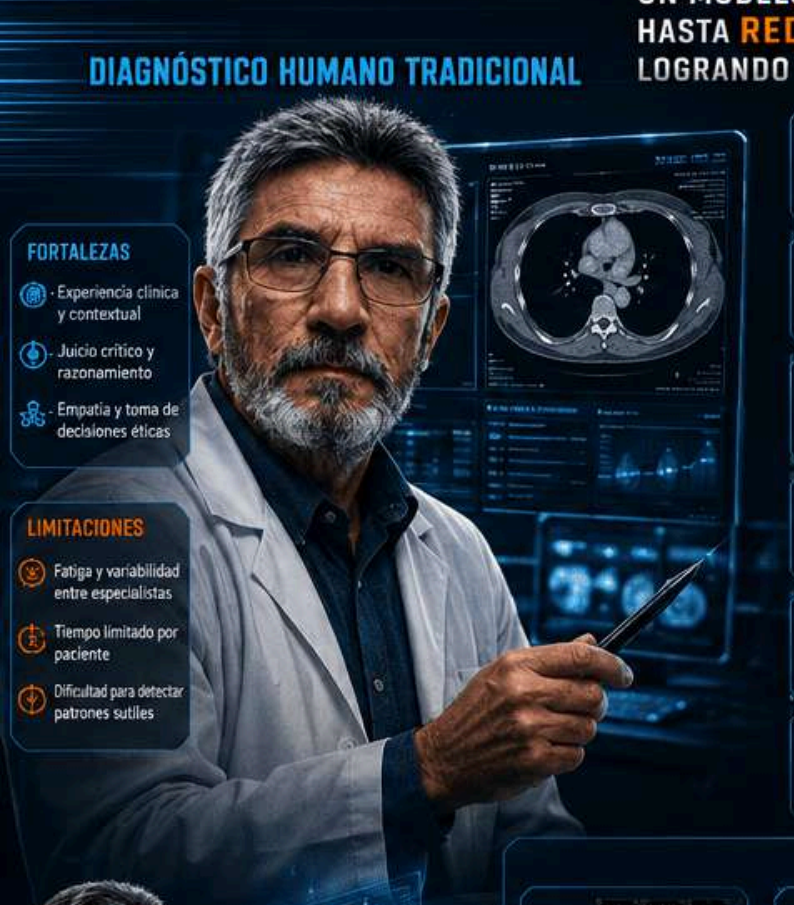
CIENCIA DE DATOS
DIVULGACIÓN • EDUCACIÓN • INNOVACIÓN

COMPARATI

EL DIAGNÓSTICO MÉDICO ASISTIDO

UN MODELO PUEDE ANALIZAR MILES DE TOM
HASTA REDUCIR EL VALOR DE $\mathcal{L}$ A NIVE
LOGRANDO IDENTIFICAR PATOLOGÍAS CON UN

DIAGNÓSTICO HUMANO TRADICIONAL



FORTALEZAS

- Experiencia clínica y contextual
- Juicio crítico y razonamiento
- Empatía y toma de decisiones éticas

LIMITACIONES

- Fatiga y variabilidad entre especialistas
- Tiempo limitado por paciente
- Dificultad para detectar patrones sutiles

FUENTE DE CONOCIMIENTO
Experiencia, años de estudio y casos previos limitados. 1

VELOCIDAD DE ANÁLISIS
Minutos por estudio (15–30 min promedio). 2

CONSISTENCIA
Variable entre radiólogos y dependiente de la fatiga. 3

DETECCIÓN DE PATRONES
Limitada a lo visible y a la experiencia. 4

APRENDIZAJE
Lento y basado en la experiencia individual. 5

OBJETIVO
Diagnóstico preciso para cada paciente. 6

CÓMO UN MODELO DE IA

1. DATOS
Miles de tomografías etiquetadas por expertos.

2. ENTRENAMIENTO
El modelo aprende patrones complejos.

3. FUNCIÓN DE PÉRDI
El modelo predice y co
con la etiqueta real.

$$L = \frac{1}{n} \sum (y -$$

> La IA no reemplaza al radiólogo, lo potencia. Juntos, logramos diagnósticos más rápidos, precisos y con mayor impacto en la vida de los pacientes.

RENDIMIENTO COMPARATIVO EN DETECCIÓN DE PATOLOG

NÓDULOS PULMONARES	HEMORRAGIA CEREBRAL	TUMORES HEPÁTICOS
IA: 97.6%	IA: 96.1%	IA: 94.8%
Humano: 88.3%	Humano: 85.7%	Humano: 83.2%

* Resultados basados en estudios publicados y benchmarks de modelos d

DATOS + ALGORITMOS + EXPERIEN

IVO

O POR IMAGEN,

IOGRAFÍAS ETIQUETADAS
LES ÓPTIMOS,
IA PRECISIÓN SOBREHUMANA.

PROPÓSITO: DIVULGAR

Comprender cómo la inteligencia artificial entrenada con miles de tomografías puede superar las limitaciones humanas y detectar patologías con precisión sobrehumana.

FUENTE DE CONOCIMIENTO
Miles de tomografías etiquetadas y aprendizaje profundo.

VELOCIDAD DE ANÁLISIS
Segundos por estudio (< 5 seg promedio).

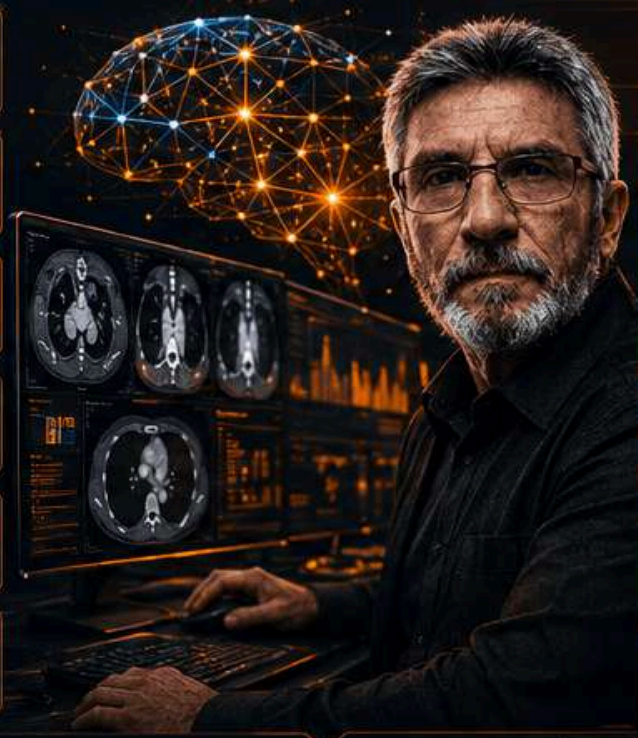
CONSISTENCIA
Alta consistencia y sin fatiga.

DETECCIÓN DE PATRONES
Detecta patrones sutiles y complejos invisibles.

APRENDIZAJE
Mejora continua al reducir el valor de L en cada iteración.

OBJETIVO
Minimizar el valor de L y maximizar la precisión diagnóstica.

DIAGNÓSTICO ASISTIDO POR IA



FORTALEZAS
- Analiza miles de tomografías en segundos
- Alta precisión y consistencia
- Detecta patrones imperceptibles al ojo humano

LIMITACIONES
- Dependencia de datos etiquetados
- Requiere validación clínica
- Menor comprensión del contexto clínico

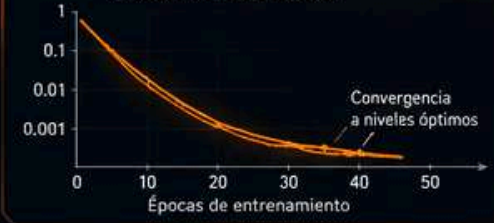
REDUCE EL VALOR DE L

DA (L)
mpara
$\hat{y})^2$

4. OPTIMIZACIÓN
El algoritmo ajusta los pesos y reduce el valor de L.

5. DESEMPEÑO ÓPTIMO
L se aproxima a 0
Precisión sobrehumana.

CURVA DE PÉRDIDA (L)



ÍAS

LESIONES ÓSEAS
IA: 95.5%
Humano: 87.1%

e IA en radiología.

VS

HUMANO			IA		
Precisión promedio	Tiempo por estudio	Casos analizados/día	Precisión promedio	Tiempo por estudio	Casos analizados/día
~86%	15–30 min	~30–60	~95%+	< 5 seg	Miles+

CIA HUMANA = SALUD DE PRECISIÓN

entrenamiento de un modelo de IA

Una aventura interactiva

El Nacimiento de una IA

Profesor:
¡Hola clase! Hoy crearemos una IA. Primero, necesitamos datos. ¡Muchos datos!

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

Ensayo y error automatizado -

¿Cómo funciona la evaluación de un modelo?

Evaluar un modelo de aprendizaje automático consiste en medir qué tan bien generaliza sus predicciones ante datos completamente nuevos que no utilizó durante la fase de entrenamiento. El **concepto fundamental** radica en simular el comportamiento que tendrá el algoritmo en el mundo real. Para lograrlo, el **principio de funcionamiento** se basa en la división del conjunto de datos original en subconjuntos independientes, típicamente en una proporción de 80% para entrenamiento (*training set*) y 20% para evaluación (*test set*). El algoritmo ajusta sus parámetros internos únicamente con los datos de entrenamiento; posteriormente, se le presentan las entradas del conjunto de prueba para observar sus predicciones sin permitirle ver las respuestas correctas de antemano.

Figura 20. ¿Cómo funciona la evaluación de un modelo?

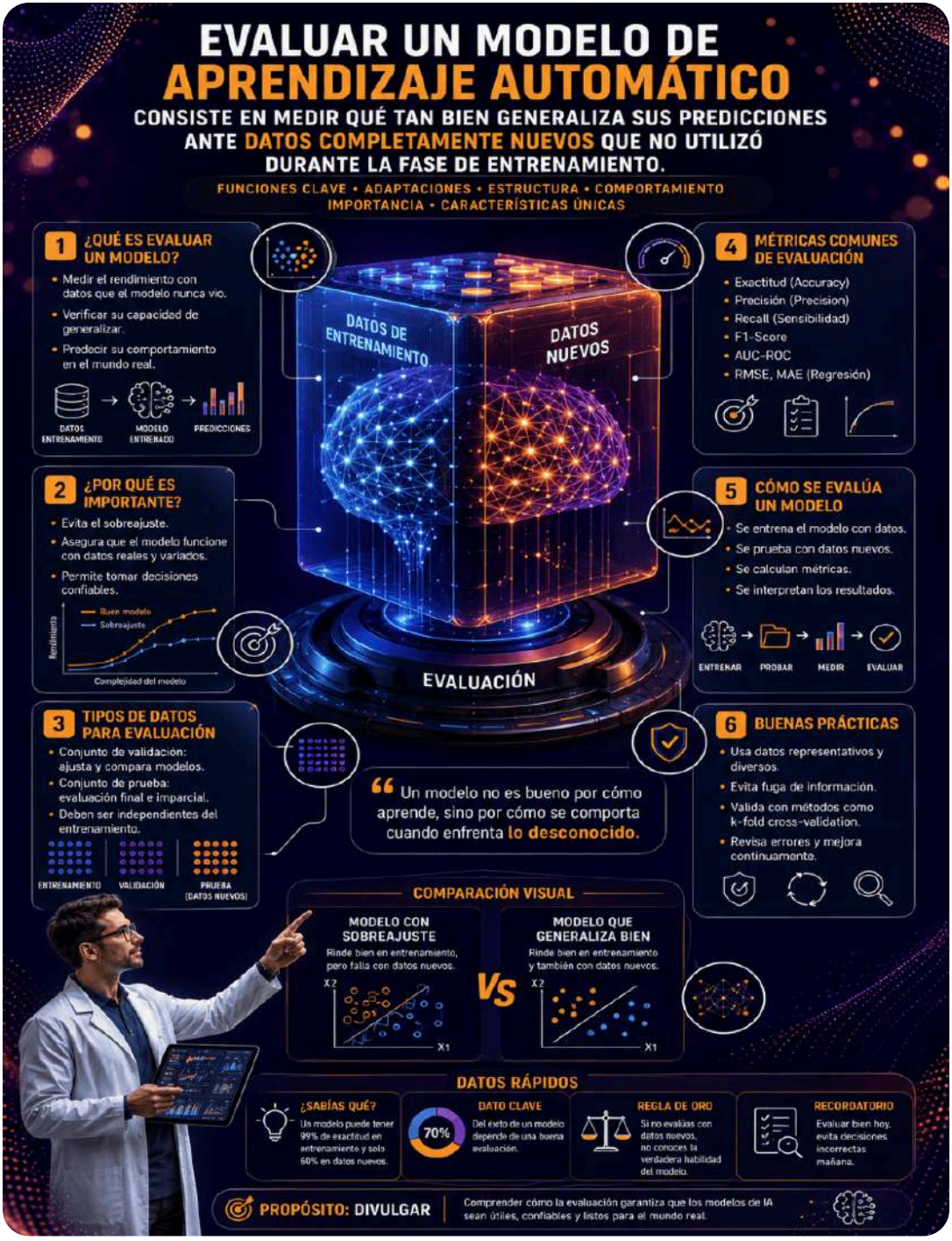


Figura 21. La evaluación consiste en medir qué tan bien hace sus predicciones.

El **proceso interno** de evaluación compara de forma sistemática los valores reales con las predicciones generadas. Los **componentes principales** de esta fase varían según la naturaleza del problema planteado:

Métricas de Regresión: Evalúan la desviación numérica continua entre el valor real y la predicción.

Métricas de Clasificación: Miden la tasa de aciertos y errores al asignar categorías discretas (mediante herramientas como la matriz de confusión).

Por ejemplo, en un problema de estimación numérica, la métrica estándar por excelencia es el **Error Cuadrático Medio** (MSE), expresado matemáticamente como:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Donde n es el número total de muestras de prueba, y_i representa el valor real observado y $\hat{y}_i$ es la predicción calculada por el modelo. Al elevar las diferencias al cuadrado, esta función penaliza con mayor severidad los errores grandes.

Para maximizar los **beneficios y ventajas** del proceso, se utiliza la técnica de **validación cruzada** (k-fold cross-validation), la cual divide los datos en k particiones y rota iterativamente el conjunto de prueba. Esto evita la fuga de información (*data leakage*) y garantiza que las métricas obtenidas sean estadísticamente sólidas. En **casos reales**, como en los sistemas de diagnóstico médico asistido por inteligencia artificial para la detección de tumores, una simple métrica de exactitud global resulta insuficiente. En este escenario, la evaluación se centra en optimizar la **sensibilidad** (minimizar los falsos negativos),

ya que clasificar a un paciente enfermo como sano tiene consecuencias drásticamente más graves que la situación inversa.

No obstante, el proceso de evaluación enfrenta severos **desafíos y limitaciones**, siendo el principal el equilibrio entre el sesgo y la varianza (*bias-variance tradeoff*). Un modelo excesivamente complejo puede memorizar el ruido de los datos de entrenamiento (sobreajuste u *overfitting*), mostrando un rendimiento perfecto en el entrenamiento pero colapsando durante la evaluación. La descomposición del error esperado en el modelo se formula de la siguiente manera:

$$\text{Error Total} = \text{Sesgo}^2 + \text{Varianza} + \sigma^2$$

Donde σ^2 representa el error irreducible o ruido inherente a los datos. En las **aplicaciones prácticas**, como los motores de recomendación en plataformas de *streaming* o la detección de fraudes bancarios en tiempo real, comprender estas métricas permite a los desarrolladores ajustar los hiperparámetros con precisión antes de desplegar los modelos a producción.

Figura 22. ¿Que tan inteligente es realmente tu modelo?

CIENCIA DE DATOS
DIVULGACIÓN • EDUCACIÓN • INNOVACIÓN

COMPAR

EN LAS APLICACIONES PRÁCTICAS, COMO L
PLATAFORMAS DE STREAMING O LA DETECCIÓN
COMPRENDER ESTAS MÉTRICAS PERMITE A L
HIPERPARÁMETROS CON PRECISIÓN ANTES DE

MOTORES DE RECOMENDACIÓN
EN PLATAFORMAS DE STREAMING

- EXPERIENCIA PERSONALIZADA
- MAYOR RETENCIÓN
- AUMENTO DE INTERACCIÓN
- DIVERSIDAD DE CONTENIDO
- ESCALABILIDAD GLOBAL

1 OBJETIVO PRINCIPAL
Maximizar relevancia y satisfacción del usuario.

2 TIPO DE PROBLEMA
Predicción de preferencia (ranking y recomendación).

3 DATOS UTILIZADOS
Historial de interacciones, visualizaciones, ratings y metadatos.

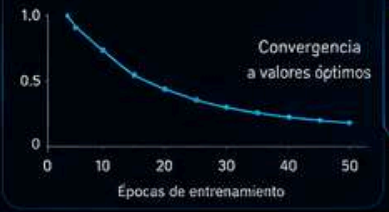
4 MÉTRICA CLAVE
NDCG, Precisión@K, Recall@K para relevancia.

5 DESAFÍOS PRINCIPALES
Escalabilidad, diversidad, novedad y cold start.

6 IMPACTO EN EL NEGOCIO
Mayor retención, tiempo de visualización y fidelización del usuario.

MÉTRICAS CLAVE
- Precisión@K
- Recall@K
- NDCG (Discounted Cumulative Gain)
- MAP (Mean Average Precision)
- Cobertura
- Diversidad
- Tiempo de respuesta

EJEMPLO: NDCG@10



EXPERTO EN DATOS
Entender y monitorear las métricas adecuadas es lo que transforma un modelo prometedor en una solución confiable y escalable.

DE LA EXPERIMEN

1. RECOLECCIÓN DE DATOS
Miles de ejemplos etiquetados.

2. ENTRENAMIENTO DEL MODELO
Aprende patrones y relaciones complejas.

MOTORES DE RECOMENDACIÓN (STREAMING)

NDCG@10	PRECISIÓN@10	RECALL@10	COBERTURA	DIVERSIDAD	TIEMPO RESPUES
0.87	0.78	0.65	72%	0.68	85 ms
ALTO	ALTO	MEDIO	ALTO	ALTO	ÓPTIMO

MÉTRICAS BIEN ENTENDIDAS, DECISIONES INT

RATIVO

OS MOTORES DE RECOMENDACIÓN EN
DE FRAUDES BANCARIOS EN TIEMPO REAL,
LOS DESARROLLADORES AJUSTAR LOS
DESPLEGAR LOS MODELOS A PRODUCCIÓN.

PROPÓSITO: DIVULGAR

Comprender y divulgar cómo las métricas en aplicaciones reales permiten mejorar el rendimiento de los modelos y tomar decisiones basadas en datos antes de su despliegue en producción.

1. **OBJETIVO PRINCIPAL**
   Minimizar fraudes y proteger al usuario y a la institución.
2. **TIPO DE PROBLEMA**
   Clasificación binaria (fraude / no fraude).
3. **DATOS UTILIZADOS**
   Transacciones, ubicación, dispositivo, comportamiento y patrones históricos.
4. **MÉTRICA CLAVE**
   AUC-ROC, F1-Score, Precisión, Recall para detección efectiva.
5. **DESAFÍOS PRINCIPALES**
   Datos desbalanceados, fraudes nuevos y adaptación a patrones cambiantes.
6. **IMPACTO EN EL NEGOCIO**
   Menores pérdidas económicas, confianza del cliente y cumplimiento regulatorio.

DETECCIÓN DE FRAUDES BANCARIOS EN TIEMPO REAL

- PROTECCIÓN EN TIEMPO REAL
- REDUCCIÓN DE PÉRDIDAS
- MENOR FALSO POSITIVO
- CUMPLIMIENTO NORMATIVO
- MEJORA DE CONFIANZA

MÉTRICAS CLAVE

- Precisión
- Recall (Sensibilidad)
- F1-Score
- AUC-ROC
- Matriz de confusión
- Tasa de falsos positivos
- Tiempo de detección

EJEMPLO: AUC-ROC

TACIÓN AL DESPLIEGUE

3. **EVALUACIÓN CON MÉTRICAS**
   Se calculan métricas clave (L, AUC, NDCG, F1).
4. **AJUSTE DE HIPERPARÁMETROS**
   Se minimiza L y se optimiza el desempeño.
5. **VALIDACIÓN FINAL**
   Verificación en datos nuevos y pruebas A/B.
6. **DESPLIEGUE A PRODUCCIÓN**
   Modelo listo para generar valor en el mundo real.

TA
S
VS

DETECCIÓN DE FRAUDES (BANCARIO)

AUC-ROC	F1-SCORE	PRECISIÓN	RECALL	FALSO POSITIVO	TIEMPO DETECCIÓN
0.98	0.92	0.93	0.91	0.7%	120 ms
EXCELENTE	EXCELENTE	EXCELENTE	EXCELENTE	BAJO	ÓPTIMO

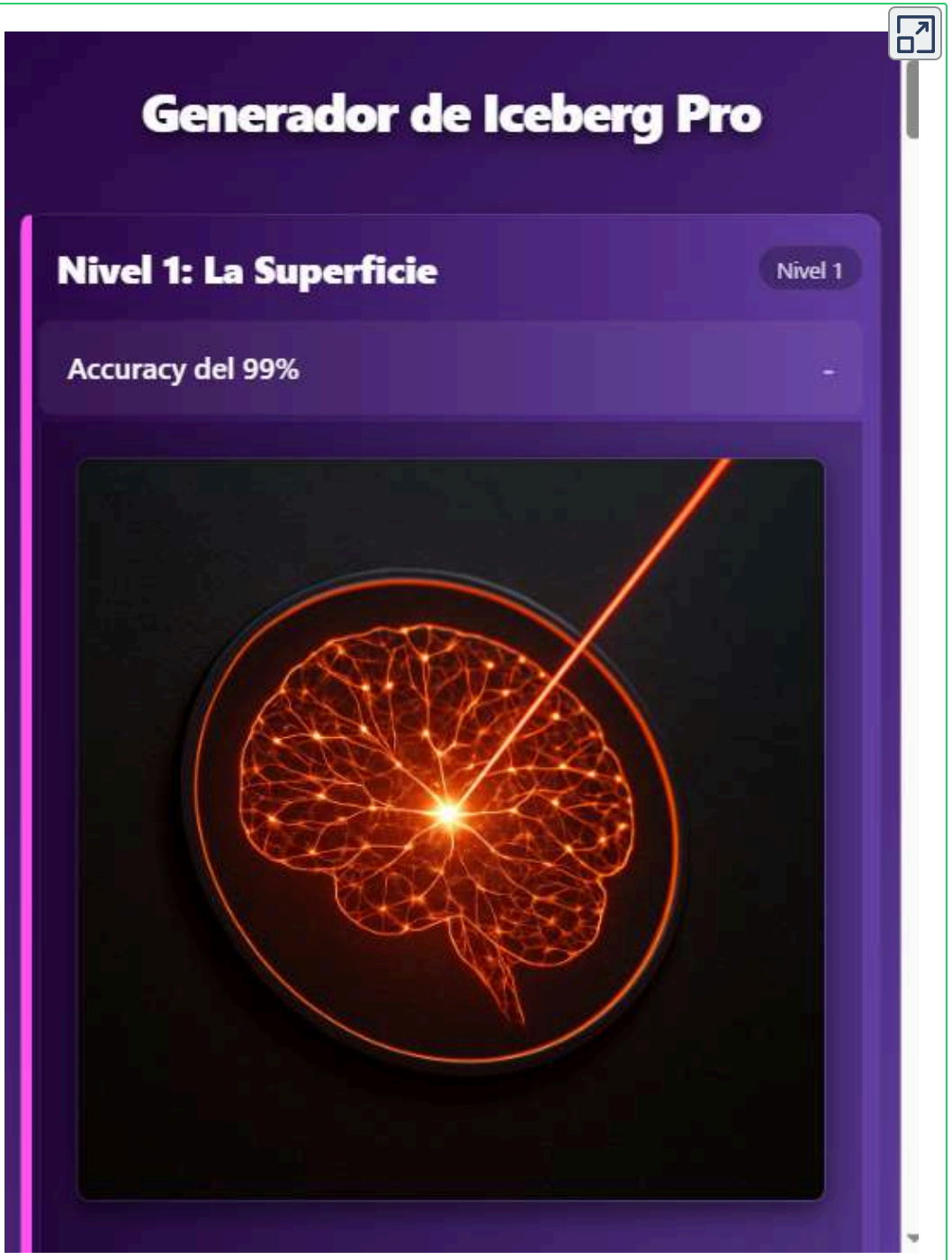
ELIGENTES, MODELOS QUE GENERAN IMPACTO REAL.

evaluación de un modelo de IA

Una aventura interactiva

Panel 1: ¡El Nuevo Ayudante!

Estudiante:
¡Guau! Este nuevo modelo de IA puede escribir mis ensayos. ¡Va a ser súper útil!


Generador de Iceberg Pro
Nivel 1: La Superficie
Nivel 1
Accuracy del 99%

¿Cómo funcionan las aplicaciones del aprendizaje automático?

Las aplicaciones del aprendizaje automático (**Machine Learning**) no siguen instrucciones rígidas programadas línea por línea, sino que operan mediante el reconocimiento de patrones ocultos en grandes volúmenes de datos. Su **principio de funcionamiento** radica en encontrar una función matemática capaz de transformar una serie de datos de entrada en una salida o predicción deseada. De manera simplificada, la aplicación busca modelar una relación del tipo:

$$y = f(\mathbf{x}; \boldsymbol{\theta})$$

Donde $\mathbf{x}$ representa el vector de características de entrada (por ejemplo, los píxeles de una imagen o el historial de compras de un usuario), y es la respuesta o predicción esperada, y $\boldsymbol{\theta}$ representa el conjunto de parámetros ajustables que el modelo optimiza automáticamente durante su fase de aprendizaje.

Figura 23. Del dato a la decisión

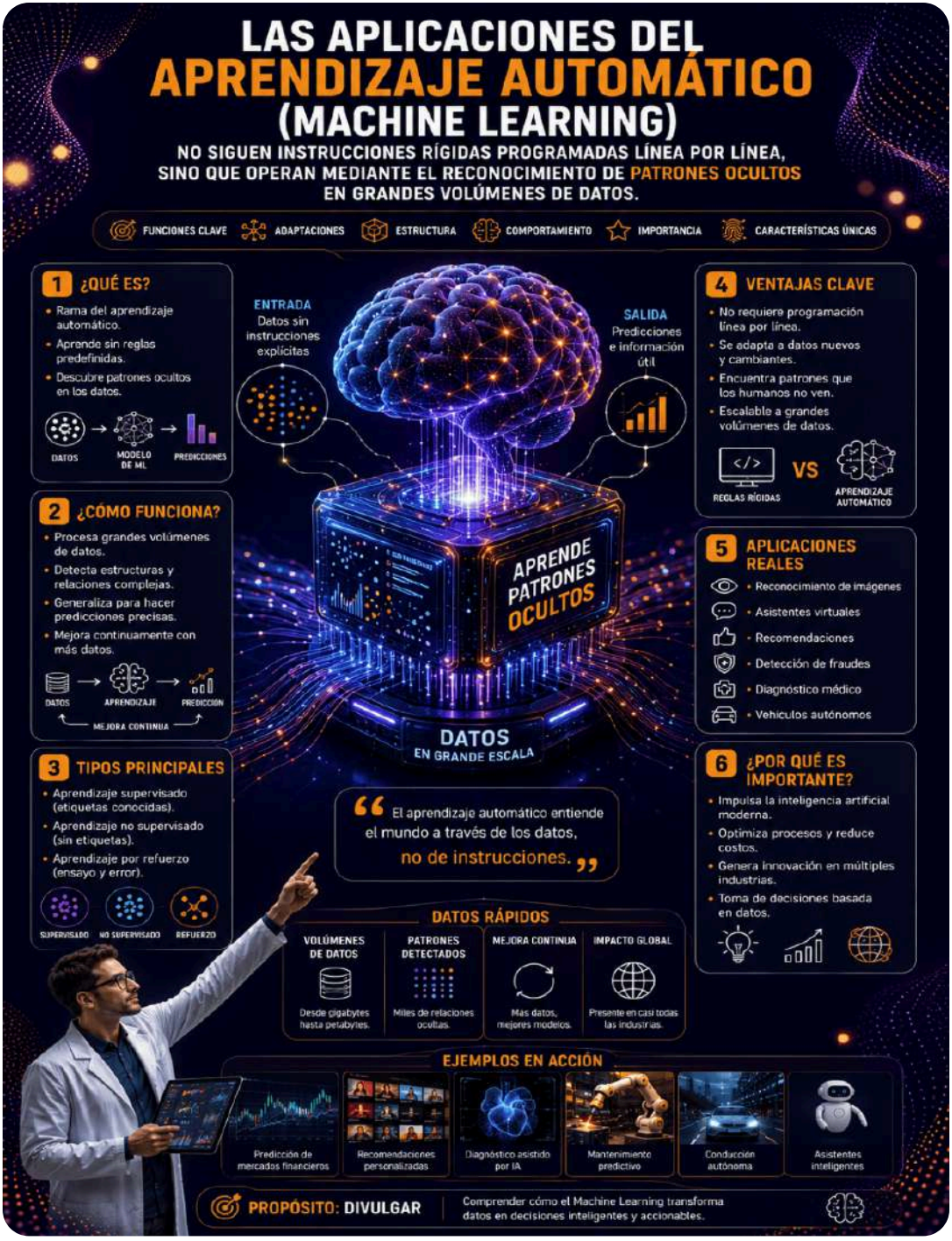


Figura 24. Las aplicaciones operan mediante el reconocimiento de patrones ocultos.

Para lograr esto, la arquitectura de una aplicación de aprendizaje automático integra varios **componentes y procesos internos** clave: la recolección y preprocesamiento de datos, la ingeniería de características, el entrenamiento del modelo y su posterior despliegue. Durante el entrenamiento, la aplicación mide qué tan lejos está su predicción del valor real mediante una **función de pérdida** $\mathcal{L}$. A través de algoritmos de optimización como el **descenso del gradiente**, los parámetros se actualizan iterativamente siguiendo la regla:

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_\theta \mathcal{L}(\theta^{(t)})$$

Donde η es la tasa de aprendizaje y $\nabla_\theta \mathcal{L}$ representa el gradiente de la función de pérdida. Este proceso iterativo permite que la aplicación "aprenda" de sus propios errores y minimice las desviaciones en predicciones futuras.

En la práctica, este mecanismo impulsa diversas **aplicaciones prácticas y casos reales** que impactan nuestra vida cotidiana:

Visión por computadora: Detección de anomalías en radiografías médicas y reconocimiento de obstáculos en vehículos autónomos.

Procesamiento del Lenguaje Natural (PLN): Sistemas de traducción en tiempo real, asistentes de voz y detectores de spam en el correo electrónico.

Sistemas de recomendación: Algoritmos en plataformas como Netflix, Spotify o Amazon que predicen los gustos del usuario analizando patrones de comportamiento globales.

Entre los principales **beneficios y ventajas** de estas aplicaciones destacan su capacidad de **escalabilidad** y su **adaptabilidad**. A diferencia del software convencional, estas herramientas pueden procesar

millones de variables simultáneamente y mejorar progresivamente su precisión a medida que interactúan con nuevos datos, sin necesidad de ser rediseñadas desde cero por un programador.

Sin embargo, el desarrollo de estas tecnologías implica serios **desafíos y limitaciones**. El rendimiento depende estrictamente de la calidad de los datos recopilados; datos sesgados o incompletos generan modelos deficientes (Garbage in, Garbage out). Además, existe el problema del **sobreajuste** (overfitting), donde el modelo memoriza los datos de entrenamiento pero falla al generalizar ante información nueva, así como la falta de explicabilidad en modelos complejos (fenómeno de la "caja negra"), lo que dificulta entender cómo la aplicación llegó a una decisión específica.

Figura 25. ¿Cómo funcionan las aplicaciones del aprendizaje automático?

COLMAYOR
IA LAB

A. RADIOGRAFÍAS MÉDICAS

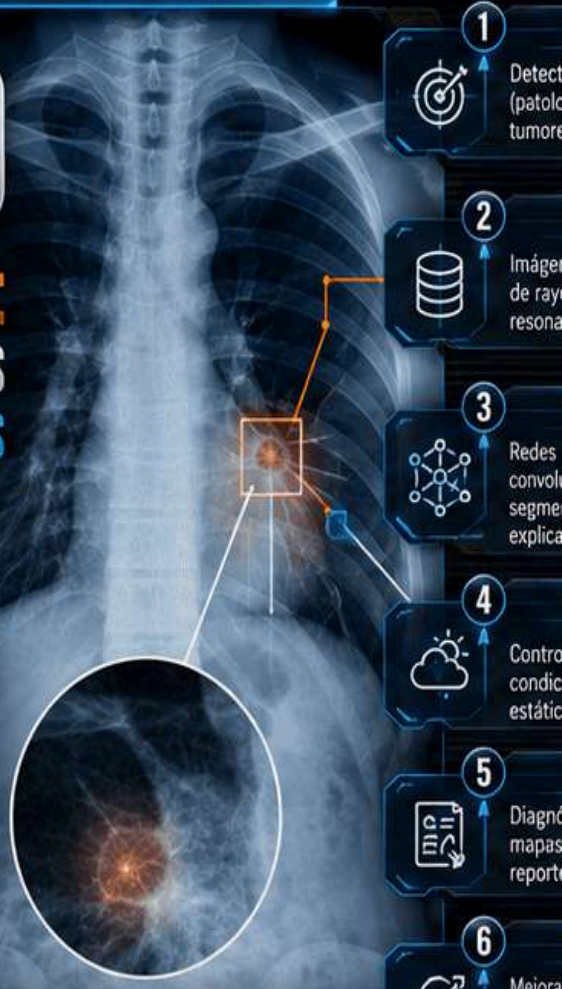
COMPARATIVO

VISIÓN POR COMPUTADORA:

DETECCIÓN DE ANOMALÍAS EN RADIORAFÍAS MÉDICAS Y RECONOCIMIENTO DE OBSTÁCULOS EN VEHÍCULOS AUTÓNOMOS.

PROPÓSITO: DIVULGAR

Comparar cómo la visión por computadora aplica técnicas avanzadas para detectar anomalías en imágenes médicas y reconocer obstáculos en vehículos autónomos, destacando diferencias clave, tecnologías y aplicaciones.



1. Detect (patolo tumore
2. Imáger de ray resona
3. Redes convolu segme explica
4. Contro condic estátic
5. Diagnó mapas report
6. Mejora tempra diagnó

ENFRENTAMIENTO DIRECTO

TIPO DE DATOS

A Imágenes médicas (grises)

B Datos multimodales (imágenes, LIDAR, RADAR)

PRECISIÓN PROMEDIO

A 92%

B 89%

TIEMPO DE RESPUESTA

A Segundos a minutos

B Milisegundos

VOL

A Medio (GBs)

FUENTE: IEEE Xplore, Nature Machine Intelligence, arXiv, NHTSA, OMS

IA QUE VE, ENTIEND

B. VEHÍCULOS AUTÓNOMOS

OBJETIVO PRINCIPAL

ar anomalías
gías, lesiones,
es, fracturas).

Reconocer y localizar obstáculos (peatones, vehículos, objetos).

TIPOS DE DATOS

nes 2D/3D
os X, TAC,
ncias.

Imágenes RGB, LIDAR, RADAR, cámaras multiespectrales.

TÉCNICAS CLAVE

neuronales
ucionales (CNN),
ntación, IA
ble.

Detección de objetos, fusión sensorial, redes neuronales profundas.

ENTORNO DE OPERACIÓN

lado,
iones
as.

Dinámico, entornos cambiantes y no estructurados.

RESULTADO / SALIDA

óstico asistido,
de calor,
es médicos.

Decisiones en tiempo real para evitar colisiones.

IMPACTO

en la detección
na y precisión
stica.

Mayor seguridad vial y reducción de accidentes.

UMEN DE DATOS

B Muy alto (TBs-PBs)

ADAPTABILIDAD

A Alta en datos históricos

B Alta en tiempo real

APLICACIÓN PRINCIPAL

A Salud humana

A Movilidad autónoma

VEREDICTO FINAL

Ambas aplicaciones usan visión por computadora para resolver problemas críticos, pero en distintos mundos: médico y automotriz. La IA transforma datos en decisiones que salvan vidas y construyen el futuro.

aplicaciones del aprendizaje automático

Una aventura interactiva

Descubriendo el ML

Alex:
¡Mira esto, Maya! Mi teléfono recomienda canciones que me encantan. ¿Cómo lo hace?

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

El algoritmo de TikTok

¿Cómo será el futuro del aprendizaje automático?

El futuro del aprendizaje automático se encamina hacia una era de **automatización total, sostenibilidad y computación cuántica**, transformando el **concepto fundamental** de cómo los sistemas procesan la información. En lugar de depender de una supervisión humana intensiva para ajustar hiperparámetros o estructurar conjuntos de datos, los modelos del mañana utilizarán el **aprendizaje automático automatizado (AutoML)** y algoritmos autosuficientes capaces de optimizarse en tiempo real.

Figura 26. ¿Cómo será el futuro del aprendizaje automático?

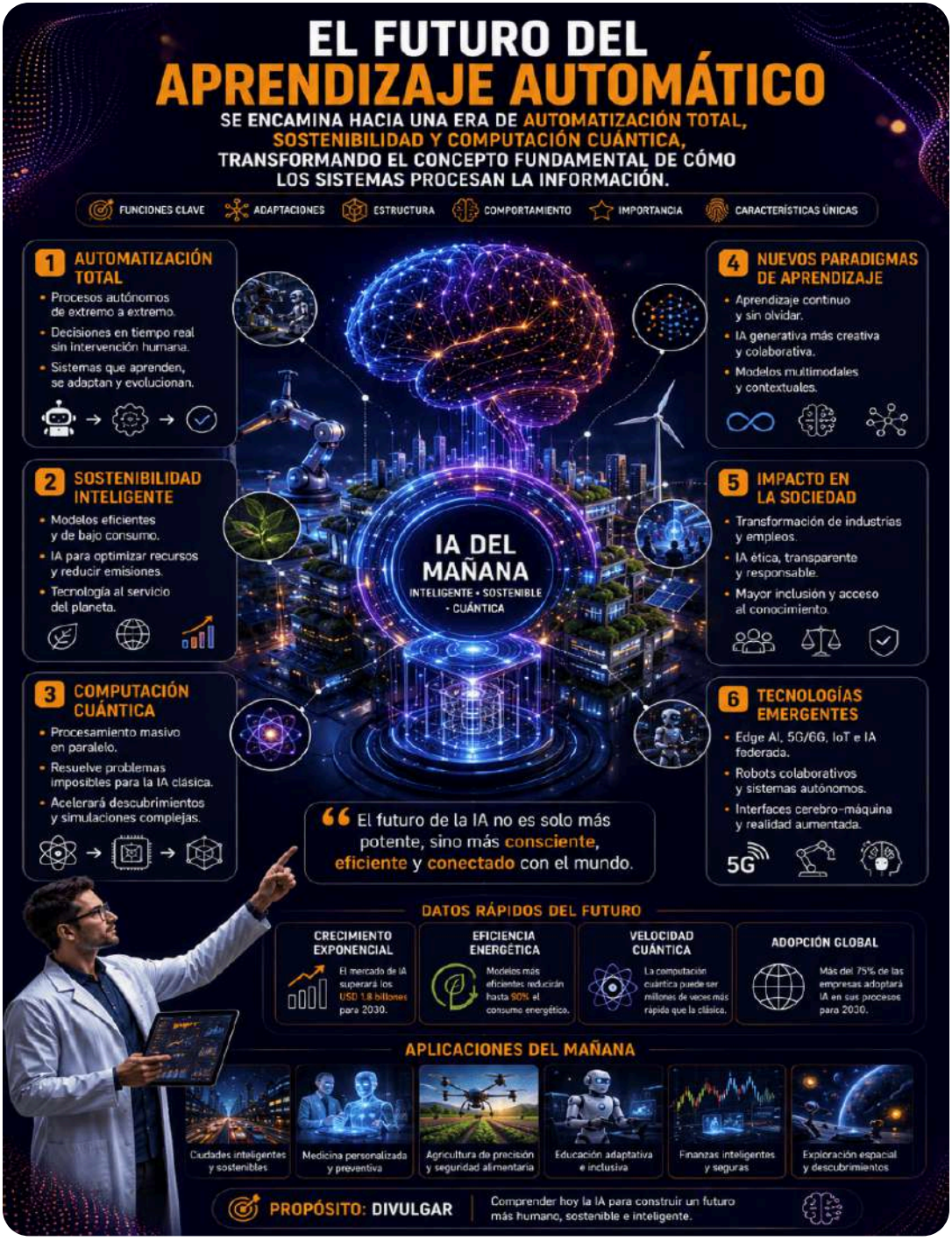


Figura 27. El futuro de la IA es más consciente, eficiente y conectado con el mundo.

En esta nueva etapa, el objetivo matemático evolucionará desde la simple minimización del error empírico hacia una optimización multiobjetivo consciente de los recursos, la cual se puede expresar mediante la función:

$$\theta^* = \arg\min_{\theta} \left(\mathcal{L}_{\text{tarea}}(\theta) + \lambda \cdot \mathcal{C}_{\text{eficiencia}}(\theta) \right)$$

Donde θ representa los parámetros ajustables del modelo, $\mathcal{L}_{\text{tarea}}$ mide la imprecisión en la tarea asignada, y $\mathcal{C}_{\text{eficiencia}}$ es una función de costo que penaliza el consumo energético y la huella de carbono, regulada por el hiperparámetro λ.

Entre los **componentes principales** se destacan los procesadores neuromórficos —diseñados para imitar la estructura física de las neuronas biológicas—, las computadoras cuánticas basadas en cúbits ($|\psi\rangle$) y los algoritmos de **aprendizaje autosupervisado**. El **principio de funcionamiento** y los **procesos internos** se basarán en la búsqueda de arquitecturas neuronales dinámicas (NAS, por sus siglas en inglés). En este proceso, el algoritmo no solo aprende los pesos de una red fija, sino que modifica activamente su propia topología, añadiendo o eliminando conexiones según la complejidad del problema sin requerir la intervención de un diseñador humano.

Las **aplicaciones prácticas** y **casos reales** proyectados transformarán industrias críticas a una velocidad sin precedentes:

Medicina de precisión: Diseño de fármacos personalizados en cuestión de horas mediante la simulación cuántica del plegamiento de proteínas.

Modelado climático global: Predicción exacta de fenómenos meteorológicos extremos con meses de anticipación al resolver ecuaciones complejas de la dinámica de fluidos.

Sistemas autónomos avanzados: Vehículos y robots capaces de navegar e interactuar en entornos totalmente desconocidos sin entrenamiento previo en esos escenarios específicos.

Figura 28. El futuro del aprendizaje automático.

Esta evolución ofrecerá **beneficios y ventajas** sustanciales, como una verdadera democratización del desarrollo de software a través de plataformas *no-code* impulsadas por inteligencia artificial, así como la capacidad de lograr un "aprendizaje de pocos disparos" (*few-shot learning*). Esto último permitirá a los algoritmos generalizar conceptos abstractos a partir de una única muestra de entrenamiento, emulando la eficiencia cognitiva del cerebro humano y reduciendo drásticamente la necesidad de recopilar petabytes de datos etiquetados.

Sin embargo, el avance del aprendizaje automático también conlleva serios **desafíos y limitaciones**. La creciente complejidad de los modelos acentúa el problema de la "caja negra", lo que exige el desarrollo de una **IA explicable (XAI)** capaz de transparentar el razonamiento matemático detrás de cada diagnóstico médico o decisión financiera. Asimismo, garantizar la seguridad frente a ataques adversarios, erradicar los sesgos discriminatorios en los datos de origen y regular el uso ético de modelos superinteligentes constituyen los obstáculos más complejos que la comunidad científica deberá resolver en los próximos años.

COLMAYOR
IA LAB

A. APRENDIZAJE AUTOMÁTICO TRADICIONAL (ACTUAL)

COMPARATIVO

EL FUTURO DEL APRENDIZAJE AUTOMÁTICO SE ENCAMINA HACIA UNA ERA DE AUTOMATIZACIÓN TOTAL, SOSTENIBILIDAD Y COMPUTACIÓN CUÁNTICA, TRANSFORMANDO EL CONCEPTO FUNDAMENTAL DE CÓMO LOS SISTEMAS PROCESAN LA INFORMACIÓN.

PROPÓSITO: DIVULGAR

Comparar las principales tendencias que definirán el futuro del aprendizaje automático, destacando sus diferencias funcionales, impactos y potencial transformrador.

CENTROS DE DATOS CONVENCIONALES

ENTRENAMIENTO EN LA NUBE

ALGORITMOS CLÁSICOS

1 Aprendiz… datos ma… estadísti…

2 Limitada… clásico … Escalabi…

3 Alto con… energéti… entrenar…

4 Horas a… entrenar… complej…

5 Huella d… por infra… intensiva…

6 Reconoci… imágenes… recomen…

ENFRENTAMIENTO RÁPIDO

POTENCIA DE CÓMPUTO

A. Tradicional — B. Cuántico

1X — 100X+

CONSUMO ENERGÉTICO

A. Tradicional: 100%

B. Cuántico: 10-20%

TIEMPO DE ENTRENAMIENTO

A. Tradicional: Semanas

B. Cuántico: Minutos

El aprendizaje automático evoluciona, pero su futuro será cuántico, sostenible y exponencial.

MENOS LÍMITES, …

EL FUTURO YA ES…

B. APRENDIZAJE AUTOMÁTICO DEL FUTURO (CUÁNTICO + SOSTENIBLE)

PARADIGMA CENTRAL

aje basado en
sivos y modelos
cos clásicos.

Aprendizaje potenciado por qubits y superposición cuántica.

CAPACIDAD DE CÓMPUTO

por hardware
(CPU/GPU).
lidad lineal.

Procesamiento cuántico paralelo. Escalabilidad exponencial.

EFICIENCIA ENERGÉTICA

sumo
co en
miento.

Potencial de reducción drástica del consumo energético.

VELOCIDAD DE APRENDIZAJE

semanas para
modelos
os.

Resolución de problemas complejos en minutos o segundos.

SOSTENIBILIDAD

e carbono alta
estructura
a.

Infraestructura más sostenible y eficiente a largo plazo.

APLICACIONES CLAVE

miento de
s, NLP, predicción,
daciones.

Descubrimiento de fármacos, optimización, simulaciones cuánticas, IA avanzada.

PROCESADORES CUÁNTICOS

INFRAESTRUCTURA SOSTENIBLE

ALGORITMOS CUÁNTICOS

	ESCALABILIDAD		IMPACTO AMBIENTAL		POTENCIAL DE INNOVACIÓN	
	A. Tradicional	B. Cuántico	A. Tradicional	B. Cuántico	A. Tradicional	B. Cuántico
	Lineal	Exponencial	Alto	Bajo	Alto	Muy alto

MÁS POSIBILIDADES.
TÁ APRENDIENDO.

- A. HOY: APRENDIZAJE AUTOMÁTICO TRADICIONAL
- B. MAÑANA: APRENDIZAJE AUTOMÁTICO CUÁNTICO + SOSTENIBLE

Simulador: el futuro del aprendizaje automático

Recursos: **500**

Prestigio: **50**

Progreso: **0**/5

El Maestro Robot

El año es 2042. La IA ha avanzado hasta el punto de que puede enseñar de manera personalizada a cada estudiante.

Un aula futurista con un robot amigable interactuando con estudiantes.

Registro

Generador de Iceberg Pro

Nivel 1: La Superficie

Nivel 1

AGI inminente -


DESCARTES
matemáticas interactivas