



Inteligencia Artificial

Inteligencia artificial.



Transformación Digital
Agencia de Transformación Digital y Telecomunicaciones



TECNOLÓGICO
NACIONAL DE MÉXICO®

Introducción

El aprendizaje automático (machine learning) es una subdisciplina fundamental de la inteligencia artificial que permite a los sistemas computacionales aprender de los datos y mejorar su desempeño sin requerir programación explícita. En lugar de seguir instrucciones predeterminadas, los algoritmos de aprendizaje automático se entrenan para identificar patrones, realizar predicciones y tomar decisiones informadas mediante el análisis de grandes volúmenes de información. Los sistemas de aprendizaje automático se clasifican en tres enfoques principales según su mecanismo de aprendizaje:

- 1)** El aprendizaje supervisado: Utiliza datos etiquetados para entrenar modelos capaces de predecir resultados o clasificar información, siendo aplicable en tareas como detección de spam y predicción de precios.
- 2)** El aprendizaje no supervisado: Descubre patrones y estructuras inherentes en datos sin etiquetar, empleándose en segmentación de clientes y análisis de asociaciones.
- 3)** El aprendizaje por refuerzo: Entrena agentes mediante interacción con un entorno, donde el sistema aprende a través de recompensas y penalizaciones para optimizar decisiones a largo plazo.

En esta lectura se describen los fundamentos para comprender cómo las máquinas procesan información, lo que permite soluciones

automatizadas y eficientes para problemas complejos que resultan imposibles de resolver mediante programación tradicional. Este material servirá de base para desarrollar actividades que permitan reconocer los fundamentos del aprendizaje automático.

2.1 Conceptos Fundamentales

El Aprendizaje Automático (*Machine Learning*) es un subcampo de la inteligencia artificial que se enfoca en desarrollar algoritmos y sistemas que permiten a las computadoras aprender patrones a partir de datos, haciendo predicciones o tomas de decisiones sin requerir programación explícita para cada caso. Esta capacidad de aprender mediante la experiencia refleja procesos análogos a cómo los humanos adquieren conocimiento.

Componentes esenciales de los modelos de Aprendizaje Automático

Todo algoritmo de aprendizaje automático comprende tres elementos fundamentales:

- **Representación:** Define qué aspecto del modelo captura el conocimiento y cómo se estructura internamente. La elección de la representación impacta directamente en la capacidad del modelo para capturar relaciones en los datos.
- **Evaluación:** Proporciona un mecanismo para diferenciar entre modelos buenos y malos. Las métricas de evaluación cuantifican qué tan bien un modelo generaliza datos no vistos.

- **Optimización:** Es el proceso para encontrar buenos modelos dentro del espacio de búsqueda definido por la representación. Implica ajustar los parámetros del modelo para minimizar los errores.

Ciclo para el desarrollo de modelos de Aprendizaje Automático

El desarrollo de modelos de aprendizaje automático sigue un proceso sistemático:

1. **Recopilación y preparación de datos.** Los datos son fundamentales para el desempeño del modelo, ya que el aprendizaje automático es impulsado por datos.
2. **División de datos.** El conjunto de datos se divide en tres subconjuntos distintos:
 - **Conjunto de entrenamiento:** Se utiliza para entrenar el modelo, donde el algoritmo aprende los patrones en los datos.
 - **Conjunto de validación:** Proporciona una evaluación imparcial del rendimiento del modelo durante el entrenamiento, lo que permite la sintonización de hiperparámetros (parámetros que controlan el comportamiento del modelo). Ayuda a identificar y prevenir el **sobreajuste** (overfitting).
 - **Conjunto de prueba:** Ofrece una evaluación final, imparcial y definitiva del rendimiento en datos completamente nuevos, reflejando la capacidad del modelo para generalizar.
3. **Selección y entrenamiento del modelo.** Se elige un algoritmo y se entrena con los datos de entrenamiento.

4. **Evaluación y sintonización.** Se evalúa el rendimiento en el conjunto de validación y se ajustan los hiperparámetros.
5. **Predicciones finales.** El modelo entrenado realiza predicciones en datos nuevos.

La Figura 1. Representa de manera resumida el ciclo de desarrollo de modelos de Aprendizaje Automático:

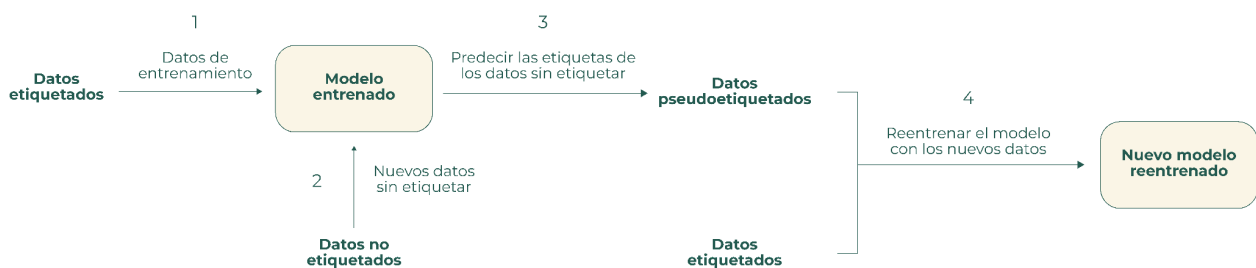


Figura 1. Representación del ciclo de desarrollo de modelos de Aprendizaje Automático

En la siguiente sección, se abordan los tipos de Aprendizaje Automático con la finalidad de proveer al lector los fundamentos necesarios para una comprensión sobre el tema.

2.2 Tipos de Aprendizaje Automático

El Aprendizaje Automático se fundamenta en el procesamiento de datos y algoritmos matemáticos para identificar relaciones complejas entre variables. Su objetivo central es desarrollar **modelos que generalicen bien** con datos no vistos previamente, lo que implica aplicar los patrones aprendidos a situaciones nuevas sin memorizar los datos de entrenamiento. Los tipos principales de Aprendizaje Automático —**Supervisado, No Supervisado y por Refuerzo**— se abordan en las siguientes subsecciones.

2.2.1 Aprendizaje Supervisado

El aprendizaje supervisado es una técnica fundamental del aprendizaje automático que utiliza conjuntos de datos etiquetados para entrenar modelos de inteligencia artificial capaces de realizar predicciones precisas. Se define como un método en el que los algoritmos aprenden a partir de datos de formación que incluyen tanto las variables de entrada, como sus correspondientes etiquetas o salidas deseadas, lo que permite que el modelo generalice y prediga resultados para datos nuevos y no etiquetados.

Definición y Conceptos Fundamentales del Aprendizaje Supervisado

En el aprendizaje supervisado, el conjunto de datos de entrenamiento consiste de pares de objetos en los que un componente representa los datos de entrada y la otra, los resultados deseados conocidos como etiquetas (labels). El objetivo principal es crear una función capaz de mapear correctamente los valores de entrada a sus correspondientes salidas, aprendiendo de los patrones presentes en los datos de entrenamiento y siendo capaz de aplicar ese conocimiento a situaciones nuevas no vistas previamente.

El aprendizaje supervisado se fundamenta en el proceso de generalización, donde el algoritmo debe ser capaz de sintetizar patrones generales a partir de ejemplos específicos. Este concepto es esencial, ya que determina si el modelo puede funcionar correctamente con datos desconocidos o si, por el contrario, ha memorizado simplemente el conjunto de entrenamiento sin aprender relaciones genuinas entre variables.

Tipos Principales de Aprendizaje Supervisado

El aprendizaje supervisado se divide en dos categorías principales según la naturaleza de la variable objetivo:

- **Clasificación:** En los **problemas de clasificación**, la variable objetivo es de tipo **categorístico o discreto**. Esto significa que el modelo predice **etiquetas o categorías predefinidas**. Los ejemplos típicos incluyen identificar correos electrónicos como *spam* o no *spam*, diagnosticar enfermedades con base en síntomas, o reconocer dígitos manuscritos.
- **Regresión:** En los problemas de regresión, la variable objetivo es continua o numérica, lo que permite predicciones de valores en una escala continua. Esto se utiliza, por ejemplo, para predecir el precio de un inmueble basándose en características como metros cuadrados y número de habitaciones, estimar salarios futuros según experiencia laboral, o pronosticar temperaturas.

Ejemplos de aplicaciones prácticas del Aprendizaje Supervisado

El aprendizaje supervisado tiene amplias aplicaciones en múltiples industrias. En medicina, se utiliza para el diagnóstico automático mediante el análisis de imágenes médicas y datos previamente etiquetados. En finanzas, se emplea en la detección de fraude y en la predicción de análisis de mercado de valores. En comunicaciones, se aplica a la detección de spam en correos electrónicos. Además, se utiliza en reconocimiento de imágenes, procesamiento de lenguaje natural,

robótica, sistemas de recomendación y muchas otras áreas en las que la precisión predictiva es crítica.

Relación con Otros Tipos de Aprendizaje

El aprendizaje supervisado se diferencia del aprendizaje no supervisado en que este último trabaja con datos sin etiquetar y busca descubrir patrones ocultos en la estructura de los datos. Existe también el aprendizaje semisupervisado, que combina técnicas de ambos enfoques utilizando pequeñas cantidades de datos etiquetados junto con grandes cantidades de datos sin etiquetar, lo que resulta particularmente útil cuando el etiquetado es costoso o impráctico.

2.2.2 Aprendizaje No Supervisado

El aprendizaje no supervisado es un tipo de aprendizaje automático que analiza y modela datos sin etiquetas predefinidas o categorías establecidas de antemano. A diferencia del aprendizaje supervisado, que utiliza datos etiquetados para entrenar modelos, el aprendizaje no supervisado trabaja con datos sin procesar buscando descubrir patrones ocultos, estructuras inherentes y relaciones intrínsecas dentro de los datos, sin intervención humana explícita.

Definición y características fundamentales del Aprendizaje No Supervisado

El aprendizaje no supervisado se define como un método de aprendizaje automático en el que los algoritmos aprenden a partir de datos sin etiquetas previas ni categorías predefinidas. El objetivo principal es

identificar patrones ocultos, estructuras, relaciones y agrupaciones en los datos, lo que permite que el sistema haga inferencias sin instrucciones explícitas. Esta aproximación promueve la autonomía en el análisis de datos, permitiendo que los algoritmos descubran información latente y segmenten datos sin intervención humana.

Una característica distintiva del aprendizaje no supervisado es que no requiere datos etiquetados, lo que ahorra tiempo y esfuerzo en la anotación de datos, aunque los resultados generalmente son menos precisos que los del aprendizaje supervisado debido a la ausencia de datos etiquetados. Sin embargo, la adquisición de datos etiquetados requiere intervención humana y puede ser tediosa, costosa e incluso imposible en algunos casos, como ocurre con datos biológicos.

Tipos principales de aprendizaje no supervisado

El aprendizaje no supervisado se subdivide en tres tareas principales:

- 1. Agrupamiento (Clustering):** El agrupamiento es la tarea más popular del aprendizaje no supervisado. Se trata de una técnica para explorar datos sin etiquetar, dividiéndolos en grupos o clusters basados en similitudes o diferencias. En términos matemáticos, para un espacio muestreado x_1 a x_n , las etiquetas de clase verdaderas no se proporcionan para cada muestra, razón por la cual se conoce como aprendizaje sin supervisor.

Los algoritmos de agrupamiento se clasifican en cuatro tipos principales:

- **Agrupamiento exclusivo:** Los datos se agrupan de manera exclusiva, de modo que si un punto de datos pertenece a un cluster definido, no puede incluirse en otro cluster. Un ejemplo simple sería separar puntos de datos mediante una línea recta en un plano bidimensional.
- **Agrupamiento superpuesto:** Utiliza conjuntos difusos (fuzzy sets) para agrupar datos, de modo que cada punto puede pertenecer a dos o más clusters con diferentes grados de membresía. Los datos se asocian con un valor de membresía apropiado.
- **Agrupamiento jerárquico:** Este tipo de agrupamiento crea una estructura de árbol conocida como dendrograma. Se basa en la unión entre los dos clusters más cercanos, comenzando con cada punto de datos como un clúster individual hasta alcanzar los clusters finales deseados.
- **Agrupamiento probabilístico:** Los clusters se crean utilizando distribuciones de probabilidad.

2. Reglas de Asociación (Association Rule Mining): Las reglas de asociación representan un enfoque basado en reglas para descubrir relaciones interesantes entre características en un conjunto de datos. Este tipo de aprendizaje no supervisado identifica patrones en forma de reglas "si-entonces", mostrando

cómo la presencia de algunos elementos en los datos implica la presencia de otros.

Se utiliza comúnmente en el análisis de canasta de mercado, una técnica de minería de datos que los minoristas utilizan para obtener una mejor comprensión de los patrones de compra de los clientes, basándose en las relaciones entre varios productos. El algoritmo más ampliamente utilizado para el aprendizaje de reglas de asociación es el algoritmo Apriori, aunque también se utilizan otros, como Eclat y FP-growth.

3. Reducción de Dimensionalidad: La reducción de dimensionalidad es una técnica que simplifica conjuntos de datos complejos mediante la reducción del número de variables o características, manteniendo la información esencial. Esta tarea es especialmente útil cuando se trabaja con datos de alta dimensionalidad, donde los datos suelen estar ubicados cerca de un subespacio de baja dimensión en el espacio de características.

Proceso de Trabajo del Aprendizaje No Supervisado

El proceso típico del aprendizaje no supervisado involucra los siguientes pasos:

1. Recopilación de datos sin etiquetar: Se reúne un conjunto de datos sin etiquetas o categorías predefinidas.

- 2. Selección del algoritmo:** Se elige un algoritmo no supervisado apropiado, como agrupamiento, minería de reglas de asociación (Apriori) o reducción de dimensionalidad (PCA) basándose en el objetivo.
- 3. Entrenamiento del modelo en datos sin procesar:** Se alimenta el conjunto de datos completo sin etiquetar al algoritmo que busca similitudes, relaciones o estructuras ocultas dentro de los datos.
- 4. Agrupamiento o transformación de datos:** El algoritmo organiza los datos en grupos (clusters), reglas o formas de menor dimensionalidad sin intervención humana.
- 5. Interpretación y uso de resultados:** Se analizan los grupos descubiertos, reglas o características para obtener información o utilizarlos para tareas posteriores como visualización, detección de anomalías o como entrada para otros modelos.

Aplicaciones del Aprendizaje No Supervisado

El aprendizaje no supervisado tiene amplia aplicación en diversos campos:

- **Segmentación de Clientes:** En marketing, el aprendizaje no supervisado agrupa a clientes según comportamientos similares, permitiendo a las empresas identificar diferentes segmentos de clientes y dirigir publicidad de manera más efectiva.
- **Análisis de Canasta de Mercado:** Esta técnica facilita la exploración de relaciones entre productos que se compran conjuntamente, permitiendo a las tiendas colocar productos relacionados más cerca entre sí.

- **Análisis Exploratorio de Datos:** Las técnicas de aprendizaje no supervisado se utilizan ampliamente para explorar datos y descubrir estructuras inherentes ocultas no comprendidas o etiquetadas previamente.
- **Análisis de Imágenes y Documentos:** El aprendizaje no supervisado agrupa imágenes o documentos similares en un grupo, lo que es muy útil para grandes bibliotecas de fotos o categorización de documentos de investigación por tema.
- **Análisis de Expresión Génica:** El aprendizaje no supervisado agrupa genes que se expresan de manera similar, ayudando a los investigadores a encontrar función génica y posibles objetivos de fármacos.
- **Astronomía:** El aprendizaje no supervisado permite el análisis de grandes conjuntos de objetos astronómicos, donde se clasifican galaxias o se detectan cúmulos de estrellas con propiedades similares.
- **Procesamiento del Lenguaje Natural (NLP):** Google News aprovecha el aprendizaje no supervisado para categorizar artículos basándose en la misma historia de varios medios de comunicación.
- **Análisis de Red Social:** El aprendizaje no supervisado se utiliza para estudiar conexiones e interacciones de usuarios, identificando comunidades que comparten intereses similares o descubriendo cómo se propagan información y tendencias en redes sociales.

Ventajas y desventajas del Aprendizaje no Supervisado

Ventajas:

- No requiere datos etiquetados, lo que ahorra tiempo y esfuerzo en anotación de datos
- Descubre patrones ocultos que podrían ser pasados por alto por humanos
- Maneja conjuntos de datos complejos y de gran escala de manera efectiva
- Útil para detección de anomalías sin ejemplos previos
- Particular eficaz para análisis exploratorio de datos

Desventajas:

- Los resultados generalmente son menos precisos que el aprendizaje supervisado debido a la ausencia de datos etiquetados
- La interpretabilidad y la transparencia pueden ser desafíos en algoritmos más complejos
- Pueden reflejar y perpetuar sesgos presentes en los datos

2.2.3 Aprendizaje por Refuerzo

El aprendizaje por refuerzo (RL, por sus siglas en inglés) es un paradigma de aprendizaje automático en el cual un agente autónomo aprende a tomar decisiones óptimas mediante la interacción con un entorno, guiado por un sistema de recompensas y penalizaciones. A diferencia del aprendizaje supervisado o no supervisado, el RL se basa en el método de prueba y error: el agente ejecuta acciones, observa las consecuencias y

ajusta su política de acciones en busca de maximizar las recompensas acumuladas a largo plazo.

¿Cómo funciona el Aprendizaje por Refuerzo?

Este proceso gira en torno a los siguientes elementos fundamentales:

- Agente: Quién toma las decisiones
- Entorno: El sistema o contexto con el que el agente interactúa
- Acción: Decisión tomada por el agente
- Estado: Situación actual del entorno
- Recompensa: Retroalimentación numérica recibida tras cada acción, positiva si se acerca al objetivo, negativa si lo aleja

El agente, a través de múltiples episodios de interacción, aprende una política (estrategia de acción) que maximiza la suma de recompensas futuras. Ejemplo clásico: una IA aprendiendo a jugar al ajedrez, donde cada movimiento se recompensa o penaliza dependiendo de su contribución a la victoria final.

Tipos de Aprendizaje por Refuerzo

Existen dos grandes tipos de aprendizaje por refuerzo, según la manera en que el agente interactúa con el entorno y aprende de su experiencia:

TIPO	DESCRIPCIÓN
Basado en modelos	El agente construye un modelo interno del entorno (predice transiciones y recompensas) y lo utiliza para planificar sus acciones antes de ejecutarlas.

Libre de modelos (model-free)	El agente aprende directamente de la interacción, sin modelar el entorno explícitamente. Mediante ensayo y error, ajusta los valores asociados a estados y acciones.
--	--

Diferencias con otros aprendizajes

A diferencia del aprendizaje supervisado (usa datos etiquetados) y del aprendizaje no supervisado (busca patrones en datos no etiquetados), el RL busca aprender a actuar: la información sobre lo correcto/incorrecto se obtiene indirectamente por las recompensas recibidas y no por etiquetas predefinidas.

Ejemplos de Aplicación

- Entrenamiento de robots autónomos
- Sistemas inteligentes para juegos como Ajedrez o Go
- Vehículos autónomos y automatización industrial

Conclusión

El aprendizaje automático constituye un paradigma fundamental de la inteligencia artificial que transforma la manera en que los sistemas computacionales procesan información y toman decisiones. A través de los conceptos, ciclos de desarrollo y tipologías presentados, se evidencia que el aprendizaje automático trasciende la programación tradicional al permitir que los sistemas aprendan de datos, identifiquen patrones complejos y generalicen conocimiento a situaciones nuevas sin requerir instrucciones explícitas para cada caso específico.

La estructura tripartita del aprendizaje automático —supervisado, no supervisado y por refuerzo— proporciona soluciones versátiles para distintos contextos y desafíos. El aprendizaje supervisado se posiciona como la solución ideal cuando se dispone de datos etiquetados y se requiere precisión predictiva en aplicaciones críticas como medicina, finanzas y seguridad. El aprendizaje no supervisado se revela como herramienta indispensable para descubrir patrones ocultos y estructuras inherentes en datos complejos, particularmente cuando el etiquetado es costoso o impracticable. Finalmente, el aprendizaje por refuerzo emerge como el enfoque óptimo para sistemas que requieren tomar decisiones secuenciales y optimizar resultados a largo plazo mediante interacción continua con su entorno.

La implementación exitosa del aprendizaje automático depende de un ciclo de desarrollo riguroso y sistemático: desde la recopilación y preparación de datos hasta la evaluación y sintonización de modelos. Este proceso garantiza que los sistemas desarrollados no solo memoricen datos de entrenamiento, sino que adquieran la capacidad de generalización necesaria para funcionar efectivamente con información nunca vista anteriormente.

Finalmente, la versatilidad de aplicaciones del aprendizaje automático en medicina, comercio, comunicaciones, astronomía, análisis genómico y redes sociales, por mencionar solo algunos, demuestra su impacto transformador en múltiples disciplinas. Comprender los fundamentos del aprendizaje automático resulta esencial en el contexto actual, ya que estos principios habilitan a profesionales, educadores y estudiantes a

desarrollar soluciones innovadoras que resuelvan problemas complejos imposibles de abordar mediante programación convencional, posicionando al aprendizaje automático como una competencia crítica para el siglo XXI.

Elaboró contenido: Dr. José Luis Sánchez Cervantes; Dra. Thelma Leonor Estevez Dorantes; Dr. Giner Alor Hernández. Diciembre, 2025.

Referencias:

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. <https://doi.org/10.1117/1.2819119>
- Crespí, A., Mesquida, A., Monserrat, M., & Mas, A. (2025). Lifecycle models in machine learning development. *Expert Systems*, 42(4), e70029. <https://doi.org/10.1111/exsy.70029>
- Deisenroth, M. P., Faisal, A. A., & Ong, C. S. (2020). *Mathematics for machine learning*. Cambridge University Press. <https://doi.org/10.1017/9781108679930>
- Guan, L. (2014). *Unsupervised learning*. Wiley. <https://doi.org/10.1002/9781118875568>
- Ivanov, S. (2022). *Reinforcement learning textbook*. Springer. <https://doi.org/10.48550/arxiv.2201.09746>
- Kolosova, T., & Berestizhevsky, S. (2020). *Supervised machine learning*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429297595>

- Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31(3), 249–268. <https://doi.org/10.48550/arXiv.0711.0458>
- Kristen, K., & Caleb, K. (2023). Rounding out the ML life cycle. En *Machine learning upgrade* (Cap. 6). Wiley. <https://doi.org/10.1002/9781394319763.ch6>
- Lorenzoni, G., Palacios, H., & García, J. (2021). Machine learning model development from a software engineering perspective: A systematic literature review. *Journal of Systems and Software*, 180, 111010. <https://doi.org/10.1016/j.jss.2021.111010>
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill. <https://www.cs.cmu.edu/~tom/mlbook.html>
- Polyzotis, N., Roy, S., Whang, S. E., & Zinkevich, M. (2018). Data lifecycle challenges in production machine learning: A survey. *Proceedings of the VLDB Endowment*, 11(12), 1–13. <https://doi.org/10.14778/3229863.3229877>
- Tammy J, Jaimie L. G, Anthony J. R. (2020). Supervised machine learning: A brief primer. *BMJ*, 369, m1786. <https://doi.org/10.1016/j.beth.2020.05.002>
- Shetty, S. H., Shetty, S., Singh, C., & Rao, A. (2022). Supervised machine learning: Algorithms and applications. En *Machine learning in engineering* (Cap. 1). Wiley. <https://doi.org/10.1002/9781119821908.ch1>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press. <https://doi.org/10.5555/3312046>

- Spjuth, O., Frid, J., & Hellander, A. (2021). The machine learning life cycle and the cloud: implications for drug discovery. *Expert Opinion on Drug Discovery*, 16(9), 1071–1079.
<https://doi.org/10.1080/17460441.2021.1932812>
- Rolf, B., Beier, A., Jackson, I., Müller, M., Reggelin, T., Stuckenschmidt, H., & Lang, S. (2025). A review on unsupervised learning algorithms and applications in supply chain management. *International Journal of Production Research*, 63(5), 1933–1983.
<https://doi.org/10.1080/00207543.2024.2390968>