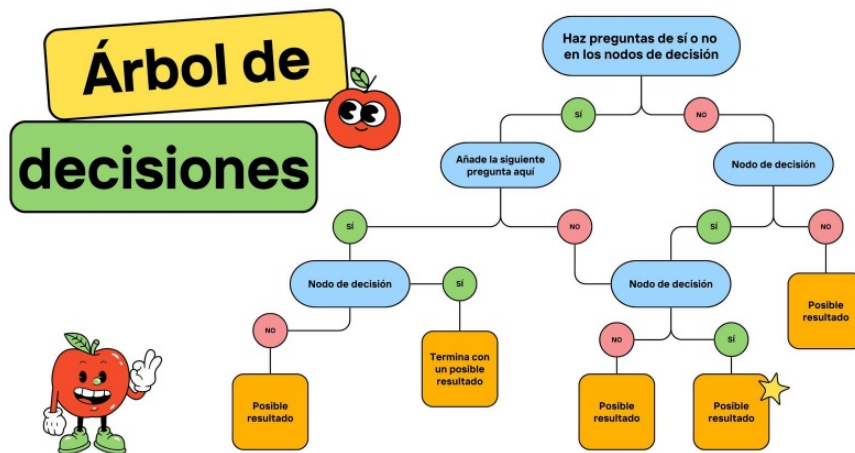


MODELO DE ARBOLES DE DECISIÓN



<https://share.google/images/IPB8t051llqS5gsZE>

Un árbol de decisión es una herramienta analítica que representa de forma esquemática las posibles alternativas para facilitar una toma de decisiones más informada, especialmente en contextos donde intervienen riesgos, costos, beneficios y múltiples opciones. Este modelo se utiliza ampliamente en áreas como la estadística y el aprendizaje automático, con el propósito principal de clasificar información o realizar predicciones a través de una estructura jerárquica de decisiones.



Su funcionamiento se basa en una secuencia lógica de condiciones, donde cada decisión depende del cumplimiento o no de ciertos criterios, lo que conduce a diferentes resultados. Además, los árboles de decisión se consideran algoritmos de aprendizaje supervisado no paramétrico, lo que implica que no requieren asumir una distribución específica de los datos y pueden adaptarse fácilmente a relaciones complejas entre múltiples variables.

Elementos que conforman un árbol de decisión

Elemento	Función
Nodo raíz	Punto de partida del análisis, con la condición inicial de mayor importancia.
Nodos internos	Contienen divisiones adicionales según otras variables.
Ramas	Representan los diferentes caminos según las decisiones tomadas.
Nodos hoja	Resultado o predicción final después de evaluar todas las condiciones previas.



Pasos para realizar un árbol de decisiones.

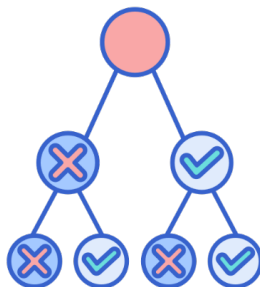
- Definir el objetivo
- Seleccionar variables predictoras
- Recopilar y preparar los datos
- Elegir criterio de división
- Construir el árbol
- Establecer criterios de parada
- Poda del árbol (opcional)
- Evaluar el modelo (Mide su rendimiento con métricas como precisión, F1-score o R^2)
- Aplicar el modelo (usa el árbol para tomar decisiones o hacer predicciones con nuevos datos).

Ejemplo:

Supongamos que queremos saber si un cliente realizará una compra:

1. ¿El cliente tiene tarjeta de lealtad?

- Sí → ¿Ha comprado en el último mes?
 - Sí → Resultado: Compra
 - No → Resultado: No compra
- No → Resultado: No compra



Ventajas

- Interpretación sencilla y visual.
- Manejo de datos mixtos (categóricos y numéricos).
- Adaptabilidad a relaciones no lineales entre variables.
- Útiles en contextos donde hay **muchas variables independientes** y se busca entender su impacto combinado.



Desventajas

- Tienen a **sobreaajustarse** si no se regulan (poda o profundidad limitada).
- **Sensibles a pequeñas variaciones** en los datos.
- Menor precisión que modelos más complejos cuando se usan de forma individual.

Referencia:

Breiman, L. (1984). *Classification and regression trees*. Estados Unidos. Wadsworth.
Redacción IBM. (2021) ¿Qué es un árbol de decisión? IBM. Recuperado de: <https://www.ibm.com/mx-es/think/topics/decision-trees>