



INTRODUCCIÓN AL ANÁLISIS SUPERVISADO

EQUIPO

JOSE ALBERTO



ALEXIS ALATORRE



MARIO IMANOL



ANTHONY LAMAS



¿QUÉ ES EL ANÁLISIS SUPERVISADO?

DEFINICIÓN

El análisis supervisado es una técnica de machine learning que usa datos etiquetados para entrenar modelos predictivos.

Cada ejemplo en los datos de entrenamiento incluye características (entradas) y etiquetas (salidas deseadas).

CLASIFICACIÓN

Clasificación: Predecir categorías (ej. spam vs. no spam)

OBJETIVO DEL ANÁLISIS SUPERVISADO

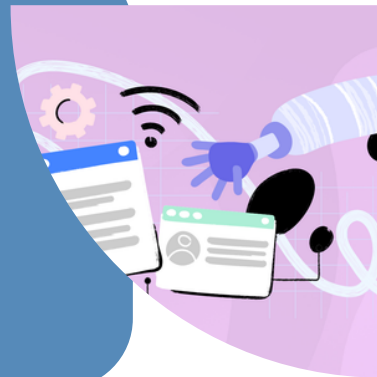
Construir modelos que puedan hacer predicciones precisas sobre nuevos datos



Generalizar bien a datos no vistos, evitando el sobreajuste



Aprender la relación entre características de entrada y etiquetas de salida



Identificar las características más importantes para la predicción



PASOS BÁSICOS DEL APRENDIZAJE SUPERVISADO

PASO 1

Recopilar datos etiquetados: Reunir datos con características y etiquetas conocidas

PASO 2

Preparación y preprocesamiento de datos:

- Limpiar datos, manejar valores faltantes, eliminar outliers

PASO 3

Elegir un modelo apropiado: Seleccionar un algoritmo adecuado (ej. regresión logística, árboles de decisión, redes neuronales)

PASO 4

Entrenar el modelo: Alimentar datos de entrenamiento, el modelo aprende parámetros

PASO 5

Evaluar el modelo: Hacer predicciones sobre datos de prueba, calcular métricas de rendimiento

PASO 6

Ajustar y afinar el modelo: Analizar errores, probar diferentes algoritmos y parámetros

PASO 7

Hacer predicciones sobre nuevos datos: Aplicar el modelo final a datos no vistos

VENTAJAS DEL ANÁLISIS SUPERVISADO

- Puede aprender relaciones complejas y no lineales de los datos
- Proporciona predicciones precisas y cuantitativas para problemas bien definidos
- Muchos algoritmos bien estudiados y bibliotecas de software disponibles

IDEA SOSTENIBLE

- Requiere datos etiquetados, que pueden ser costosos y llevar tiempo de obtener
- El rendimiento depende en gran medida de la calidad y cantidad de los datos de entrenamiento
- Puede tener un mal rendimiento en datos muy diferentes a los datos de entrenamiento
- Algunos algoritmos pueden ser difíciles de interpretar (modelos de "caja negra")

Caja Negra: ni siquiera los programadores y administradores de la máquina o el algoritmo saben o entienden cómo se llegó al resultado. Solo el algoritmo mismo es consciente de exactamente cómo se tomaron las decisiones.

APLICACIONES COMUNES

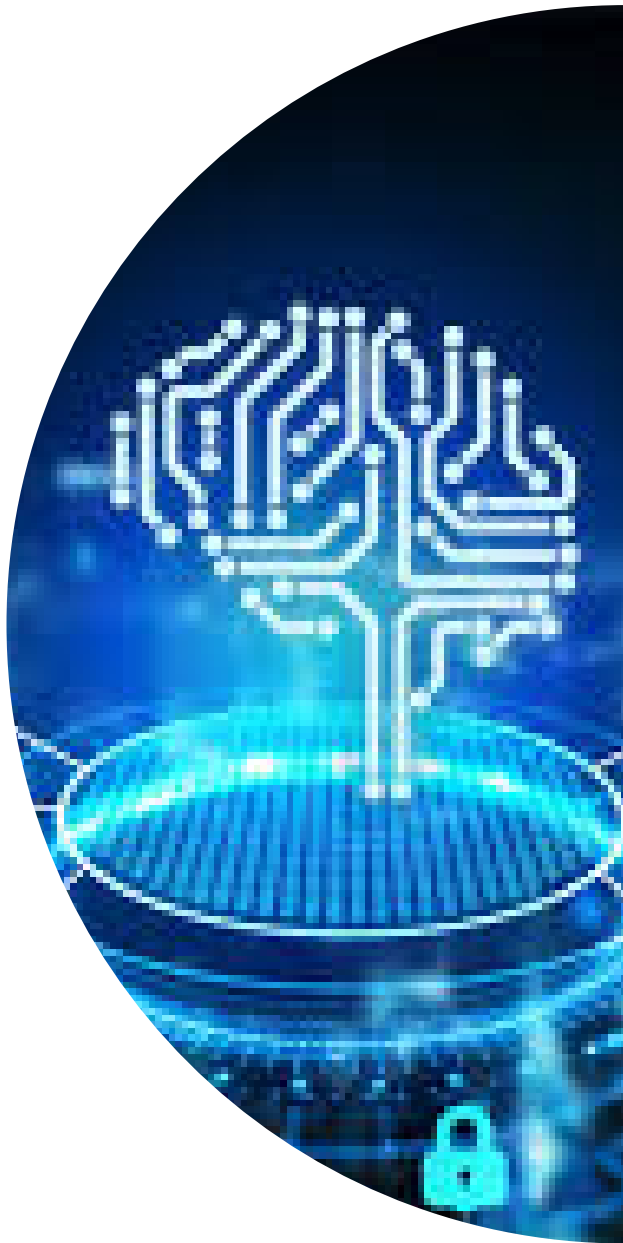
Detección de fraude:
Clasificar transacciones
como fraudulentas o
legítimas

Diagnóstico médico:
Predecir enfermedades
basándose en síntomas y
pruebas

Reconocimiento de
imágenes: Clasificar
imágenes en categorías
(perros, gatos, etc.)

Procesamiento del
lenguaje natural: Análisis
de sentimiento, detección
de spam, etc.

CONCLUSIONES



- El análisis supervisado es una poderosa técnica de machine learning para construir modelos predictivos a partir de datos etiquetados.
- Requiere una cuidadosa recopilación y preprocesamiento de datos, selección y ajuste del modelo, y evaluación.
- Ampliamente aplicable en diversos dominios, pero es importante ser consciente de sus limitaciones y desafíos.



¡MUCHAS GRACIAS!

CONÓCENOS

<https://www.utdelacosta.edu.mx/>

Orgullosamente UTC

