INTELIGENCIA ARTIFICIAL Y MACHINE LEARNING EN BASES DE DATOS

INTEGRANTES:

- JHON CABALLERO 4-829-832
- JOSÉ CHAVEZ 4-827-1936
- JACKSON DEL CID 4-843-160
- TEDDY SOTO 4- 885-1695

BASE DE DATOS I

RUTA DEL TEMA:
DEL DATO AL IMPACTO

01



Contexto y fundamentos

– el qué

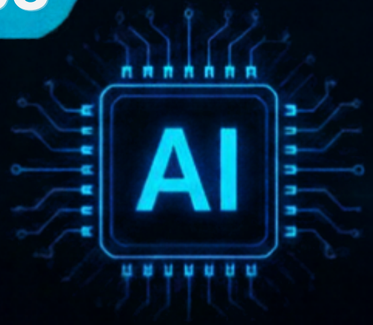
02



Preparación y punte de datos

– el cómo

03



IA dentro de la base de datos

– el núcleo

04



Aplicaciones, ética y conclusión

– el impacto



La calidad del dato condiciona la calidad del aprendizaje

Los datos:
combustible de la IA

> Sin datos confiables, no existe aprendizaje confiable

1. **Datos organizados**
2. **Modelo de Machine Learning**
3. **Decisiones inteligentes**



- calidad
- volumen
- integridad

NOTA TÉCNICA: La base de datos aporta el material de aprendizaje y el contexto operativo

ML y diseño de la base de datos

A SOFTWARE TRADICIONAL

- Reglas explícitas
- Entrada + lógica
- Salida predefinida

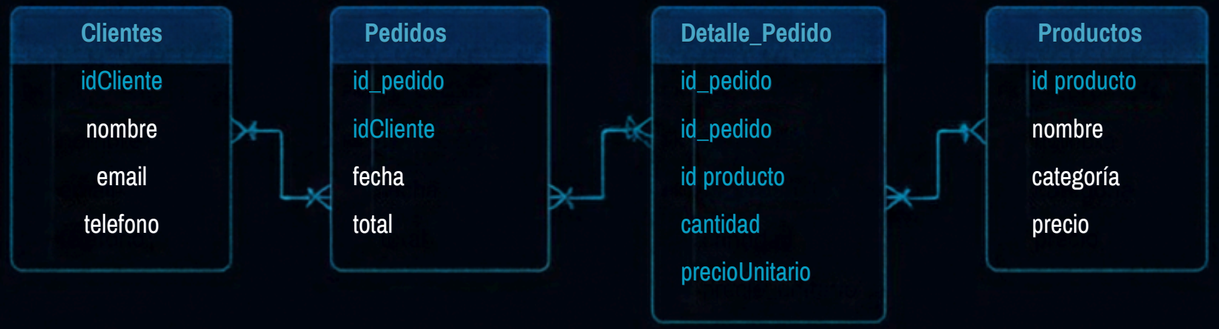
VS

B MACHINE LEARNING

- Datos + algoritmo
- Aprendizaje de patrones
- Predicción

LA BASE OBLIGATORIA

TABLAS | RELACIONES | CLAVES | NORMALIZACIÓN



Antes de entrenar un modelo, hay que diseñar datos que puedan entenderse y relacionarse

ETL: preparar datos para aprender

JOSÉ CHAVEZ 4-827-1936

```
-- CONEXIÓN A FUENTES
SELECT *
FROM CLIENTES;
SELECT *
FROM VENTAS;
SELECT *
FROM LOGSAPP;
-- API EXTERNA
GET /DATOS
STATUS: 200 OK
>_
```

01



Extraer

obtener datos desde el SGBD y otras fuentes

02



Transformar

limpiar, integrar y estandarizar

03



Cargar

entregar un dataset listo para el modelo

```
-- TRANSFORMACIONES
LIMPIAR (KLOS):
ELIMINAR DUPLICADOS(;
ESTANDARDAR, FORMATOS(;
INTEGRAR, TABLAS(;

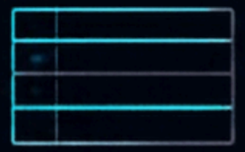
-- CARGA FINAL
EXPORT DATASET
TO /DATASET/LISTO_M1.CSV;
>_
```

CONTROL DE CALIDAD

- Valores nulos



- Duplicados



- Formatos inconsistentes

01/02/25	→	2025-02-01
$1,250.00	→	1250.00

- Datos atípicos



Un dataset limpio reduce ruido y mejora la confiabilidad del modelo

SQL: el puente hacia Machine Learning

JOSÉ CHAVEZ 4-827-1936

JOINs

combinar información relacionada

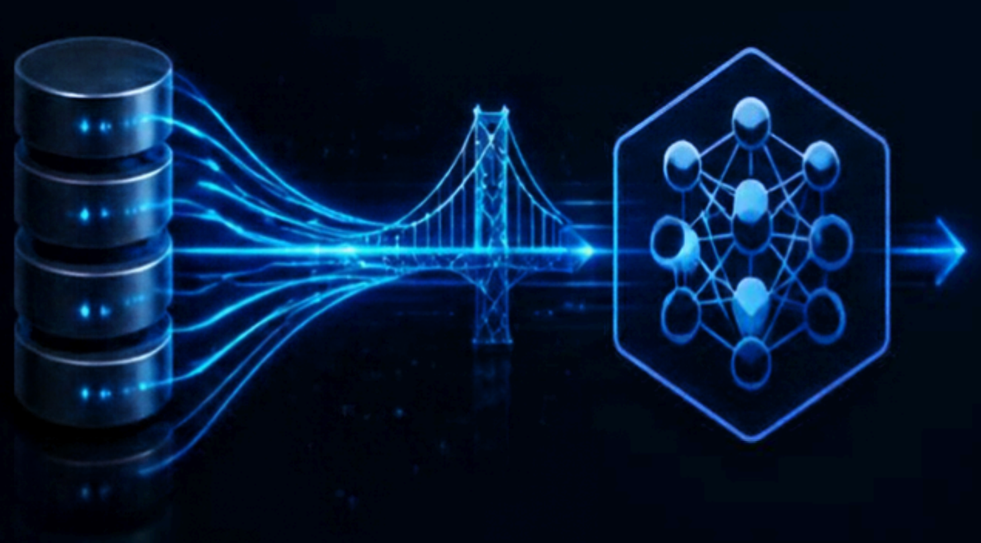
Vistas

encapsular datasets reutilizables

GROUP BY

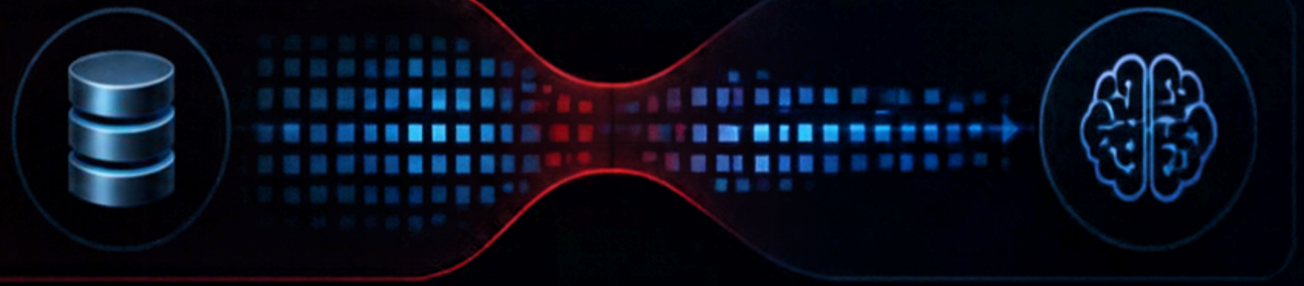
resumir y encontrar patrones

```
SELECT cliente_id, AVG(monto) AS promedio
FROM transacciones
GROUP BY cliente_id;
```



Cuello de botella

Mover grandes volúmenes desde el SGBD hacia el entorno de ML aumenta tiempo, costo y latencia

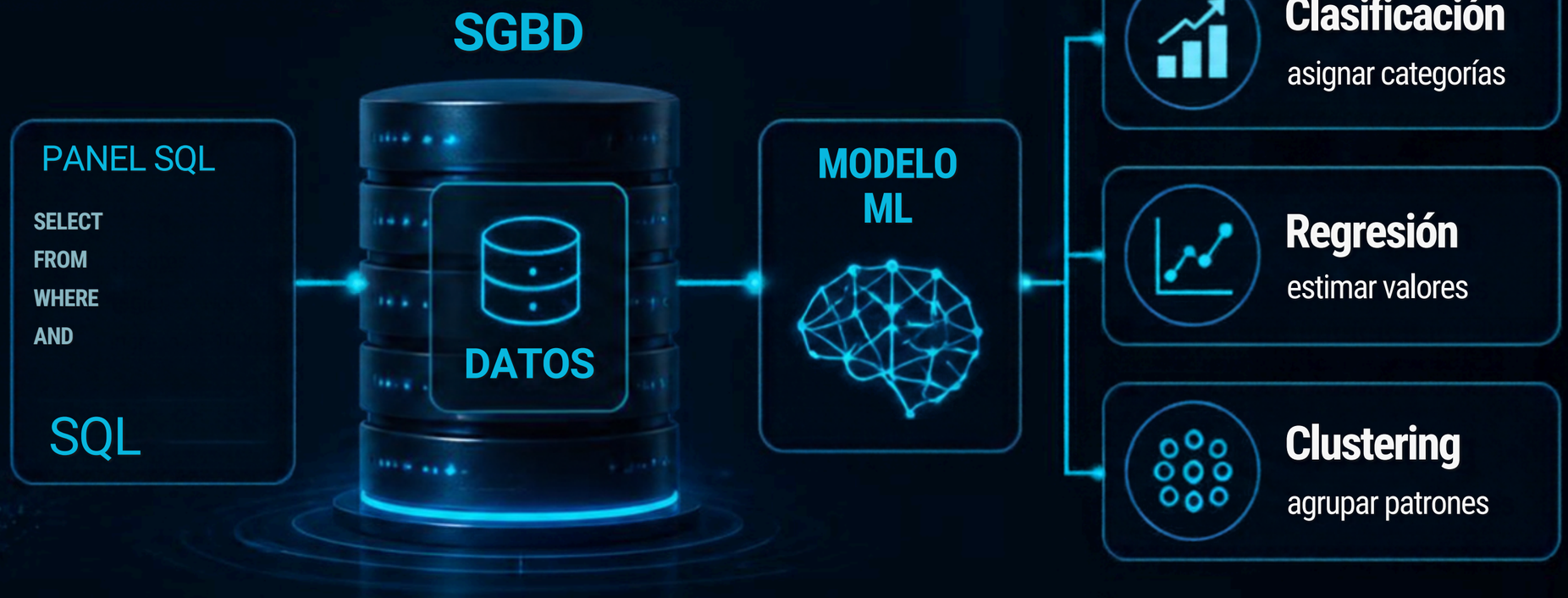


In-Database Machine Learning

TEDDY SOTO 4-885-1695

IDEA CENTRAL

El modelo se ejecuta cerca de los datos



Menos movimiento de datos, menor latencia y mayor control

PLATAFORMAS DE EJEMPLO	ORACLE	DB2	POSTGRESQL

Ventajas y Limitaciones

TEDDY SOTO 4-885-1695

VENTAJAS

- Los datos se mantienen bajo el control de la base de datos, sin extraerlos a motores de ML separados, que agregan latencia y generan inquietudes de seguridad y almacenamiento.
- Los algoritmos son rápidos y escalables, y pueden puntuar datos por lotes o en tiempo real.

LIMITACIONES

- Integrar ML en el SGBD promete ventajas para producción, pero el rendimiento y la capacidad de las bases de datos para distintos escenarios de ML todavía no están del todo claros. Por eso conviene decir que no reemplaza siempre a las herramientas de ML especializadas.

PLATAFORMAS DE EJEMPLO	ORACLE	DB2	POSTGRESQL

Bases de Datos Inteligentes

TEDDY SOTO 4-885-1695

IDEA CLAVE

Aquí la IA no analiza los datos del usuario, sino que administra la propia base de datos.

- Oracle Autonomous Database usa ML para reducir el trabajo humano y los errores.
- Se describen en tres ideas: self-driving, self-securing y self-repairing.
- Índices automáticos: La indexación automática detecta cambios en la carga de trabajo, analiza continuamente el comportamiento de la aplicación para encontrar oportunidades de indexación y crea los índices recomendados.
- Optimización de consultas: Oracle usa ML para monitorear de forma continua el uso y las cargas de trabajo.
- No es solo Oracle: HeatWave Autopilot de MySQL usa ML para predecir índices secundarios según la carga de trabajo, considerando tanto el rendimiento de las consultas como el costo de mantener el índice.
- Control humano: Las acciones de indexación automática se pueden ver, controlar y auditar, y se pueden activar o desactivar.



Bases de Datos Vectoriales

TEDDY SOTO 4-885-1695

- Qué son los embeddings: Convierten datos no estructurados (texto, imágenes, audio) en representaciones numéricas que permiten búsqueda semántica y recomendaciones, más allá de las palabras clave.
- Cómo se usan: Un fragmento de texto o una imagen pasa por un modelo de embeddings, el vector se guarda en una columna VECTOR junto a los datos originales, y un índice vectorial ayuda a encontrar los vecinos más cercanos.
- Dónde existen: PostgreSQL (con pgvector) y Oracle, entre otros

- Ventaja de tenerlo en el SGBD: Se puede combinar la búsqueda por similitud con las búsquedas relacionales tradicionales usando SQL, y se integra con RAG y modelos de lenguaje.
- Ejemplo simple: Buscar “computadora portátil” y que el sistema también encuentre “laptop”, aunque la palabra no coincida.

PLATAFORMAS DE EJEMPLO	ORACLE	DB2	POSTGRESQL

JHON CABALLERO 4-829-832

APLICACIONES



ORACLE AI Database 26ai

AI Y ML EN ACCIÓN
CON EL PODER DE LOS DATOS

DETECCIÓN DE FRAUDES
EN TIEMPO REAL

RECOMENDACIONES
PERSONALIZADAS

PREDICCIÓN DE INVENTARIOS
PARA OPTIMIZACIÓN LOGÍSTICA



BASES DE DATOS
DATOS CONFIABLES



MACHINE LEARNING
PATRONES E INTELIGENCIA

DETECCIÓN DE FRAUDES
EN TIEMPO REAL



RECOMENDACIONES
PERSONALIZADAS



PREDICCIÓN DE INVENTARIOS
PARA LOGÍSTICA



JHON CABALLERO 4-829-832

CALIDAD DE DATOS, ÉTICA Y TENDENCIAS



SESGOS:

DERIVADOS DE DATOS INCOMPLETOS O DE BAJA CALIDAD

RETOS ÉTICOS:

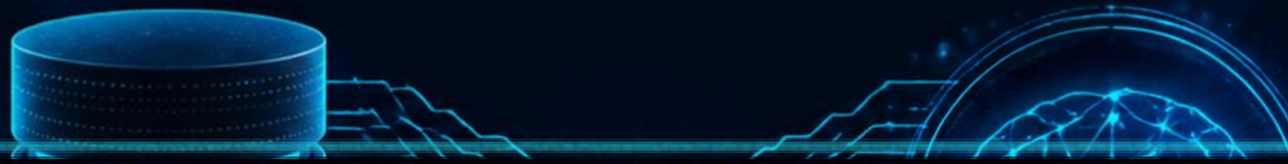
PRINCIPIOS:

- BENEFICIAR A LAS PERSONAS
- RESPETAR LOS DERECHOS HUMANOS
- GARANTIZAR TRANSPARENCIA
- OPERAR CON SEGURIDAD Y ASUMIR RESPONSABILIDAD.

TENDENCIAS FUTURAS:

- BASES DE DATOS NATIVAS DE IA
- AGENTES DE IA

CONCLUSIÓN



La unión de la IA y las bases de datos convierte a los sistemas tradicionales en entornos inteligentes y autónomos. Mientras los datos impulsan el aprendizaje automático, la IA optimiza la administración, el análisis predictivo y la toma de decisiones, marcando el futuro de la gestión de la información pese a los retos éticos y de calidad.

El diseño correcto de la base de datos
no es reemplazable por IA; es su requisito fundamental.

BIBLIOGRAFÍA

• GOOGLE CLOUD. (S. F.). ¿QUÉ ES EL APRENDIZAJE AUTOMÁTICO? TIPOS Y USOS. HTTPS://CLOUD.GOOGLE.COM/LEARN/WHAT-IS-MACHINE-LEARNING?HL=ES-419

• GOOGLE CLOUD. (S. F.-B). FRAUD DEFENSE. HTTPS://CLOUD.GOOGLE.COM/SECURITY/PRODUCTS/FRAUD-DEFENSE?HL=ES-419

• IBM. (2022, 23 DE SEPTIEMBRE). ¿QUÉ ES LA CALIDAD DE LOS DATOS? HTTPS://WWW.IBM.COM/ES-ES/THINK/TOPICS/DATA-QUALITY

• IBM. (2025, 15 DE MAYO). ¿QUÉ ES LA GESTIÓN DE LA CALIDAD DE LOS DATOS? HTTPS://WWW.IBM.COM/ES-ES/THINK/TOPICS/DATA-QUALITY-MANAGEMENT

• MICROSOFT. (S. F.). INTELIGENCIA ARTIFICIAL Y APRENDIZAJE AUTOMÁTICO: SERVICIOS DE AZURE AI. HTTPS://AZURE.MICROSOFT.COM/ES-ES/PRODUCTS/CATEGORY/AI

• MICROSOFT LEARN. (2026, 13 DE AGOSTO). PRODUCTOS DE INTELIGENCIA ARTIFICIAL Y APRENDIZAJE AUTOMÁTICO. HTTPS://LEARN.MICROSOFT.COM/ES-ES/AZURE/ARCHITECTURE/AI-ML/GUIDE/DATA-SCIENCE-AND-MACHINE-LEARNING

• MICROSOFT LEARN. (S. F.). DETECCIÓN DE FRAUDE EN LA BANCA MÓVIL. HTTPS://LEARN.MICROSOFT.COM/ES-ES/AZURE/ARCHITECTURE/GUIDE/AI/BANK-FRAUD-SOLUTION

• ORACLE. (S. F.). AUTONOMOUS AI DATABASE. HTTPS://WWW.ORACLE.COM/ES/AUTONOMOUS-DATABASE/

• ORACLE. (S. F.-B). FUNCIONES DE AUTONOMOUS AI DATABASE. HTTPS://WWW.ORACLE.COM/ES/AUTONOMOUS-DATABASE/FEATURES/

• ORACLE. (2024, 10 DE SEPTIEMBRE). MACHINE LEARNING SERVICES. HTTPS://WWW.ORACLE.COM/ES/ARTIFICIAL-INTELLIGENCE/MACHINE-LEARNING/

• ORACLE. (2025, 19 DE MAYO). MACHINE LEARNING EN ORACLE DATABASE. HTTPS://WWW.ORACLE.COM/ES/ARTIFICIAL-INTELLIGENCE/DATABASE-MACHINE-LEARNING/

• ORACLE. (2026, 8 DE AGOSTO). BASES DE DATOS DE IA. HTTPS://WWW.ORACLE.COM/ES/DATABASE/

•ORACLE AMÉRICA LATINA. (2026, 11 DE AGOSTO). DESARROLLO DE APLICACIONES DE BASE DE DATOS. HTTPS://WWW.ORACLE.COM/LATAM/DATABASE/TECHNOLOGIES/APPLICATION-DEVELOPMENT.HTML.

• ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS [OCDE]. (2019, 1 DE SEPTIEMBRE). LOS PRINCIPIOS DE INTELIGENCIA ARTIFICIAL DE LA OCDE. DATOS.GOB.ES. HTTPS://DATOS.GOB.ES/ES/BLOG/LOS-PRINCIPIOS-DE-INTELIGENCIA-ARTIFICIAL-DE-LA-OCDE

• UNESCO. (2026, 19 DE ENERO). ÉTICA DE LA INTELIGENCIA ARTIFICIAL. HTTPS://WWW.UNESCO.ORG/ES/ARTIFICIAL-INTELLIGENCE/RECOMMENDATION-ETHICS• UNESCO. (2026, 6 DE OCTUBRE). INTELIGENCIA ARTIFICIAL Y TECNOLOGÍAS EMERGENTES. HTTPS://WWW.UNESCO.ORG/ES/ARTIFICIAL-INTELLIGENCE