

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA
DESARROLLO Y GESTIÓN DE SOFTWARE



**DESARROLLO PARA DISPOSITIVOS
INTELIGENTES**

I.3. Reporte de investigación de los tipos de aplicaciones, procesamiento y herramientas para inteligencia artificial, machine learning, data mining y big data.

DOCENTE:

ING. LUIS ENRIQUE MASCOTE CANO

PRESENTA:

DARON TARÍN GONZÁLEZ

GRUPO:

IDGS91N

Chihuahua, Chih., 25 de septiembre de 2025

Contenido

Introducción	3
1. Inteligencia Artificial (IA).....	3
Definición.....	3
Tipos de aplicaciones	4
Procesamiento	5
Herramientas y tecnologías	6
2. Machine Learning (ML)	7
Definición.....	7
Tipos de aplicaciones	7
Procesamiento	9
Herramientas y tecnologías	9
3. Data Mining (DM)	9
Definición.....	9
Tipos de aplicaciones	9
Procesamiento	11
Herramientas y tecnologías	11
4. Big Data	11
Definición.....	11
Tipos de aplicaciones	11
Procesamiento	12
Herramientas y tecnologías	13
Conclusión	14
Referencias	15

Introducción

El avance tecnológico de las últimas décadas ha estado marcado por el auge de la **Inteligencia Artificial (IA)**, el **Machine Learning (ML)**, el **Data Mining (DM)** y el **Big Data**, cuatro dominios que han cambiado profundamente el panorama de la investigación, la industria y la vida cotidiana. Cada uno cumple un rol específico dentro del ecosistema de los datos:

- **La IA** busca desarrollar sistemas que imiten o incluso superen ciertas capacidades humanas, como la percepción visual o la comprensión del lenguaje.
- **El ML** se centra en construir algoritmos que aprenden patrones a partir de los datos, permitiendo realizar predicciones sin reglas explícitas.
- **El DM** tiene como objetivo descubrir relaciones ocultas en los datos que puedan ser útiles para la toma de decisiones estratégicas.
- **El Big Data** aporta la infraestructura y las herramientas necesarias para manejar volúmenes masivos de información de forma eficiente.

Estos dominios no actúan de manera aislada. Por el contrario, se complementan: el Big Data suministra los datos y la capacidad de procesarlos, el Data Mining los analiza para encontrar patrones, el ML los convierte en modelos predictivos y la IA los integra en aplicaciones inteligentes.

En las siguientes secciones se describen en detalle las aplicaciones más frecuentes de cada dominio, las modalidades de procesamiento **batch** y **streaming**, y las herramientas más importantes utilizadas actualmente.

1. Inteligencia Artificial (IA)

Definición

La Inteligencia Artificial es la disciplina de la informática que busca crear sistemas capaces de ejecutar tareas que requieren inteligencia humana: desde percibir el entorno y tomar decisiones, hasta planificar acciones y comunicarse en lenguaje natural. Existen diversos tipos de IA, como la IA estrecha (diseñada para tareas específicas, como un asistente de voz), la IA general (con capacidad de adaptarse a distintos contextos, aún en desarrollo) y la IA generativa, que puede producir contenido nuevo como texto o imágenes (IBM, 2025).

Tipos de aplicaciones

1. Visión por computadora

- Permite que las máquinas interpreten y analicen imágenes o videos.
- *Ejemplos:* detección facial en aeropuertos, reconocimiento de matrículas de vehículos en sistemas de tráfico, clasificación de imágenes médicas para detectar tumores.



2. Procesamiento de lenguaje natural (NLP)

- Se centra en la comprensión y generación de lenguaje humano.
- *Ejemplos:* traductores automáticos como Google Translate, chatbots de atención al cliente y modelos como GPT-4 que responden preguntas o redactan documentos.



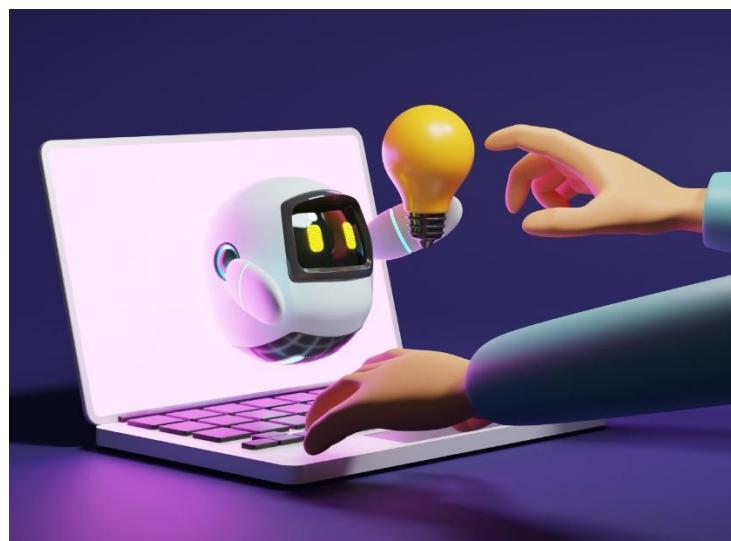
3. Sistemas de recomendación

- Personalizan el contenido según las preferencias del usuario.
- *Ejemplos:* Netflix sugiere series, Amazon recomienda productos y Spotify genera listas de reproducción personalizadas.



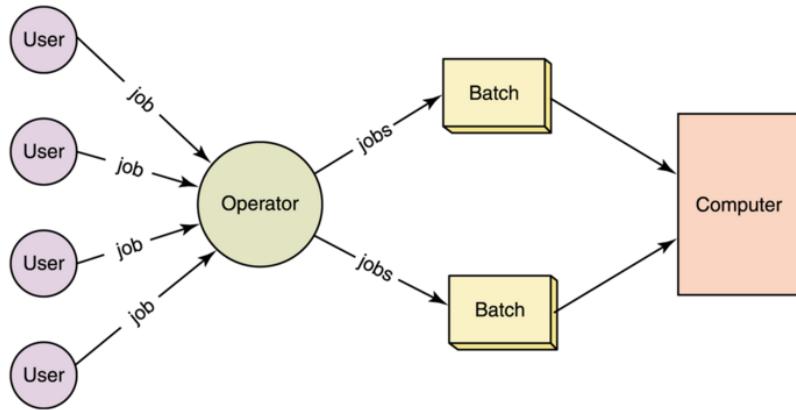
4. IA generativa

- Capaz de crear contenido nuevo a partir de patrones aprendidos.
- *Ejemplos:* DALL·E o Stable Diffusion generan imágenes desde texto, y ChatGPT crea redacciones académicas o guiones.



Procesamiento

- **Batch:** usado principalmente para **entrenar modelos de IA**. Por ejemplo, entrenar un sistema de visión requiere millones de imágenes que se procesan en lotes. La latencia no es crítica, pero sí lo es la precisión.



- **Streaming:** empleado en la **inferencia en tiempo real**. Ejemplos claros son la traducción simultánea en videollamadas, la detección de intrusos en cámaras de seguridad y los sistemas de asistencia en automóviles autónomos.



ORACLE
Cloud

Herramientas y tecnologías

- **TensorFlow:** creado por Google, ampliamente usado en la industria y la investigación, con soporte para despliegue móvil y web.
- **PyTorch:** popular en la comunidad científica por su flexibilidad y sencillez.

- **BigDL**: framework que integra deep learning en clústeres de Big Data, facilitando el entrenamiento a gran escala.
- **TensorFlow Serving y TorchServe**: herramientas de despliegue de modelos en producción.
- **Aceleradores de hardware (GPU, TPU, FPGA)**: permiten entrenar y ejecutar modelos de IA con mayor velocidad y eficiencia energética.

2. Machine Learning (ML)

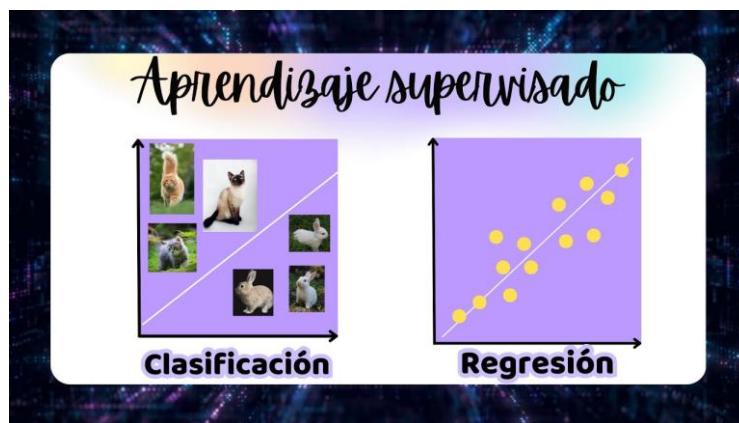
Definición

El Machine Learning es un subcampo de la IA que desarrolla algoritmos que **aprenden de los datos** para hacer predicciones o decisiones sin estar programados con reglas fijas. Es la base de aplicaciones modernas que van desde la predicción de demanda hasta el reconocimiento facial.

Tipos de aplicaciones

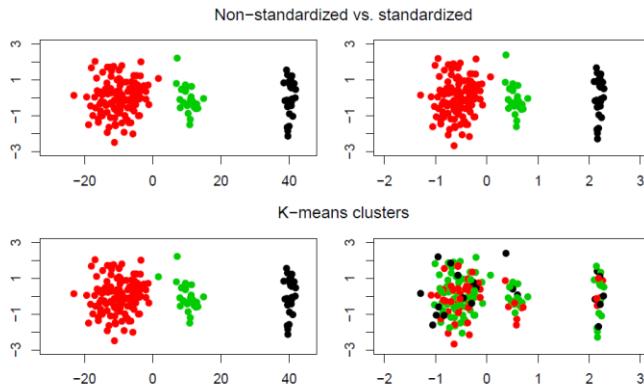
1. Clasificación y regresión supervisada

- Usada en predicciones de churn, diagnóstico médico, reconocimiento de imágenes o predicción de precios de viviendas.



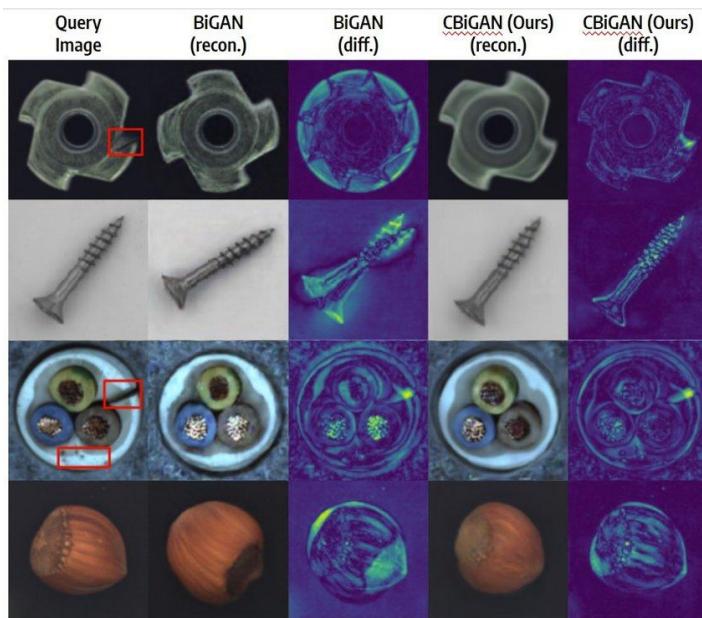
2. Clustering y segmentación (aprendizaje no supervisado)

- Agrupa datos sin etiquetas, revelando patrones ocultos.
- *Ejemplo*: segmentación de clientes en marketing para diseñar campañas personalizadas.



3. Detección de anomalías

- Identifica patrones inusuales.
- *Ejemplo:* sistemas bancarios que bloquean transacciones sospechosas en tiempo real.



4. Redes neuronales profundas

- Permiten avances en visión, audio y lenguaje.
- *Ejemplo:* asistentes virtuales como Siri o Alexa.

Procesamiento

- **Batch:** entrenamiento de modelos con grandes datasets, validación y ajuste de hiperparámetros.
- **Streaming:** predicciones en tiempo real, como detección de fraudes en tarjetas de crédito o recomendaciones dinámicas en e-commerce.
- **Estrategia híbrida:** sistemas que reentrenan el modelo en batch (ej. cada noche) pero hacen inferencias en streaming (ej. cada transacción).

Herramientas y tecnologías

- **scikit-learn:** librería de referencia en Python para modelos clásicos.
- **Apache Spark MLlib:** algoritmos distribuidos para clústeres de datos masivos.
- **TensorFlow/Keras:** entrenamiento de redes neuronales con alto rendimiento.
- **PyTorch:** muy usado en investigación por su facilidad de uso.
- **Kubeflow y MLflow:** plataformas que gestionan el ciclo de vida completo de los modelos, desde el prototipado hasta el despliegue.

3. Data Mining (DM)

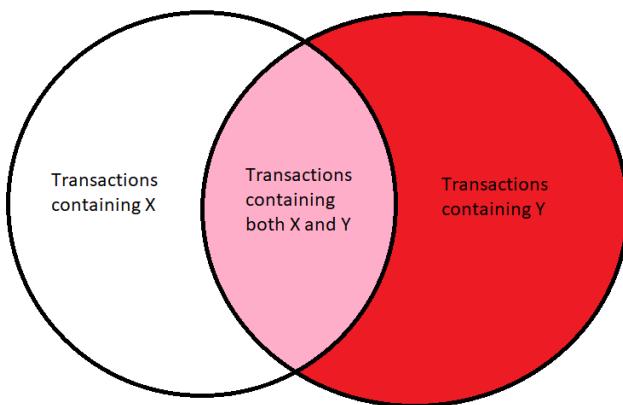
Definición

El Data Mining, o minería de datos, es la disciplina que **descubre patrones, correlaciones y conocimiento oculto** en grandes volúmenes de información, con el objetivo de apoyar la toma de decisiones. Es ampliamente usado en investigación científica, marketing y seguridad informática.

Tipos de aplicaciones

1. Reglas de asociación

- Descubren relaciones entre productos o eventos.
- *Ejemplo:* “los clientes que compran pan suelen comprar mantequilla”.



2. Segmentación o clustering

- Agrupa datos para identificar perfiles de usuario.
- *Ejemplo:* análisis de hábitos de consumo para diseñar campañas específicas.

3. Detección de anomalías

- Descubre comportamientos inesperados o riesgosos.
- *Ejemplo:* detección de fraudes en tarjetas de crédito.

4. Análisis de series temporales

- Identifica tendencias y patrones repetitivos.
- *Ejemplo:* pronóstico de ventas mensuales en un supermercado.



Procesamiento

- **Batch:** predominante en minería exploratoria, se aplica sobre datasets históricos.
- **Streaming:** importante para detectar anomalías en tiempo real, como monitoreo de redes o transacciones bancarias.

Herramientas y tecnologías

- **RapidMiner:** entorno visual para minería de datos con flujos reproducibles.
- **KNIME:** plataforma de análisis con módulos reutilizables.
- **Weka:** muy usada en universidades y laboratorios para enseñanza e investigación.
- **Orange:** software con interfaz visual para análisis exploratorio.
- **SQL avanzado y extensiones analíticas:** muchas tareas de minería pueden realizarse en bases de datos modernas.

4. Big Data

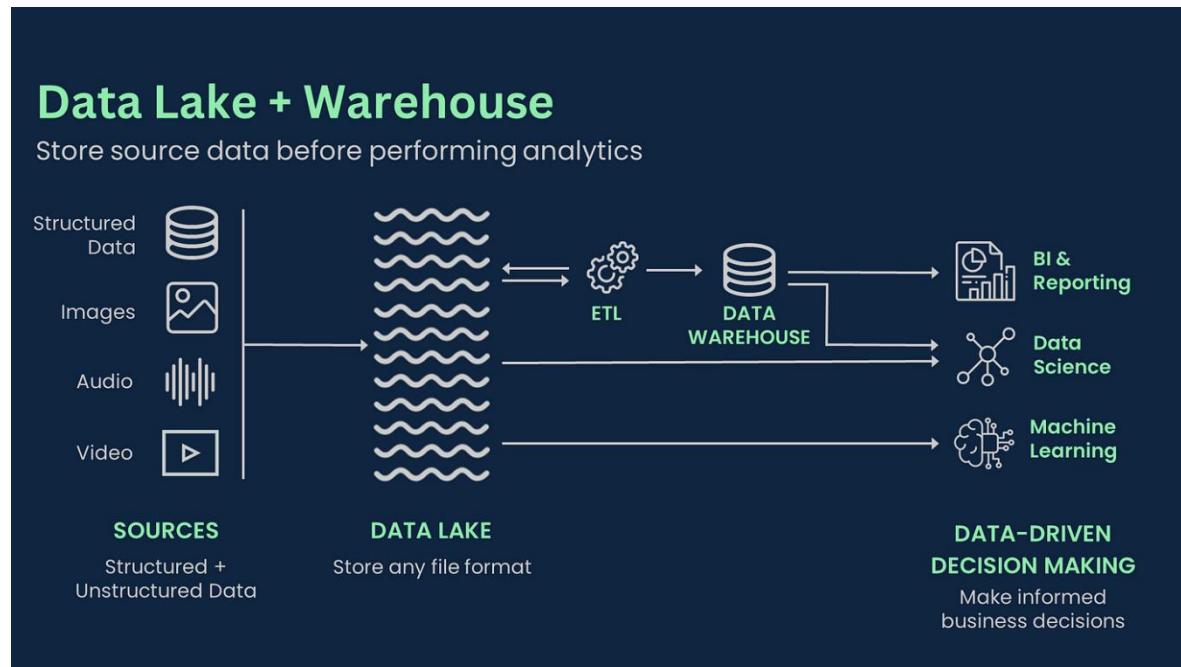
Definición

El Big Data se refiere al conjunto de **métodos y tecnologías diseñados para manejar datos con gran volumen, velocidad y variedad**, que superan las capacidades de los sistemas tradicionales. Es la base que soporta a la IA, el ML y el DM a gran escala.

Tipos de aplicaciones

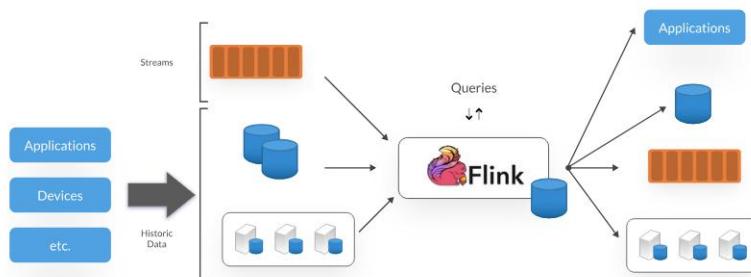
1. **Procesamiento batch (ETL masivo)**
 - Limpieza, transformación y carga de datos en data warehouses.
 - *Ejemplo:* análisis diario de logs de servidores en Google BigQuery.
2. **Procesamiento en streaming**
 - Analiza datos en tiempo real.
 - *Ejemplo:* monitoreo de sensores en una fábrica para prevenir fallas.
3. **Data lakehouse**
 - Combina la flexibilidad de los data lakes y la estructura de los data warehouses.

- *Ejemplo:* Databricks Lakehouse integra BI, ML y analítica avanzada.



Procesamiento

- **Batch:** útil cuando se requiere análisis global y no importa la latencia (cierres contables, entrenamientos).
- **Streaming:** indispensable en casos críticos como monitoreo de sistemas, detección de fraudes o gestión de tráfico.
- **Comparación:** Spark procesa en micro-batches mientras que Flink maneja flujos continuos con menor latencia (DataCamp, 2025).



Herramientas y tecnologías

- **Apache Hadoop (HDFS, MapReduce, YARN)**: pionero en procesamiento distribuido.
- **Apache Spark**: motor unificado de datos para batch y streaming.
- **Apache Flink**: especializado en flujos en tiempo real.
- **Google BigQuery**: servicio serverless de análisis masivo en la nube.
- **Snowflake**: data warehouse en la nube con escalabilidad automática.
- **Databricks Lakehouse**: unificación de almacenamiento, ML y BI.
- **Calcite**: framework de optimización de consultas.
- **Thrill**: framework batch en C++ con alto rendimiento.
- **BigDL**: deep learning sobre clústeres de Spark.

Conclusión

Los dominios de la Inteligencia Artificial (IA), el Machine Learning (ML), el Data Mining (DM) y el Big Data se encuentran profundamente interconectados, aunque cada uno cumple funciones específicas dentro del ecosistema digital. La IA tiene como propósito principal simular las capacidades cognitivas humanas, como la percepción, el razonamiento y la comunicación en lenguaje natural. El ML, por su parte, constituye el conjunto de algoritmos predictivos que hacen posible que los sistemas aprendan a partir de los datos y realicen inferencias sin necesidad de programación explícita. El DM se enfoca en descubrir patrones no evidentes y reglas útiles dentro de grandes volúmenes de información, lo cual resulta fundamental para la toma de decisiones estratégicas. Finalmente, el Big Data aporta la infraestructura técnica necesaria para procesar y gestionar datos masivos de manera eficiente, asegurando que las demás disciplinas puedan operar a escala.

A pesar de sus diferencias, existe un importante solapamiento entre estos dominios. El ML, por ejemplo, funciona como el núcleo metodológico de la IA y del DM, ya que proporciona las técnicas que permiten tanto la construcción de modelos predictivos como el hallazgo de patrones en los datos. Asimismo, herramientas como Apache Spark y Apache Flink se utilizan transversalmente en diferentes ámbitos, actuando como motores de procesamiento que soportan desde aplicaciones de inteligencia artificial hasta análisis de minería de datos y soluciones de Big Data.

En cuanto a sus diferencias, puede afirmarse que la IA se orienta a tareas cognitivas complejas; el ML se centra en el desarrollo de modelos de predicción; el DM tiene como objetivo principal el descubrimiento exploratorio de patrones; y el Big Data se especializa en la administración de arquitecturas e infraestructuras capaces de manejar grandes volúmenes, velocidades y variedades de datos.

El modo de procesamiento también marca una distinción esencial. El enfoque por lotes o *batch* resulta idóneo para análisis históricos, entrenamientos masivos de modelos y procesos que toleran cierta latencia. Por otro lado, el *streaming* es vital en escenarios donde la inmediatez de la información es crucial, como en la detección de fraudes, el monitoreo de sensores en tiempo real o la personalización dinámica de servicios digitales.

En conclusión, la elección adecuada tanto de las herramientas como de los modos de procesamiento constituye un factor determinante para el éxito de los proyectos modernos basados en datos. Un sistema mal diseñado puede volverse lento, costoso e ineficaz, mientras que una arquitectura correctamente planificada y optimizada es capaz de transformar la información en un activo estratégico que impulse la innovación y la competitividad en las organizaciones.

Referencias

- IBM. (2025). *Tipos de Inteligencia Artificial*. <https://www.ibm.com/mx-es/think/topics/artificial-intelligence-types>
- Wikipedia. (2025). Aplicaciones de la inteligencia artificial. https://es.wikipedia.org/wiki/Aplicaciones_de_la_inteligencia_artificial
- TensorFlow. (2025). API Documentation. https://www.tensorflow.org/api_docs
- PyTorch. (2025). Documentation. <https://pytorch.org/docs/>
- Google Cloud Blog. (2019). Batch vs. stream processing. <https://cloud.google.com/blog/products/data-analytics/how-cloud-batch-and-stream-data-processing-works>
- KNIME. (2025). Documentation. <https://docs.knime.com/>
- DataCamp. (2025). Flink vs Spark: A Comprehensive Comparison. <https://www.datacamp.com/blog/flink-vs-spark>
- Apache Hadoop. (2025). Overview. <https://hadoop.apache.org/>