

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN



EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS

**REPORTE DE INVESTIGACIÓN DE LOS CONCEPTOS,
BENEFICIOS, RESTRICCIONES Y RETOS DE INTELIGENCIA
ARTIFICIAL, MACHINE LEARNING, DATA MINING Y BIG
DATA**

IDGS81N

PRESENTA:

REGINA CHÁVEZ TAMAYO

DOCENTE:

LUIS ENRIQUE MASCOTE CANO

Chihuahua, Chih., 26 de septiembre de 2025

Índice

| | |
|------------------------------|---|
| Introducción | 3 |
| Definiciones..... | 3 |
| Inteligencia Artificial..... | 3 |
| Machine Learning..... | 3 |
| Data Mining | 3 |
| Big Data | 4 |
| Beneficios comparados | 4 |
| Restricciones y retos..... | 5 |
| Inteligencia Artificial..... | 5 |
| Machine Learning..... | 5 |
| Data Mining | 5 |
| Big Data | 5 |
| Conclusión | 6 |
| Referencias..... | 7 |

Introducción

Este reporte describe y compara cuatro dominios clave del ecosistema de datos: Inteligencia Artificial, Machine Learning, Data Mining y Big Data. Se presentan definiciones reconocidas, un resumen histórico, beneficios, restricciones y retos, para ofrecer una visión clara y accionable que permita distinguir alcances y límites entre ellos.

Definiciones

Inteligencia Artificial

McCarthy describe la IA como la ciencia e ingeniería de construir máquinas inteligentes, en particular programas capaces de realizar tareas que requieren inteligencia.

Origen y evolución:

El campo se consolida en 1956 (conferencia de Dartmouth). Tras ciclos de entusiasmo e "inviernos", la disponibilidad de datos y cómputo acelerado impulsó el auge reciente (especialmente aprendizaje profundo) y su adopción transversal.

Machine Learning

Mitchell lo define como la capacidad de un programa para mejorar su desempeño en una tarea con experiencia medida por una métrica.

Origen y evolución:

Desde perceptrones y métodos estadísticos hasta redes profundas actuales, ML se volvió el motor práctico de la IA moderna gracias a big data y hardware especializado.

Data Mining

En el proceso KDD, DM es la etapa que extrae patrones válidos, novedosos, útiles y comprensibles a partir de datos.

Origen y evolución:

En los 90, KDD unificó estadística, bases de datos y ML en un flujo: selección -> preprocesamiento -> minería -> evaluación, con foco en utilidad e interpretabilidad de hallazgos.

Big Data

NIST caracteriza Big Data como datos a escala (volumen, velocidad y variedad) que requieren arquitecturas y patrones de interoperabilidad para su captura, almacenamiento y análisis.

Origen y evolución:

A inicios de los 2000 se formaliza el fenómeno y, más tarde, frameworks distribuidos (Hadoop/Spark) y la nube habilitan analítica y ML a gran escala.

Beneficios comparados

| Dominio | Beneficios |
|--------------------------------|---|
| Inteligencia Artificial | <ol style="list-style-type: none"> 1. Automatiza tareas cognitivas complejas (p. ej., visión, lenguaje). 2. Soporta decisiones en contextos inciertos mediante agentes inteligentes. 3. Habilita nuevos productos/servicios basados en capacidades "inteligentes". |
| Machine Learning | <ol style="list-style-type: none"> 1. Predice tendencias con mejora continua basada en datos. 2. Reduce errores manuales al aprender patrones de manera estadística. 3. Escala a múltiples dominios gracias a datos y cómputo accesibles. |
| Data Mining | <ol style="list-style-type: none"> 1. Descubre patrones útiles y comprensibles para negocio. 2. Estructura el proceso KDD para calidad y trazabilidad. 3. Aporta criterios de "interesabilidad" para seleccionar hallazgos relevantes. |
| Big Data | <ol style="list-style-type: none"> 1. Permite analizar volúmenes y velocidades antes inviables. 2. Integra variedad de fuentes (logs, multimedia, sensores). 3. Fomenta arquitecturas interoperables y escalables (marcos NIST). |

Restricciones y retos

Inteligencia Artificial

Restricciones:

- Dependencia de grandes volúmenes de datos y cómputo (coste/energía).
- Dificultad de explicabilidad en modelos complejos (caja negra).

Retos:

- Gobernanza y cumplimiento (marcos regulatorios y auditoría técnica).
- Mitigación de sesgos y garantía de equidad y seguridad.

Machine Learning

Restricciones:

- Calidad/disponibilidad de datos (ruido, desbalance, faltantes).
- Fragilidad fuera de distribución y riesgo de overfitting/underfitting.

Retos:

- Interpretabilidad útil para negocio sin sacrificar desempeño.
- MLOps robusto (despliegue, monitoreo, costos e infraestructura).

Data Mining

Restricciones:

- Preprocesamiento intensivo y selección de métricas de "interesabilidad".
- Riesgos de privacidad al combinar fuentes y reidentificar.

Retos:

- Ética y finalidad legítima del descubrimiento de patrones.
- Valorar hallazgos (útiles, comprensibles) en contextos cambiantes.

Big Data

Restricciones:

- Costos de almacenamiento/transmisión y latencia en tiempo real.
- Complejidad de interoperabilidad entre plataformas y formatos.

Retos:

- Gobernanza/estandarización (catálogo, linaje, seguridad).
- Cumplimiento normativo y gestión de metadatos para auditoría.

Conclusión

La Inteligencia Artificial, el Machine Learning, el Data Mining y el Big Data no compiten entre sí, si no que se necesitan. Big Data aporta los grandes volúmenes de información, Data Mining ayuda a encontrar patrones útiles, Machine Learning transforma esos datos en modelos que aprenden y mejoran, y la Inteligencia Artificial usa todo eso para construir aplicaciones que “entienden” y actúan.

Los beneficios son claros, automatizar tareas, predecir mejor y personalizar experiencias. Pero no basta con elegir un algoritmo. Para que funcione de verdad, se requiere buena calidad de datos, transparencia para explicar cómo se toman decisiones, cuidado para evitar sesgos, seguridad de la información y cumplir con las normas.

En los próximos años, las herramientas estarán más integradas (desde los datos hasta el uso en producción) y habrá reglas más claras para usarlas con responsabilidad. Si preparo desde el inicio mis procesos, organizo bien los datos y dejo evidencia de cómo trabajo, reduzco riesgos y logro resultados útiles más rápido.

Referencias

IBM. (s. f.). ¿Qué es machine learning? Recuperado el 26 de septiembre de 2025, de <https://www.ibm.com/mx-es/think/topics/machine-learning>

Laney, D. (2012, agosto 25). Deja VVVu: Gartner's original "Volume–Velocity–Variety" definition of big data. AIIM. <https://community.aiim.org/blogs/doug-laney/2012/08/25/deja-vvvu-gartners-original-volume-velocity-variety-definition-of-big-data>

National Institute of Standards and Technology. (2019, diciembre 13). Big Data at NIST. NIST. <https://www.nist.gov/itl/big-data-nist>

OpenAI. (s. f.). ChatGPT [Modelo de lenguaje grande]. Recuperado el 26 de septiembre de 2025, de <https://chat.openai.com/>

Oracle. (2022, mayo 2). ¿Qué es la minería de datos? Oracle. <https://www.oracle.com/latam/big-data/what-is-data-mining/>

UNESCO. (s. f.). Inteligencia artificial. UNESCO. Recuperado el 26 de septiembre de 2025, de <https://www.unesco.org/es/artificial-intelligence>