

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN



Extracción de Conocimiento en Bases de Datos

I.1. Reporte de investigación de los conceptos, beneficios, restricciones y retos de inteligencia artificial, machine learning, data mining y big data

IDGS91N

PRESENTA:

JORGE ALEJANDRO HERNANDEZ CONTRERAS

DOCENTE:

Enrique Mascote

Chihuahua, Chih., 25 de Septiembre de 2025

Introducción.....	3
Definición de cada dominio y origen.....	3
Inteligencia Artificial.....	3
Machine Learning.....	3
Data Mining.....	3
Big Data.....	4
Beneficios comparados.....	4
Restricciones.....	5
IA.....	5
Machine Learning.....	5
Data Mining.....	5
Big Data.....	5
Retos actuales.....	5
IA.....	5
Machine Learning.....	5
Data Mining.....	6
Big Data.....	6
Conclusiones y perspectivas futuras.....	6
Bibliografía.....	7

Introducción

La inteligencia artificial, el machine learning, el data mining y el big data son disciplinas interrelacionadas que hoy impulsan decisiones automatizadas, análisis masivos y productos inteligentes. Aunque comparten técnicas, cada dominio tiene objetivos, limitaciones y retos propios: IA se enfoca en crear sistemas capaces de tareas “inteligentes”, ML en que los sistemas aprendan de datos, DM en descubrir patrones útiles en conjuntos de datos; y Big Data en gestionar y procesar volúmenes, velocidades y variedades de datos a escala.

Definición de cada dominio y origen

Inteligencia Artificial

- **Definición:**
Tecnología que permite a máquinas realizar tareas que normalmente requieren inteligencia humana como percepción, razonamiento, toma de decisiones y lenguaje, construida sobre datos, algoritmos y capacidad de cómputo.
- **Origen y evolución:**
El término se consolidó en la conferencia de Dartmouth en 1956 y la disciplina pasó de enfoques simbólicos y heurísticos en los años 50 - 70 a métodos estadísticos y redes neuronales, con un salto reciente gracias al deep learning y mayor capacidad de cómputo.

Machine Learning

- **Definición:**
Subcampo de la IA que estudia algoritmos capaces de aprender patrones y generalizar desde datos de entrenamiento para hacer predicciones o tomar decisiones sin programación explícita.
- **Origen y evolución:**
Sus raíces están en estadística y cibernética; métodos como redes neuronales, árboles de decisión y SVM evolucionaron desde mediados del siglo XX. En las últimas dos décadas, el deep learning y la disponibilidad de datos masivos han ampliado su éxito y aplicación.

Data Mining

- **Definición:**
Proceso de extraer patrones, relaciones o conocimiento útil a partir de grandes conjuntos de datos mediante técnicas estadísticas, de ML y de bases de datos.

- **Origen y evolución:**

El análisis de datos existe desde la estadística clásica, “data mining” creció con la capacidad computacional y bases de datos entre los 1990s y los 2000s, integrando técnicas de ML para descubrir patrones no evidentes. Hoy se usa en marketing, detección de fraude, bioinformática y más.

Big Data

- **Definición:**

Conjunto de datos caracterizados por su alto volumen, alta velocidad y/o alta variedad que requieren formas innovadoras y coste-efectivas de procesamiento para obtener información y automatizar decisiones.

- **Origen y evolución:**

El concepto surgió con el crecimiento exponencial de datos digitales. Desde principios de 2010s se consolidó con tecnologías como Hadoop, sistemas distribuidos y arquitecturas de stream processing para procesar datos a escala.

Beneficios comparados

Dominio	Beneficio 1	Beneficio 2	Beneficio 3
Inteligencia Artificial	Automatización de decisiones complejas	Mejora en eficiencia operativa	Nuevas capacidades como visión, lenguaje, creatividad
Machine Learning	Predicción de tendencias y comportamientos	Automatización de tareas repetitivas	Mejora continua con datos nuevos
Data Mining	Descubrimiento de patrones ocultos	Soporte a la toma de decisiones basada en evidencia	Segmentación y personalización
Big Data	Análisis a escala	Integración de fuentes heterogéneas en tiempo/near-real	Habilita ML/IA con volumen suficiente para modelos robustos

Restricciones

IA

- Requerimiento de grandes recursos computacionales y energía para modelos avanzados.
- Dependencia de datos de calidad, rendimiento débil si los datos son pobres o no representativos.

Machine Learning

- Necesidad de datos etiquetados y tiempo/coste para etiquetado.
- Riesgo de overfitting y mala generalización si el diseño y validación no son rigurosos.

Data Mining

- Calidad y preprocesamiento de datos: sin limpieza, los resultados son poco fiables.
- Interpretabilidad de patrones: no todo patrón es útil o estadísticamente significativo.

Big Data

- Costes de almacenamiento, transferencia y procesamiento a gran escala.
- Complejidad en integración y gobernanza de datos heterogéneos.

Retos actuales

IA

- Explicabilidad y transparencia, justificar decisiones en sectores regulados.
- Mitigar sesgos y garantizar equidad en modelos entrenados con datos históricos.

Machine Learning

- Robustez frente a ataques y seguridad del modelo en producción.
- MLOps y reproducibilidad, gestionar ciclo de vida del modelo, despliegue, monitorización y actualización.

Data Mining

- Privacidad y ética: evitar reidentificación y usos indebidos de patrones personales.
- Separar señal de ruido a medida que los conjuntos crecen y aumentar la validez estadística.

Big Data

- Procesamiento en tiempo real sobre flujos masivos de datos.
- Gobernanza, cumplimiento y protección de datos al combinar fuentes.

Conclusiones y perspectivas futuras

Complementariedad: IA, ML, DM y Big Data funcionan como un ecosistema: Big Data aporta la materia prima, DM descubre patrones, ML formaliza predictores y la IA aplica decisiones autónomas. La integración efectiva potencia el valor de negocio.

Priorizar gobernanza y ética: Ante el avance técnico, las organizaciones deben priorizar calidad de datos, explicabilidad y gobernanza para asegurar confianza y cumplimiento regulatorio. Los problemas éticos y de sesgo pueden minar adopciones si se ignoran.

Inversión en MLOps y capacidades operativas: El éxito no está solo en el modelo sino en su mantenimiento, monitorización y ciclo de vida, un área en rápido crecimiento y demanda.

Perspectiva futura: Esperamos mayor hibridación: arquitecturas más eficientes, mejores herramientas de XAI, y plataformas que integren en tiempo real Big Data con ML/IA para decisiones automatizadas seguras y auditables.

Bibliografía

OECD. (2025). Artificial intelligence. OECD AI policy resources. Recuperado de <https://www.oecd.org/en/topics/policy-issues/artificial-intelligence.html>

UNESCO. (s. f.). Artificial Intelligence. UNESCO. Recuperado de <https://www.unesco.org/en/artificial-intelligence>

IBM. (s. f.). What is Machine Learning (ML)? IBM. Recuperado de <https://www.ibm.com/think/topics/machine-learning>

Gartner. (s. f.). Definition of Big Data. Gartner IT Glossary. Recuperado de <https://www.gartner.com/en/information-technology/glossary/big-data>

TechTarget. (2024). What is data mining? SearchBusinessAnalytics / TechTarget. Recuperado de <https://www.techtarget.com/searchbusinessanalytics/definition/data-mining>