

# **UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA**

## **DESARROLLO Y GESTIÓN DE SOFTWARE**



### **EVIDENCIA 1: REPORTE DE INVESTIGACIÓN DE IA, ML, DM Y BIG DATA**

### **EXTRACCIÓN DE CONOCIMIENTO EN BASES DE DATOS**

**PRESENTA:**

**KARLA ALEJANDRA DE LA CRUZ ZEA**

**DOCENTE:**

**ING. LUIS ENRIQUE MASCOTE CANO**

22 de septiembre de 2025

## Contenido

Introducción .....	2
Definiciones y evolución histórica .....	2
Comparación de beneficios.....	3
Restricciones y retos.....	3
Conclusión .....	4
Referencias.....	4

## Introducción

Las organizaciones generan cada vez más datos y buscan convertirlos en información útil para la toma de decisiones. En este contexto, la Inteligencia Artificial, Machine Learning, Data Mining y Big Data son áreas clave que, aunque relacionadas, tienen objetivos y técnicas distintas.

## Definiciones y evolución histórica

Dominio	Definición	Origen y evolución
<b>Inteligencia Artificial (IA)</b>	“Campo de estudio que busca crear sistemas capaces de realizar tareas que requieren inteligencia humana” (Russell & Norvig, 2021).	Nació en 1956 en la Conferencia de Dartmouth. Al inicio se centró en reglas heurísticas; hoy se basa en redes neuronales profundas y sistemas de aprendizaje automático.
<b>Machine Learning (ML)</b>	“Subcampo de la IA que proporciona a los sistemas la capacidad de aprender automáticamente y mejorar a partir de la experiencia sin ser programados explícitamente” (Mitchell, 1997).	Sus raíces datan de los 50 (perceptrón). En los 90 se popularizan algoritmos estadísticos. Con el auge del big data y GPUs se consolida en la década de 2010.
<b>Data Mining (DM)</b>	“Proceso de descubrir patrones significativos y relaciones en grandes conjuntos de datos” (Han, Kamber & Pei, 2012).	Surgió en los 80-90 como evolución del análisis de bases de datos. Hoy se integra con ML y Big Data para análisis avanzado.
<b>Big Data</b>	“Conjunto de tecnologías y prácticas para procesar y analizar grandes volúmenes de datos caracterizados por volumen, velocidad y variedad” (Gartner, 2011).	El término se populariza a mediados de los 2000 con Hadoop. Evoluciona hacia ecosistemas en la nube y análisis en tiempo real.

## Comparación de beneficios

Dominio	Beneficio 1	Beneficio 2	Beneficio 3
<b>IA</b>	Automatiza procesos complejos	Mejora la toma de decisiones basada en datos	Habilita nuevas experiencias (chatbots, asistentes virtuales)
<b>ML</b>	Predicción de tendencias y comportamientos	Reducción de errores humanos	Mejora continua del modelo con datos nuevos
<b>DM</b>	Descubre patrones ocultos en grandes datos	Segmenta clientes y mercados	Identifica oportunidades y riesgos antes no visibles
<b>Big Data</b>	Manejo de volúmenes masivos de datos	Procesamiento veloz en batch y streaming	Integración de fuentes diversas para análisis unificado

## Restricciones y retos

Restricciones y retos actuales de cada dominio

Dominio	Restricciones (2)	Retos actuales (2)
<b>IA</b>	Alto consumo de recursos de cómputo; Necesidad de datos de calidad	Explicabilidad y transparencia de modelos complejos; Mitigación de sesgos en datos de entrenamiento
<b>ML</b>	Dependencia de grandes volúmenes de datos etiquetados; Dificultad para trasladar modelos del laboratorio a producción	Privacidad y manejo ético de datos; Entrenamiento continuo sin degradar rendimiento
<b>DM</b>	Limitación cuando los datos son desestructurados o dispersos; Puede generar hallazgos espurios si no se valida bien	Integrar minería sobre datos en streaming; Escalar algoritmos clásicos a volúmenes Big Data
<b>Big Data</b>	Coste elevado de almacenamiento e infraestructura; Latencia y complejidad en el procesamiento en tiempo real	Garantizar seguridad y privacidad en grandes entornos distribuidos; Formar talento especializado para administrar ecosistemas complejos

## Conclusión

Los cuatro dominios se complementan. Big Data proporciona la infraestructura y el volumen; DM extrae patrones; ML aprende de esos patrones para generar predicciones; y la IA integra todo para tomar decisiones y automatizar tareas de forma “inteligente”. Estos aunque comparten beneficios como mejorar decisiones y automatizar procesos, presentan restricciones técnicas (recursos de cómputo, calidad de datos) y retos éticos (privacidad, sesgos). La tendencia actual es integrar estos dominios en plataformas unificadas en la nube que permitan análisis híbridos batch y streaming. Comprender estos dominios permite elegir la combinación correcta de tecnologías y procesos para cada proyecto, maximizando resultados y minimizando riesgos.

## Referencias

- Gartner. (2011). Big Data definition. Recuperado de <https://www.gartner.com>
- Han, J., Kamber, M., & Pei, J. (2012). Data Mining: Concepts and Techniques (3ª ed.). Elsevier.
- Mitchell, T. M. (1997). Machine Learning. McGraw-Hill.
- Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4ª ed.). Pearson.
- Apache Software Foundation. (s.f.). Apache Hadoop. Recuperado de <https://hadoop.apache.org/>
- TensorFlow developers. (s.f.). TensorFlow. Recuperado de <https://www.tensorflow.org/>