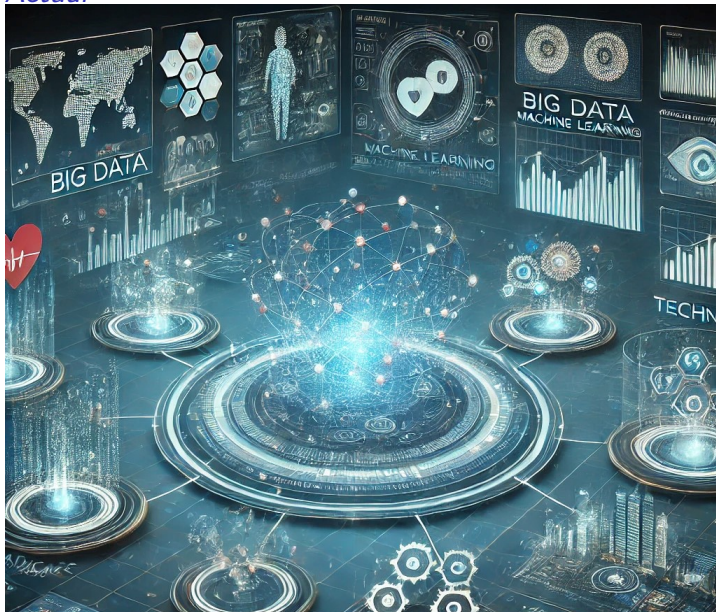


# Introducción a la Ciencia de Datos

Facultad de Ciencias Basicas - UTB

# Introducción a la Ciencia de Datos

La Ciencia de Datos *Su Impacto y Aplicaciones en el Mundo Actual*



# ¿Qué es la Ciencia de Datos?

**Definición:** La Ciencia de Datos es un campo interdisciplinario que utiliza métodos, algoritmos y sistemas científicos para extraer conocimientos y obtener valor a partir de grandes volúmenes de datos.

**Elementos Clave:**

- **Datos:** Crudos, estructurados o no estructurados.
- **Procesamiento:** Limpieza, análisis y modelado.
- **Valor:** Información útil para la toma de decisiones.

# Importancia de la Ciencia de Datos

## Transformación de Sectores:

- ▶ **Negocios:** Toma de decisiones basada en datos.
  - ▶ **Salud:** Predicción de enfermedades, tratamientos personalizados.
  - ▶ **Marketing:** Análisis de patrones de compra y comportamiento del cliente.
- **Tecnología:** Inteligencia artificial y aprendizaje automático (Machine Learning).

## Ventaja Competitiva:

Empresas que usan la ciencia de datos mejoran la eficiencia y la innovación.

# Componentes de la Ciencia de Datos

1. **Estadística:** Fundamenta la interpretación de datos.
2. **Programación:** Uso de lenguajes como Python, R y SQL.
3. **Aprendizaje Automático:** Algoritmos que aprenden de los datos.
4. **Visualización:** Comunicación efectiva de hallazgos a través de gráficos.
5. **Manejo de Grandes Volúmenes de Datos (Big Data):** Herramientas y tecnologías para almacenar y procesar datos masivos.

# Proceso de Ciencia de Datos

1. **Recolección de Datos:** Datos provienen de múltiples fuentes.
2. **Preparación de Datos:** Limpieza y transformación de los datos.
3. **Análisis Exploratorio de Datos (EDA):** Identificación de patrones, tendencias y anomalías.
4. **Modelado:** Creación de modelos predictivos o descriptivos.
5. **Validación:** Evaluación de la precisión de los modelos.
6. **Comunicación de Resultados:** Presentación de hallazgos a través de informes y visualizaciones.

# Herramientas Populares en Ciencia de Datos

- ▶ **Lenguajes de Programación:** Python, R, SQL.
- ▶ **Bibliotecas:** Pandas, NumPy, Scikit-learn, TensorFlow.
- ▶ **Visualización:** Matplotlib, Seaborn, Tableau.
- ▶ **Big Data:** Apache Spark, Hadoop.

# Aplicaciones en el Mundo Real

1. **Recomendación de Productos** (Amazon, Netflix): Basado en el comportamiento del usuario.
2. **Diagnóstico Médico**: Algoritmos que ayudan a predecir enfermedades.
3. **Análisis Financiero**: Predicción de riesgos y tendencias en mercados.
4. **Reconocimiento de Voz e Imagen**: Aplicaciones como Siri o Google Photos.



# Retos en la Ciencia de Datos

- ▶ **Calidad de Datos:** Datos incompletos o inconsistentes.
- ▶ **Privacidad y Ética:** Manejo de datos sensibles.
- ▶ **Interpretación de Modelos:** Comprender cómo funcionan los algoritmos complejos.
- ▶ **Escalabilidad:** Procesar datos masivos en tiempo real.

# Futuro de la Ciencia de Datos

- ▶ **Automatización de Modelos:** Uso de AutoML para reducir la intervención humana.
- ▶ **Aumento del Aprendizaje Automático:** Expansión del uso de AI y modelos predictivos.
- ▶ **Ética en la Ciencia de Datos:** Creciente interés en el uso responsable de datos.
- ▶ **Integración con IoT:** Expansión de datos provenientes de dispositivos conectados.

# Conclusión

*La Ciencia de Datos está transformando industrias y cambiando la forma en que las organizaciones toman decisiones. Con una demanda creciente de profesionales capacitados, es un campo dinámico con aplicaciones en casi todos los sectores.*

**Pregunta final:** ¿Cómo puede la ciencia de datos transformar tu sector?

