

## **Relación con otras disciplinas:**

**Bases de Datos:** Conceptos como el Data Warehouse y el procesamiento analítico en línea (OLAP). Técnicas de indización para el diseño de algoritmos eficientes de minería de datos.

**Recuperación de Información:** Uso de medidas de similitud entre documentos y consultas sobre clasificaciones de documentos en función en sus palabras claves, para obtener información de datos textuales.

**Estadística:** Ha proporcionado muchos de los conceptos, algoritmos y técnicas que se utilizan en minería de datos, como lo son la media, la varianza, las distribuciones, el análisis univariante y multivariante, la regresión lineal y no lineal, la teoría del muestreo, la validación cruzada, la modelización paramétrica y no paramétrica, las técnicas bayesianas, etc.

**Machine Learning:** Desarrollo de algoritmos capaces de aprender. La máquina aprende un modelo a partir de ejemplos y lo usa para resolver el problema.

**Sistemas para la toma de decisión:** Herramientas y sistemas informatizados que asisten a los directivos en la resolución de problemas y en la toma de decisiones, con el objetivo de proporcionar la información necesaria para realizar decisiones efectivas en el ámbito empresarial o en tareas de diagnóstico.

**Visualización de datos:** Uso de técnicas de visualización que permiten al usuario descubrir, intuir o entender patrones que de otra manera serían difíciles de entender a partir de descripciones matemáticas o textuales de los resultados, tales como las gráficas, las icónicas, las basadas en píxeles, las jerárquicas, etc.

**Computación paralela y distribuida:** El coste computacional de las tareas más complejas de la minería de datos se reparte entre diferentes procesadores o computadores.

**Otras disciplinas:** Dependiendo del tipo de datos a ser minados o del tipo de aplicación, la minería de datos puede hacer uso de otras disciplinas como el lenguaje natural, el análisis de imágenes, el procesamiento de señales, los gráficos por computadora, etc.

---

## **Aplicaciones:**

- **Aplicaciones financieras y banca:** Obtención de patrones de uso fraudulento de tarjetas de crédito, análisis de riesgos en créditos.
- **Análisis de mercado:** Análisis de la cesta de compra (compras conjuntas, secuenciales, etc.); evaluación de campañas publicitarias; análisis de la fidelidad de los clientes (Reducción de la fuga); estimación de stocks, costes, ventas, etc.
- **Seguros:** Determinación de los clientes que podrían ser potencialmente caros, análisis de procedimientos médicos solicitados conjuntamente.

- **Procesos industriales:** Modelos de calidad, predicción de fallos y accidentes, extracción de modelos de coste y producción.
- **Medicina:** Identificación de patologías y diagnóstico de enfermedades, detección de pacientes con riesgo a sufrir una patología concreta, recomendación de fármacos para una isma patología.
- **Biología:** Análisis de secuencias de genes, análisis de secuencias de proteínas, predecir si un compuesto químico causa cáncer, clasificación de cuerpos celestes.
- **Telecomunicaciones:** Establecimiento de patrones de llamadas, detección de fraude.
- **Correo electrónico:** Clasificación y distribución automática de correo, detección de spam.
- **Recreativos:**
- **Turismo:** Identificar patrones de reservas, determinar las características socioeconómicas de los turistas en un determinado destino.
- **Asuntos fiscales:** Detección de evasión fiscal.
- **Educación:** Selección o capacitación de estudiantes, detección de abandonos y fracasos, estimación de tiempo de estancia en la institución.

---

## Sistemas y herramientas de la Minería de Datos:

- **Tipo de bases de datos minada:** Teniendo en cuenta los diferentes modelos de datos, se pueden considerar: sistemas de minería de datos relacionales, multidimensionales, orientados a objetos, etc. Asimismo, considerando el tipo de datos manejados, se consideran también: sistemas textuales, multimedia, espaciales o web.
- **Tipo de conocimiento minado:** Pueden tenerse en cuenta los niveles de abstracción del conocimiento minado: conocimiento generalizado (alto nivel de abstracción), a nivel primitivo (a nivel de filas de datos), o conocimiento a múltiples niveles de abstracción. Podemos distinguir también entre los sistemas que buscan regularidades en los datos (patrones), frente a los que analizan las irregularidades (excepciones).
- **Tipo de funcionalidad y técnica:** Los sistemas de Minería de Datos se pueden clasificar basándose en su funcionalidad (clasificación, agrupamiento, etc.) o por los métodos de análisis de los datos empleados (técnicas estadísticas, redes neuronales, etc.)
- **Tipo de aplicación:** Se puede distinguir entre dos clases de sistemas según la aplicación para la que se usan: sistemas de propósito general y sistemas específicos (aplicaciones financieras, web, e-mail, análisis de stocks, etc.).