

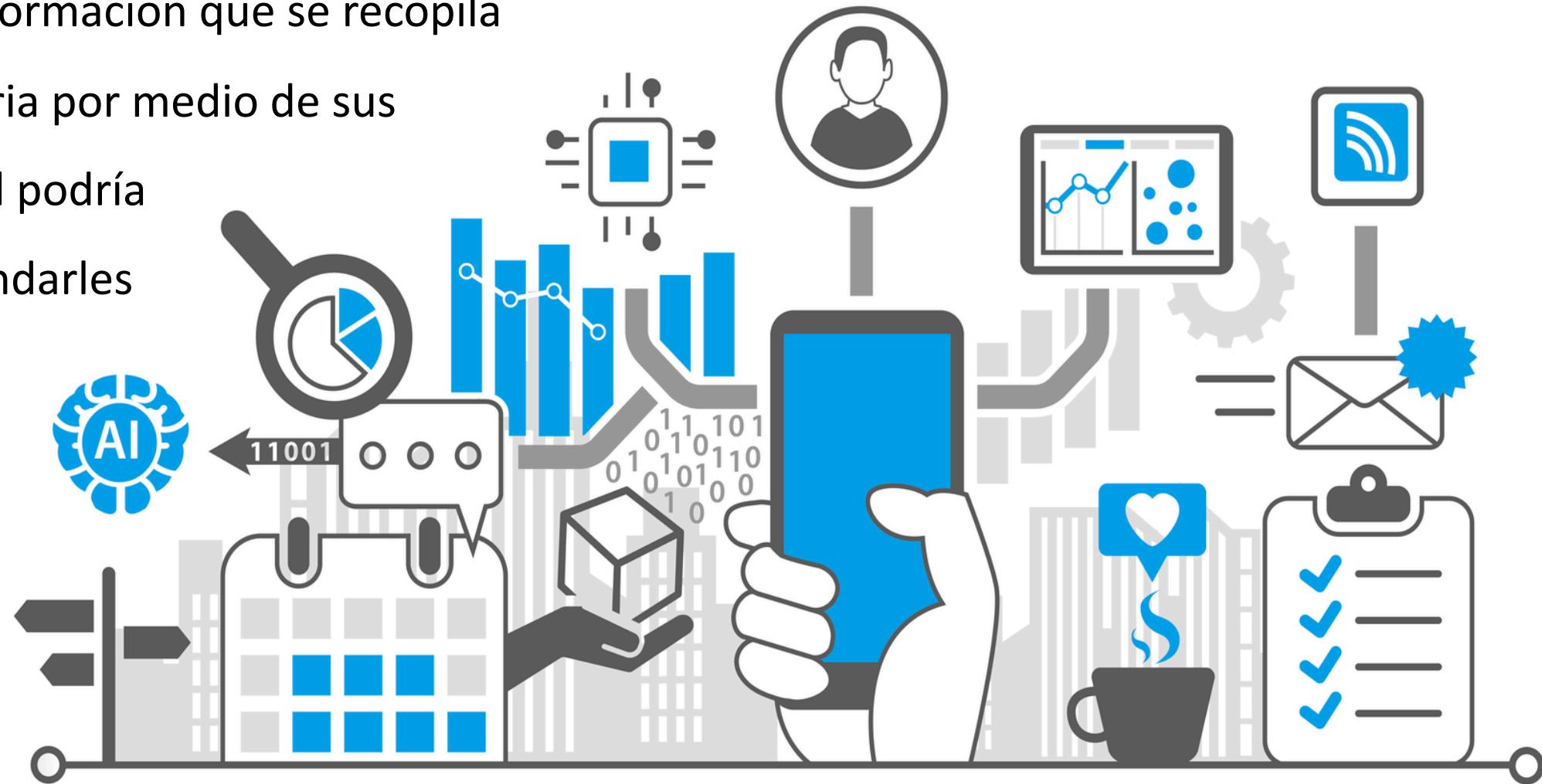


MINERIA DE DATOS

INTRODUCCIÓN

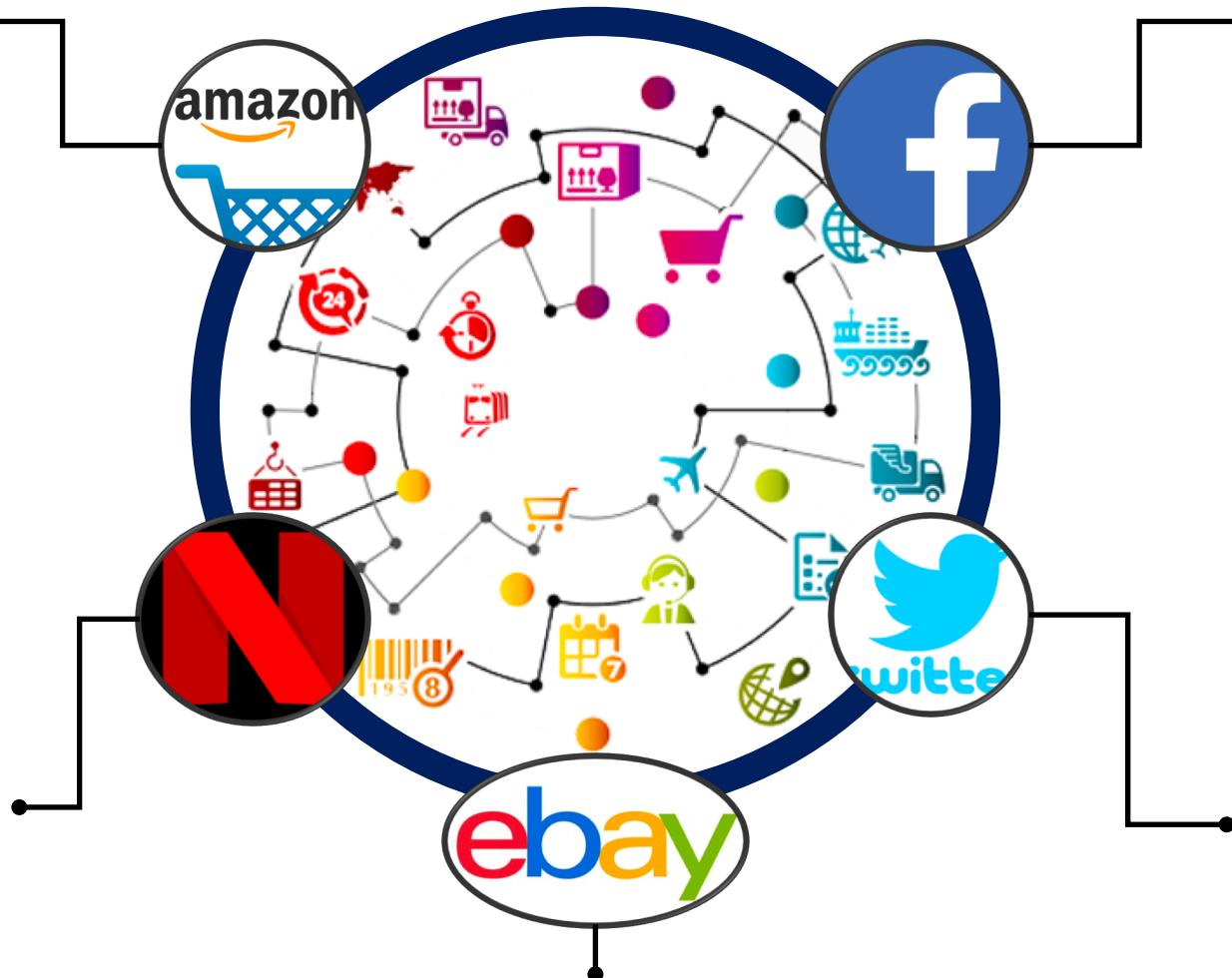
1

En los últimos años las empresas se dieron cuenta de la oportunidad que presenta la cantidad de información que se recopila de manera diaria por medio de sus clientes, la cual podría permitirles brindarles un mejor servicio.



- **199** Millones promedio mensual usuarios.
 - **100** Millones suscriptores Prime.
 - **3er** lugar dispositivos inteligentes.

EJEMPLOS



- **180** Millones usuarios activos.
 - **2** Billones de transacciones por día.

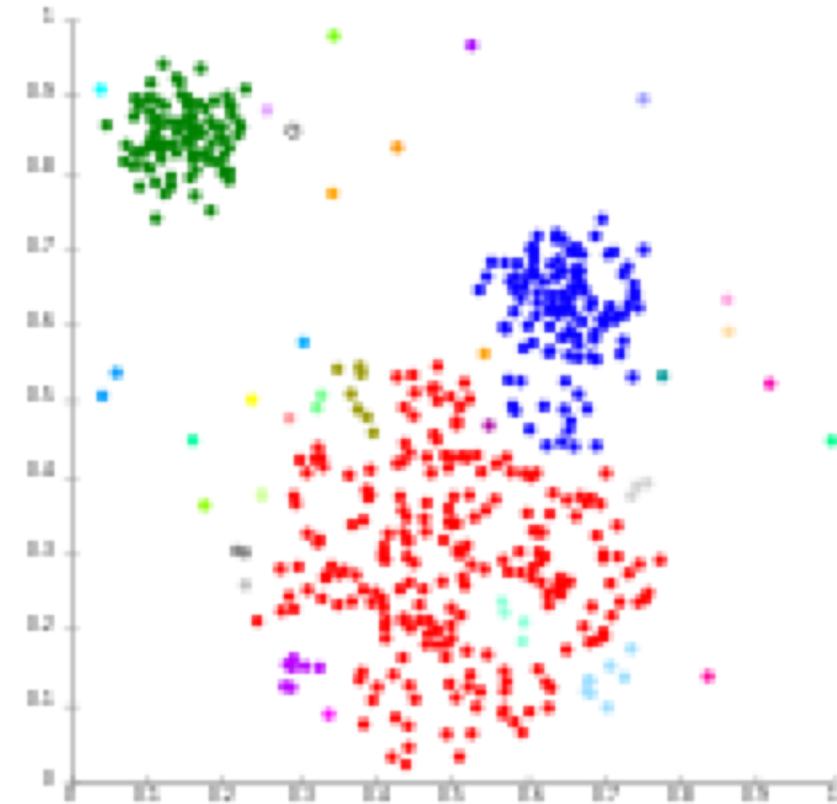
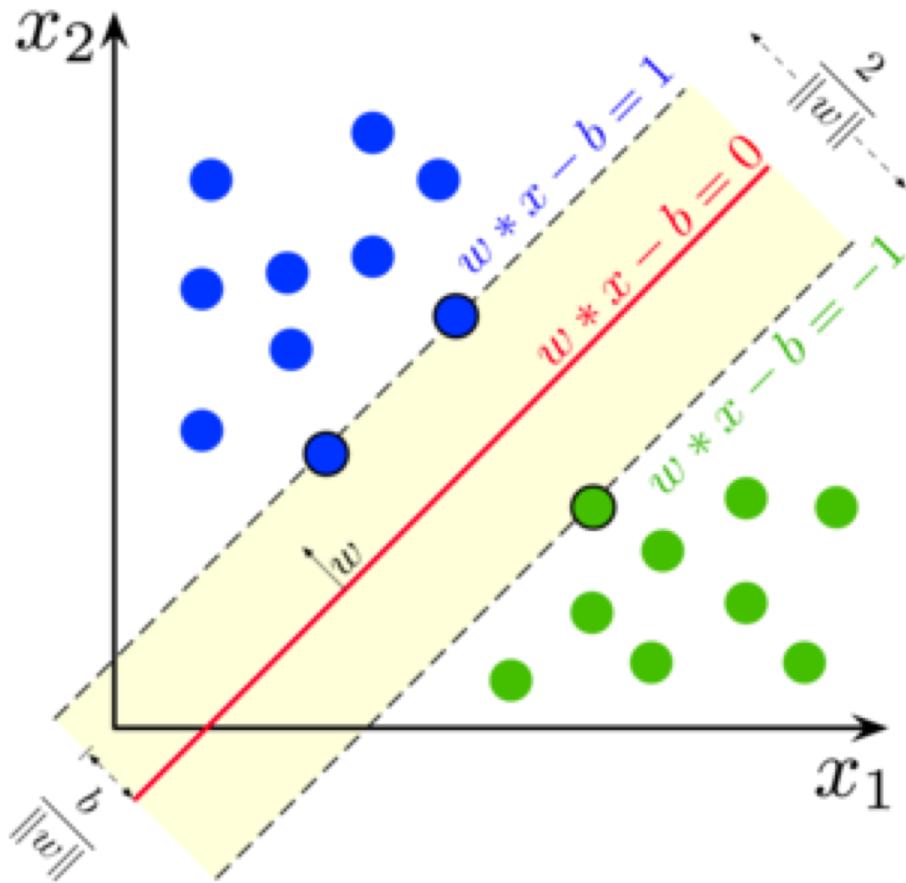
Extraer información útil de estos datos es una tarea exhaustiva que ha creado la necesidad de un nuevo campo de investigación. Ciencia de datos.

CONCEPTO:

La minería de datos es el proceso de detectar la información procesable de los conjuntos grandes de datos. Utiliza el análisis matemático para deducir los patrones y tendencias que existen en los datos.



Normalmente, estos patrones no se pueden detectar mediante la exploración tradicional de los datos porque las relaciones son demasiado complejas o porque hay demasiados datos.



El objetivo es encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.

¿CÓMO APLICAR UN MODELO DE MINERÍA DE DATOS?

- Pronóstico
- Recomendaciones
- Búsqueda de secuencias
- Riesgo y probabilidad
- Agrupación

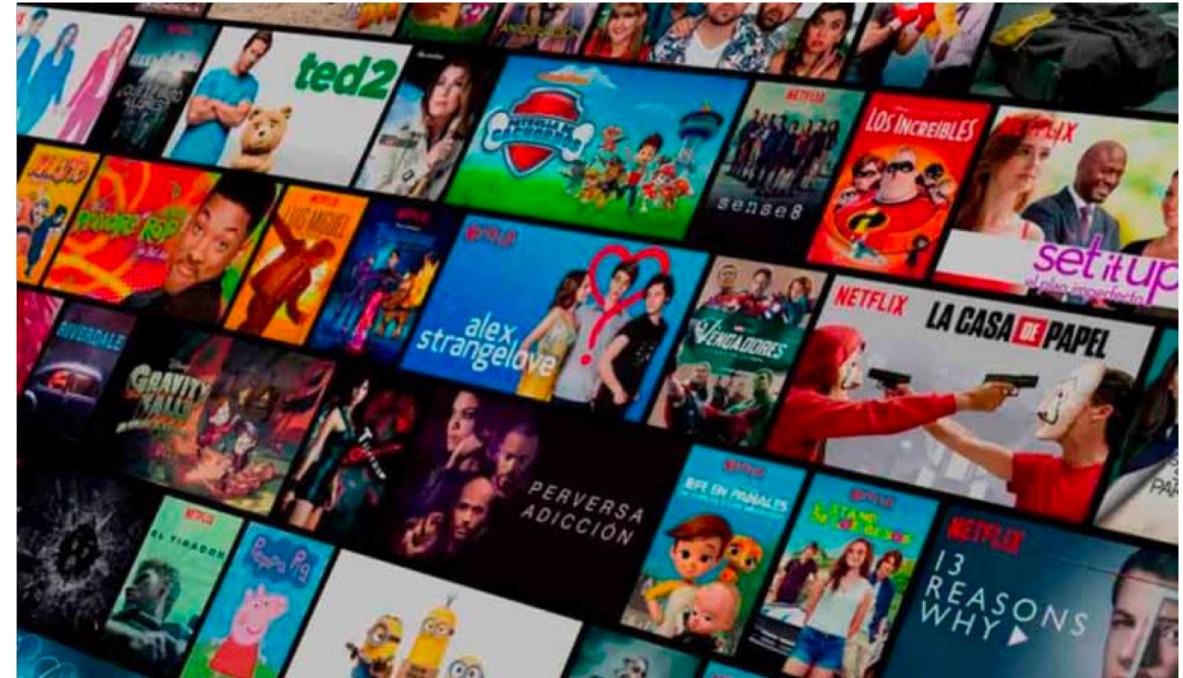


- **Pronóstico**



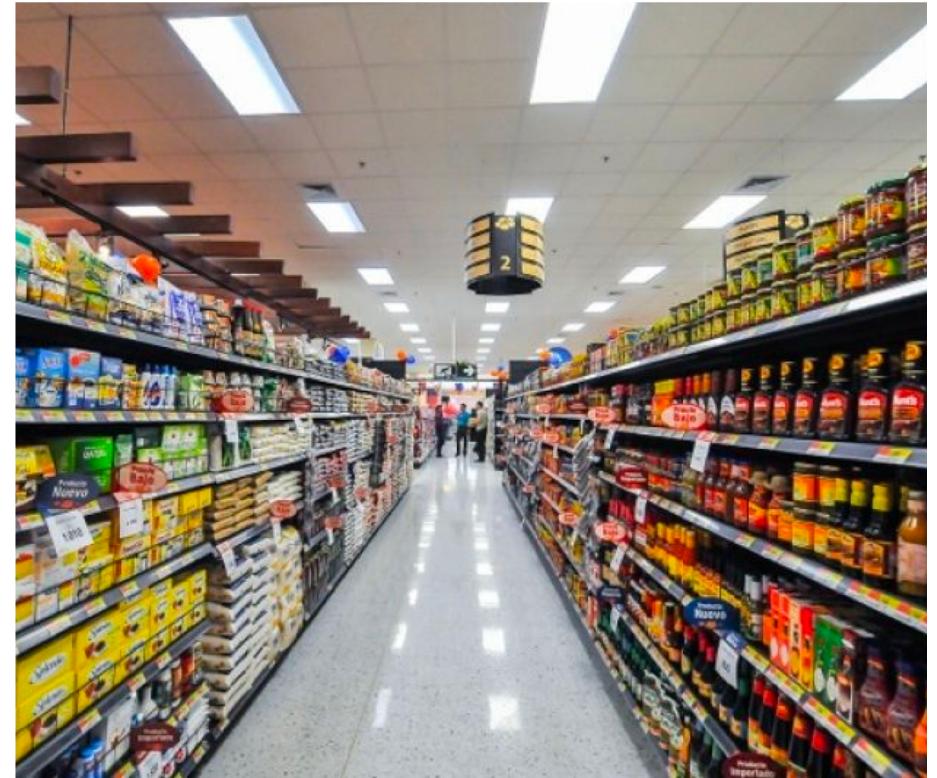
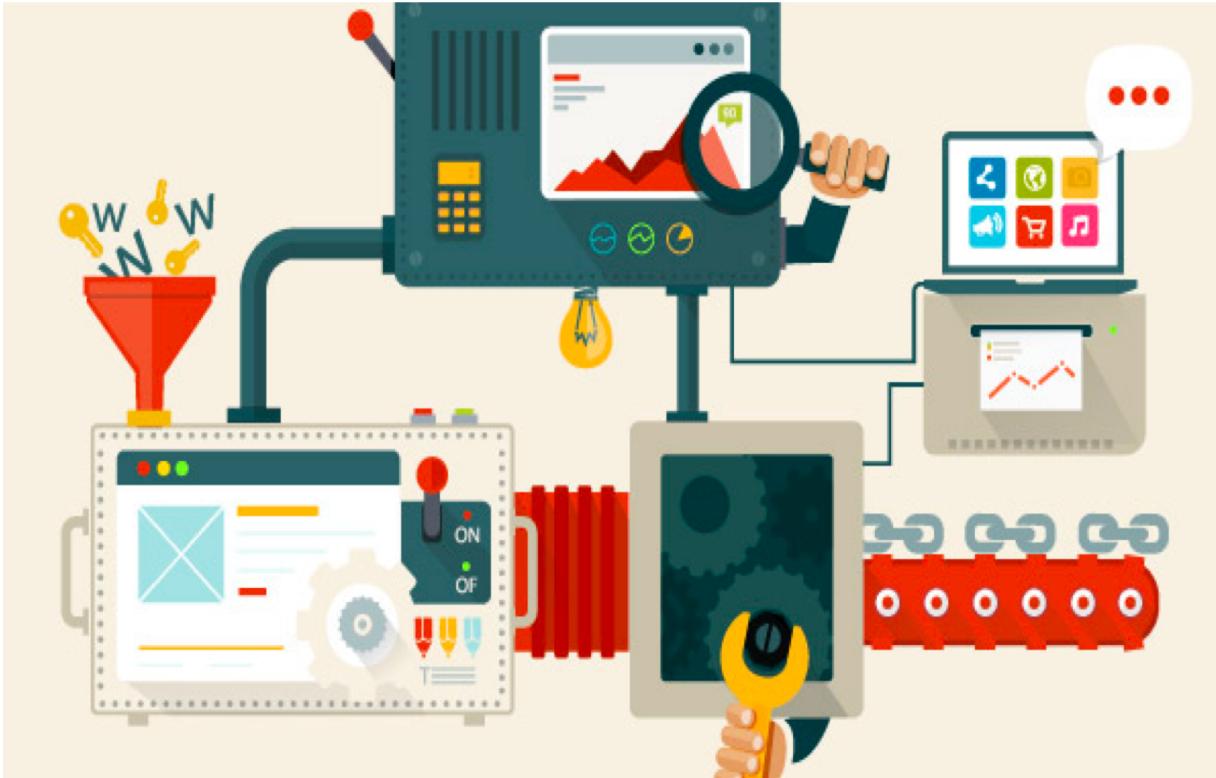
Cálculo de las ventas y predicción de las cargas del servidor o del tiempo de inactividad del servidor.

• Recomendaciones



Determinación de los productos que se pueden vender juntos y generación de recomendaciones.

- **Búsqueda de secuencias**



Análisis de los artículos que los clientes han introducido en el carrito de la compra y predicción de posibles eventos.

- **Riesgo y probabilidad**



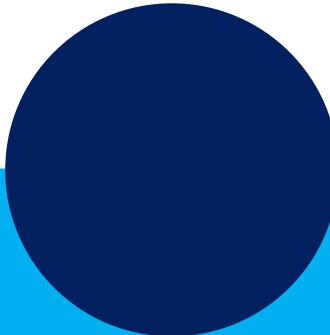
Elección de los mejores clientes para la distribución de correo directo, determinación del punto de equilibrio probable para los escenarios de riesgo, y asignación de probabilidades a diagnósticos y otros resultados.

• Agrupación



Distribución de clientes o eventos en grupos de elementos relacionados, y análisis y predicción de afinidades.

EJERCICIO



Mapa Mental

Realizar un mapa mental con al menos un ejemplo diferente de los mencionados en las seis formas de aplicar los modelos de minería de datos: Pronóstico, Recomendaciones, Búsqueda de secuencias, Riesgo y probabilidad y Agrupación.



TÉCNICAS DE MINERÍA DE DATOS

2

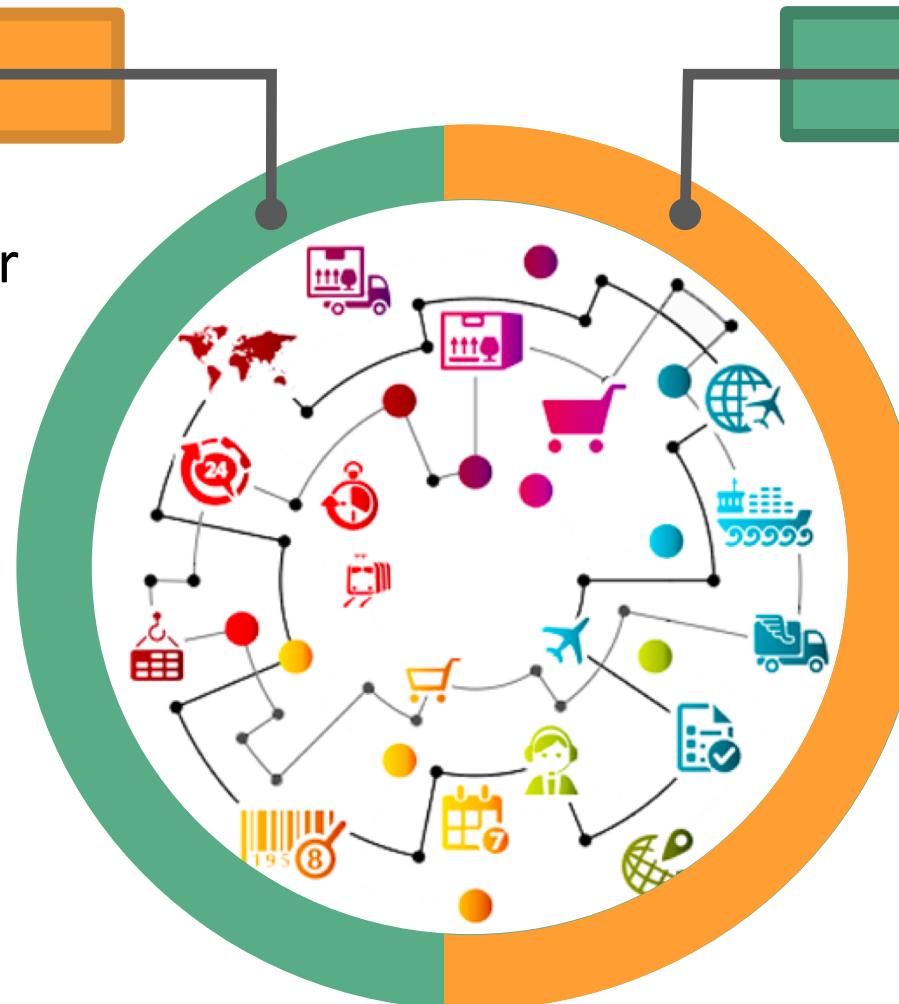
Las tareas de minería de datos se dividen generalmente en dos categorías

Descriptivo

El objetivo es encontrar patrones que den un resumen de las relaciones ocultas dentro de los datos. Descubre las características más importantes de la base de datos.

Predictivo

Predecir el valor de un atributo en particular basándome en los datos recolectados de otros atributos.



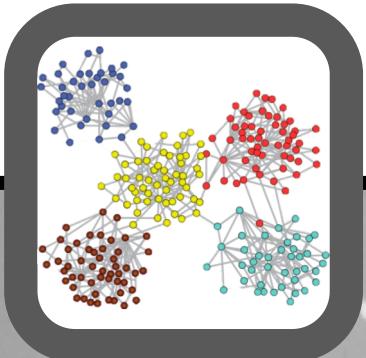
Agrupar elementos en grupos con características similares

Descubrir hechos que ocurren en común dentro de un conjunto de datos. Relación relevante entre los elementos.

Estudia el comportamiento de valores extremos que difieren del patrón general de una muestra.

Representación gráfica de los elementos más importantes de la base de datos.

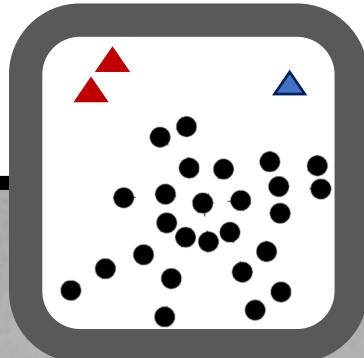
CLUSTERING



REGLAS DE ASOCIACIÓN



DETECCIÓN DE OUTLIERS



VISUALIZACIÓN



DESCRIPTIVAS

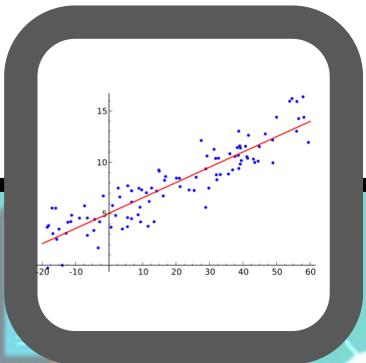
Analiza el vínculo entre una variable dependiente y una o varias independientes, encontrando una relación matemática.

Ordenamiento o disposición por clases tomando en cuenta las características de los elementos dentro de estas.

Se especializa en analizar datos, encontrar subsecuencias interesantes dentro de un grupo de secuencias.

Analizar datos actuales históricos reales para darnos información sobre acontecimientos no conocidos o futuros.

REGRESIÓN



CLASIFICACIÓN



PATRONES SECUENCIALES



PREDICCIÓN

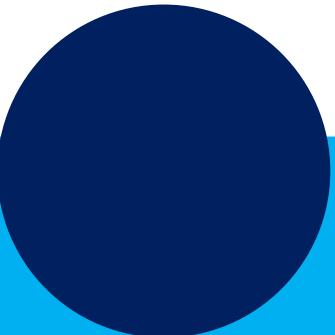


PREDICTIVAS

- Solución eficiente comparada con otras aplicaciones estadísticas de datos.
- Ayuda a las empresas a hacer ajustes en producción y operaciones para mejorar su beneficio.
- Facilita las predicciones automáticas de comportamientos y tendencias, ayudando a descubrir patrones escondidos.
- Es un proceso veloz que facilita a los usuarios a analizar grandes cantidades de información en el menor tiempo posible.

- En ciertas condiciones, las técnicas de minería de datos pueden resultar no ser muy exactas, lo que puede causar serias consecuencias.
- El diseño y uso de las herramientas de minería de datos tiene diferentes resultados dependiendo del algoritmo que se utilice para realizarlo. Esto hace que encontrar los parámetros adecuados se vuelva una tarea difícil.

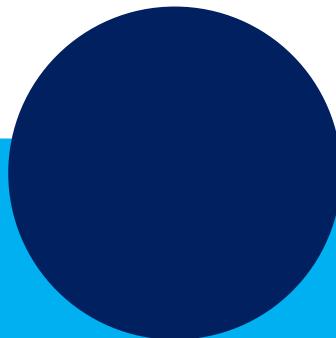
**TRABAJO
EN EQUIPO**



- **Presentación**

Por equipos, se expondrá una presentación describiendo una técnica de minería de datos por equipo.

TRABAJO INDIVIDUAL



- **RESUMEN**

Después de cada presentación por equipo, se entregara un resumen de al menos media cuartilla de los conceptos más importantes vistos en la presentación. En caso de ser encargado un ejercicio en cualquiera de las presentaciones, este se entregará dentro del resumen.

- **MAPA MENTAL**

Al término de todas las exposiciones por equipo, se entregará un mapa mental que contenga todas las técnicas de minería de datos.