



Universidad Autónoma del Estado de México



Centro Universitario UAEM Valle de México

Ingeniería en Sistemas y Comunicaciones

Unidad de Aprendizaje

Temas Selectos de Sistemas

Unidad de Competencia

Aplicaciones

Elaboró: Saturnino Job Morales Escobar

Fecha de elaboración: Agosto de 2018



Programa de Estudios por Competencias
TEMAS SELECTOS DE SISTEMAS

I. IDENTIFICACIÓN DEL CURSO

ORGANISMO ACADÉMICO: CENTRO UNIVERSITARIO UAEM VALLE DE MEXICO								
Programa Educativo: INGENIERIA EN SISTEMAS Y COMUNICACIONES					Área de docencia: CIENCIAS DE LA INGENIERÍA			
Aprobación por los H. H. Consejos Académico y de Gobierno			Fecha: Fecha de actualización:		Programa elaborado por: SATURNINO JOB MORALES ESCOBAR, IVONNE RODRÍGUEZ PÉREZ		Fecha de actualización: Noviembre 2010	
Clave	Horas de teoría	Horas de práctica	Total de horas	Créditos	Tipo de Unidad de Aprendizaje	Carácter de la Unidad de Aprendizaje	Núcleo de formación	Modalidad
L32295	3	1	4	7	CURSO	OBLIGATORIA	SUSTANTIVO	PRESENCIAL
Prerrequisitos (Conocimientos Previos): PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS, MATEMÁTICAS DISCRETAS, ALGORITMOS Y ESTRUCTURAS DE DATOS.					Unidad de Aprendizaje Antecedente NINGUNA		Unidades de Aprendizaje Consecuentes: NINGUNA	
Programas educativos en los que se imparte: INGENIERIA EN SISTEMAS Y COMUNICACIONES								

CIENCIAS BÁSICAS Y MATEMÁTICAS

ALGEBRA Y GEOMETRÍA ANALÍTICA	3 3 9	ALGEBRA LINEAL	2 2 6	ECUACIONES DIFERENCIALES	3 1 7	CÁLCULO VECTORIAL	3 1 7
ESTÁTICA Y DINÁMICA	3 3 9	CÁLCULO DIFERENCIAL E INTEGRAL	3 1 7	LÓGICA MATEMÁTICA	3 1 7	ELECTROMAGNETISMO	2 2 6
QUÍMICA	2 2 6	MATEMÁTICAS DISCRETAS	3 1 7	MÉTODOS NUMÉRICOS	4 2 10	PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS	4 2 10

CIENCIAS DE LA INGENIERÍA

INTRODUCCIÓN A LA INGENIERÍA	2 4 8	FUNDAMENTOS DE PROGRAMACIÓN	3 3 9	FUNDAMENTOS DE BASES DE DATOS	3 1 7	SISTEMAS OPERATIVOS	4 2 10	LENGUAJES FORMALES Y AUTÓMATAS	2 2 6	ELECTRÓNICA ANALÓGICA	4 2 10	COMUNICACIÓN POR MEDIOS ÓPTICOS 226	TEMAS SELECTOS DE SISTEMAS 317
LENGUAJES DE BAJO NIVEL	2 4 8	ALGORITMOS Y ESTRUCTURAS DE DATOS	2 4 8	BASES DE DATOS	2 4 8	PROGRAMACIÓN ORIENTADA A OBJETOS	2 4 8	INGENIERÍA DE SOFTWARE	2 4 8	COMUNICACIÓN VÍA MICROONDAS Y SATELITAL 226	SISTEMAS ELECTRÓNICOS DE COMUNICACIÓN 226		
CIRCUITOS ELÉCTRICOS	2 2 6	SISTEMAS DE INFORMACIÓN	2 4 8	PROGRAMACIÓN AVANZADA 226		COMPLADORES 226							
INVESTIGACIÓN DE OPERACIONES	2 4 8	ADMINISTRACIÓN DE BASES DE DATOS 226		TEORÍA DEL CONTROL* 226		GRAFICACIÓN 226							
SISTEMAS OPERATIVOS PARA RED	4 2 10	CALIDAD DEL SOFTWARE 226											

INGENIERÍA APLICADA

INTRODUCCIÓN A LA COMPUTACIÓN	2 4 8	ARQUITECTURA DE COMPUTADORAS	2 4 8	REDES	4 2 10	PROTOSCOLOS DE COMUNICACIÓN DE DATOS	3 3 9	DESARROLLO DE PROYECTOS	2 2 6	FORMULACIÓN Y EVALUACIÓN DE PROYECTOS	3 3 9	RESIDENCIA PROFESIONAL	0 30 30
SEMINARIO DE TITULACIÓN* 226		INTELIGENCIA ARTIFICIAL* 226		SISTEMAS DIGITALES	2 4 8	SISTEMAS DISTRIBUIDOS 226		INTERCONEXIÓN Y SEGURIDAD EN REDES	2 2 6	TRANSMISIÓN Y COMUNICACIÓN DE DATOS 226			
SISTEMAS DE TIEMPO REAL 226				SISTEMAS EXPERTOS 226		SISTEMAS DE INSTRUMENTACIÓN Y CONTROL 226		TALLER DE INVESTIGACIÓN 226					

CIENCIAS SOCIALES Y HUMANIDADES

ADMINISTRACIÓN 317	CONTABILIDAD 317	ECOLOGÍA, ÉTICA Y NORMATIVIDAD 317	INGLÉS C1 226	INGLÉS C2 226	PSICOLOGÍA ORGANIZACIONAL 226	PLANEACIÓN ESTRATÉGICA 226	METODOLOGÍA DE LA INVESTIGACIÓN 226
CIENCIA Y HUMANISMO 317	PROBLEMAS SOCIOECONÓMICOS DE MÉXICO 317	TÉCNICAS DE COMUNICACIÓN 317			ADMINISTRACIÓN DE CENTROS DE CÓMPUTO 226	AUDITORIA Y SEGURIDAD INFORMÁTICA 226	

Temas Selectos de Sistemas

Unidades de Competencia

- 1. Conceptos Básicos del Reconocimiento de Patrones*
- 2. Selección de Variables*
- 3. Clasificación Supervisada*
- 4. Clasificación no Supervisada*
- 5. Procesamiento de Lenguaje Natural*
- 6. Aplicaciones*

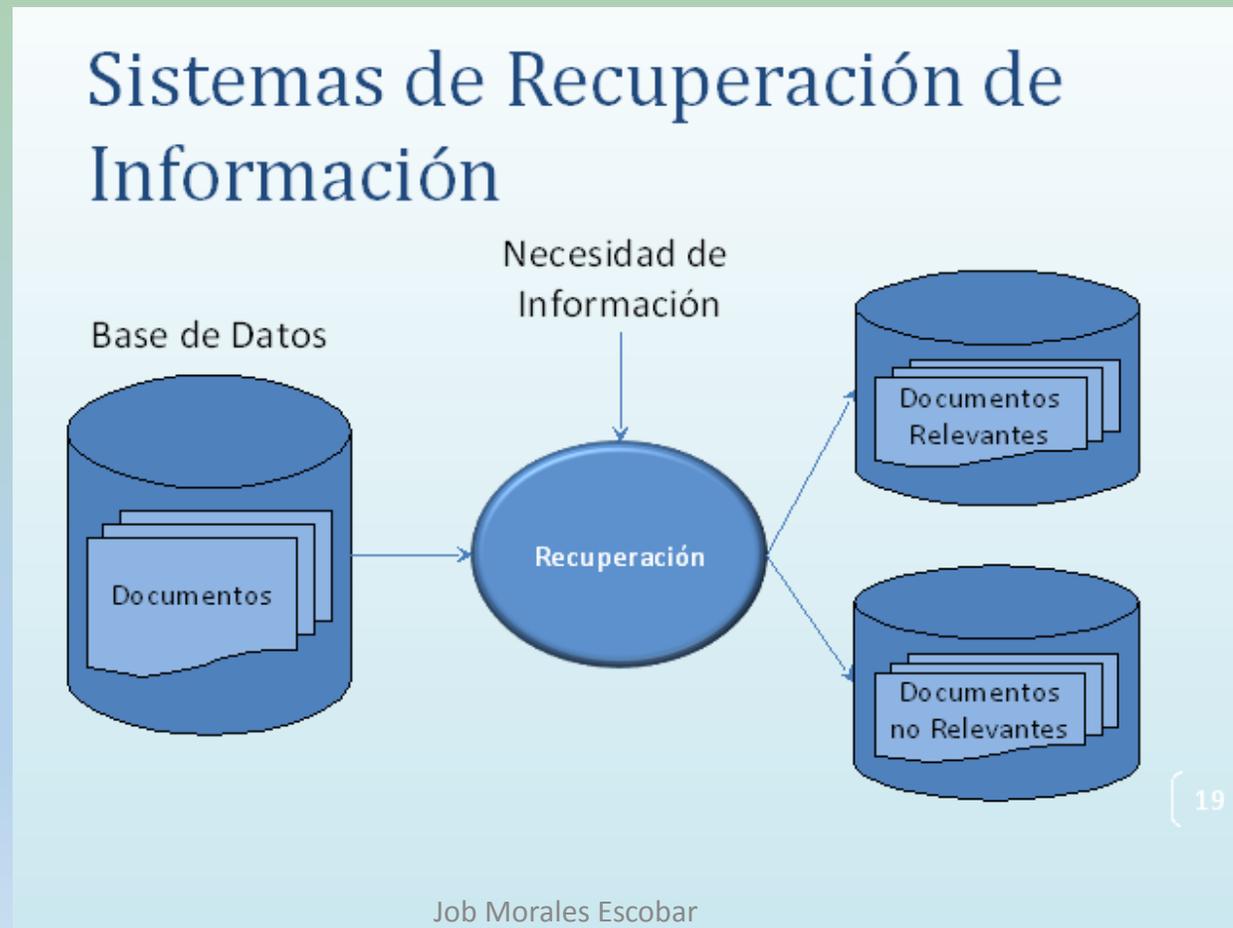
Evaluación de la unidad de competencia

<i>Exámenes</i>	<i>50%</i>
<i>Exposición de una aplicación</i>	<i>10%</i>
<i>Ejercicios</i>	<i>40%</i>

Objetivo de la unidad de competencia

Analizar la aplicación de los algoritmos y técnicas del Reconocimiento de Patrones y el procesamiento del lenguaje natural a la minería de datos y a los sistemas de recuperación de información.

Posteriormente se tratarán los sistemas de recuperación de información



Contenido

- *Introducción*
- *Conceptos de Minería de Datos*
- *Presentar un panorama de la aplicación de la Minería de Datos*
- *Métodos utilizados en la minería de datos*
- *Productos orientados a la Minería de Datos*
- *Aplicación de Minería de datos*

Introducción

Ante la gran cantidad de datos, documentos, información, imágenes, etcétera, que se generan de manera digital, la capacidad de procesarla manualmente es sobrepasada y su acumulación da como resultado un problema humanamente intratable.

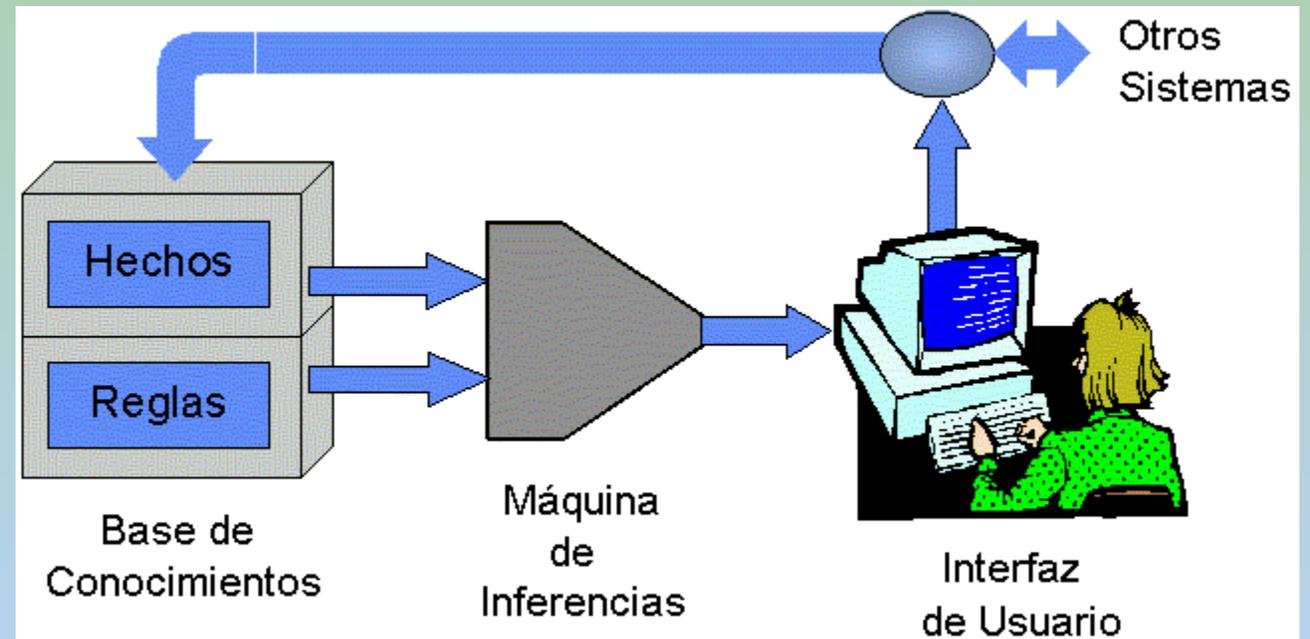
Una solución, es desarrollar herramientas computacionales que realicen ese trabajo.



Introducción

continuación

Así, se han desarrollado diversos tipos de sistemas, desde sistemas de información hasta sistemas expertos, sin embargo, el usuario necesita proporcionar datos suficientes para obtener el resultado.



Introducción

continuación

De manera ideal, el sistema debería hacer la búsqueda de relaciones interesantes entre los datos sin la intervención del usuario. Con esta premisa, se han desarrollado diferentes tipos de mineros.

En este curso, se estudiará la aplicación a la minería de datos, en el entendido que existen aplicaciones a la minería de textos y de imágenes, entre otras.

Busco en datos ...

Secuencias

Archivos

Consumos

Hábitos de Consumo

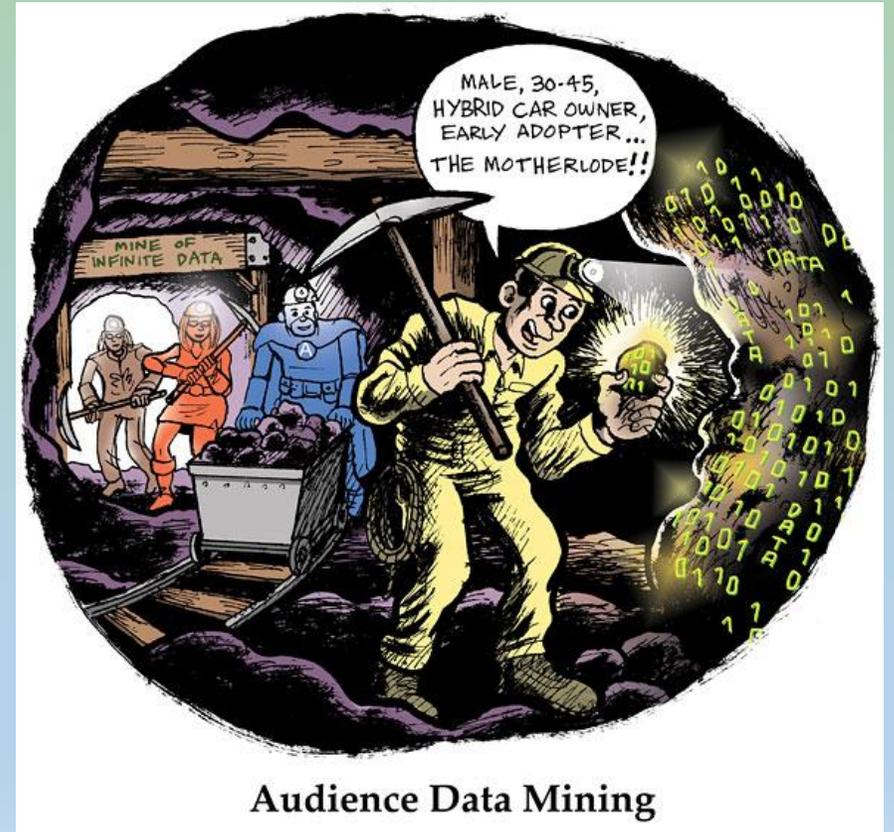
Internet

Dispositivos

Informes

Definición de Minería de Datos

El proceso de extracción de conocimiento útil y comprensible, previamente desconocido, a partir de grandes cantidades de datos almacenados en distintos formatos.



Otra definición de Minería de Datos

*Extracción no trivial
de información
implícita, previamente
desconocida y
potencialmente útil a
partir de datos.*



Una definición más

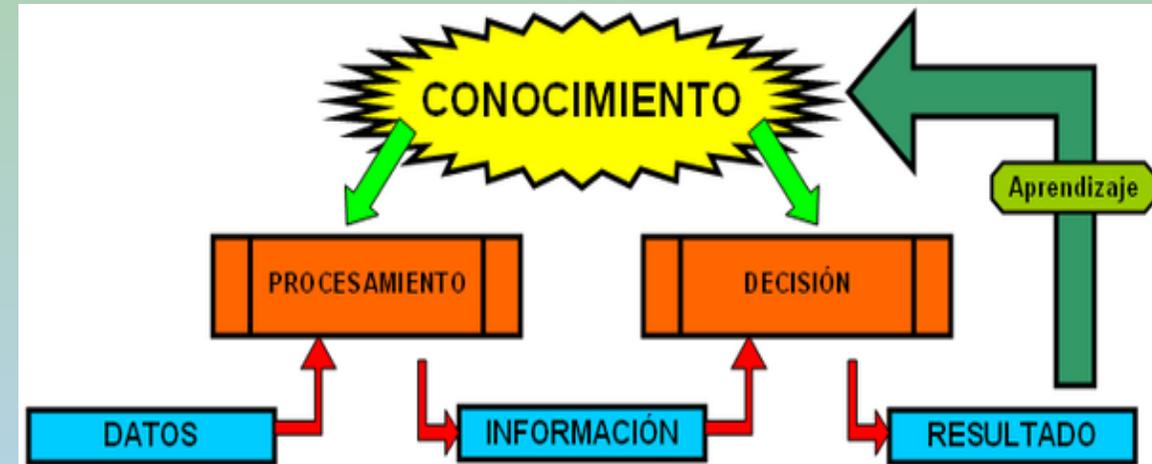
La exploración y el análisis -por medios automáticos o semiautomáticos- de grandes cantidades de datos con el fin de descubrir patrones con significado.



Figura 48. Reconocimiento de patrones

KDD, Minería de datos y Campos Relacionados

- *Extracción de patrones útiles en datos*
 - *Extracción de conocimiento*
 - *Descubrimiento de información*
 - *Procesamiento de patrones de datos*
 - *Data mining*
 - *Estadística*
 - *Análisis de datos*
 - *Sistemas de manejo de información*



Tarea fundamental de la Minería de Datos

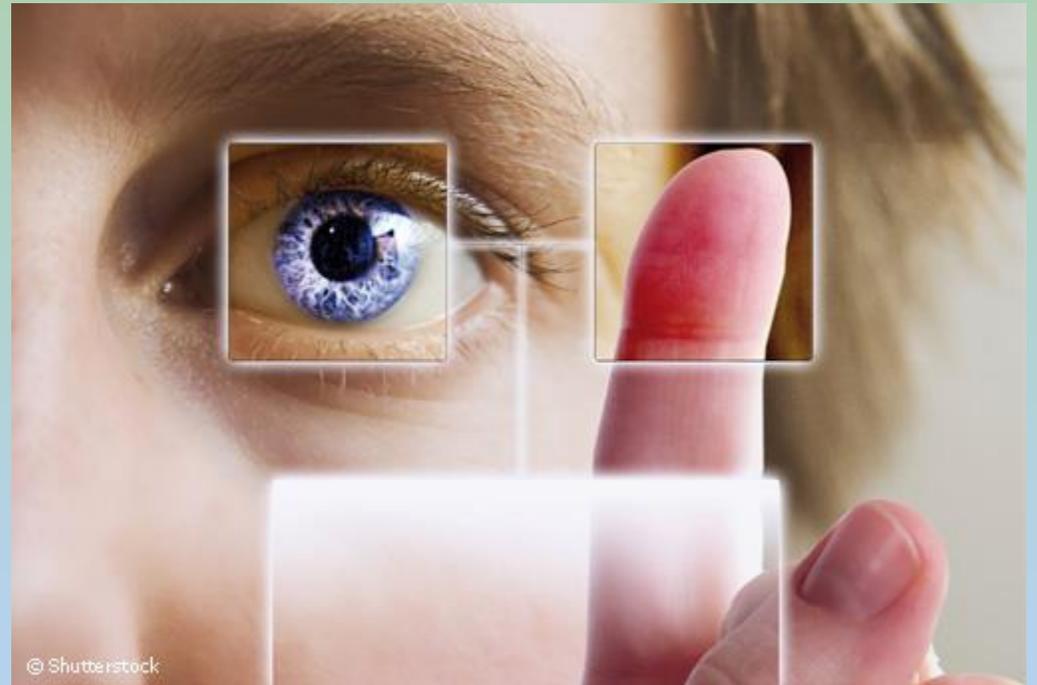
- *Encontrar modelos inteligibles a partir de los datos.*
- *Desarrollar un proceso automático o semiautomático*

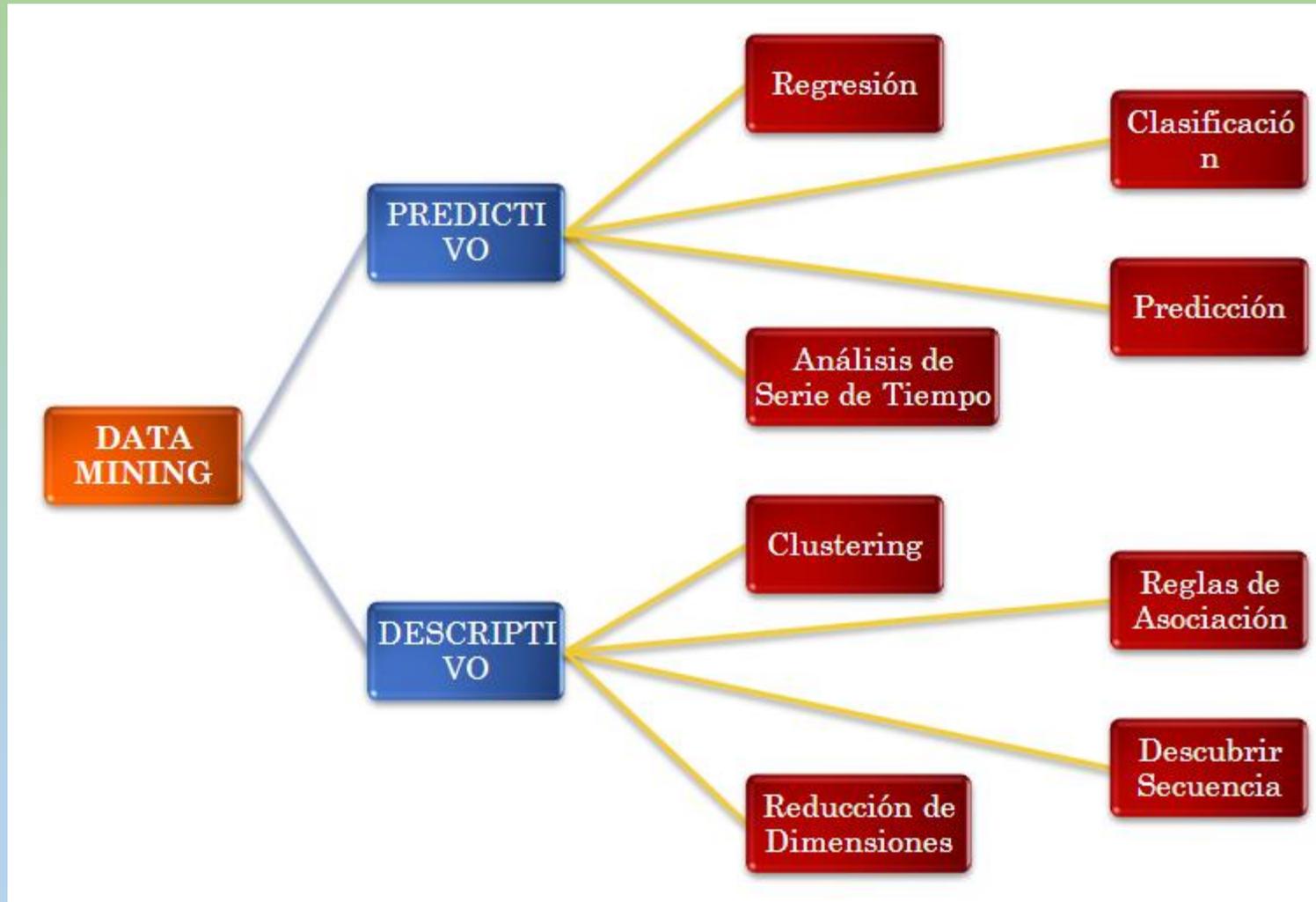
De tal forma que los patrones descubiertos deben ayudar a la toma de decisiones, y por tanto, ser de beneficio a la organización.



KDD y Data Mining (Minería de Datos MD)

- *KDD es todo el proceso de descubrir conocimiento útil a partir de datos*
- *MD es un paso en el proceso*
 - *Aplicación de algoritmos para extraer patrones*
- *Intersección con otros campos*
 - *Inteligencia artificial*
 - *Reconocimiento de patrones*
 - *Bases de datos*
 - *Estadística*
 - *Visualización de datos, etc.*

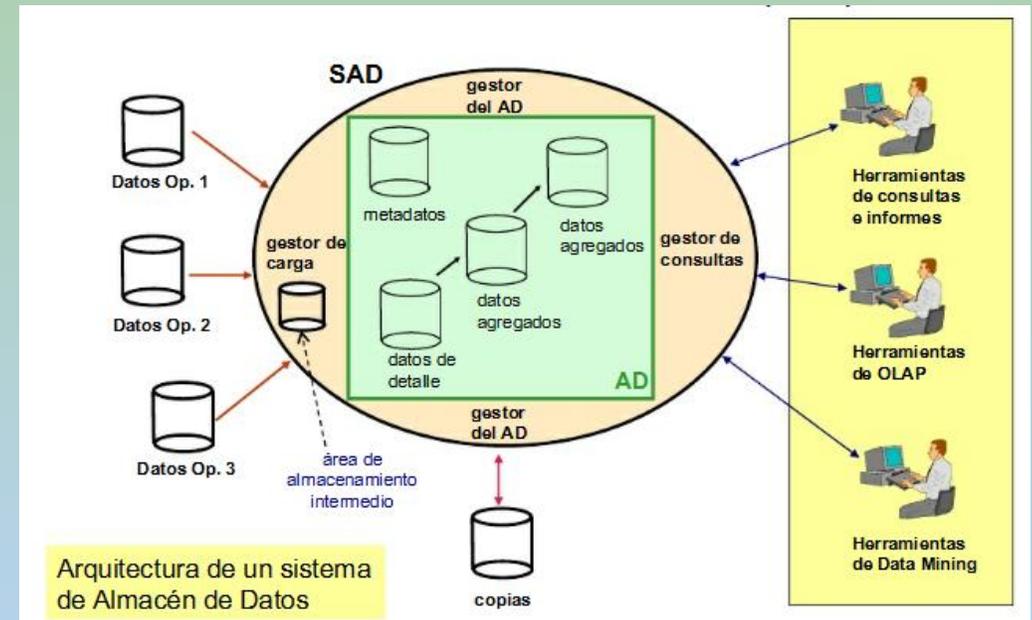




Definiciones Básicas

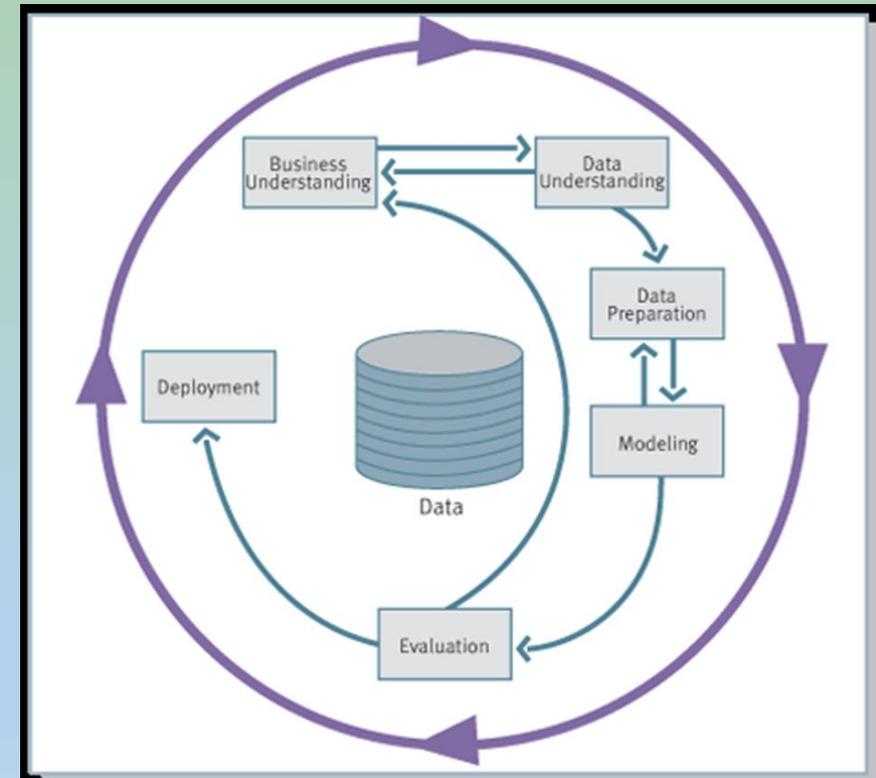
(*Fayyad, Piatetsky-Shapiro y Smith 1996*)

- *Kdd, es el proceso no trivial de identificar patrones en datos que sean válidos, nuevos, potencialmente útiles y entendibles*
- *Datos, es un conjunto de hechos*
- *Patrones, es una expresión en algún lenguaje que describe un subconjunto de datos o un modelo aplicable a ese subconjunto*

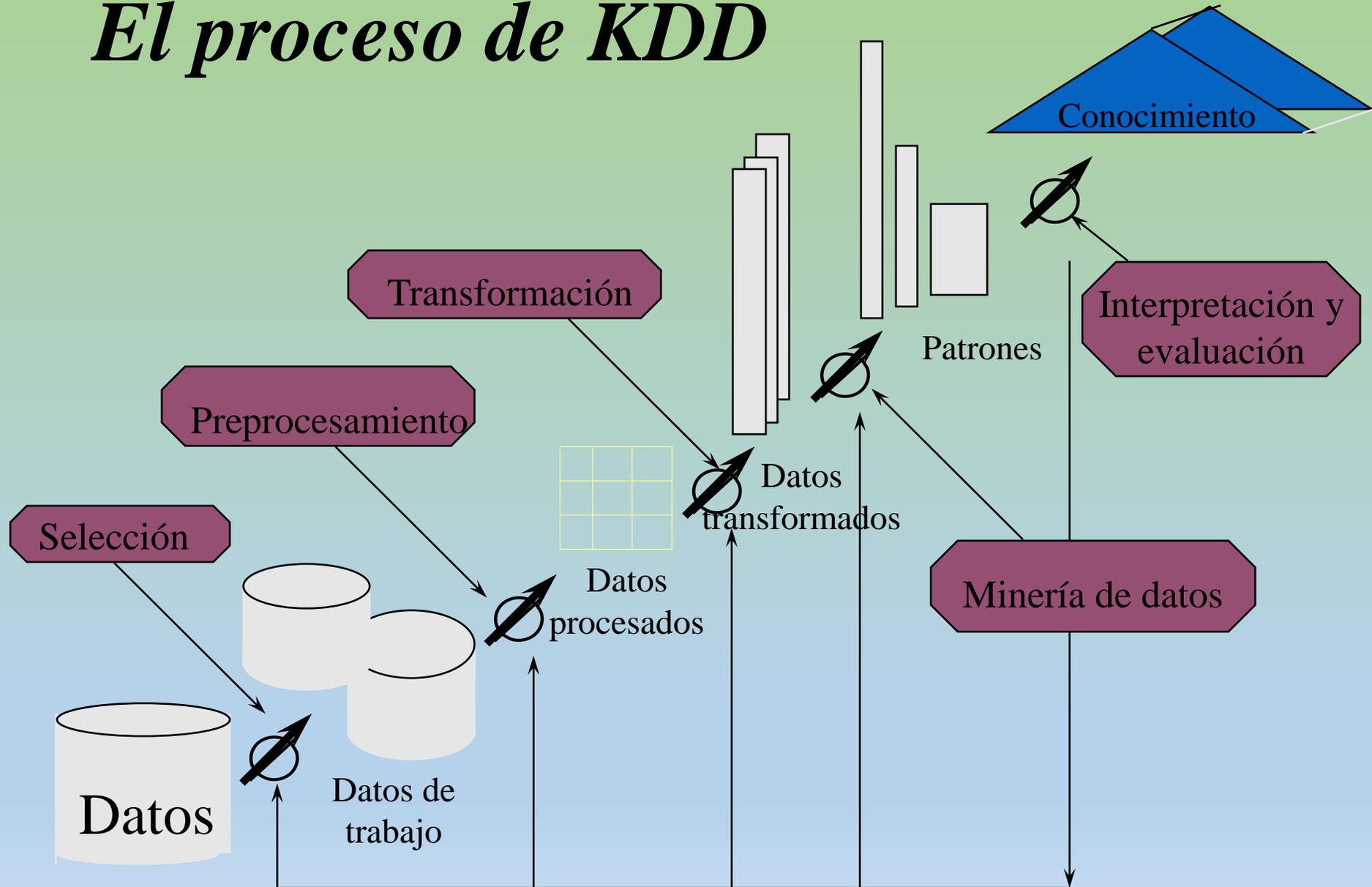


Características del KDD

- *Proceso no trivial*
 - *Preparación de datos*
 - *Búsqueda de patrones*
 - *Evaluación del conocimiento*
 - *Refinamiento*
 - *Involucra búsquedas o inferencias*
- *Los patrones deben ser válidos sobre nuevos datos, con algún grado de certeza*



El proceso de KDD



Arquitectura de los datos



Dr. Francisco J. Mata

2

Minería de Datos

- *Los Mineros, utilizan diferentes técnicas*
- *Procesamiento automático de grandes cantidades de datos*
- *Encontrar relaciones (reglas) que no son visibles utilizando técnicas tradicionales*
- *Reconocimiento de Patrones ofrece alternativas de solución*

Aplicación de KDD

- *En Gobierno*
 - *En el Órgano Interno de Control*
- *En la Ciencia*
- *Empresas Financieras*
- *Manufactura*
- *Telecomunicaciones*
- *Comercio electrónico*



**GOBIERNO
FEDERAL**

Aplicación de KDD

- *Empresas*
 - *Toma de decisiones*
 - *Optimización de recursos*
 - *Herramientas de procesamiento automático*
- *Descubrimiento de Conocimiento en datos (KDD)*
 - *Métodos de minería de datos*
 - *Descubrimiento de patrones*



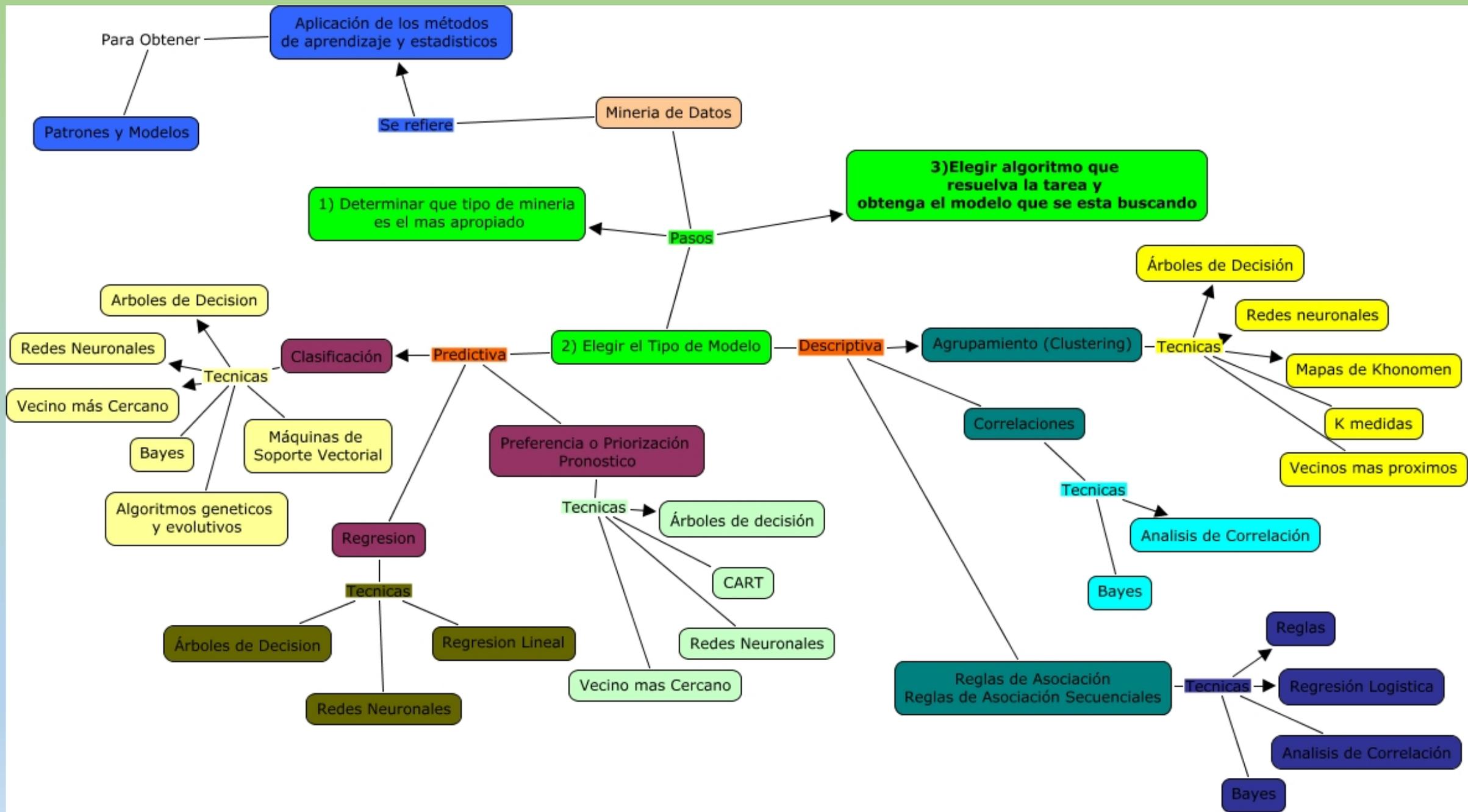
Algunos Métodos Usados en Minería

- *Clasificación supervisada*

- *Árboles de decisión*
- *Regresión*
- *Series temporales*
- *Clasificadores bayesianos*
- *Máquinas de soporte vectorial*

- *Clasificación no supervisada*

- *Segmentación*
- *Agrupamiento*
- *Reglas de asociación*
- *Patrones secuenciales*



Aplicación del Enfoque Lógico Combinatorio del Reconocimiento de Patrones

- *Uso de variables cualitativas y cuantitativas.*
- *Criterios de comparación por rasgo*
- *Funciones de semejanza*

Permiten la generación de reglas basadas en los patrones encontrados en los datos con base en la analogía.

Como recordatorio: En la clasificación supervisada

- *Se conocen las clases presentes en el problema.*
- *Se tienen muestras de objetos de cada clase a partir de las cuales se obtiene un clasificador “óptimo”.*
- *El objetivo es clasificar nuevos objetos con el clasificador obtenido.*



En la clasificación no supervisada

- *Dos subtipos*
 - *Libre (no se sabe el número de grupos).*
 - *Restringida (Este número forma parte del problema)*
- *Se tiene una muestra de objetos*
- *Se obtiene relaciones entre los objetos de tal forma que se obtienen los grupos presentes y los objetos de cada grupo*



En selección de variables

- *Determinación de los factores que inciden en la aparición de un fenómeno*
- *Reducir el número de variables involucradas en un problema*

Problema resuelto parcialmente en problemas de clasificación supervisada

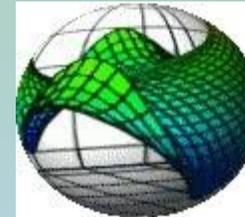
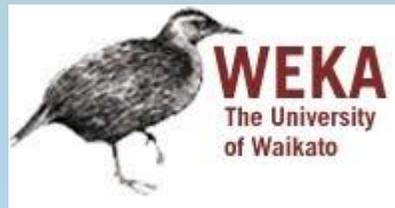


Se recomienda revisar alguno de los siguientes mineros comerciales

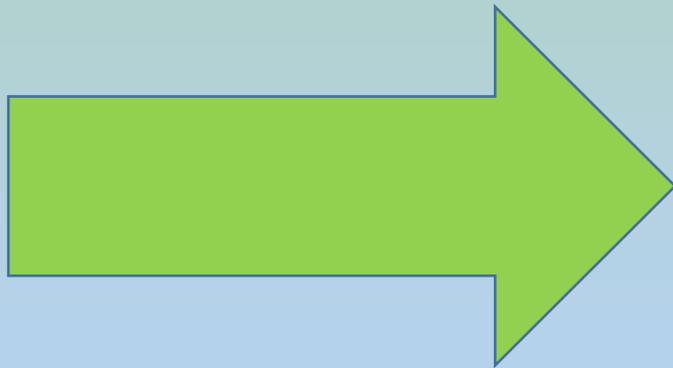
- *SPSS Clementine*
- *SAS Enterprise miner*
- *DataEngine*
- *Oracle Data mining suit*
- *IBM SPSS Mining Suite*
- *Entre otros*



Entre los desarrollados con software libre para minería de datos se pueden consultar:



Una Aplicación en negocios



Inteligencia de negocios



Introducción a Business Intelligence

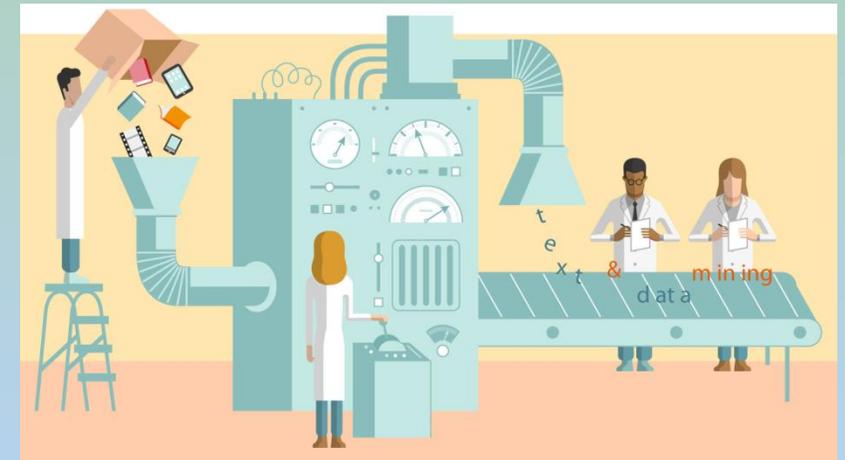


Descripción y pronóstico de patrones de comportamiento

- *Detección de diferencias significativas en los registros de los sistemas automatizados*
- *Análisis dinámico de las bases de datos*
- *Pronóstico de unidades con anomalías*
- *Descripción de unidades con anomalías*
- *Mecanismos para Control Interno*

Algunos retos

- *Tratamiento a grandes Bases de Datos*
- *Gran cantidad de variables*
- *Cómo determinar tamaño de las muestras*
- *Cómo determinar de manera general la validez de resultados de los resultados*
- *Mejorar la calidad de los datos*
- *Interacción con usuarios*
- *Integración con otros sistemas, ...*



Bibliografía

- *Soumen Chakrabarti.*

“Mining the Web : discovering knowledge from hypertext data / Soumen Chakrabarti.”

Morgan Kaufmann Publisher, 2003

- *Mehmed Kantar*

“Data Mining Concepts Models Methods and Algorithms”

Hoboken, N.J. : John Wiley : IEEE Press, 2011

- *Anne Kao and Stephen R. Poteet (eds)*

“Natural language processing and text mining”.

London : Springer, 2010.

Bibliografía

- *Ian H. Witten, Eibe Frank, Mark A. Hall.*
“Data mining: practical machine learning tools and techniques”, 3rd ed.
Amsterdam ; Boston, MA : Morgan Kaufman : Elsevier, c2011.
- *Hernández Orallo José, Ramírez Quintana M. José, Ferri Ramírez César*
“Introducción a la Minería de Datos”
Pearson, 2004
- *Ruiz Schulcloper J./Martínez Trinidad J.,*
“Enfoque Lógico Combinatorio al Reconocimiento de Patrones”, Instituto
Politécnico Nacional, 1999.

Guión Explicativo

- *Este material está desarrollado como apoyo al curso presencial de la unidad de competencia “Aplicaciones” de la UDA “Temas Selectos de Sistemas”, correspondiente a la carrera de Ingeniería en Sistemas y Comunicaciones.*
- *Esta unidad de competencia es la continuación de la unidad de competencia Procesamiento de Lenguaje Natural.*

Guión Explicativo

- *Se recomienda estudiar el tema antes de la sesión presencial.*
- *Se recomienda seguir la secuencia en la que se presenta el material para mayor comprensión de los temas.*
- *Los temas se pueden consultar en extenso en las referencias proporcionadas en la bibliografía.*

Guión Explicativo

- *Se recomienda realizar ejercicios adicionales para complementar lo visto en clase y el material presentado.*
- *Al final de la unidad, se expondrá un artículo científico de reciente publicación donde se ilustre el desarrollo de alguna técnica o su aplicación.*