



La minería de datos usa los siguientes algoritmos durante el estudio de los datos:

Algoritmos supervisados

Predicen el comportamiento de un conjunto de datos. Se clasifican en:

Árboles de decisión

Evalúan varias condiciones para pronosticar el comportamiento de los datos.

Ejemplo de uso:

La selección de los posibles compradores de un producto a partir del análisis de las compras que han realizado en cierta tienda.

ALGORITMO Partición (N :nodo, E :cjto. de ejemplos)

SI todos los ejemplos E son de la misma clase c **ENTONCES**

Asignar la clase c al nodo N

SALIR; //Esta rama es pura, ya no hay que seguir partiendo. N es hoja.

SI NO:

Particiones := generar posibles particiones.

MejorPartición := seleccionar la mejor partición según el criterio de la partición.

PARA CADA condición i de la partición elegida.

Añadir un nodo hijo i a N y asignar los ejemplos consistentes a cada hijo (E_i)

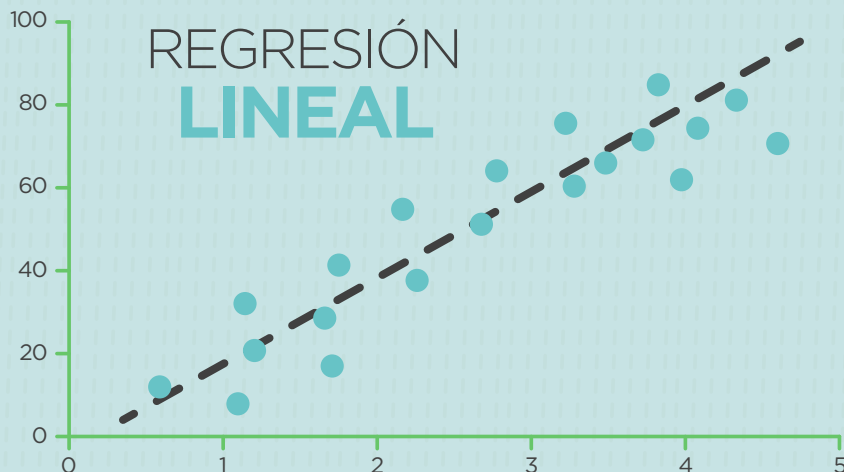
Partición(i , E_i). // Realizar el mismo procedimiento global con cada.

FIN PARA

FIN SI

FIN ALGORITMO

Para clasificar un cjto de ejemplos E , se invoca con la llamada Partición (R, E), donde R es un nodo raíz de un árbol por empezar.



Regresión

Relaciona las variables que pueden tomar más de un valor en un rango para predecir el siguiente elemento de un conjunto de datos.

Ejemplo de uso:

El pronóstico del alza o baja del dólar en un periodo corto.



Algoritmos no supervisados

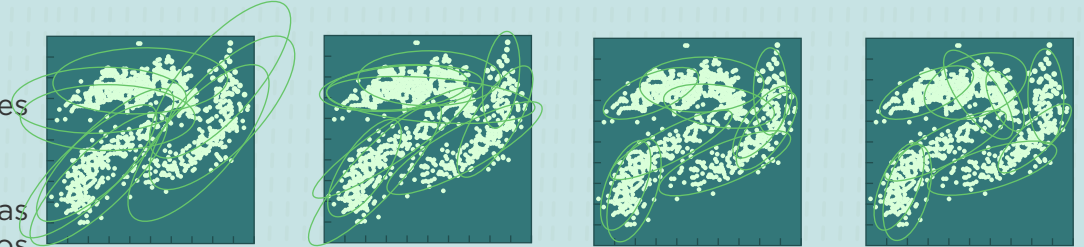
Descubren patrones y tendencias con un conjunto de datos actual. Se clasifican en:

Segmentación

Dividen los datos en grupos que poseen propiedades similares para después analizarlos.

Ejemplo de uso:

La clasificación de usuarios que realizan algunas instituciones educativas para ofrecerles cursos o diplomados afines a su perfil profesional.

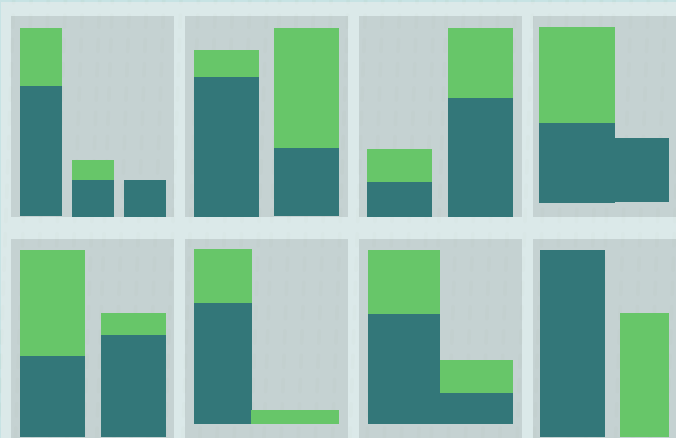


Asociación

Determinan las correlaciones que existen entre un conjunto de datos.

Ejemplo de uso:

Las sugerencias de compra que una tienda en línea realiza a sus usuarios y que tienen una relación con los productos que le interesan en ese momento.



Análisis de secuencias

Obtiene la frecuencia con la que se repite un elemento en un conjunto de datos.

Ejemplo de uso:

La obtención de los productos más vendidos en cierto mes en una tienda departamental.

