



Centro de
INNOVACIÓN y DESARROLLO
Tecnológico en Cómputo



Reconocimiento de Patrones

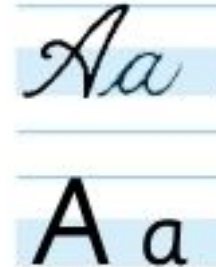
Lección 1: Introducción (parte 1)

El Reconocimiento de Patrones (RP) es una tarea fundamental realizada de forma cotidiana por todos los seres vivos.

- Identificar fuentes de alimento, depredadores o parejas
- Reconocer rostros y voces
- Diferenciar comida en mal estado



En psicología se le nombra al fenómeno de que un conjunto de estímulos son asociados a un mismo patrón.



a a A

Reconocemos la letra “a”, sea mayúscula, minúscula, cursiva, escrita a mano (por uno u otras personas).

Psicología Cognitiva

El reconocimiento de patrones es un procesamiento de información que consiste en codificar el estímulo de alguna manera y compararlo con un patrón ya existente en la memoria.

El mayor reto:

¿Cómo se codifica este estímulo?



Computación

“Pattern Recognition is the assignment of visual or logical patterns to classes based on the features of these patterns and their relationships” [1].

“Pattern recognition — the act of taking in raw data and taking an action based on the category of the pattern” [2].

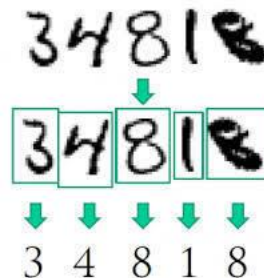
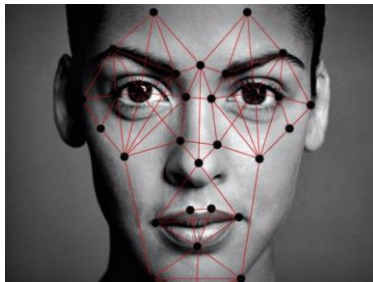
“Pattern Recognition is the scientific discipline dealing with methods for object description and classification” [3]

“Pattern recognition is about assigning labels to objects” [4].



Reconocimiento de Patrones (Computación)

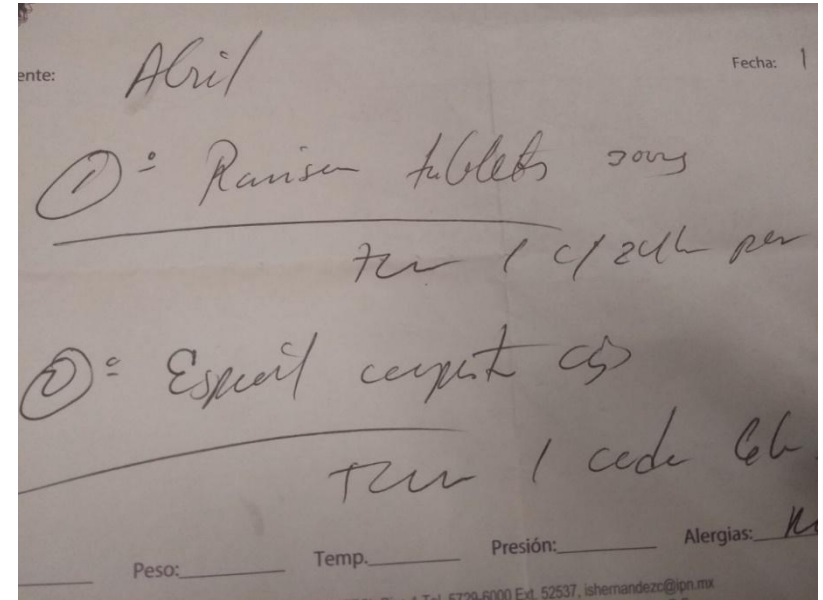
Resulta natural, además de muy útil, el diseñar sistemas que nos permitan emular esta capacidad de reconocer patrones.

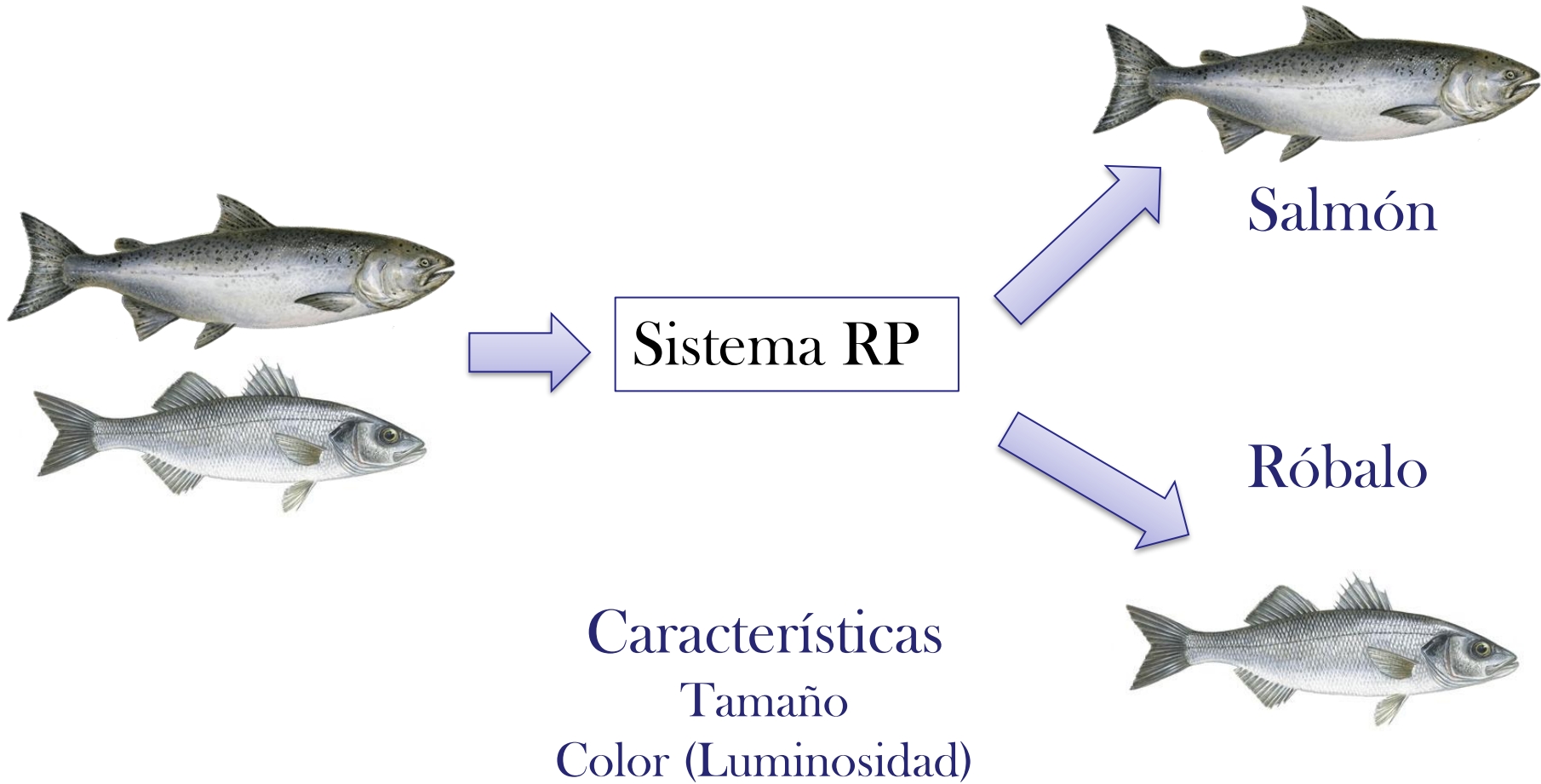


Una computadora podría realizar estas tareas de forma más eficiente.

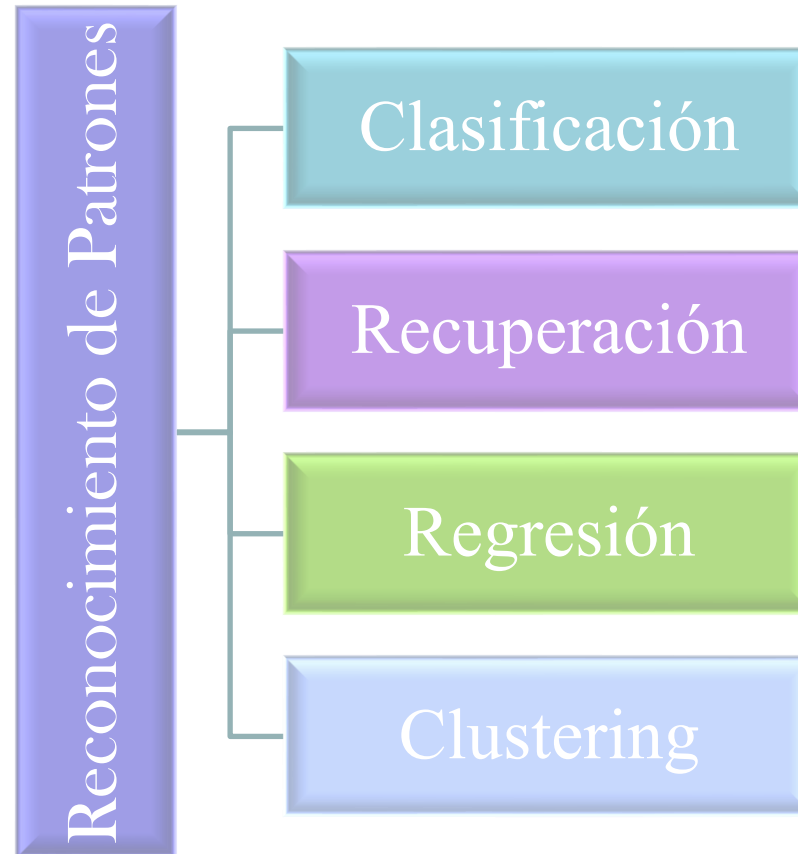
Reconocimiento de Patrones (Desafíos)

- ✦ La naturaleza intrínseca del fenómeno analizado.
- ✦ La calidad de los datos.
- ✦ Diseño del conjunto de datos.
- ✦ Generalización.
- ✦ Evaluación

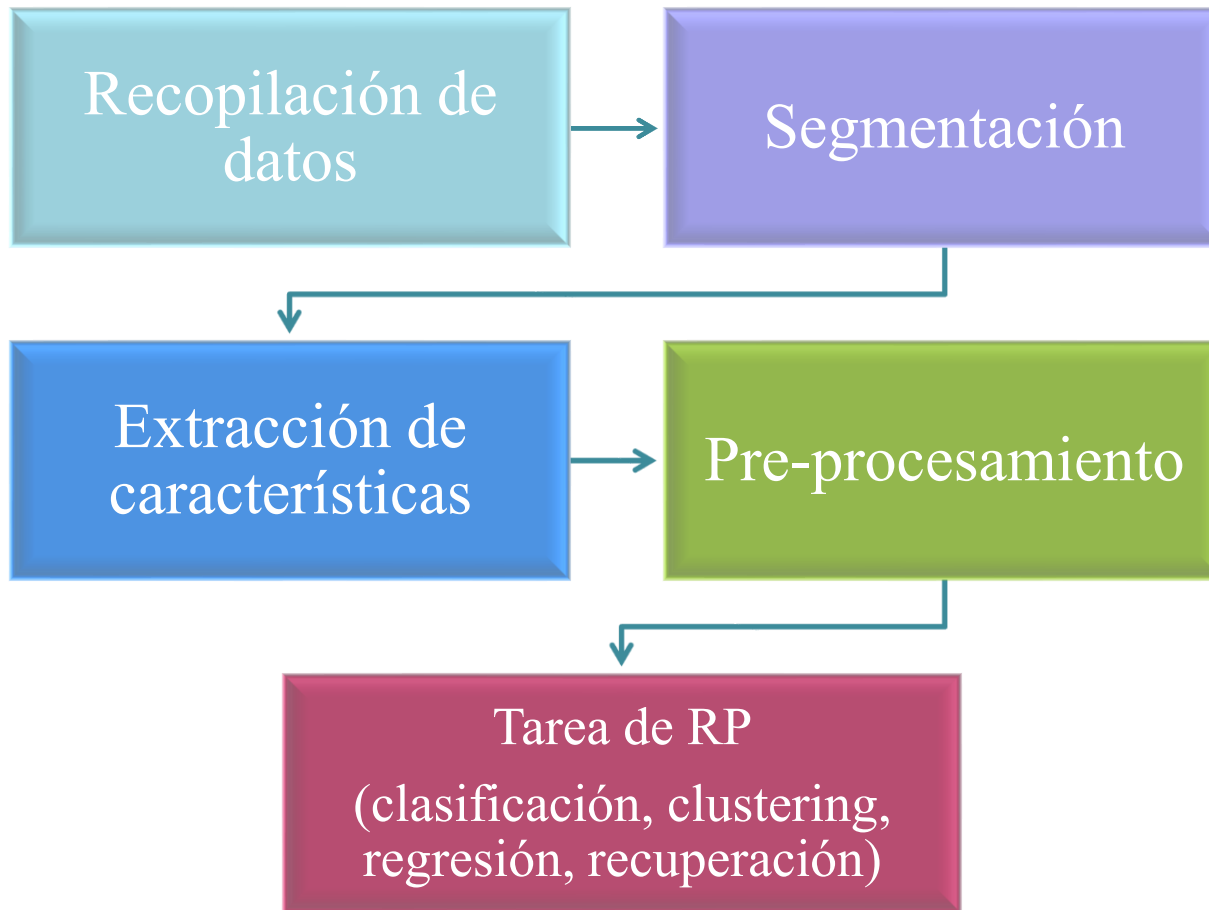


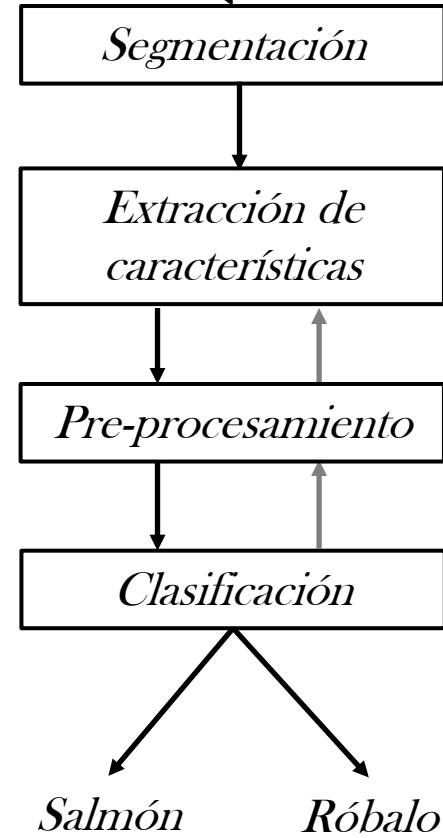


Tareas del Aprendizaje supervisado



Etapas en un sistema de RP





Ciclo de Diseño



Conceptos Básicos

Patrón

RAE: Defensor, protector; santo titular de una iglesia; persona que emplea trabajadores; modelo que sirve de muestra para sacar otra cosa igual ...

Wikipedia: sucesos u objetos recurrentes, estos elementos se repiten de una manera predecible; conjunto de reglas que pueden ser usadas para crear o generar entidades; conjunto de rasgos esenciales en un diseño gráfico, mapa o escrito.

Conceptos Básicos

Patrón

“Patterns are *physical* representations of the objects. Usually signals, images or simple tables of values. Often, we will refer to patterns as objects, cases or samples” [3].

“Instance is an individual, independent example of the concept to be learned. Instances are characterized by the values of a set of predetermined attributes” [5].

“A pattern is a physical object or an abstract notion” [6].



Conceptos Básicos

Patrón

Patrón: un conjunto de características que describen un ente (real o abstracto) que pertenece a una categoría (clase). Estos patrones pueden estar determinados por factores deterministas o estocásticos.

Sinónimos: instancia, ejemplo, tupla, vector característico.

Conceptos Básicos

Rasgos: valores que describen las características de un ente (real o abstracto), es decir los elementos que conforman un patrón. Estos pueden ser numéricos (discretos o continuos) o categóricos (genero, estado civil, entre otros).

Sinónimos: atributo, propiedad, variable, características.

Espacio de características: espacio d dimensional, donde d es el número de características.

Conceptos Básicos

Clases: son categorías de objetos asociados con conceptos o prototipos [3].

Objetos de una misma clase son similares entre ellos, mientras que objetos de clases diferentes no muestran similitud.

Ejemplo:

Clases de fruta: naranja, limón, manzana, pera, fresa.

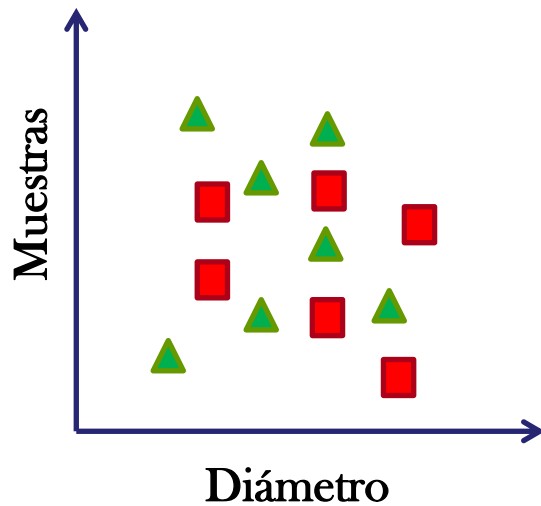
Atributos: color, forma, peso, tamaño.

Asignar una fruta a una clase con base en sus atributos.

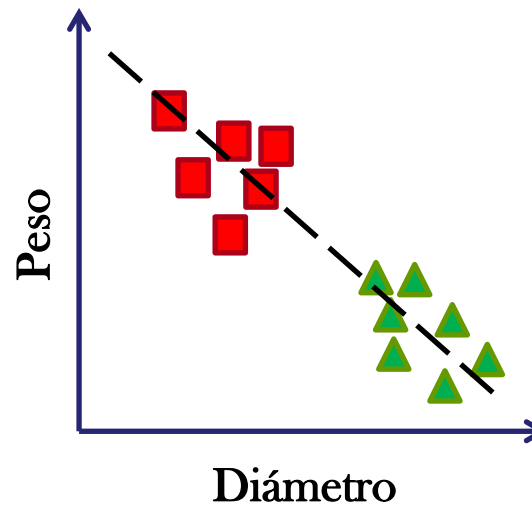
Conceptos Básicos

Importancia de los atributos

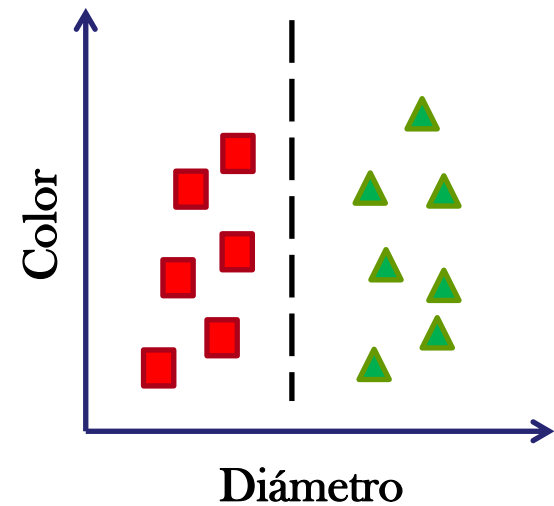
¿Qué atributos son más útiles? (Criterio de eficacia)



No es suficiente



Correlacionadas



Color: NO
Diámetro: SI

Conceptos Básicos

Importancia de los atributos

Los rasgos que nos permiten diferenciar un patrón (objeto) de otro.

Una buena selección de atributos nos permite:

- ❖ Enfocarnos en información relevante.
- ❖ Reducción de datos
- ❖ Mejoras en los resultados.

Generalmente no sabemos: cual seleccionar, que representan y como ajustarlos.

Los esquemas de RP tratan de aprovechar al máximo las características disponibles.



Introducción

Conceptos Básicos

Banco de Datos

Iris Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Famous database; from Fisher, 1936



Data Set Characteristics:	Multivariate	Number of Instances:	150	Area:	Life
Attribute Characteristics:	Real	Number of Attributes:	4	Date Donated	1988-07-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	449872

Conceptos Básicos

Banco de Datos

Un banco de datos esta generalmente organizado como una matriz de N filas (patrones) y n columnas (rasgos), con una columna extra para la clase.

	Largo del Sépalo (n_1)	Ancho del Sépalo (n_2)	Largo del Pétalo (n_3)	Ancho del Pétalo (n_4)	Clase (n_5)
N_1	5.9	3.0	4.2	1.5	Iris-versicolor
N_2	6.9	3.1	4.9	1.5	Iris-versicolor
N_3	4.6	3.2	1.4	0.2	Iris-setosa
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
N_{150}	7.7	3.8	6.7	2.2	Iris-virginica

Introducción

Conceptos Básicos

Banco de Datos

Wine Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: Using chemical analysis determine the origin of wines



Data Set Characteristics:	Multivariate	Number of Instances:	178	Area:	Physical
Attribute Characteristics:	Integer, Real	Number of Attributes:	13	Date Donated	1991-07-01
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	355623

Conceptos Básicos

Banco de Datos

El banco de datos Wine esta conformado por los siguientes atributos:

- | | |
|-----------------------|-----------------------------------|
| 1 – Alcohol | 8 – Nonflavanoid phenols |
| 2 – Malic acid | 9 – Proanthocyanins |
| 3 – Ash | 10 – Color intensity |
| 4 – Alcalinity of ash | 11 – Hue |
| 5 – Magnesium | 12 – OD280/OD315 of diluted wines |
| 6 – Total phenols | 13 – Proline |
| 7 – Flavanoids | |

Introducción

Conceptos Básicos

Banco de Datos

1,14.23,1.71,2.43,15.6,127,2.8,3.06,.28,2.29,5.64,1.04,3.92,1065
 1,13.2,1.78,2.14,11.2,100,2.65,2.76,.26,1.28,4.38,1.05,3.4,1050
 1,13.16,2.36,2.67,18.6,101,2.8,3.24,.3,2.81,5.68,1.03,3.17,1185
 1,14.37,1.95,2.5,16.8,113,3.85,3.49,.24,2.18,7.8,.86,3.45,1480
 1,13.24,2.59,2.87,21,118,2.8,2.69,.39,1.82,4.32,1.04,2.93,735
 1,14.2,1.76,2.45,15.2,112,3.27,3.39,.34,1.97,6.75,1.05,2.85,1450
 1,14.39,1.87,2.45,14.6,96,2.5,2.52,.3,1.98,5.25,1.02,3.58,1290
 1,14.06,2.15,2.61,17.6,121,2.6,2.51,.31,1.25,5.05,1.06,3.58,1295
 1,14.83,1.64,2.17,14,97,2.8,2.98,.29,1.98,5.2,1.08,2.85,1045
 1,13.86,1.35,2.27,16,98,2.98,3.15,.22,1.85,7.22,1.01,3.55,1045
 1,14.1,2.16,2.3,18,105,2.95,3.32,.22,2.38,5.75,1.25,3.17,1510
 1,14.12,1.48,2.32,16.8,95,2.2,2.43,.26,1.57,5,1.17,2.82,1280
 1,13.75,1.73,2.41,16,89,2.6,2.76,.29,1.81,5.6,1.15,2.9,1320
 1,14.75,1.73,2.39,11.4,91,3.1,3.69,.43,2.81,5.4,1.25,2.73,1150
 1,14.38,1.87,2.38,12,102,3.3,3.64,.29,2.96,7.5,1.2,3,1547
 1,13.63,1.81,2.7,17.2,112,2.85,2.91,.3,1.46,7.3,1.28,2.88,1310
 1,14.3,1.92,2.72,20,120,2.8,3.14,.33,1.97,6.2,1.07,2.65,1280
 1,13.83,1.57,2.62,20,115,2.95,3.4,.4,1.72,6.6,1.13,2.57,1130
 1,14.19,1.59,2.48,16.5,108,3.3,3.93,.32,1.86,8.7,1.23,2.82,1680

¿Es posible entender estos datos?

Repositorios de Bancos de Datos

UCI (Machine Learning Repository)

University of California, Irvine

<http://archive.ics.uci.edu/ml/>

KEEL (Knowledge Extraction based on Evolutionary Learning)

Universidad de Granada

<http://sci2s.ugr.es/keel/datasets.php>

Kaggle

Google LLC

<https://www.kaggle.com/datasets>



Clasificación: Determinar la clase o grupo al cual pertenece una cosa

Un clasificador sencillo: permite clasificar automáticamente naranjas y limones.



¿Que necesitamos?

Un criterio de decisión: Tamaño
 Color
 Textura

Clasificación

Con medir el diámetro, es posible decidir si es una naranja o un limón.



$$x^1 = 3.9$$



$$x^2 = 4.1$$



$$x^3 = 4.4$$



$$x^4 = 4.7$$



$$x^5 = 4.5$$



$$x^6 = 6.1$$



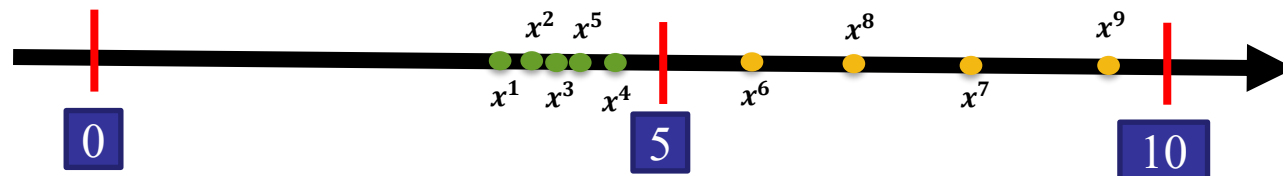
$$x^7 = 8.5$$



$$x^8 = 7.1$$



$$x^9 = 9.5$$



Introducción

Clasificación

Con más información del objeto, la clasificación puede mejorar.

$$x^1 = \begin{bmatrix} 3.9 \\ (103,158,42) \end{bmatrix}$$



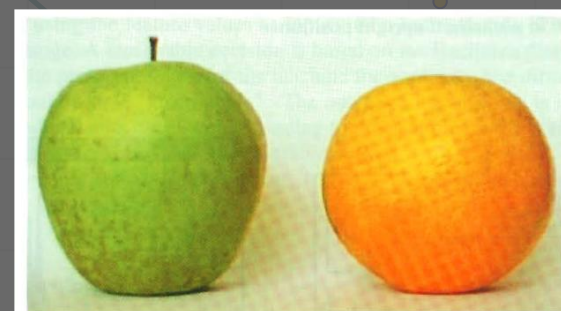
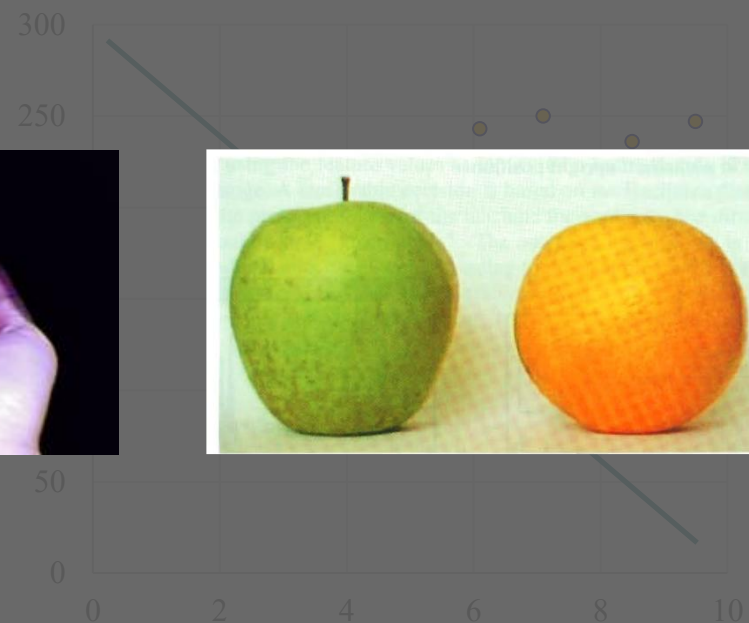
$$x^6 = \begin{bmatrix} 6.1 \\ (243,184,17) \end{bmatrix}$$



$$x^4 = \begin{bmatrix} \\ (110,150,48) \end{bmatrix}$$

$$x^9 = \begin{bmatrix} \\ (247,188,22) \end{bmatrix}$$

$$x^5 = \begin{bmatrix} 4.5 \\ (115,151,38) \end{bmatrix}$$

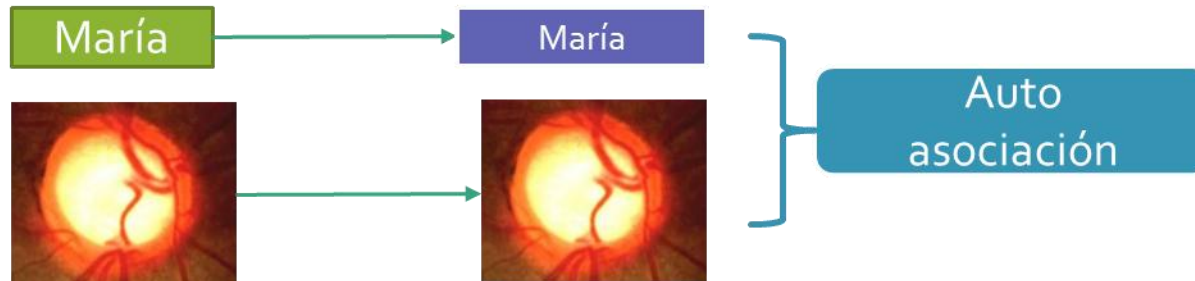
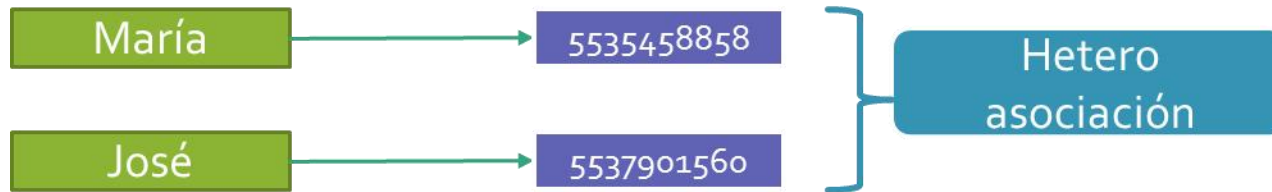


Diámetro

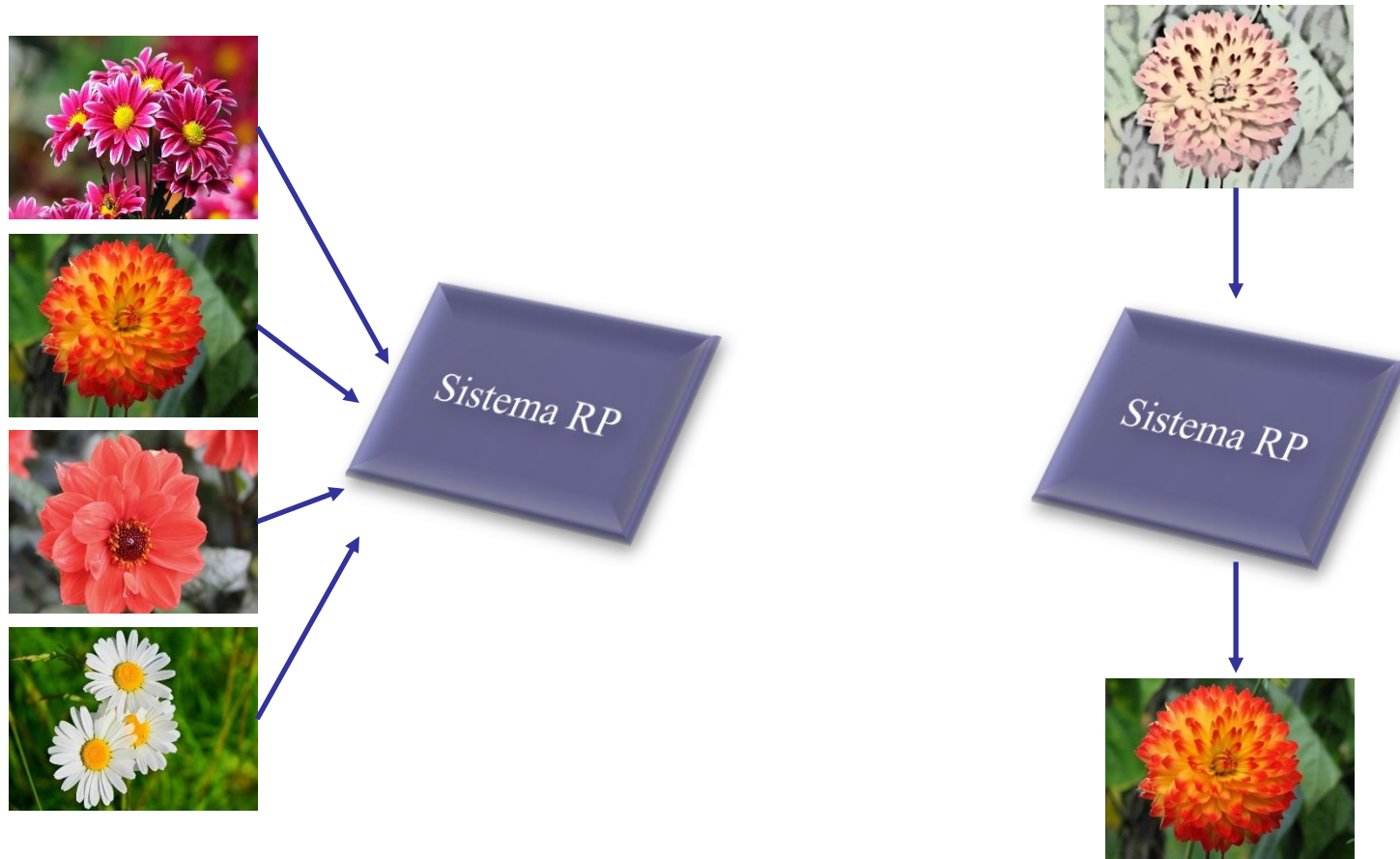
Clasificación



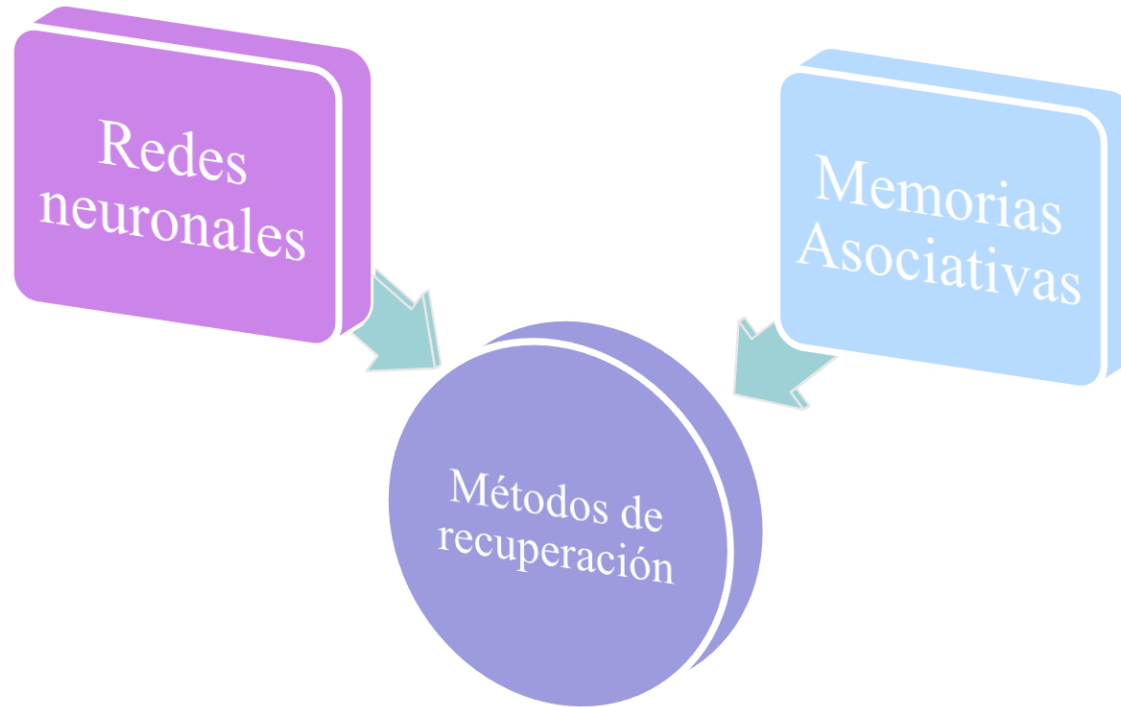
Recuperación: recuperar correctamente patrones completos a partir de patrones de entrada, los cuales pueden estar alterados.



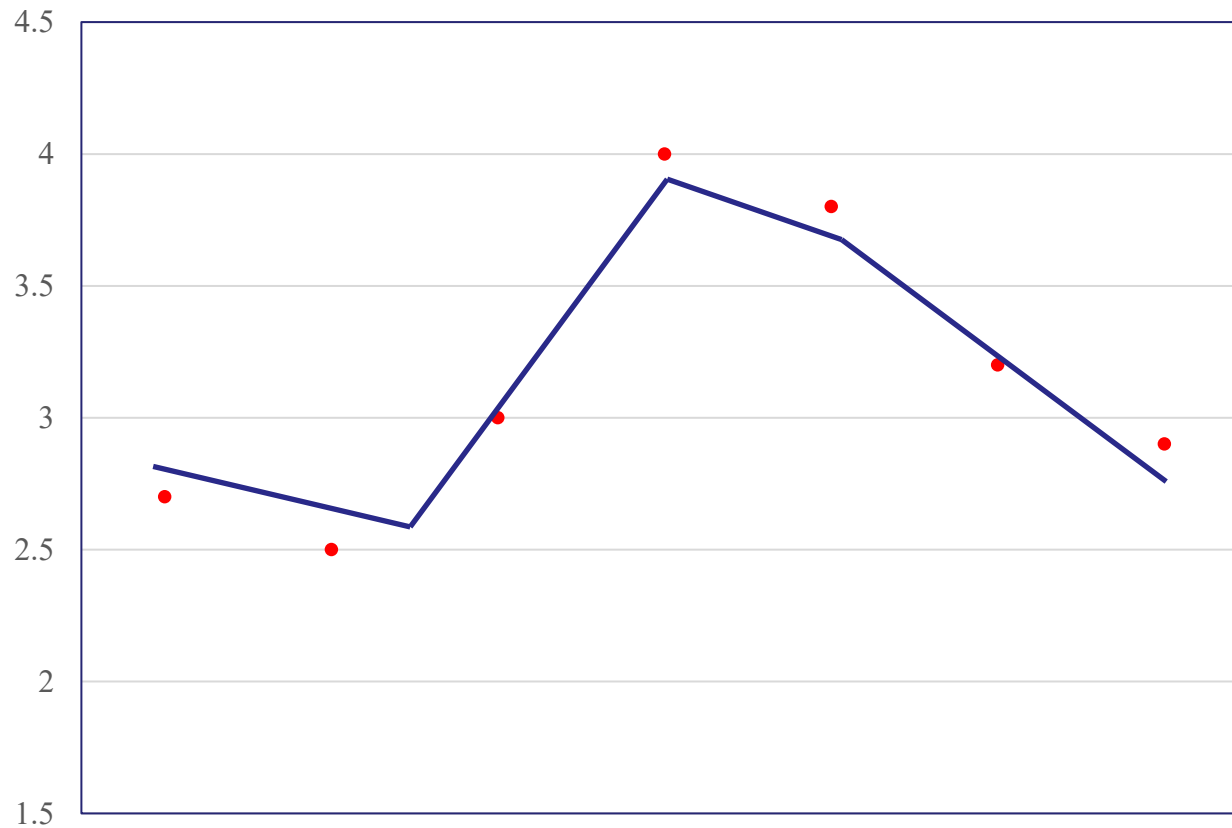
Recuperación



Recuperación

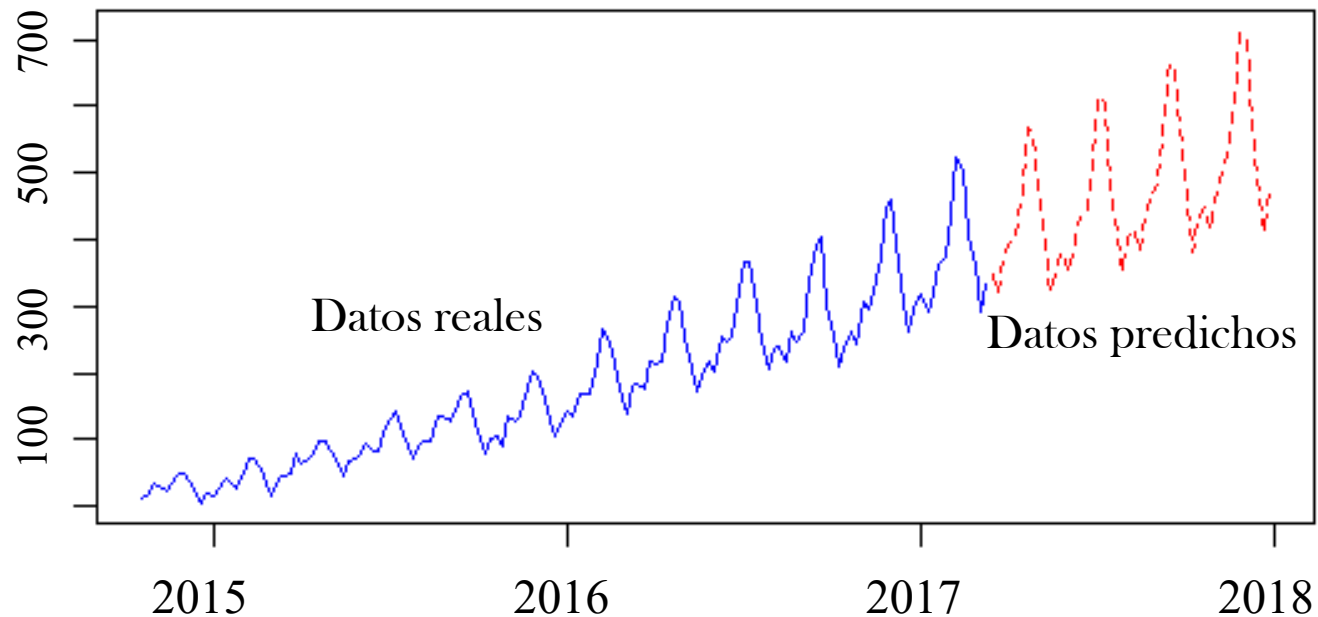


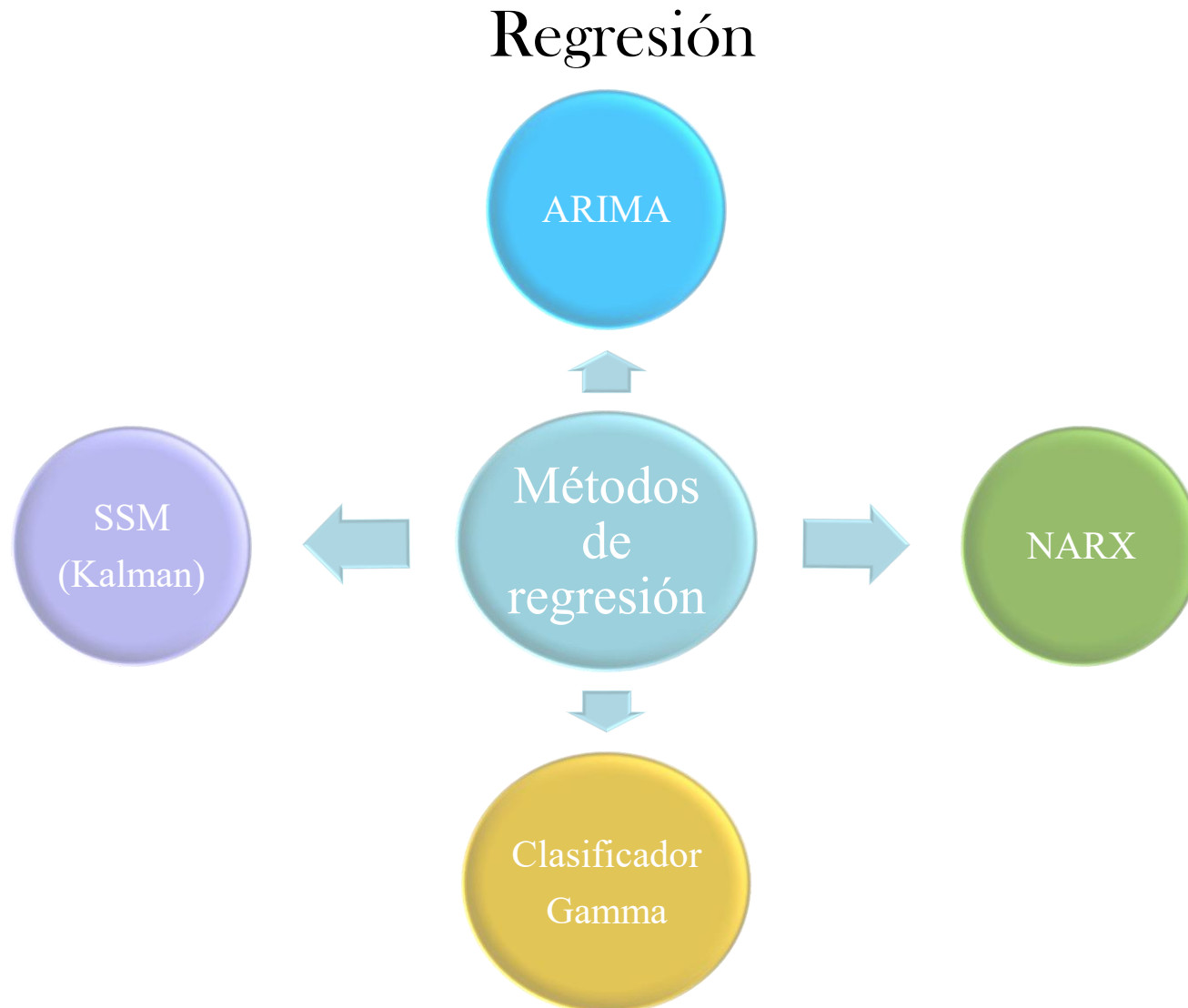
Regresión: esta tarea se relaciona principalmente con el ajuste de funciones.



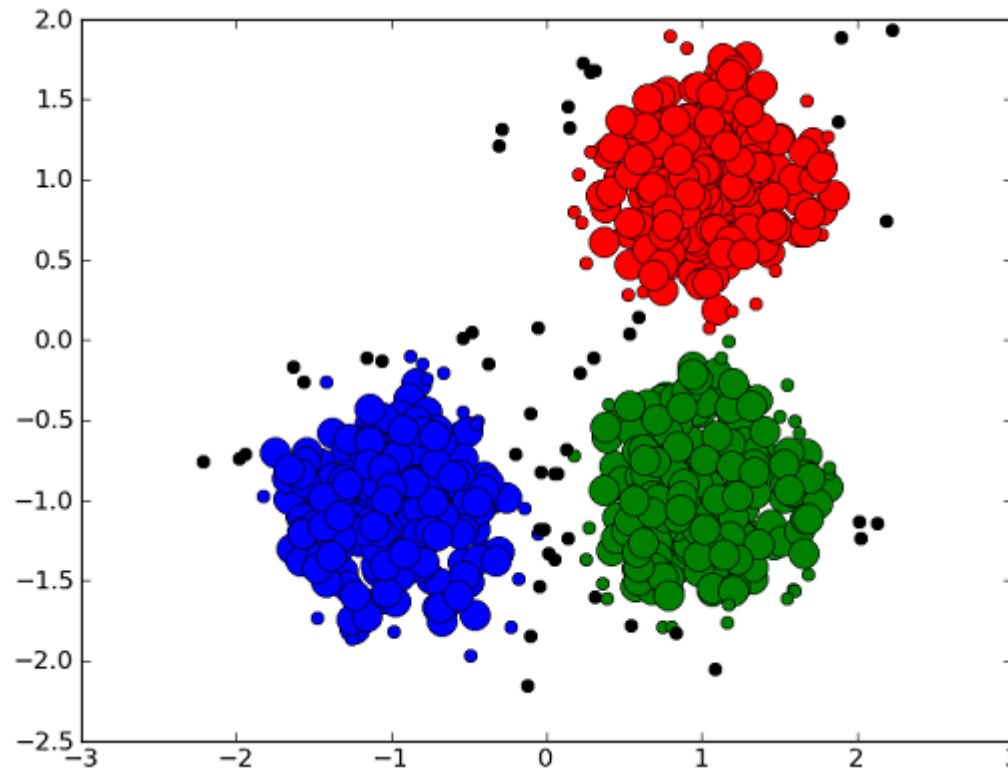
Regresión

Predicción de series de tiempo

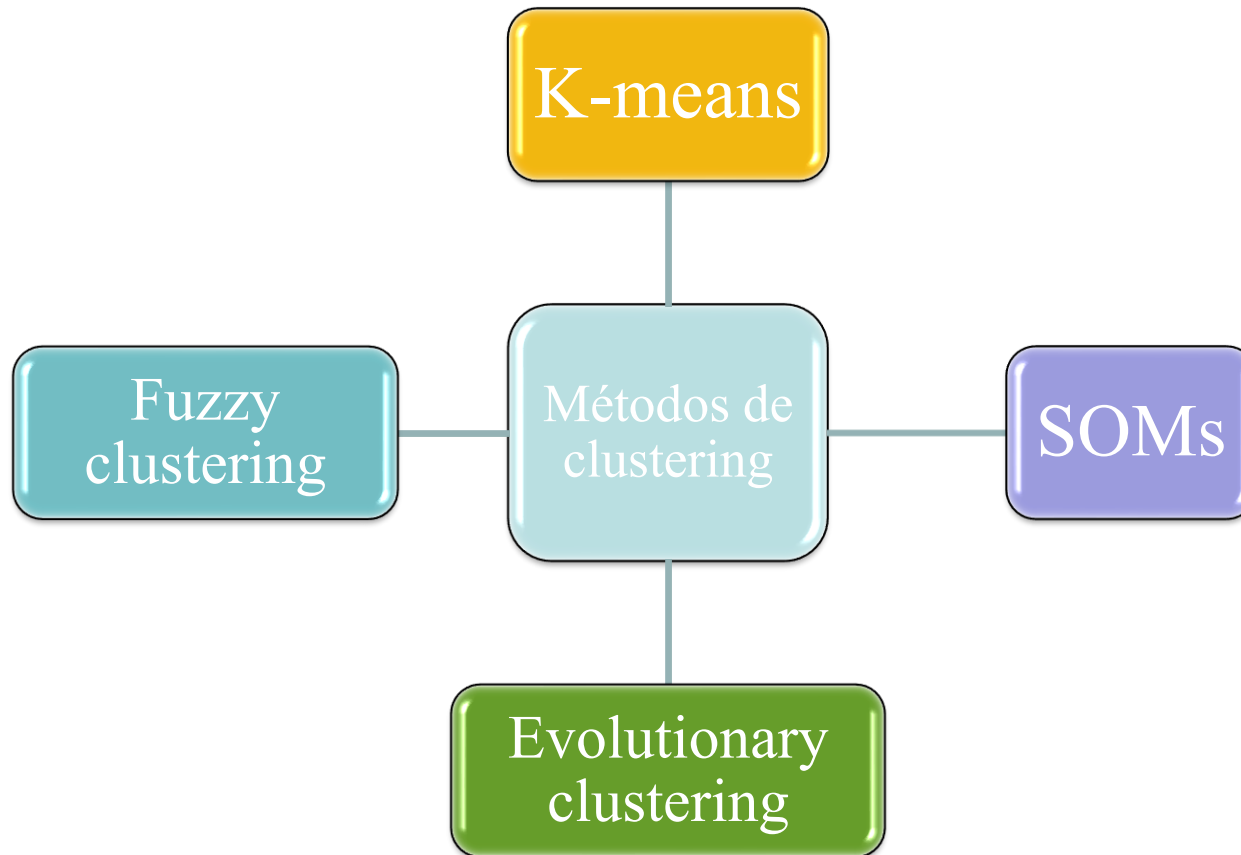




Clustering: crear agrupamiento de forma que las instancias de un mismo cluster son similares y los miembros de diferentes clusters no muestren similitud (aprendizaje no supervisado).



Clustering



Tarea 1

1. Revisar algún repositorio de bancos de datos y seleccionar 10.
2. Proporcionar la siguiente información para cada banco de datos:

✓Nombre del banco de datos

✓Particularidades:

- Número de instancias
- Número de atributos
- Tipo de los atributos
- Número de clases

- [1] **Leondes, C.T. (2018).** *Image Processing and Pattern Recognition*. California: Academic Press.
- [2] **Duda, R.O., Hart, P.E. & Stork, D.G. (2001).** *Pattern Classification*. 2nd edition. Wiley-Interscience.
- [3] **Marques de Sá, J:P. (2001).** *Pattern Recognition: Concepts, Methods and Applications*. Berlin: Springer-Verlag.
- [4] **Kuncheva, L. (2014).** *Combining Pattern Classifiers: Methods and Algorithms*. 2nd edition. USA: Wiley.
- [5] **Witten, I.H., Frank, E. & Hall, M.A. (2011).** *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd edition. USA: Elsevier.
- [6] **Murty, N.M. & Devi, V.S. (2011).** *Pattern Recognition: An Algorithmic Approach*. Springer.





¡ Gracias !

Thanks !

Obrigado

Xie xie ni

Domo arigatou

Спасибо

Merci

Grazie

Alfa Beta