

# UEA Datos a Gran Escala

## Métodos y Tecnologías para el Análisis (Modelado) de los Datos a Gran Escala



Dr. Pedro Pablo González Pérez

e-mail: [pgonzalez@correo.cua.uam.mx](mailto:pgonzalez@correo.cua.uam.mx)

<http://dcni.cua.uam.mx/division/usuario?p=31#>

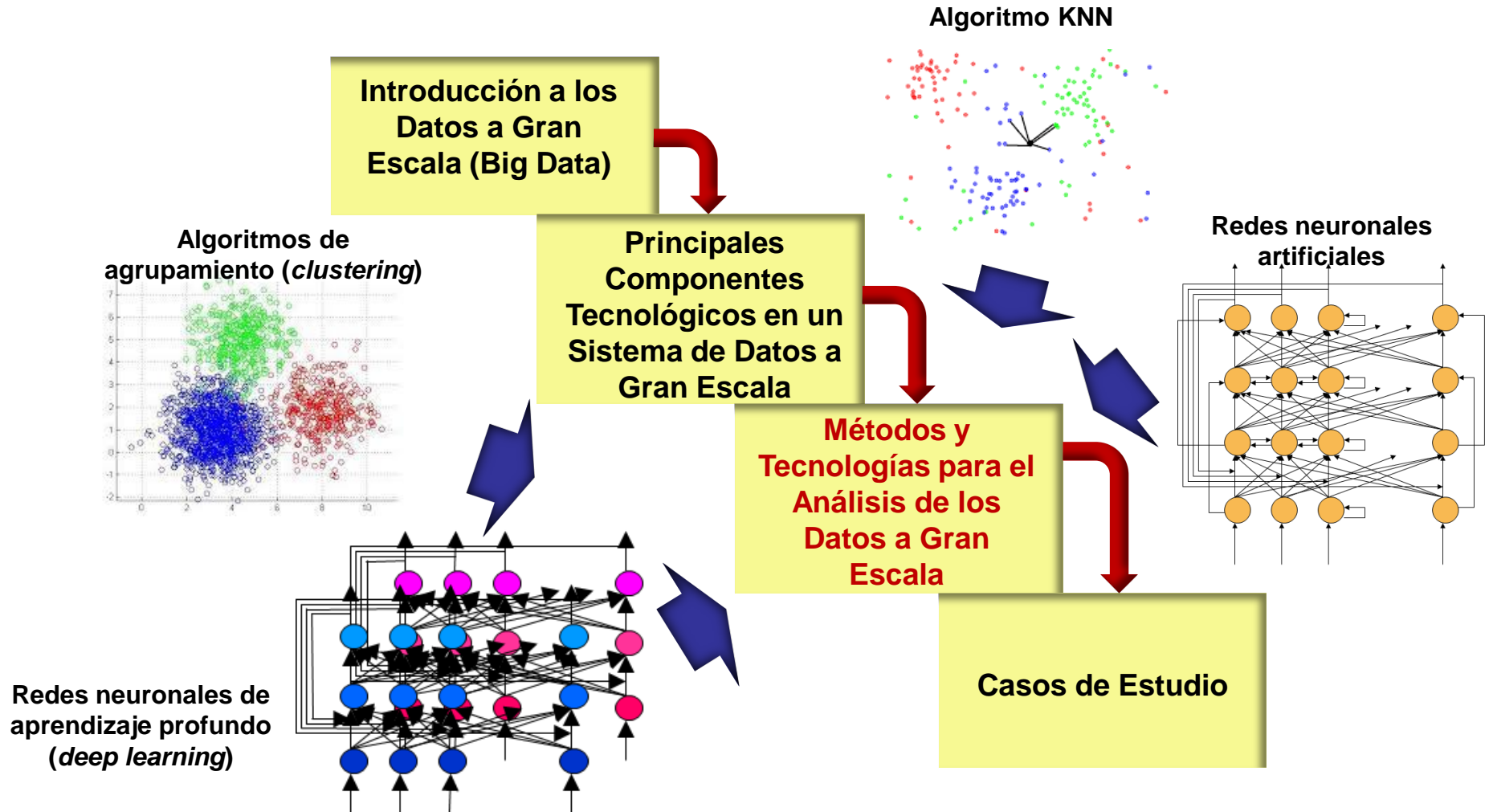
Departamento de Matemáticas Aplicadas y Sistemas



UNIVERSIDAD  
AUTÓNOMA  
METROPOLITANA  
Unidad Cuajimalpa

# UEA Datos a Gran Escala

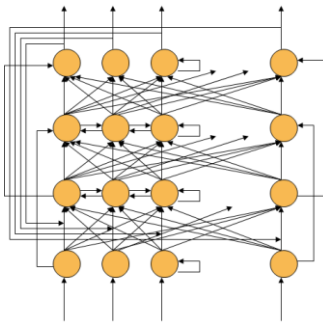
## Métodos y Tecnologías para el Análisis (Modelado) de los Datos a Gran Escala



# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

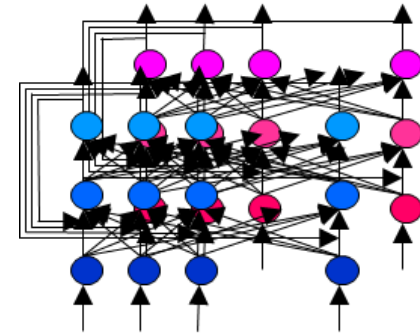
Modelos de Redes Neuronales Artificiales



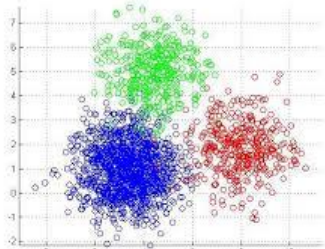
Análisis de Datos a Gran Escala (Big Data)



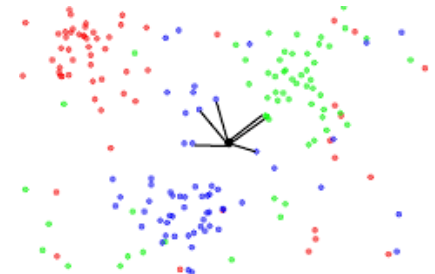
Modelos de Aprendizaje Profundo (Deep Learning)



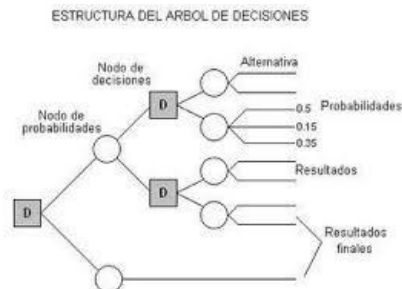
Algoritmos de Agrupamiento (Clustering)



Algoritmos KNN



Árboles de Decisión



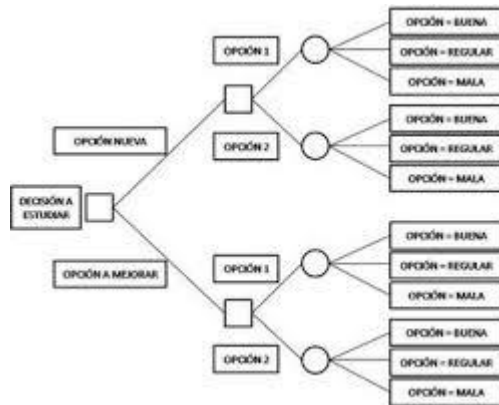
Técnicas Estadísticas

- ☐ Regresión lineal.
- ☐ Clasificación.
- ☐ Modelos no lineales.
- ☐ Reducción de dimensión.
- ☐ Selección de subconjuntos.
- ☐ Métodos basados en árboles.

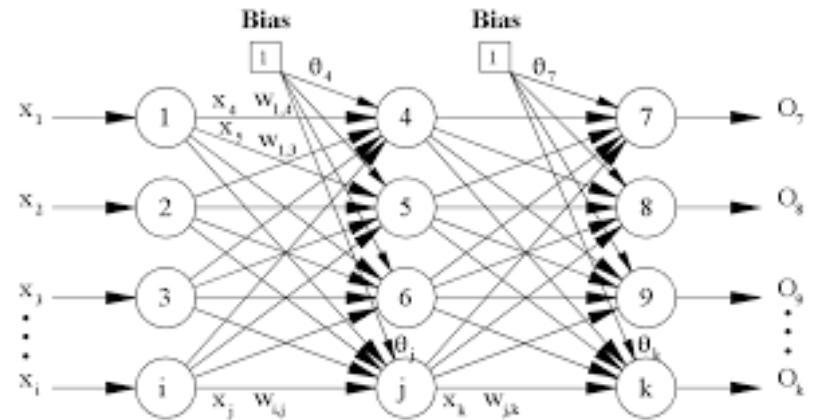
# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

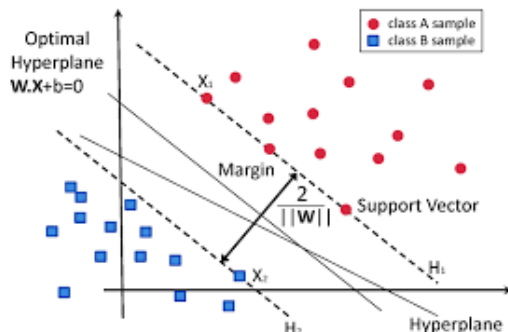
### Árboles de decisión



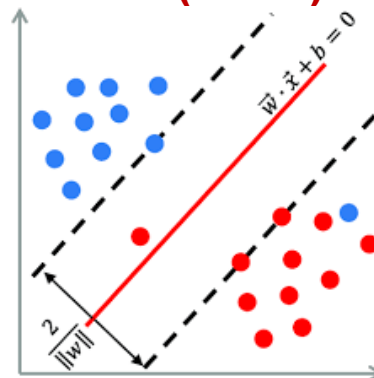
### Red neuronal *Perceptron backpropagation*



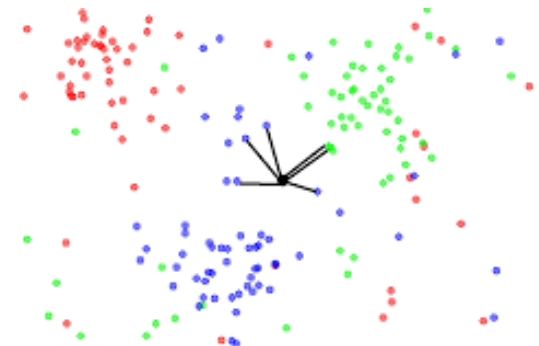
### Máquina de soporte vectorial (SVM)



### Máquina de soporte vectorial lineal (LSVM)



### Algoritmo KNN



# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

### Redes Neuronales Artificiales

#### Notation:

$L$  : layer index

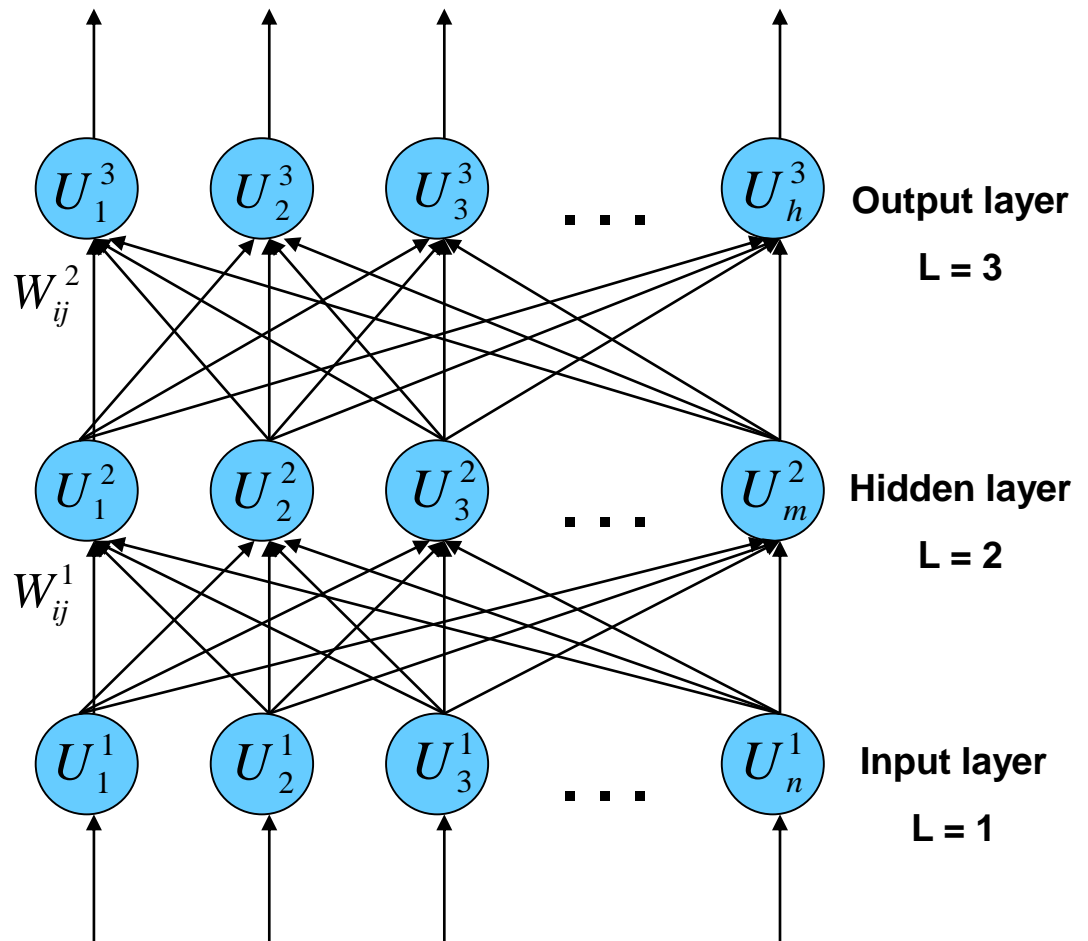
$U_i^L$  : neuron  $i$  in the layer  $L$

$W_{ij}^L$  : the weight of the connexion from  
the neuron  $i$  in the layer  $L$  to the  
neuron  $j$  in the layer  $L + 1$

$n$  : number of neurons in the  
input layer

$m$  : number of neurons in the  
hidden layer

$h$  : number of neurons in the  
output layer

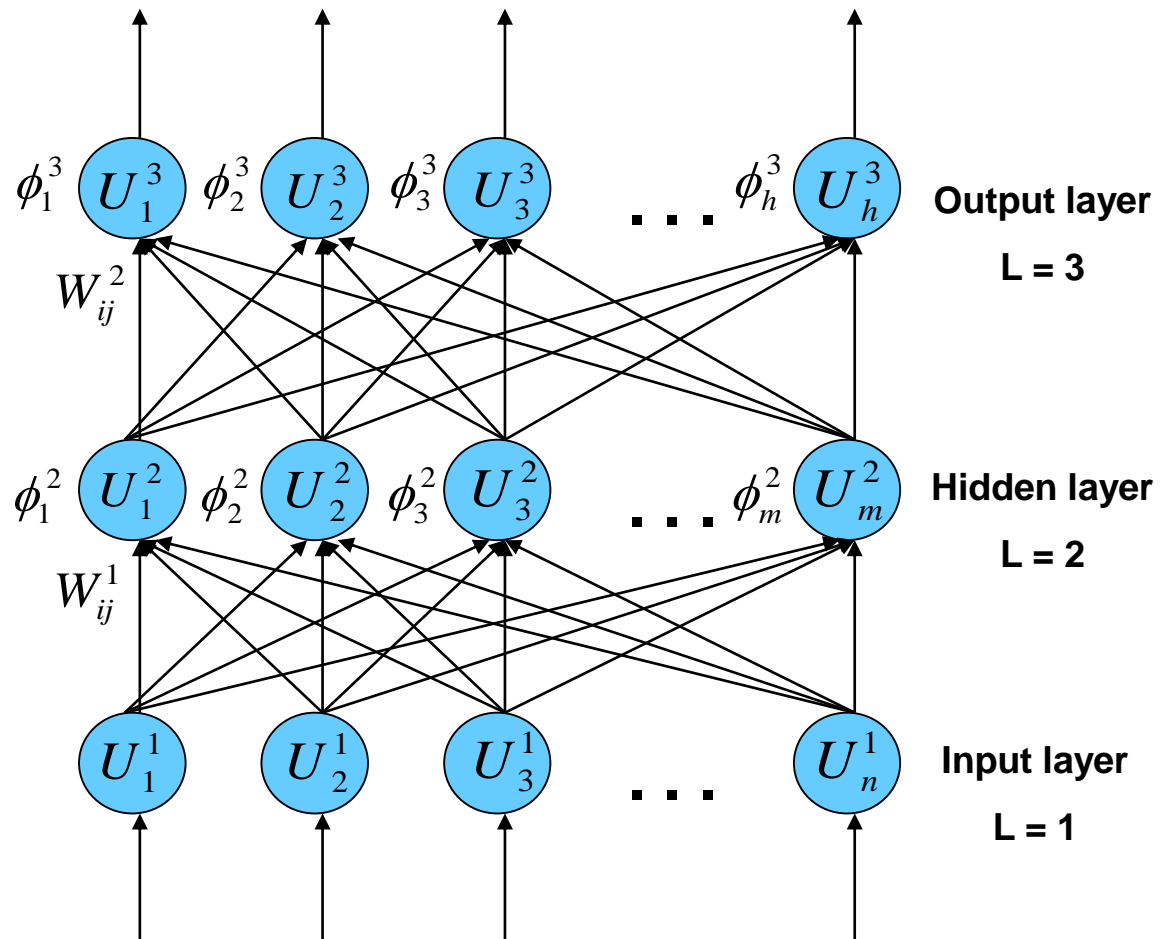


# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

### Redes Neuronales Artificiales

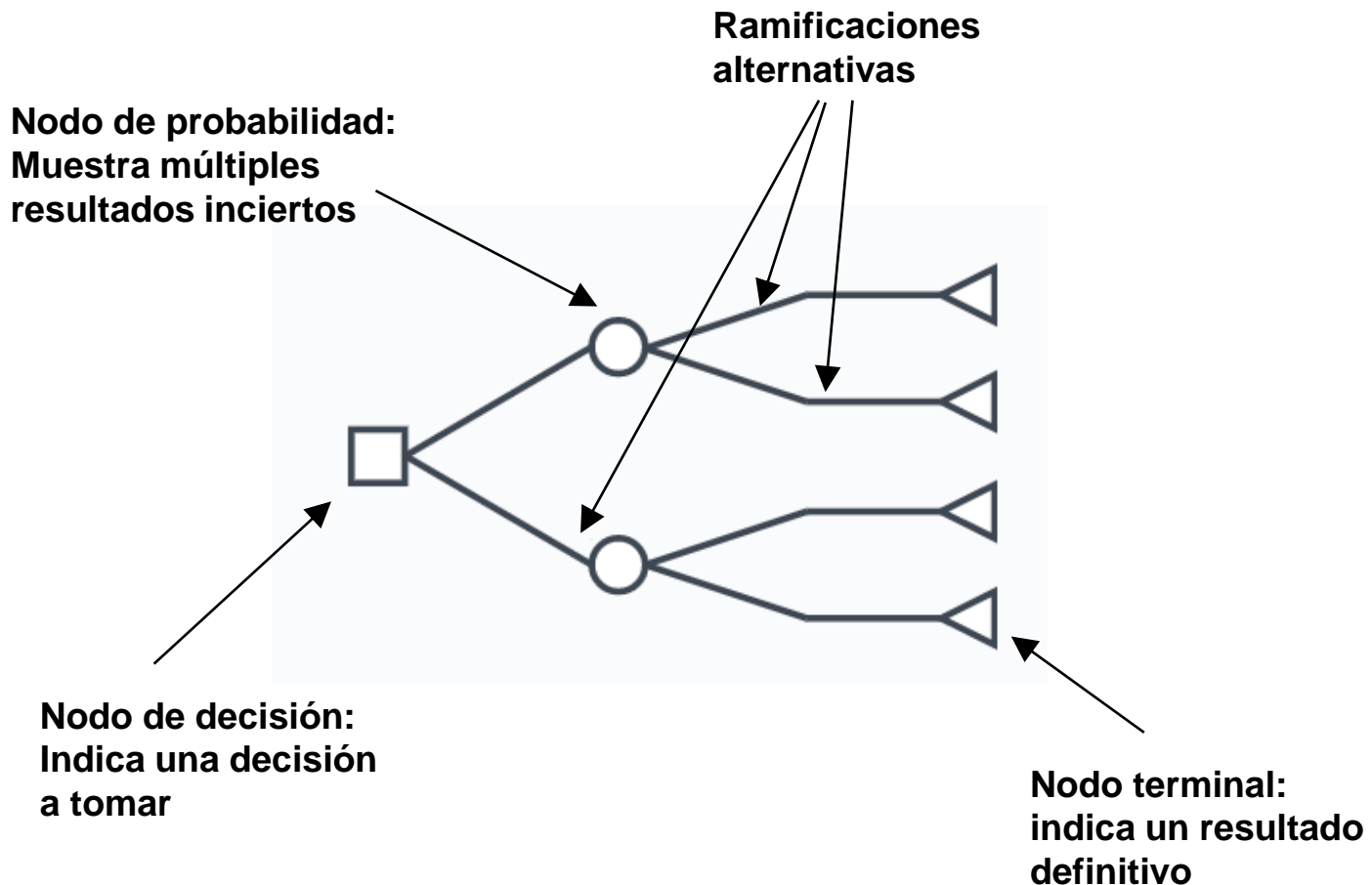
$\phi_i^L$  is iteratively computed from  $\phi_j^{L+1}$  and weights in the layer  $L + 1$



# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

### Árboles de Decisión

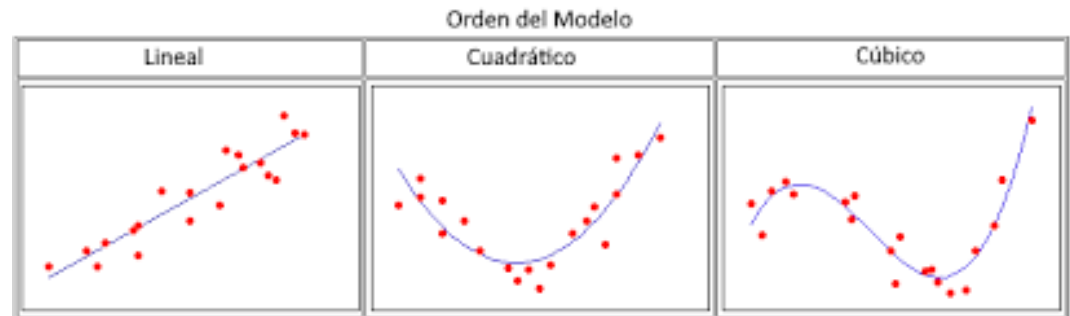
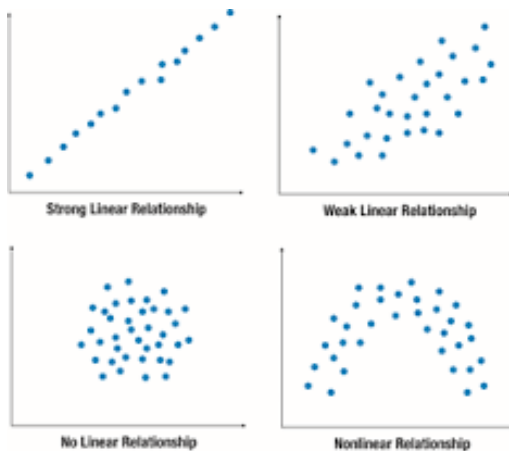
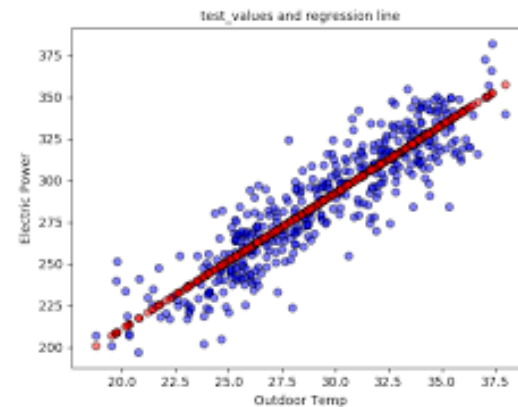
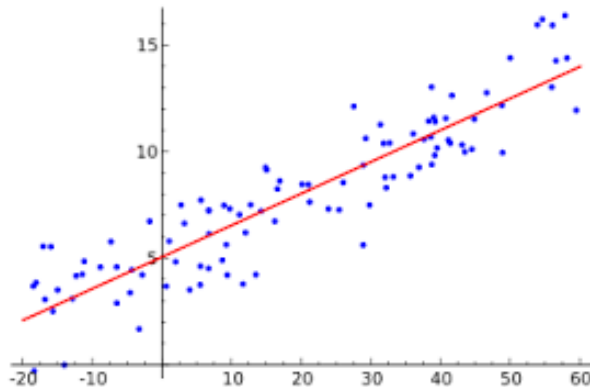




# UEA Datos a Gran Escala

## Métodos y tecnologías para el análisis de los datos a gran escala.

### Modelos Estadísticos: Modelos de Regresión

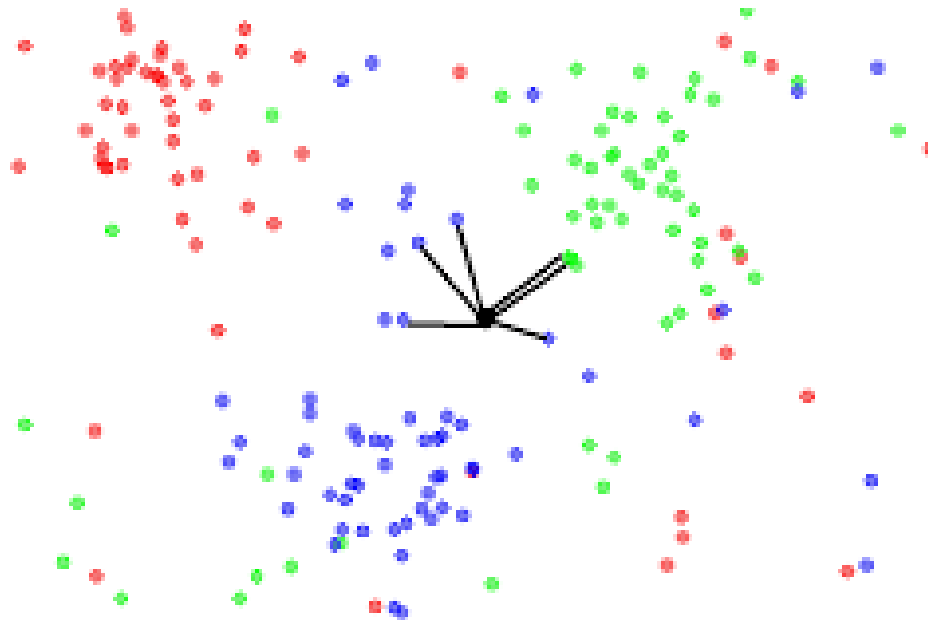




# UEA Datos a Gran Escala

**Métodos y tecnologías para el análisis de los datos a gran escala.**

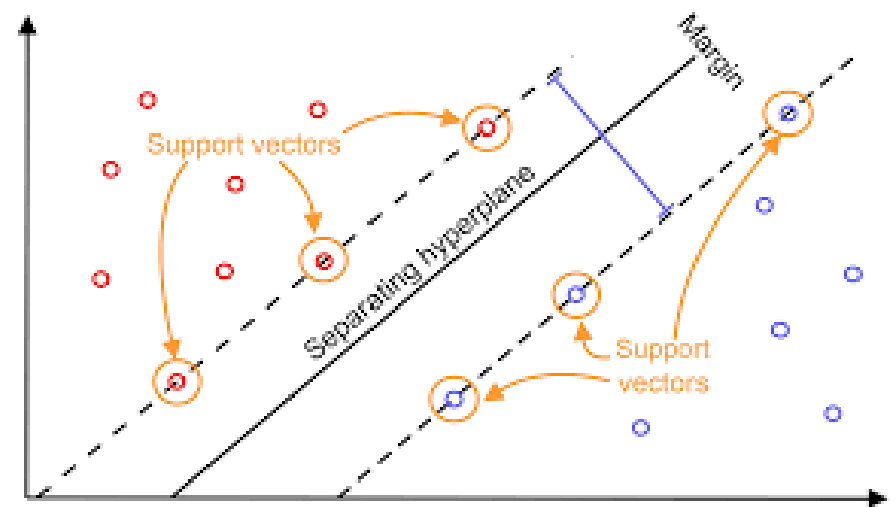
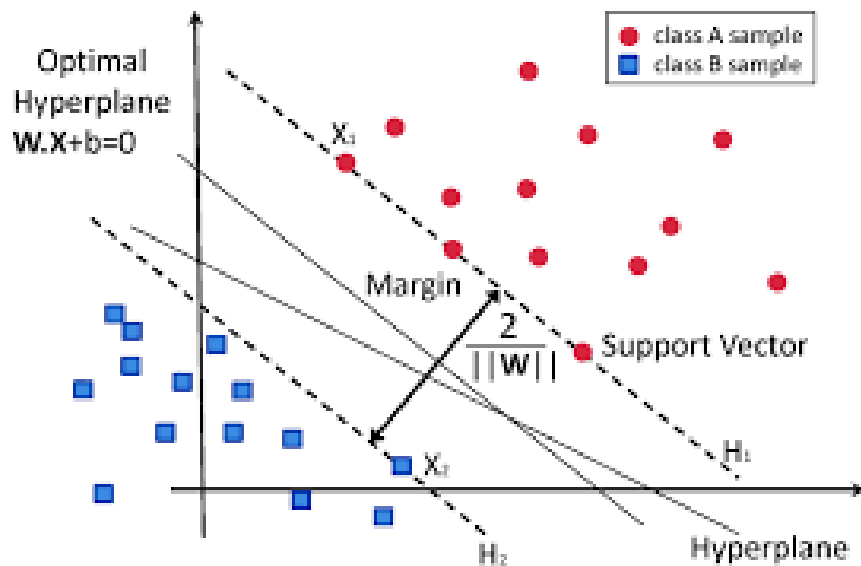
**Algoritmo de los K Vecinos Más Cercanos (KNN, *K-nearest neighbors*)**



# UEA Datos a Gran Escala

Métodos y tecnologías para el análisis de los datos a gran escala.

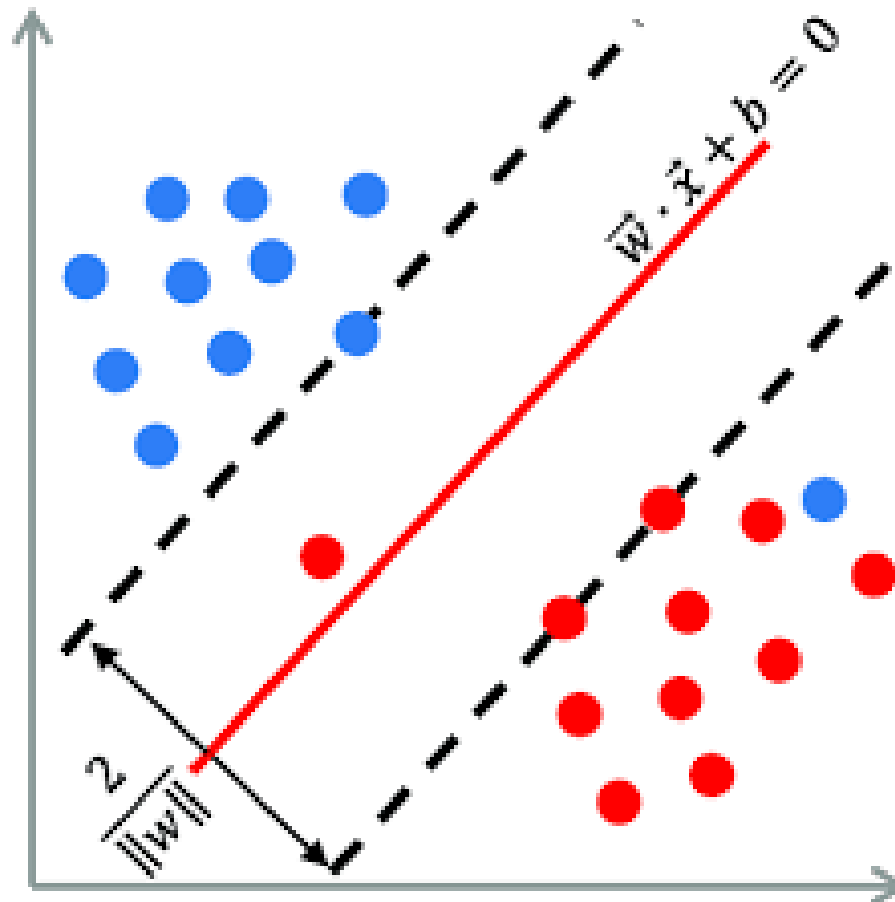
Aprendizaje Profundo (*Deep Learning*): Máquina de Soporte Vectorial



# UEA Datos a Gran Escala

Métodos y tecnologías para el análisis de los datos a gran escala.

Aprendizaje Profundo (*Deep Learning*): Máquina de Soporte Vectorial Lineal



# UEA Datos a Gran Escala

---

## **Bibliografía necesaria o recomendable:**

1. Joyanes Aguilar, L. Big data. Análisis de grandes volúmenes de datos en organizaciones. Alfaomega, 2013.
2. Marr, B. Big data in practice. How 45 successful companies used big data analytics to deliver extraordinary results. Wiley, 2016.
3. Marz, N., Warren, J. Big data: Principles and best practices of scalable realtime data systems. Manning Publications, 2015.
4. Mayer-Schönberger, V., Cukier, K. Big data. La revolución de los datos masivos. Turner Noema, 2013.
5. Mayer-Schönberger, V., Cukier, K. Big data: A revolution that will transform how we live, work, and think. John Murray, 2017.
6. Sinha, S. Making big data work for your business. Impactt Publishing, 2014.