



Modelos no supervisados

Conceptos básicos

Fuzzy C-means y Mapas de Kohonen

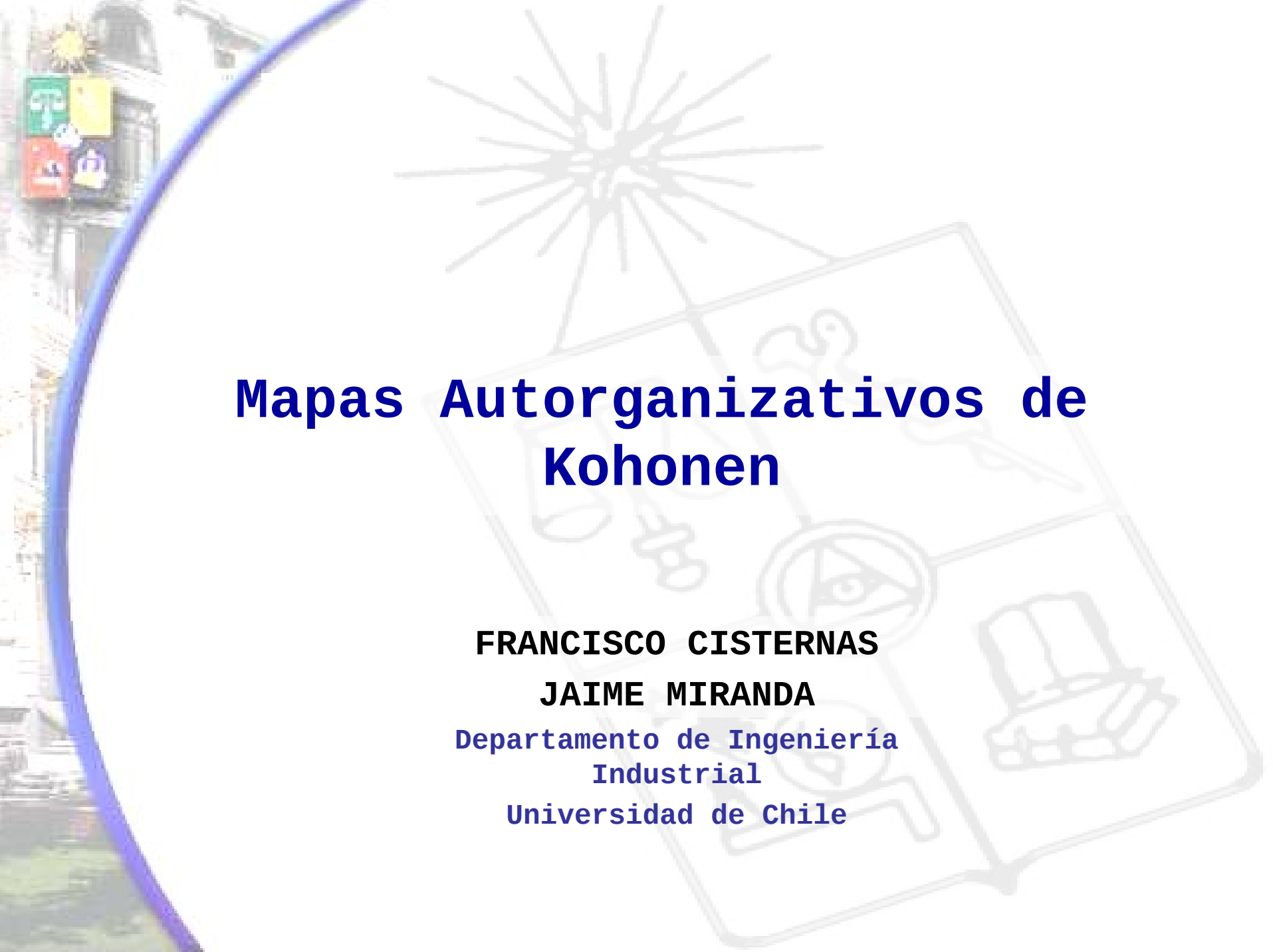
FRANCISCO CISTERNAS

**Departamento de Ingeniería
Industrial
Universidad de Chile**

Aprendizaje no supervisado

1. Algunas nociones

- ☐ Solo ocupan como información los atributos o variables de entrada.
- ☐ No se conocen las clases a priori.
- ☐ Se ocupan relaciones de similitud.



Mapas Autorganizativos de Kohonen

FRANCISCO CISTERNAS

JAIME MIRANDA

**Departamento de Ingeniería
Industrial**

Universidad de Chile

1. Proviene de la cuantización vectorial

- ☐ Aproximación a la función de densidad de probabilidad $P(x)$ de las variables de entrada (x) usando un número finito de prototipo.
- ☐ La idea es relacionar un objeto con su prototipo mas cercano.
- ☐ Trata de buscar una aplicación topológicamente ordenada del espacio de entrada.
- ☐ Los hiperplanos de separación son lineales en trozos.

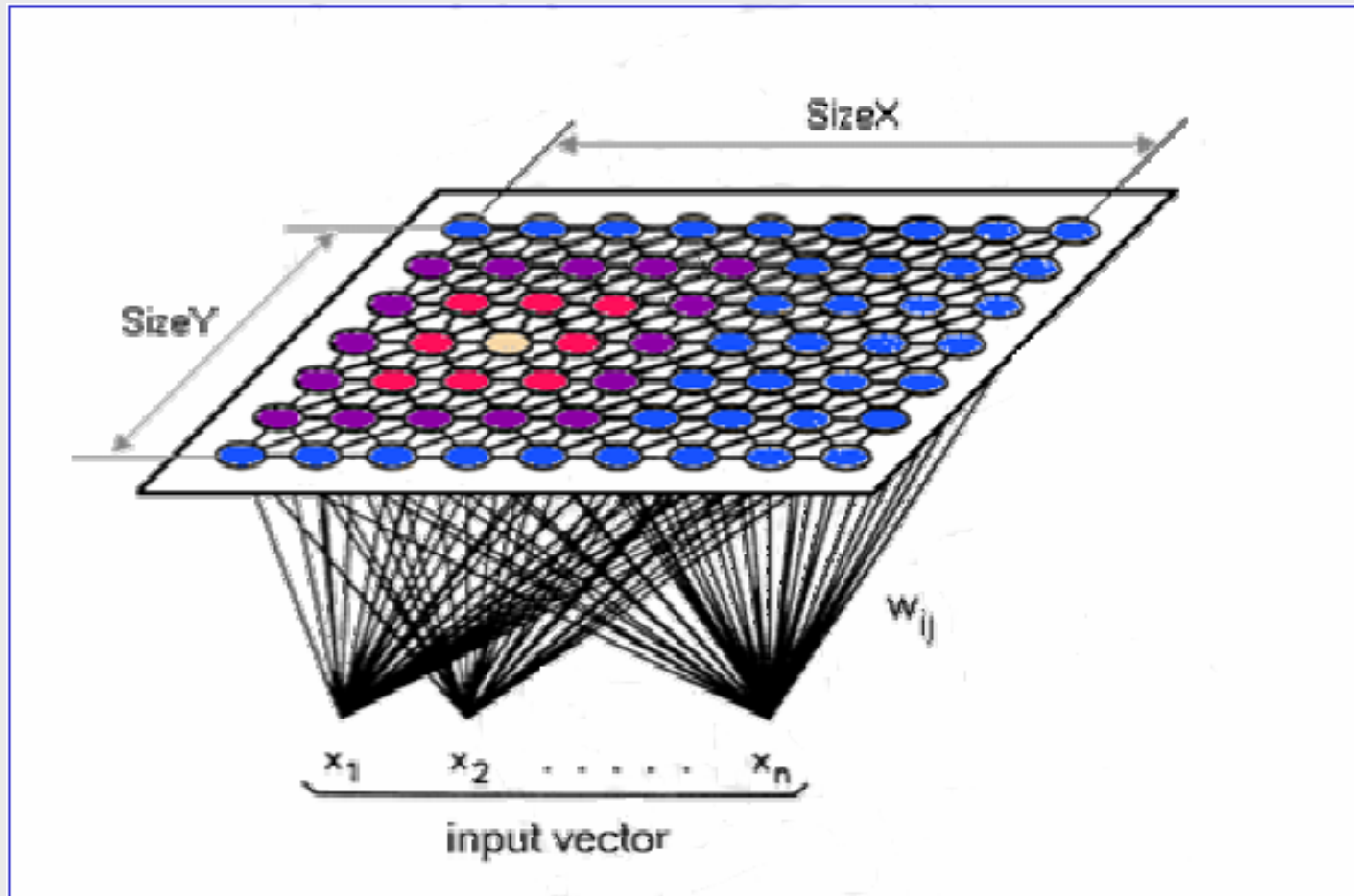
Características

1. Método no supervisado
2. Especializado para encontrar relaciones entre subconjuntos
3. Reduce la dimensionalidad del espacio de entradas
4. Tiene forma de grilla (2D) o caja (3D)

Algunos Usos

1. **Clustering:** es posible agrupar datos del conjunto de entrada, atendiendo a diferentes criterios con el fin de asignarle una clase a cada objeto del conjunto de entrada..
2. **Visualización:** este agrupamiento, trata de mostrar en forma grafica las similitudes a través de la mantención de las características topológicas.
3. **Clasificación:** una vez calibrado el mapa, o asignada algún tipo de etiqueta a cada clusters, se puede usar para clasificar sobre datos desconocidos.

Ejemplo gráfico



Algoritmo SOM

1. Inicializar los prototipos

$$m_i = m_i(0)$$

2. Encontrar neurona ganadora (BESTMATCHING)

$$\|x(t) - m_c(t)\| = \min\{\|x(t) - m_i(t)\|\}$$

3. Modificar los vectores prototipos

$$m_i(t+1) = m_i(t) + \alpha[x(t) - m_i(t)] \quad \forall i \in N_c$$

$$m_i(t+1) = m_i(t)$$

4. Volver a 3 hasta que no ocurran cambios significativos en el mapa

Tasa de
aprendizaje



1. El algoritmo básicamente realiza:

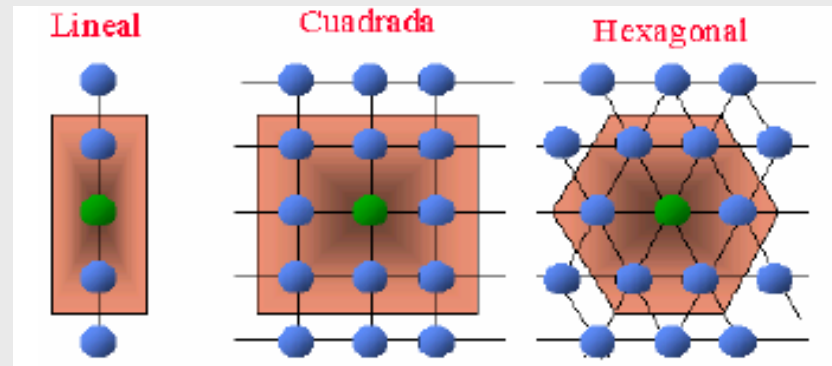
- ☐ Definir la estructura de la red (cantidad de neuronas y dimensionalidad).
- ☐ Inicializar neuronas aleatoriamente.
- ☐ Presentar un ejemplo y determinar la neurona ganadora.
- ☐ Modificar la neurona ganadora y las de la vecindad de acuerdo al ejemplo.
- ☐ Volver a 3) hasta que no se produzcan cambios significativos.

Parámetros del Modelo

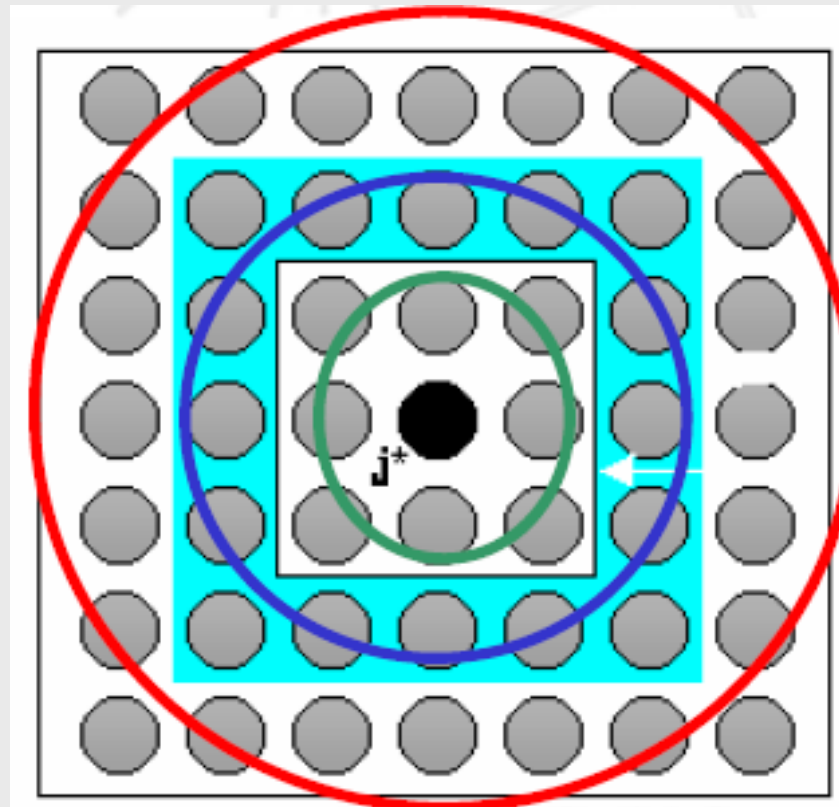
1. Dimensiones del mapa
 - ☐ 2D ó 3D.
 - ☐ Número de nodos y clases.
2. Geometría del mapa
 - ☐ Definición de la función de vecindad y radios de convergencia.
3. Tasa de aprendizaje
 - ☐ Ajuste de los pesos.
4. Épocas
 - ☐ Número de veces que se “muestran” los datos al método.

Función de Vecindad

1. Nos dice que neuronas serán modificadas en cada iteración del algoritmo
2. Esta vecindad decrece a medida que pasan la épocas.



Función de Vecindad circular



R1

R2

R3

En cada etapa de la iteración se disminuye el radio de pertenencia R_i

Ejemplo

1. Problema:

- ❑ Clasificar a un grupo de países a través de variables socioeconómicas.
- ❑ Conjunto de 23 países.
- ❑ Cada país esta caracterizado por 18 atributos (4 grupos).

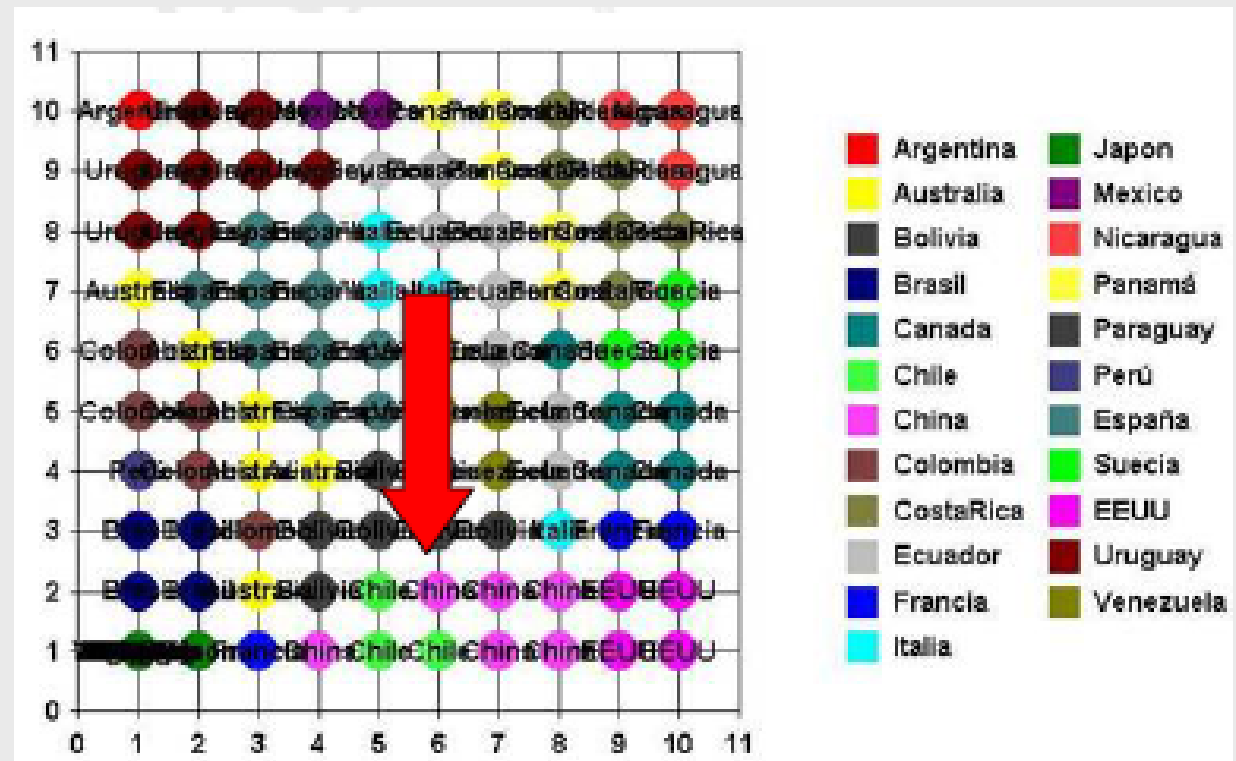
ATRIBUTOS DE LOS PAISES	
CLASES	DESCRIPCIÓN
1	Consumo energético y condicione sambientales
2	Desarrollo económico
3	Desarrollo Tecnológico
4	Población y densidad demográfica

Resultados

1. Desarrollo económico

□ •Países cercanos

- •China.
- •Francia
- •Australia



Resultados

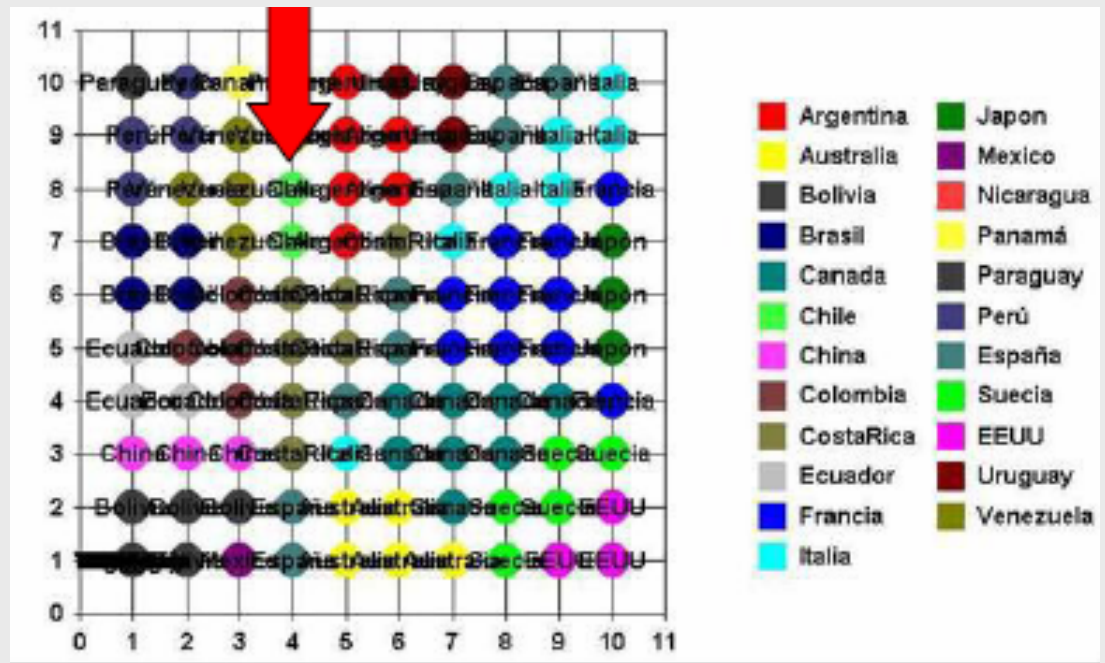
1. Desarrollo tecnológico

□ •Países vecinos.

- •Argentina.
- •Venezuela.
- •Costa Rica
- •Uruguay

□ •Otros países:

- •EEUU.
- •Suecia.
- •Australia

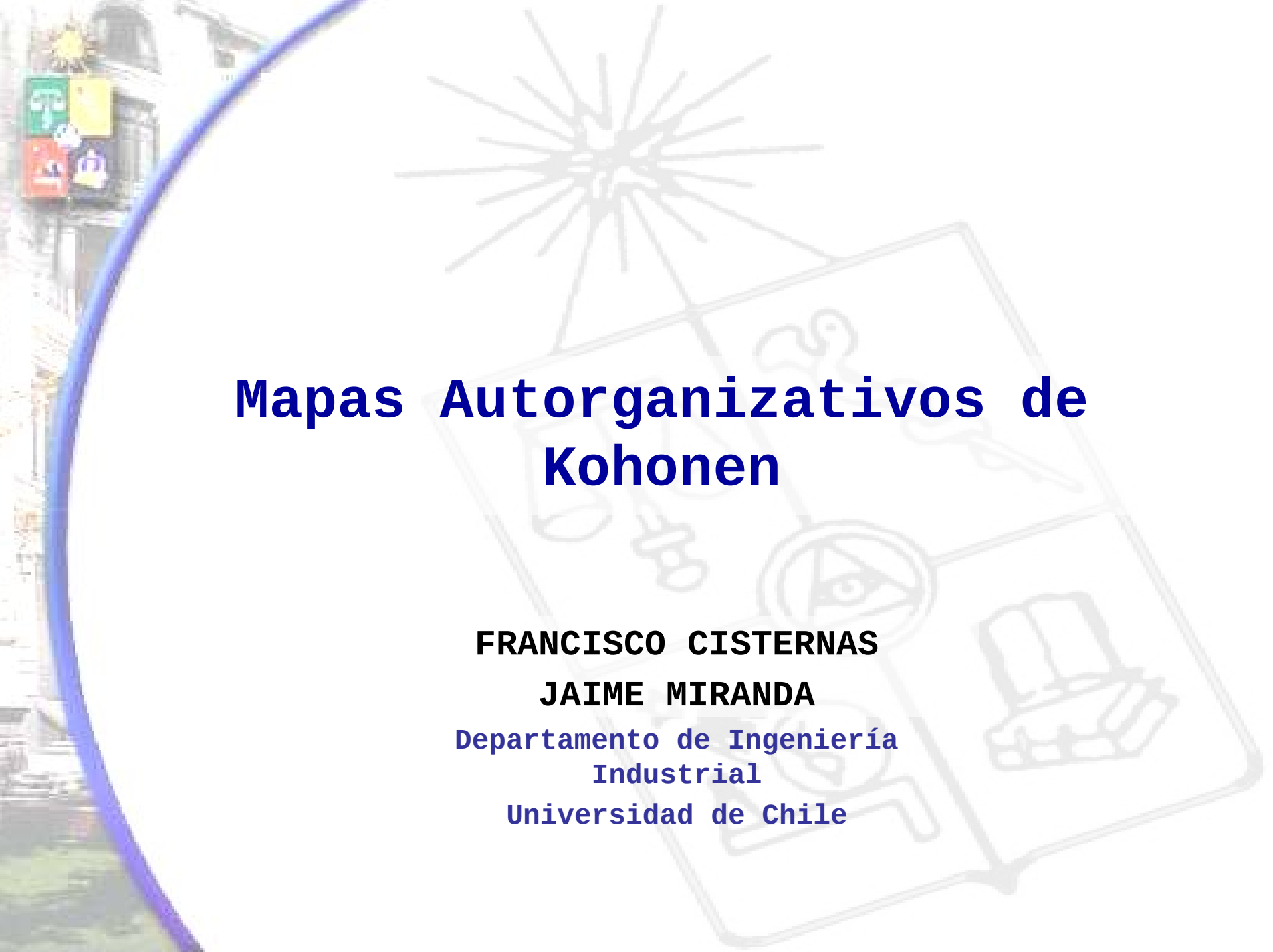


Ventajas

1. Entrenamiento no supervisado.
2. No necesita pares de Entrada/Salida, solo los de entrada.
3. Disminuye la dimensionalidad del problema a algo “representable” en un espacio visualmente representable.

Limitaciones

1. Solo entrega información sobre la ubicación del espacio.
2. Es necesario interpretar la información encontrada (Componentes principales).



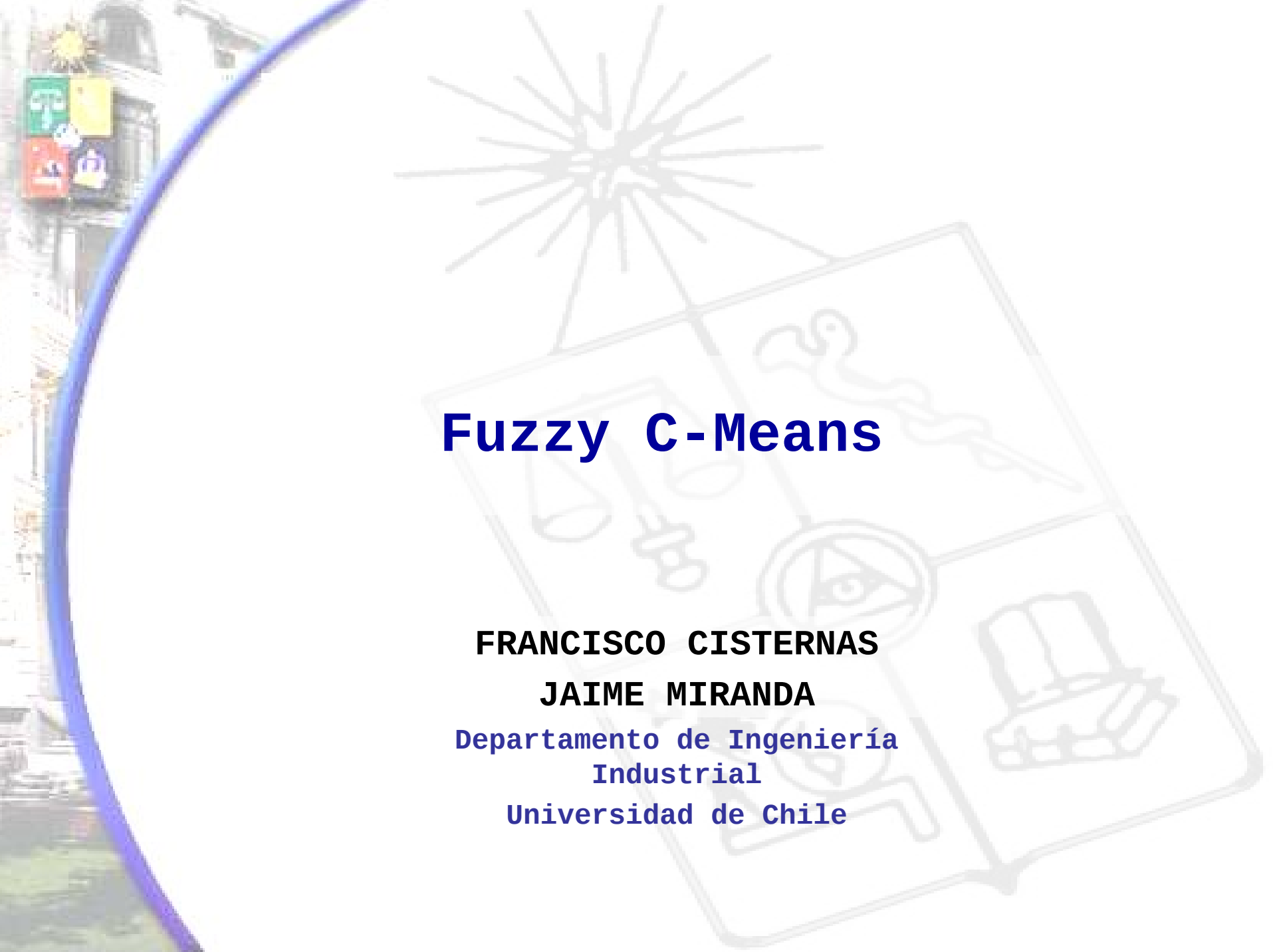
Mapas Autorganizativos de Kohonen

FRANCISCO CISTERNAS

JAIME MIRANDA

Departamento de Ingeniería
Industrial

Universidad de Chile



Fuzzy C-Means

FRANCISCO CISTERNAS

JAIME MIRANDA

Departamento de Ingeniería
Industrial

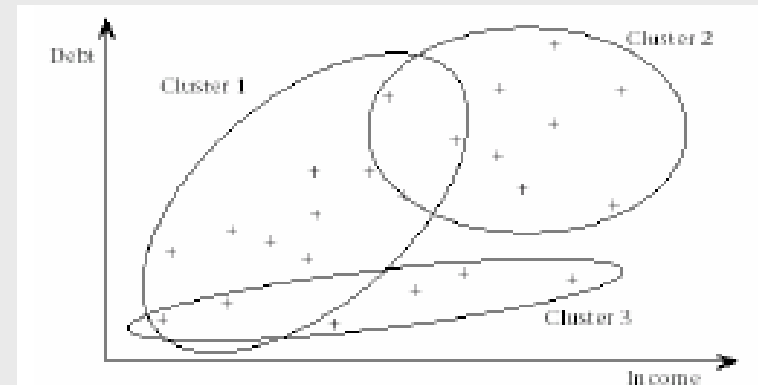
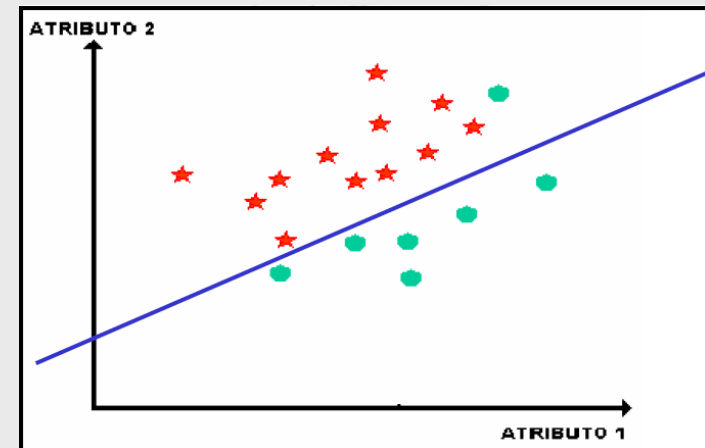
Universidad de Chile

1. Métodos supervisados
 - ☐ Redes neuronales
 - ☐ Árboles de decisión
2. Métodos no supervisados
 - ☐ Mapas de Kohonen (SOM)
 - ☐ Cluster: FuzzyC-means

Discriminantes

1. Conceptos básicos

- ❑ • **Medida de distancia**
 - ❑ • Prototipo o centro de clase más cercana.
- ❑ • **Hipersuperficies**
 - ❑ • Clasificación de acuerdo a si los objetos están a uno u otro lado de una hipersuperficie o conjunto de hiperplanos.

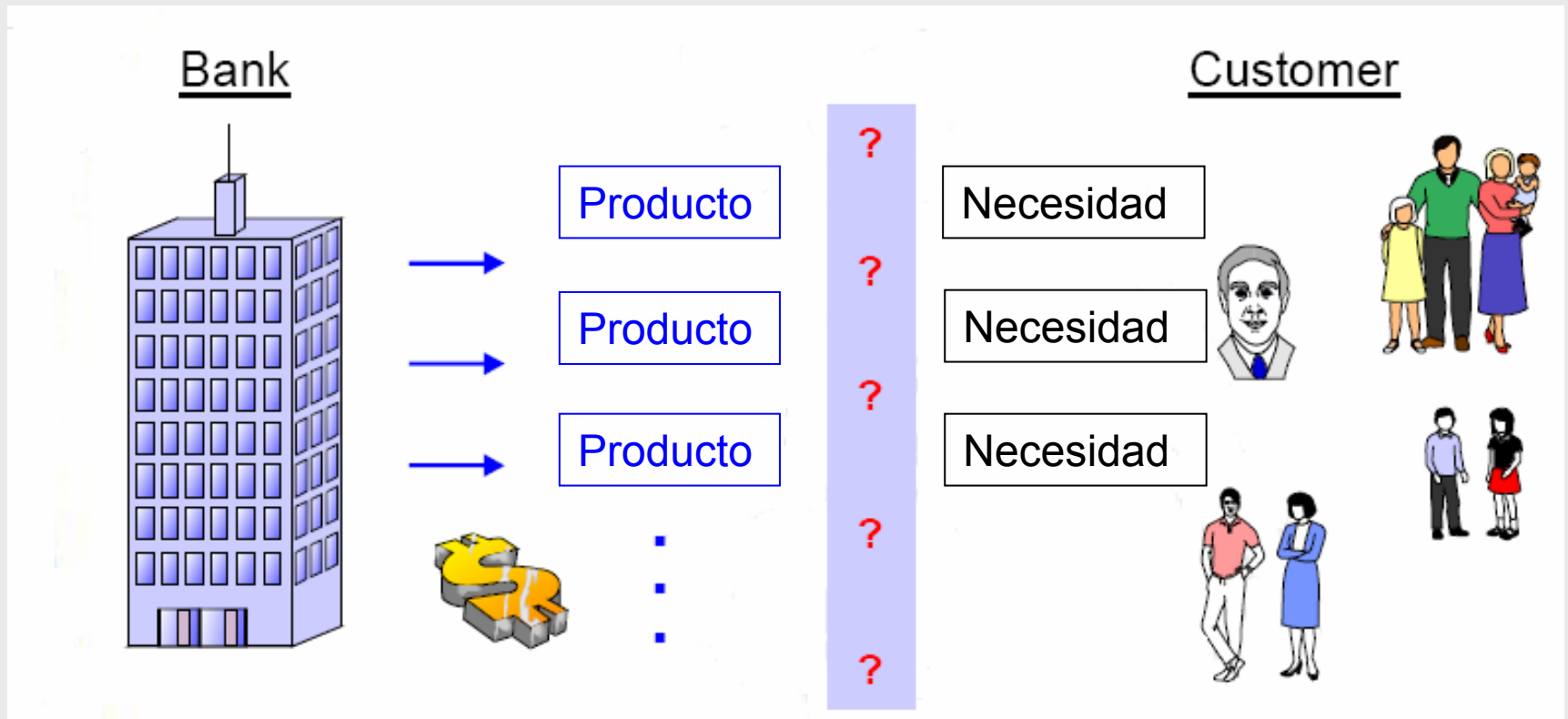


Algunas Nociones

1. Método de clustering que utiliza un algoritmo matemático.
2. A través de la iteración este algoritmo y determina el grado de pertenencia de un objeto a una clase determinada y el centro de esta.
3. Método no supervisado.

1. La base de este algoritmo es la minimización de la suma de las distancias de cada objeto al centro de la clase.
2. Se formula un problema de optimización cuya función objetivo es minimizar la suma de las distancias ponderadas por el grado de pertenencia del objeto.

Segmentación de clientes



Qué producto para que (segmento) de clientes?

1. Variables y Parámetros:

- ☐ n = Número de objetos.
- ☐ c = Numero de clases.
- ☐ X_i = Vector con valores de los atributos del objeto i , con $i = (1, \dots, n)$.
- ☐ C_j = Centro de la clase j , con $j = (1, \dots, c)$.
- ☐ $d^2(X_i, C_j)$ = Distancia entre el objeto i al centro de la clase j .

1. Variables y Parámetros:

- u_{ij} = Grado de pertenencia del objeto i a la clase j .
- $u_{i,j} \in [0,1]$

$$\sum_{j=1}^c u_{ij} = 1$$

- Matriz de pertenencia
- Función objetivo:


$$\min \sum_{i=1}^n \sum_{j=1}^c (U_{ij})^m * d^2(x_i, c_j)$$

1. PASO 1

- Determinar una matriz $U = (U_{ij})_{ij}$ que cumpla:

$$\sum_{j=1}^c u_{ij} = 1$$

2. PASO 2

- Determinar los centros de las clases:

$$c_j = \frac{\sum_{i=1}^n u_{ij}^m X_i}{\sum_{i=1}^n u_{ij}^m}$$

Iteraciones del Algoritmo

1. PASO 3

- Actualizar los grados de pertenencia:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d(x_i, c_j)}{d(x_i, c_k)} \right)^{\frac{2}{m-1}}}$$

2. PASO 4

- Criterio de detención

$$\|U^{k+1} - U^k\| < \varepsilon$$

U^k = matriz en la iteración k

1. Número de clases

- ☐ En cuantos segmentos separo mi muestra.
- ☐ “Cluster Analysis”.
- ☐ Regla de los codos.

2. Coeficiente de difusividad (m)

- ☐ Rango (1 , ∞).

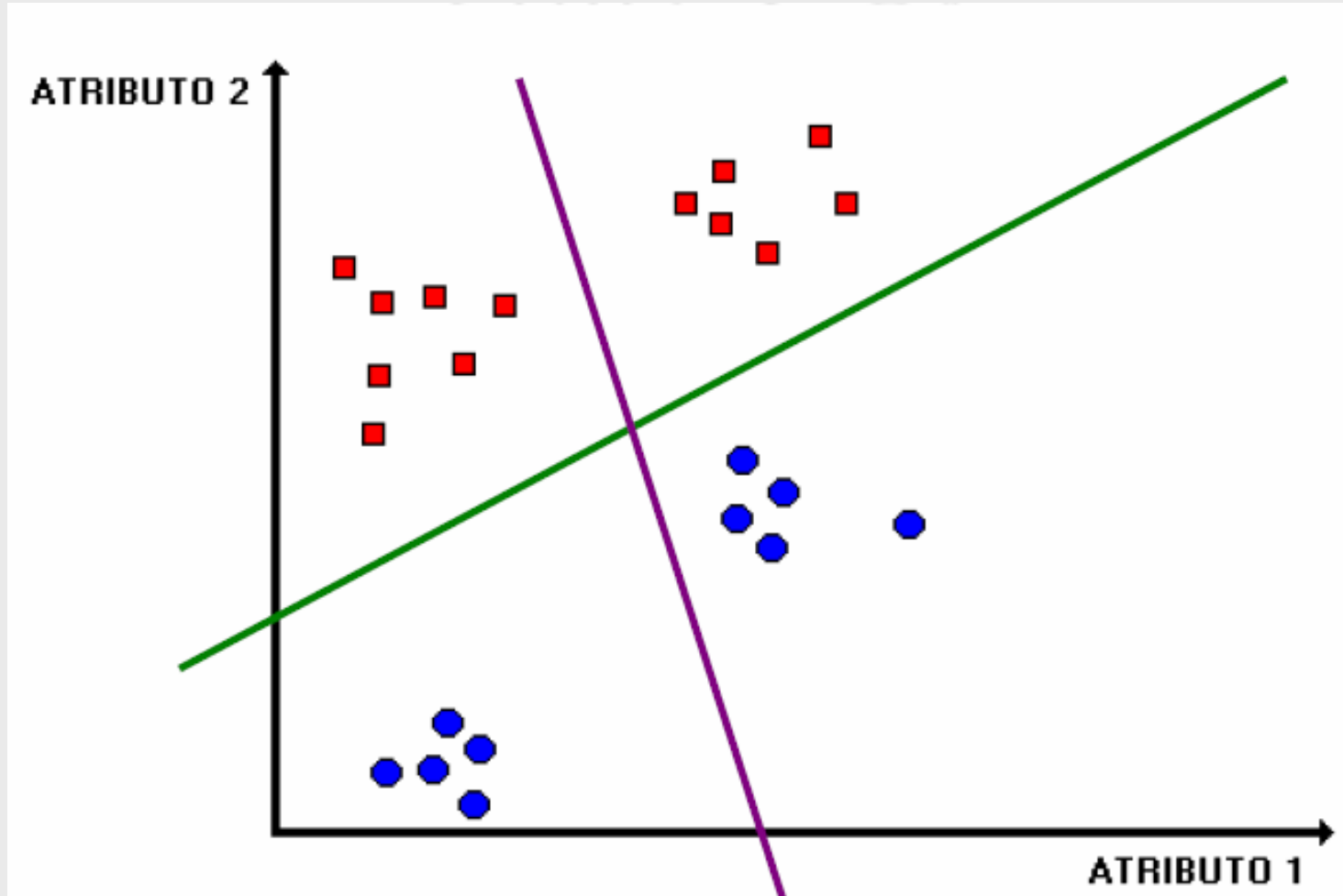
3. Umbral de detención

- ☐ El error de ajuste que el modelo acepta.

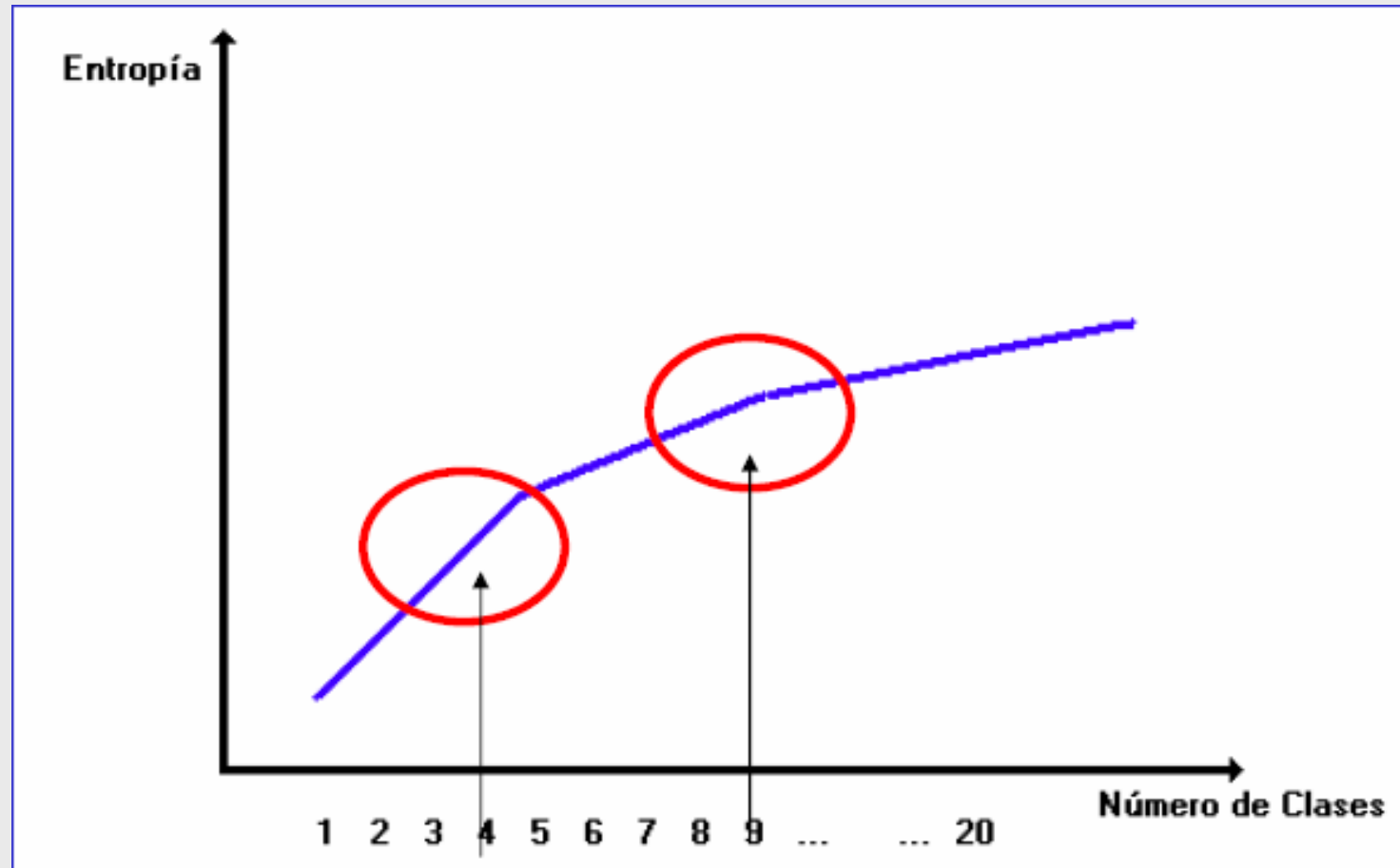
4. Valores iniciales de los centros

1. Este método a través de algún criterio especial nos indica el número óptimo de cluster's
2. Uso de la entropía (ClassificationEntropy):
 - ☐ **Regla de los codos**
 - ☐ mide la discontinuidad de la entropía respecto al numero de clases.

Ejemplo

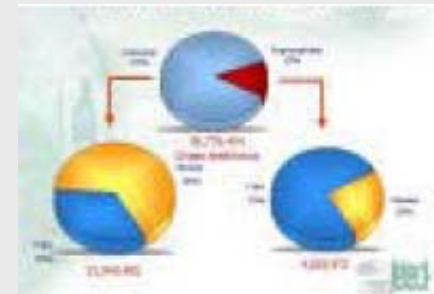


Ejemplo

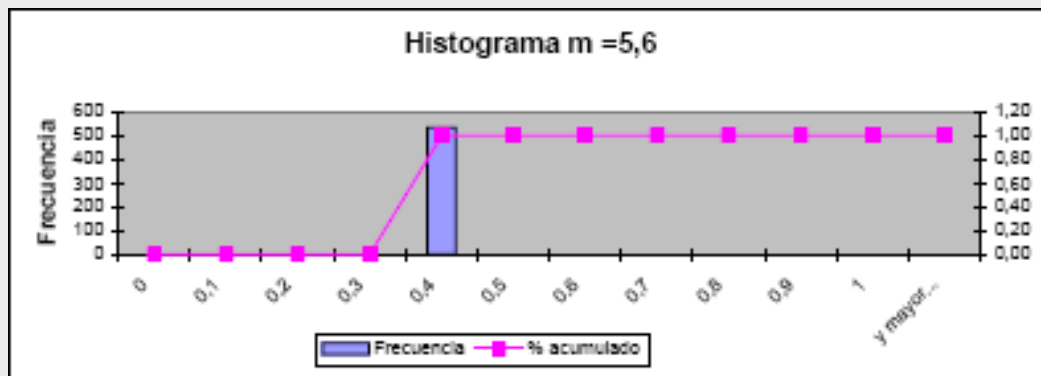
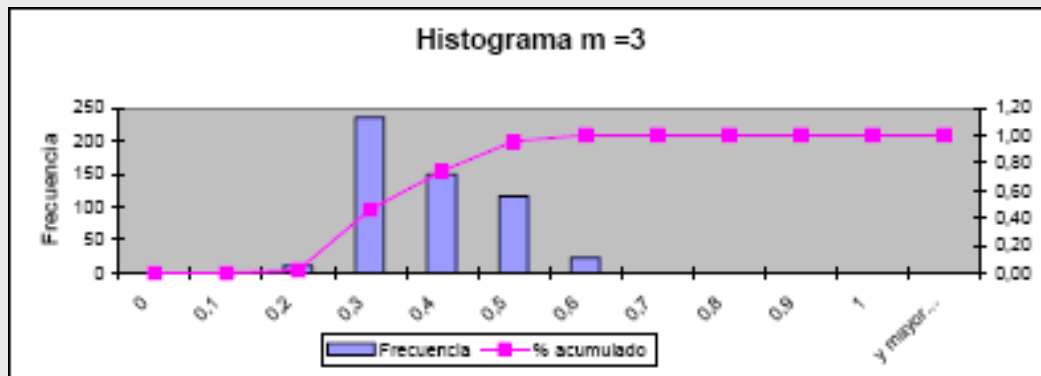
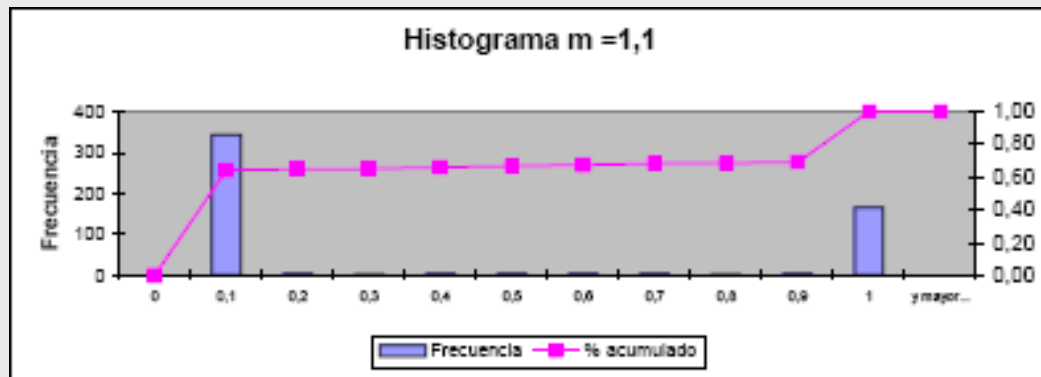


Parámetro de difusividad (m)

1. Muestra que tan difusa será la clasificación.
2. Rango $(1, \infty)$.
3. Valor mínimo = 1 , mínima difusividad (todos distintos).
4. Valor máximo = ∞ máxima difundida (todos iguales).
5. Valores pequeños da clasificaciones más discretas.
6. Al aumentar la clasificación se hace más difusa.



Ejemplos





Modelos no supervisados

Conceptos básicos

Fuzzy C-means y Mapas de Kohonen

FRANCISCO CISTERNAS

**Departamento de Ingeniería
Industrial
Universidad de Chile**