

# Reconocimiento de patrones

Clase 12: Agrupamiento y reconocimiento de patrones





# Antes de empezar...

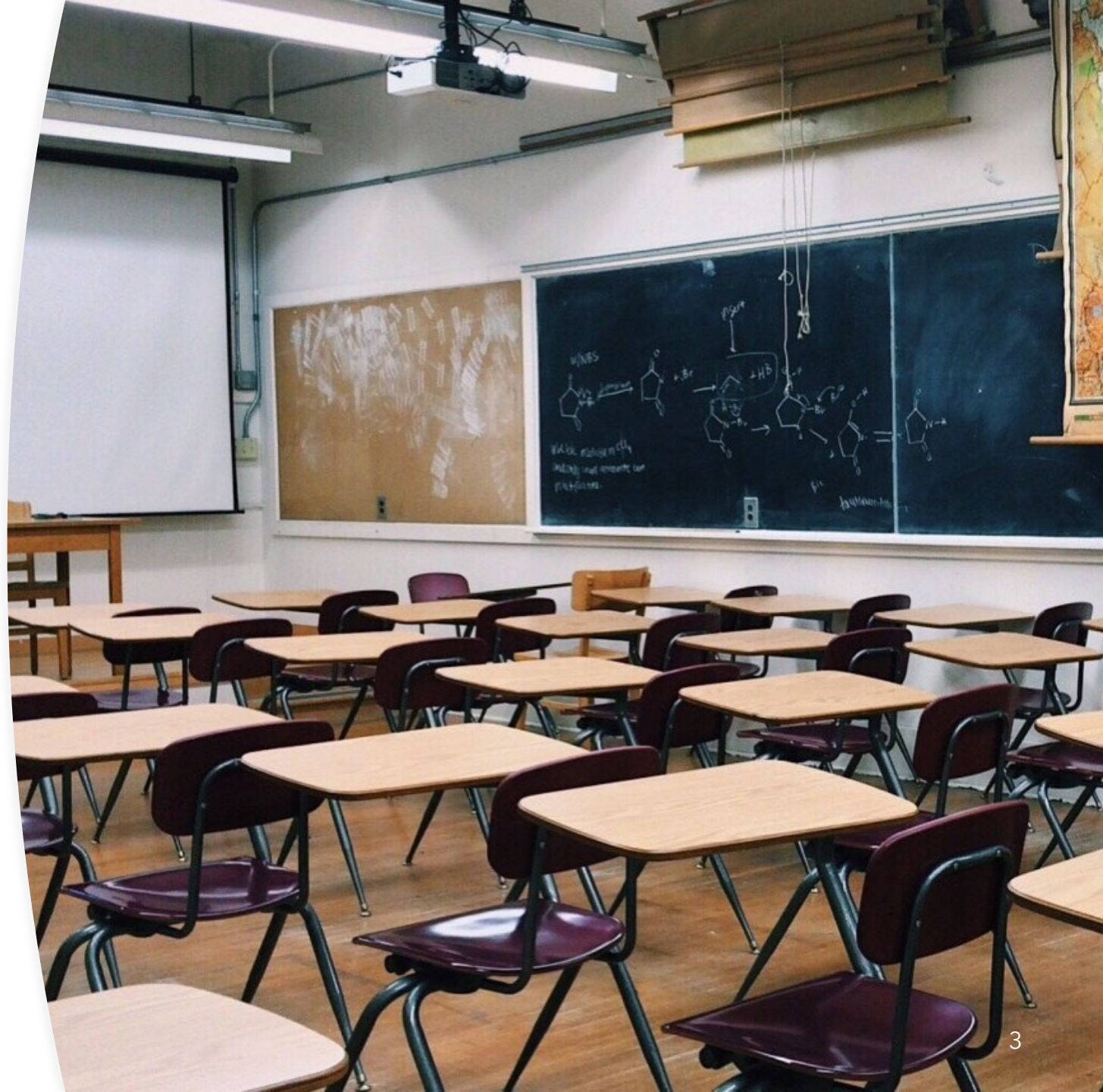
---

- Dudas de ISODATA



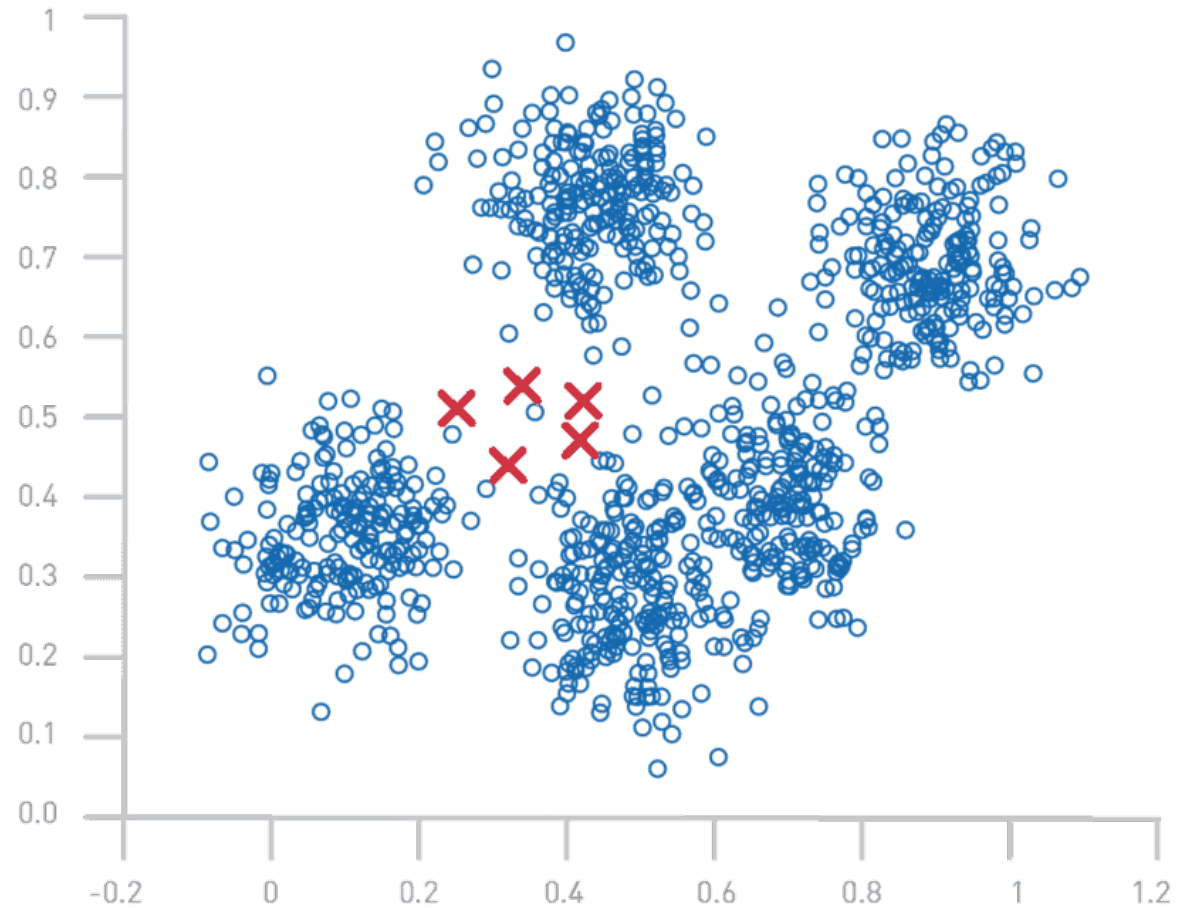
# Para el día de hoy...

- Evaluando los resultados del agrupamiento
- Reconocimiento de patrones no supervisado



# Evaluando los resultados del agrupamiento

- ¿Qué tan bueno es nuestro resultado?
- ¿Cuáles indicadores podemos usar?



# El problema de la dimensionalidad



# Algunas medidas

---

Distancia entre  
centros de los  
grupos

El número de  
grupos

La varianza de  
un grupo

Distancia  
promedio al  
centro de un  
grupo

Los puntos más  
cercanos y  
lejanos al centro  
del grupo

## Un ejemplo

- Supongamos que tenemos un conjunto de 200 muestras en  $\mathbb{R}^4$  que fueron agrupadas usando alguno de los algoritmos vistos
- Esperábamos entre 5 y 7 grupos
- ¿Qué podemos interpretar?

$i$	$\sigma_{x_1}$	$\sigma_{x_2}$	$\sigma_{x_3}$	$\sigma_{x_4}$
1	1.2	1.5	2.1	0.8
2	1.3	1.4	1.2	1.4
3	2.0	1.3	1.7	1.1
4	0.9	1.1	2.2	1.7
5	0.3	0.4	2.3	0.6
6	2.5	4.2	1.7	2.3
7	2.2	1.9	1.1	3.5

$m_j$	$i$	$j:$	1	2	3	4	5	6	7
40	1			3.5	5.6	3.7	25.1	17.3	7.8
23	2				4.8	5.8	26.3	14.5	6.0
48	3					5.0	24.0	12.7	6.3
18	4						22.2	13.8	4.9
26	5							33.6	22.4
3	6								15.1
42	7								

# Algunas ideas

- Los grupos 1,3 y 7 son grupos mayores y 2,4, y 5 son menores
- Los grupos 1-4 y 7 están “cerca” y el 5 está lejos
- El grupo 2 es “casi” una esfera
- El clúster 6 se podría eliminar

$i$	$\sigma_{x_1}$	$\sigma_{x_2}$	$\sigma_{x_3}$	$\sigma_{x_4}$
1	1.2	1.5	2.1	0.8
2	1.3	1.4	1.2	1.4
3	2.0	1.3	1.7	1.1
4	0.9	1.1	2.2	1.7
5	0.3	0.4	2.3	0.6
6	2.5	4.2	1.7	2.3
7	2.2	1.9	1.1	3.5

$m_j$	$i$	$j:$	1	2	3	4	5	6	7
40	1			3.5	5.6	3.7	25.1	17.3	7.8
23	2				4.8	5.8	26.3	14.5	6.0
48	3					5.0	24.0	12.7	6.3
18	4						22.2	13.8	4.9
26	5							33.6	22.4
3	6								15.1
42	7								



# Reconocimiento de patrones no supervisado

Supongamos nos dan un grupo de patrones sin ninguna información

El problema puede iniciar como identificar las clases dentro de esos datos

Podríamos utilizar los centros como el método de representación

Cada grupo puede ser interpretado como una clase

Posteriormente, podemos usar esas clases para determinar funciones de decisión

# Comentarios finales



Hemos revisado los conceptos principales de clasificadores basados en distancia



La distancia mínima puede ser utilizada con prototipos únicos y se puede extender la idea a múltiples prototipos



Dada suficiente capacidad de memoria se puede aproximar un desempeño óptimo



El problema clave es encontrar grupos y prototipos

# Lo que vimos

<b>3. Clasificación de patrones por medio de funciones de distancia</b>	
3.1	Clasificación de patrones por medio de distancia mínima
3.2	Algoritmos para el agrupamiento de datos
3.2.1	Medidas de similitud
3.2.2	Algoritmos de agrupamiento: k-medias, LBG e Isodata



# Para la otra vez...

- Algoritmos evolutivo



The End.