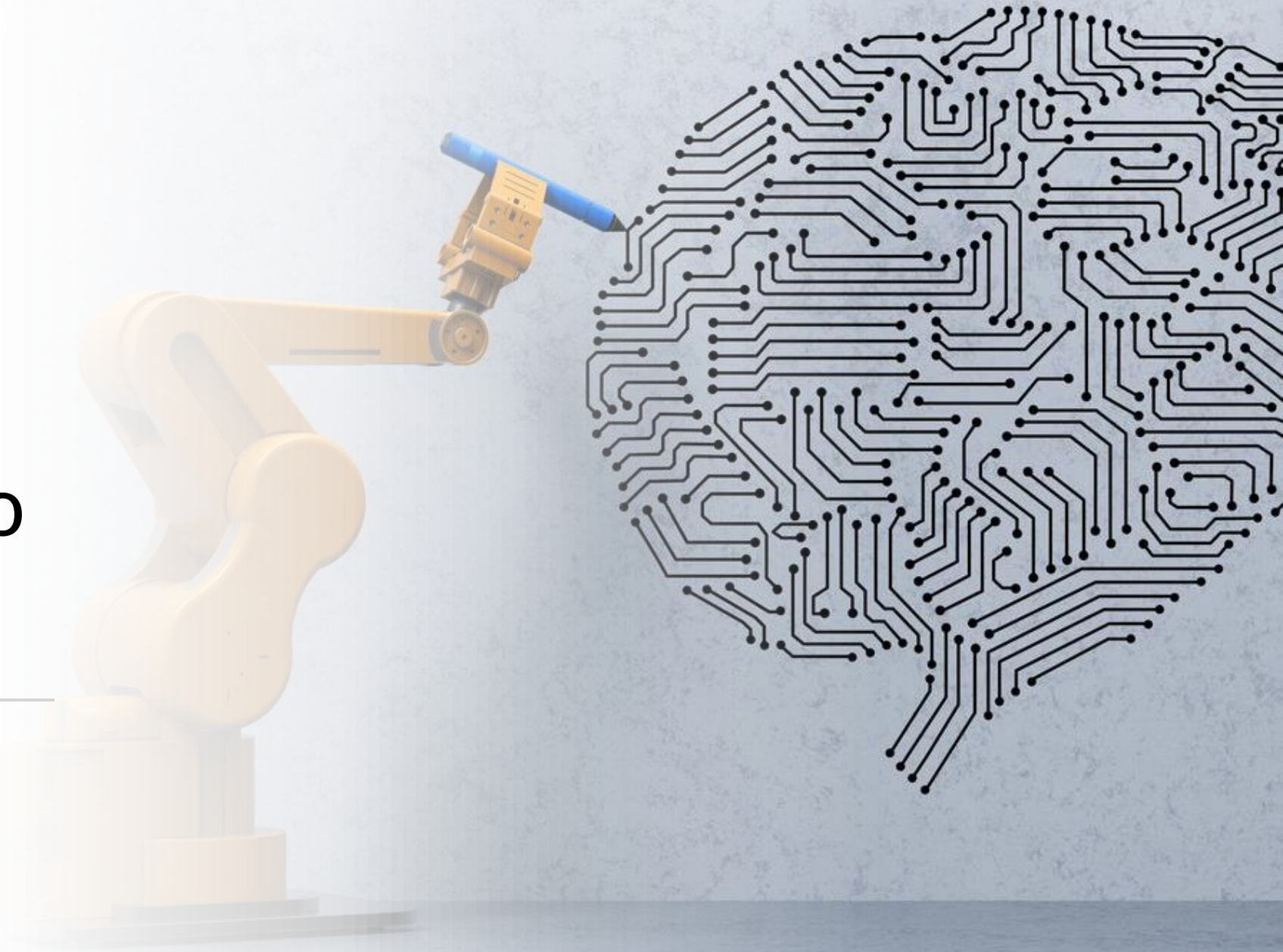




Reconocimiento de Patrones

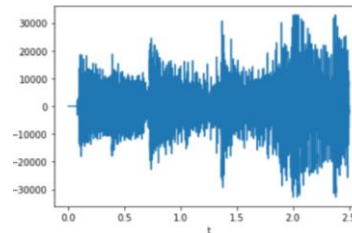
Clase 2

Extracción de características



¿Qué es la extracción de características?

- Es el proceso mediante el cual se transforma un conjunto de datos crudos (imágenes, audio, series de tiempo, etc.) en un vector de valores numéricos que describe de manera compacta la información relevante del patrón a reconocer.

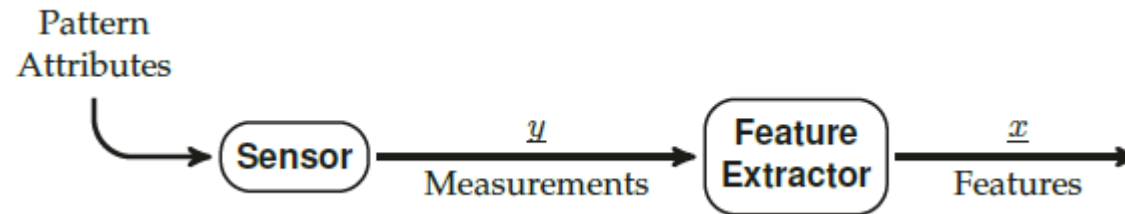


$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

En otras palabras

En lugar de trabajar directamente con los datos originales, que suelen ser de alta dimensionalidad, ruidosos y redundantes, se busca representar cada muestra mediante un conjunto reducido de atributos que conserven la información discriminante necesaria para separar clases.

$$\text{Measurements } \underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \xrightarrow{\underline{x} = f(\underline{y})} \text{Features } \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



$$y \in \mathbb{R}^N \rightarrow x \in \mathbb{R}^d, \quad d \ll N$$

¿Por qué tener menos características?



Un tamaño de muestra finito (número finito N de muestras) limita la dimensionalidad n para la cual las estimaciones estadísticas son confiables.



El tener características que están correlacionadas con alguna otra puede degradar el desempeño.



La complejidad del clasificador incrementa con el número de n características.

¿En que consiste?



Analizar la naturaleza del dato (imagen, audio, señal, etc.)



Identificar propiedades relevantes del patrón



Calcular medidas cuantitativas que describan dichas propiedades



Construir un vector de características comparable entre muestras

Se busca

Discriminabilidad (separar clases)

Compacidad (reducir dimensionalidad)

Robustez (ruido, variaciones intra-clase)

Costo computacional

Características Estadísticas

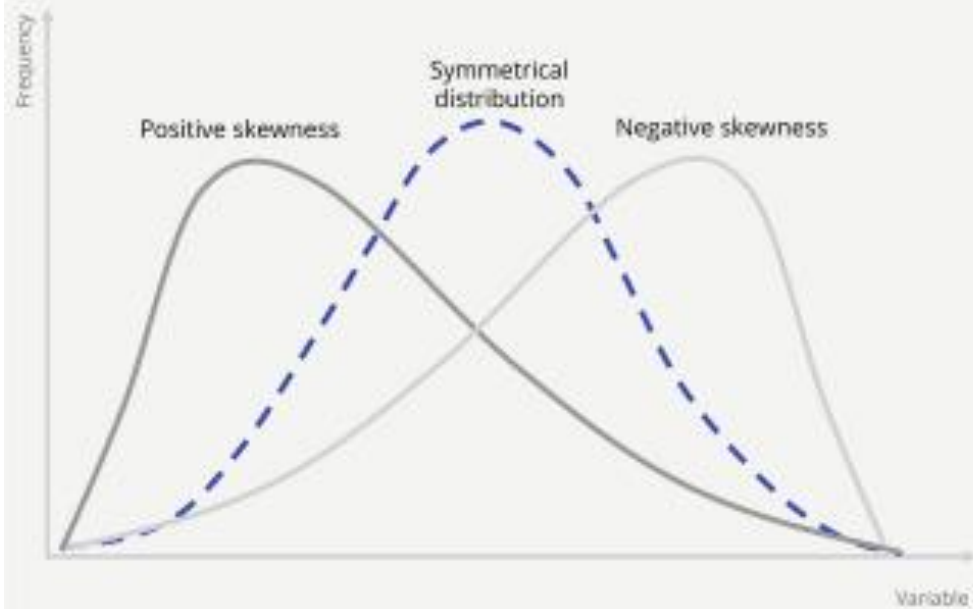
- Describen la distribución de los datos sin considerar explícitamente su estructura espacial o temporal.
- Medidas de primer y segundo orden

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \qquad \sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

- Orden superior

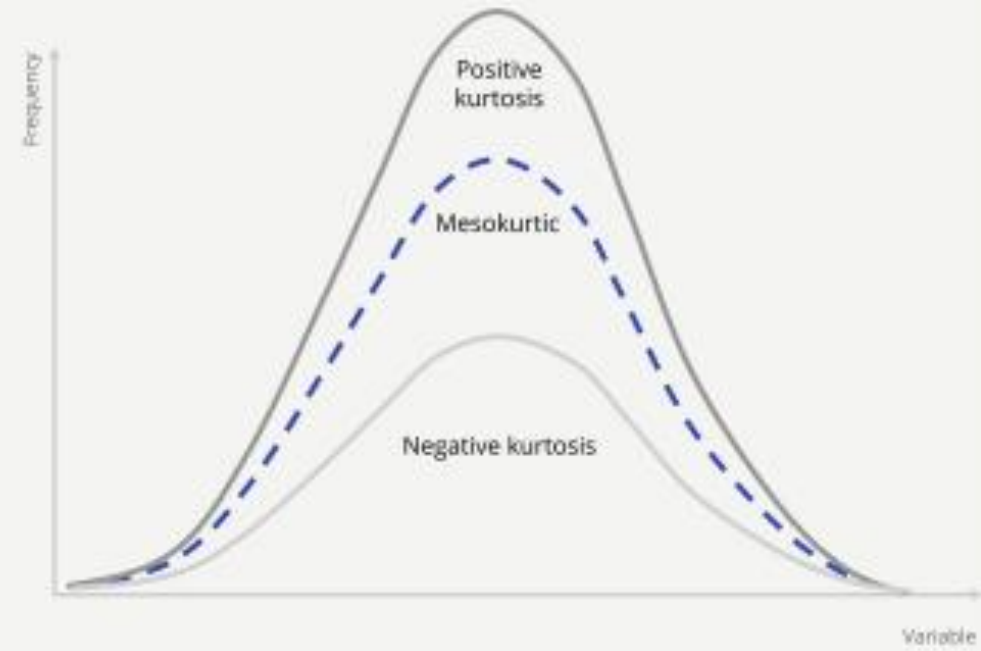
$$skewness = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_i - \mu)^3 \qquad excess\ kurtosis = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_i - \mu)^3 - 3$$

SKEWNESS



>0 cola alargada a la derecha
 $=0$ simétrica
 <0 cola alargada a la izquierda

KURTOSIS



>0 datos concentrados en la media,
colas pesadas
 $=0$ simétrica
 <0 datos más dispersos y colas más
ligeras

Ejemplo 1.

$$x_1 = [4, 5, 5, 5, 6, 5, 5, 4]$$

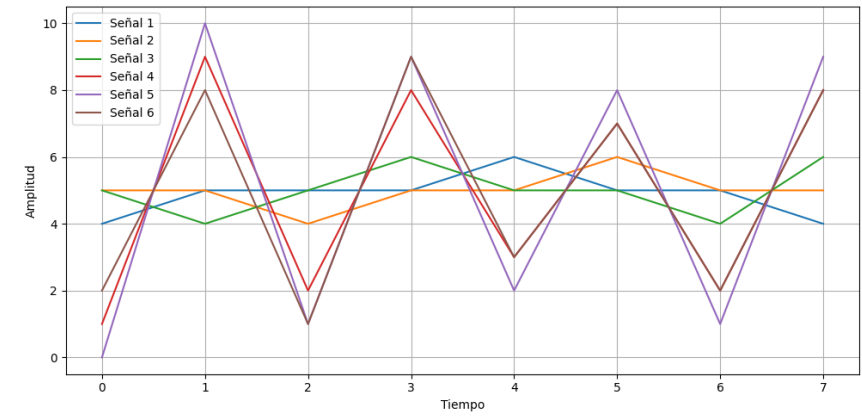
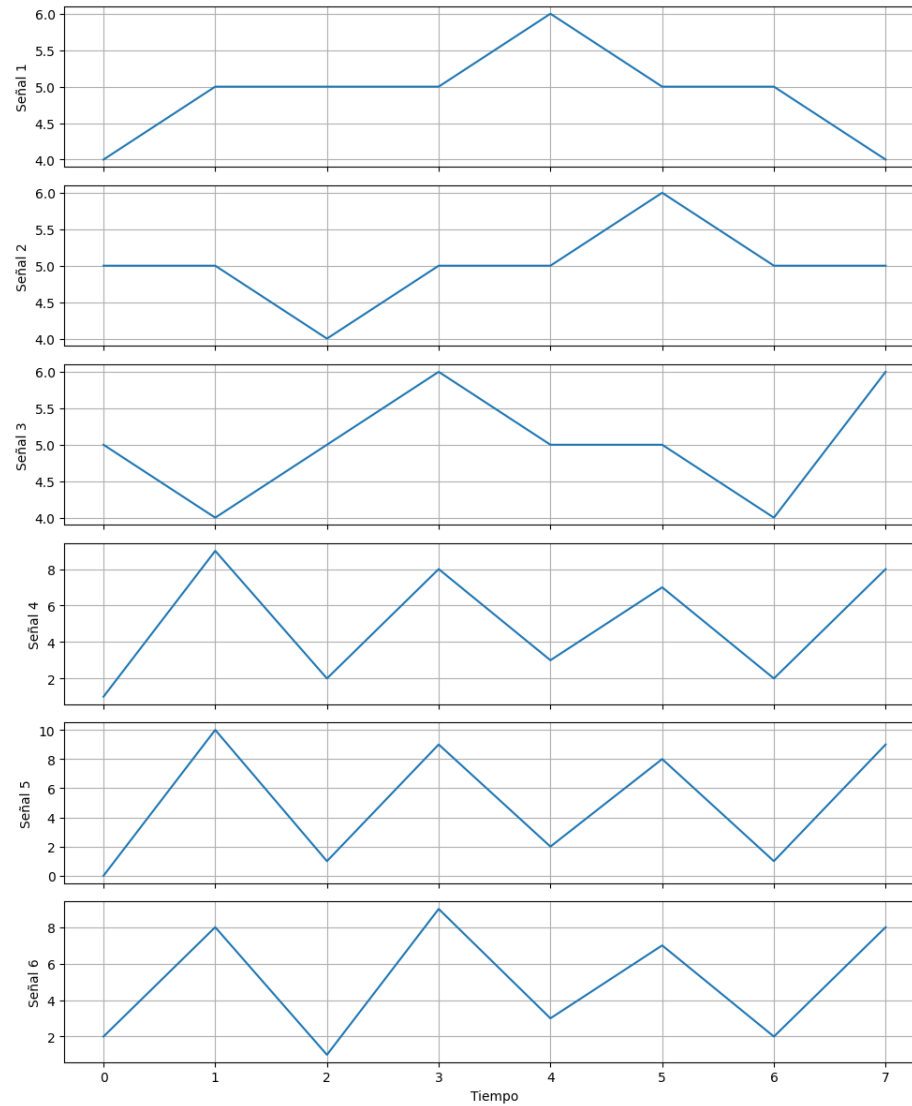
$$x_2 = [5, 5, 4, 5, 5, 6, 5, 5]$$

$$x_3 = [5, 4, 5, 6, 5, 5, 4, 6]$$

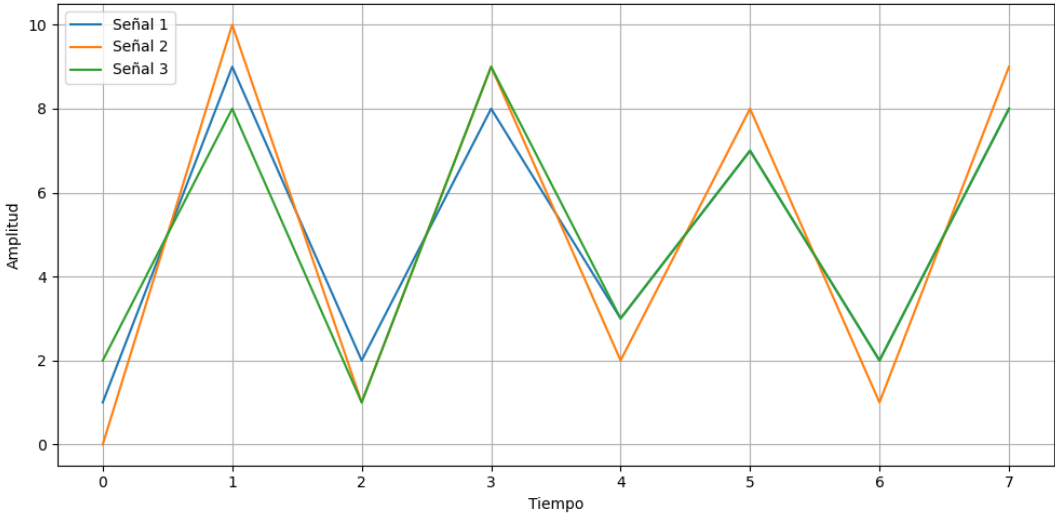
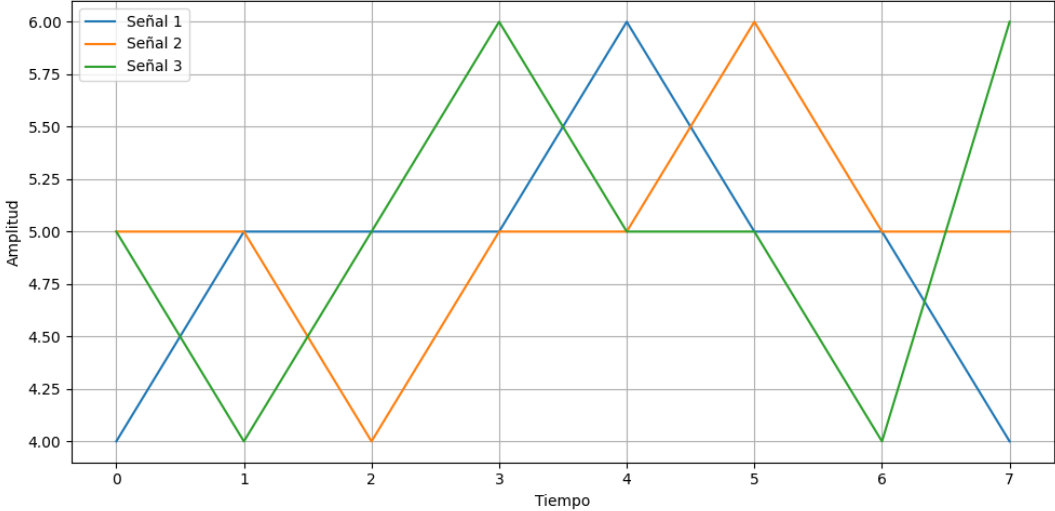
$$x_4 = [1, 9, 2, 8, 3, 7, 2, 8]$$

$$x_5 = [0, 10, 1, 9, 2, 8, 1, 9]$$

$$x_6 = [2, 8, 1, 9, 3, 7, 2, 8]$$



Señal	Media(μ)	Varianza(σ^2)
x_1	4.8	0.36
x_2	5	0.25
x_3	5	0.5
x_4	5	9.5
x_5	5	12
x_6	5	8.75



Ejemplo 2.

$x_1 = [9.5, 10.0, 10.5, 11.0, 11.3, 11.7, 12.0, 12.5, 13.0, 13.5]$

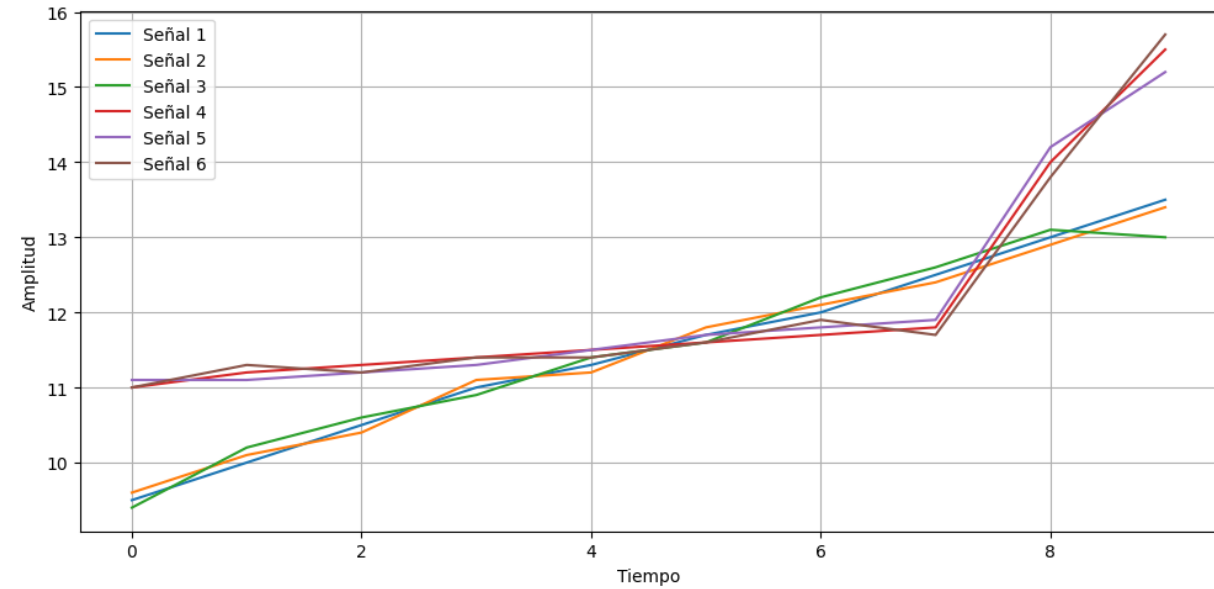
$x_2 = [9.6, 10.1, 10.4, 11.1, 11.2, 11.8, 12.1, 12.4, 12.9, 13.4]$

$x_3 = [9.4, 10.2, 10.6, 10.9, 11.4, 11.6, 12.2, 12.6, 13.1, 13.0]$

$x_4 = [11.0, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 14.0, 15.5]$

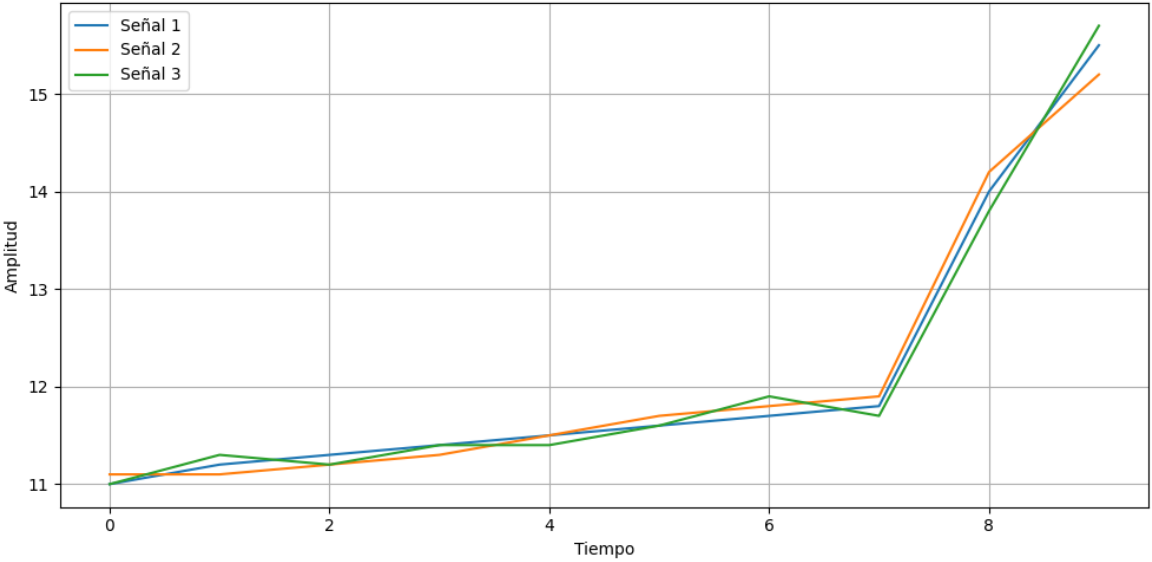
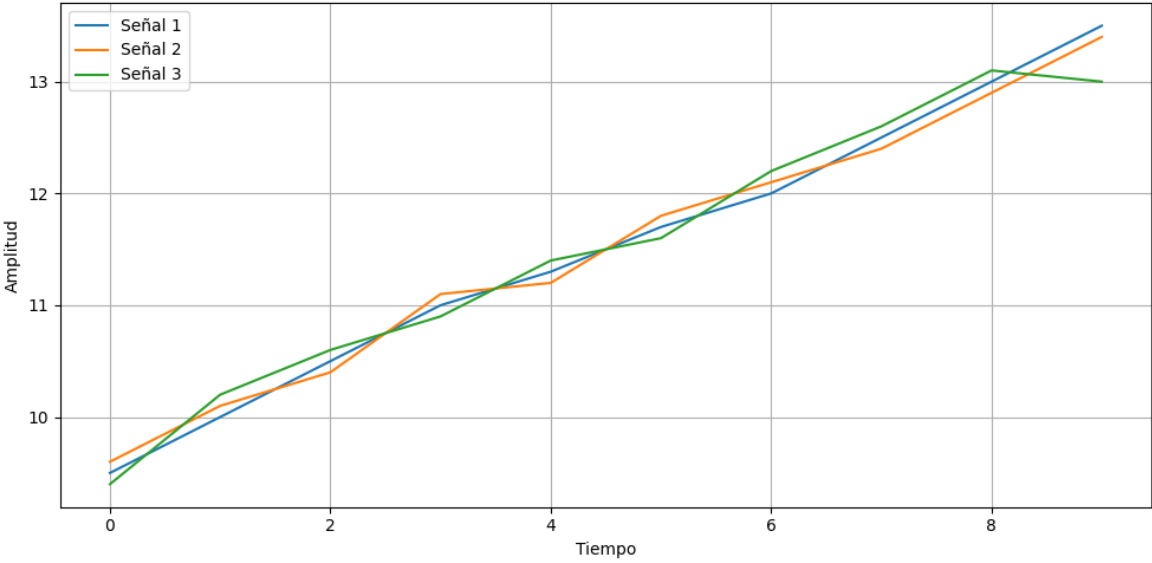
$x_5 = [11.1, 11.1, 11.2, 11.3, 11.5, 11.7, 11.8, 11.9, 14.2, 15.2]$

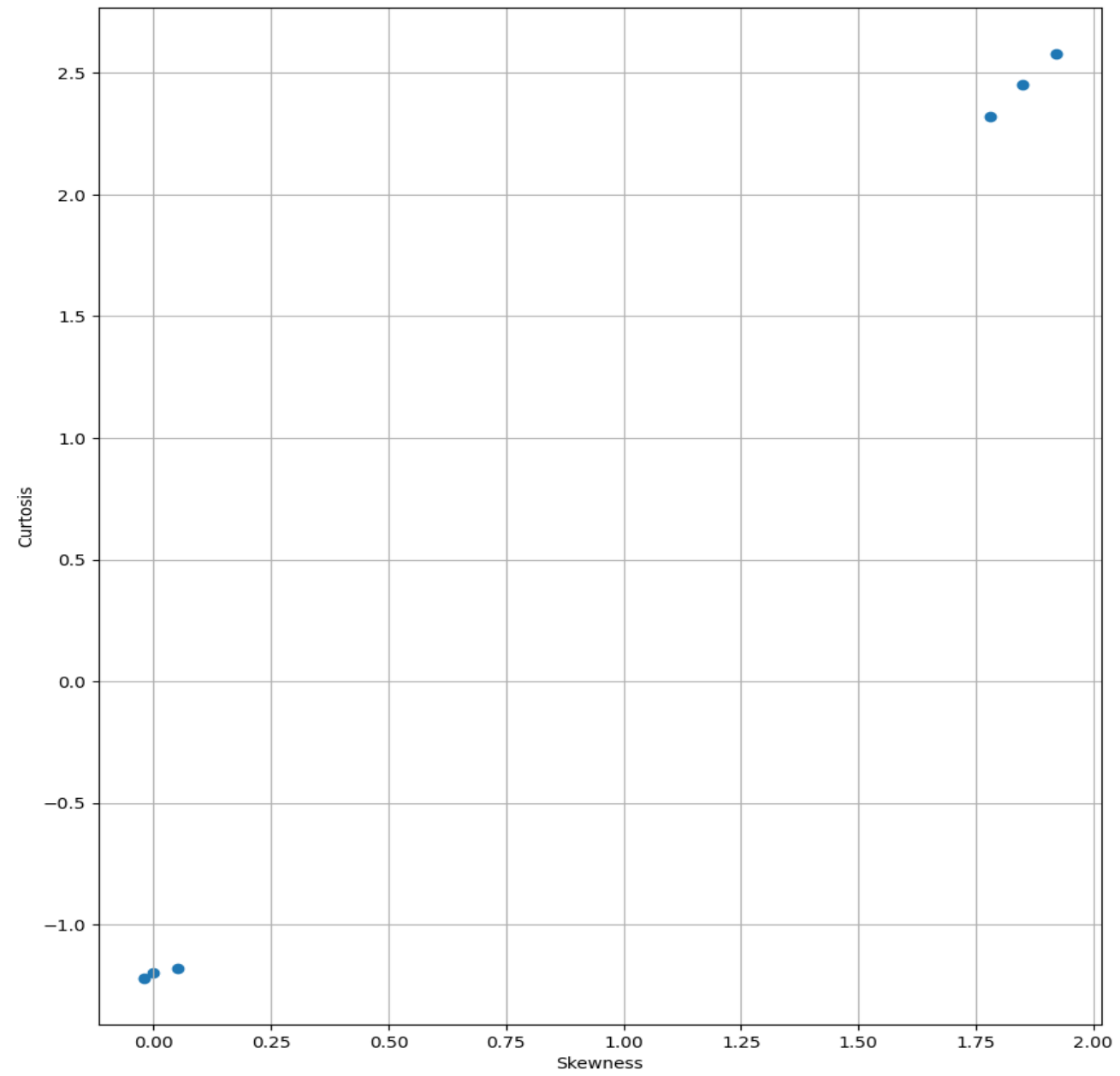
$x_6 = [11.0, 11.3, 11.2, 11.4, 11.4, 11.6, 11.9, 11.7, 13.8, 15.7]$



Señal	Media(μ)	Varianza(σ^2)
x_1	11.5	1.65
x_2	11.5	1.62
x_3	11.5	1.68
x_4	11.5	1.64
x_5	11.52	1.66
x_6	11.48	1.63

Señal	Skewness	Curtosis
x_1	0.00	-1.20
x_2	0.05	-1.18
x_3	-0.02	-1.22
x_4	1.85	2.45
x_5	1.92	2.58
x_6	1.78	2.32





Pero....

¿Qué tan bien funciona aplicarlo de forma global?

- Esto se aplica de forma global, pero también es posible analizarlo de forma local.
 - Ventanas con traslape
 - Ventanas sin traslape
- Se toma una señal de tamaño N y se divide en L ventanas.
- Las características locales capturan estructura espacial o temporal.

$$x = [2, 3, 4, 5, 6, 7, 8, 9, 8, 7, 6, 5, 4, 3, 2, 1]$$

- Sin traslape

- Tamaño de ventana $L=4$
- Traslape: 0%

$$\omega_1 = [2, 3, 4, 5]$$

$$\omega_2 = [6, 7, 8, 9]$$

$$\omega_3 = [8, 7, 6, 5]$$

$$\omega_4 = [4, 3, 2, 1]$$

Ventana	Media(μ)	Varianza(σ^2)
ω_1	3.5	1.25
ω_2	7.5	1.25
ω_3	6.5	1.25
ω_4	2.5	1.25

$$x = [2, 3, 4, 5, 6, 7, 8, 9, 8, 7, 6, 5, 4, 3, 2, 1]$$

- Con traslape

- Tamaño de ventana $L = 4$
- Traslape: 50%
- Paso: 2

$$\omega_1 = [2, 3, 4, 5]$$

$$\omega_2 = [4, 5, 6, 7]$$

$$\omega_3 = [6, 7, 8, 9]$$

$$\omega_4 = [8, 9, 8, 7]$$

$$\omega_5 = [8, 7, 6, 5]$$

$$\omega_6 = [6, 5, 4, 3]$$

$$\omega_7 = [4, 3, 2, 1]$$

Ventana	Media(μ)	Varianza(σ^2)
ω_1	3.5	1.25
ω_2	5.5	1.25
ω_3	7.5	1.25
ω_4	8.0	0.25
ω_5	6.5	1.25
ω_6	4.5	1.25
ω_7	2.5	1.25

- ¿Qué hago con todas estas ventanas?

1. Concatenación directa

$$f = [f_1, f_2, f_3, \dots]$$

$$f_{st} = [3.5, 1.25, 7.25, 1.25, 6.25, 1.25, 2.5, 1.25]$$

$$f_{ct} = [3.25, 1.25, 5.5, 1.25, 7.5, 1.25, 8.0, 0.25, 6.5, 1.25, 4.5, 1.25, 2.5, 1.25]$$

Ventaja: Conserva información local

Desventaja: Dimensión crece rápido

2. Estadísticas sobre estadísticas.

Para cada característica a lo largo de ventanas:

$$\mu_\mu, \sigma_\mu^2, \mu_{\sigma^2}, \sigma_{\sigma^2}^2, \dots$$

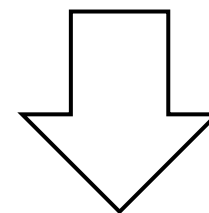
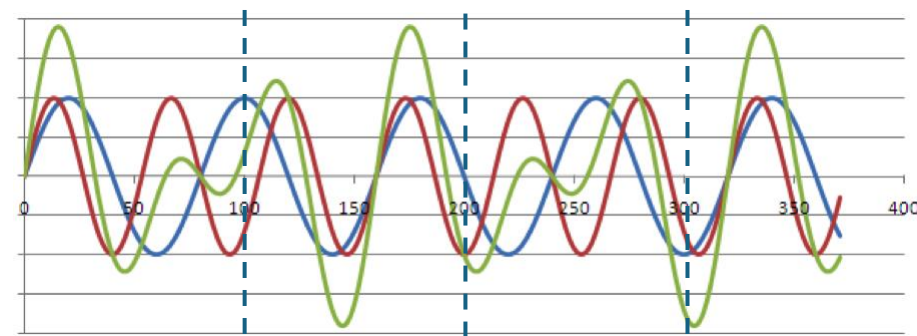
$$f_{st} = [5.0, 4.25, \dots]$$

$$f_{ct} = [5.42, 3.60, \dots]$$

3. Pooling.

$$\max_k(\sigma_k^2)$$

Global	Tendencia general
Local	Cambios y Estructura
Local + Pooling	Estructura y Robustez

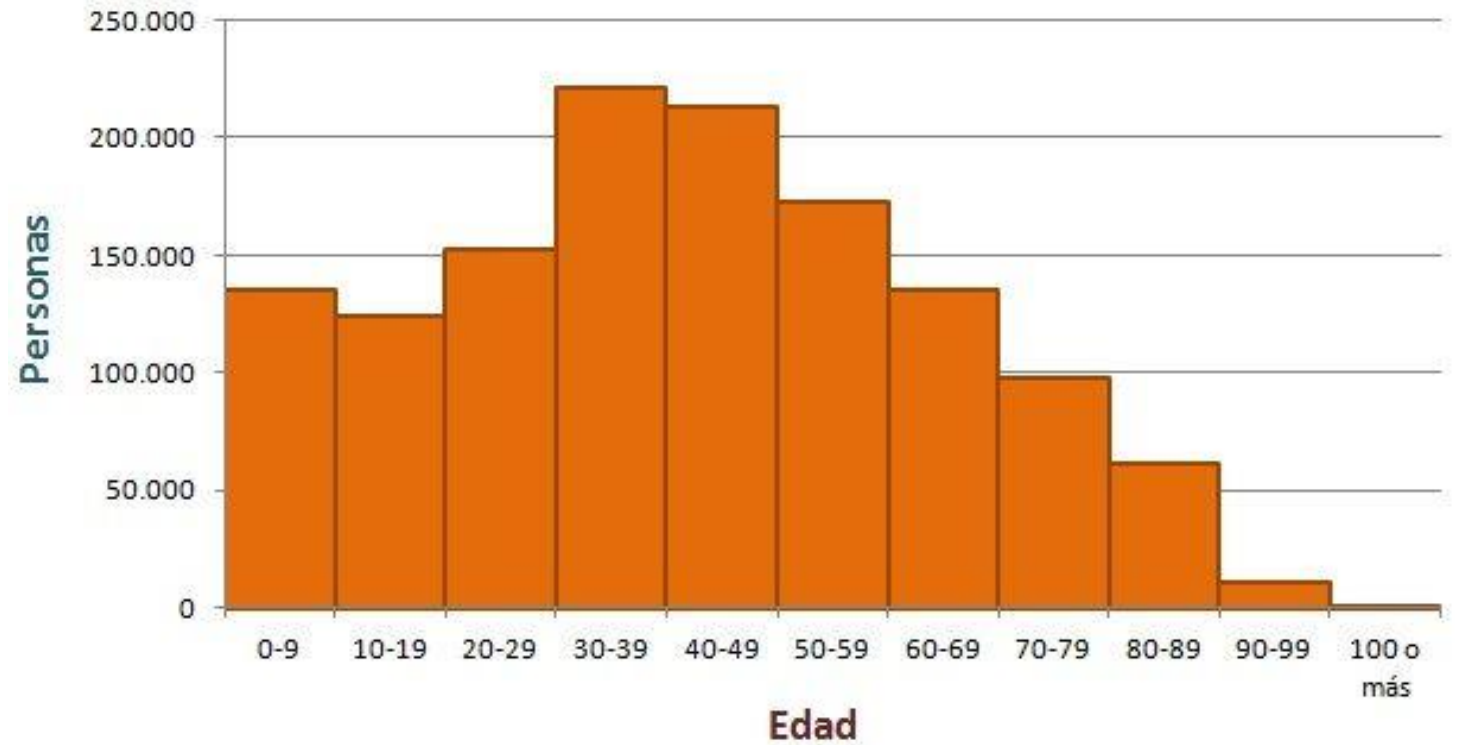


$$f_k = [f_1, f_2, f_3, \dots]$$

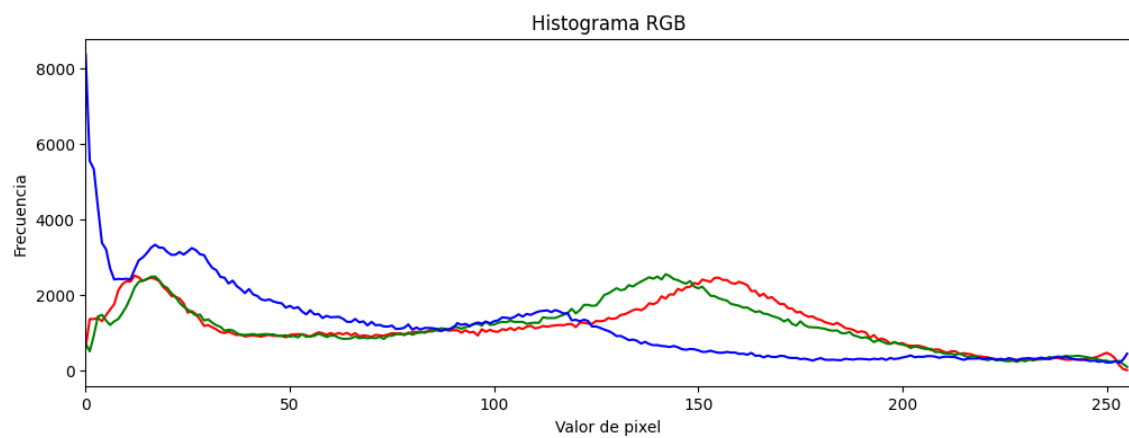
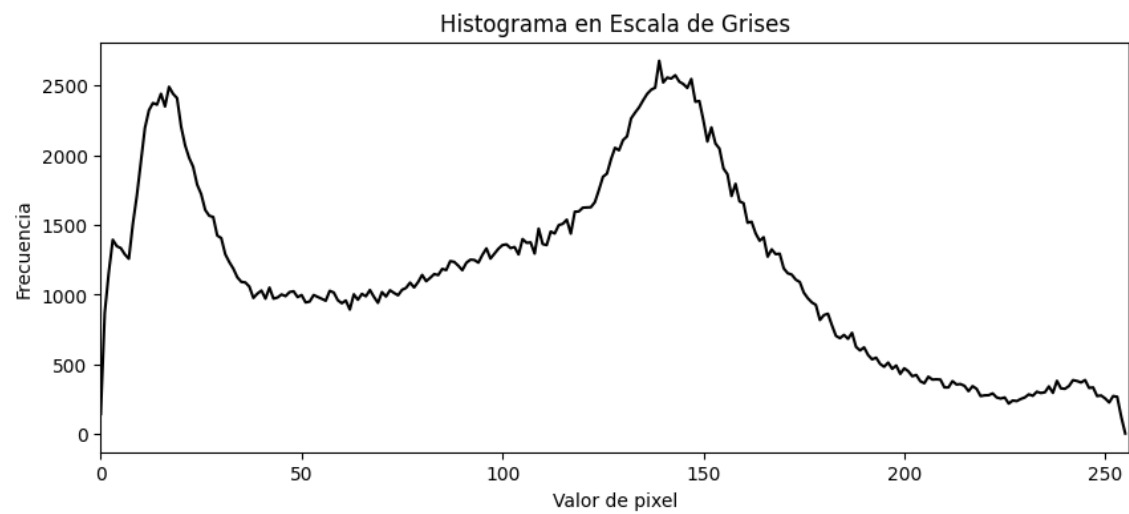
¿Cuándo no
vale la pena
usar
ventanas
locales?



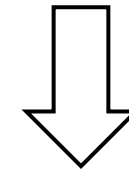
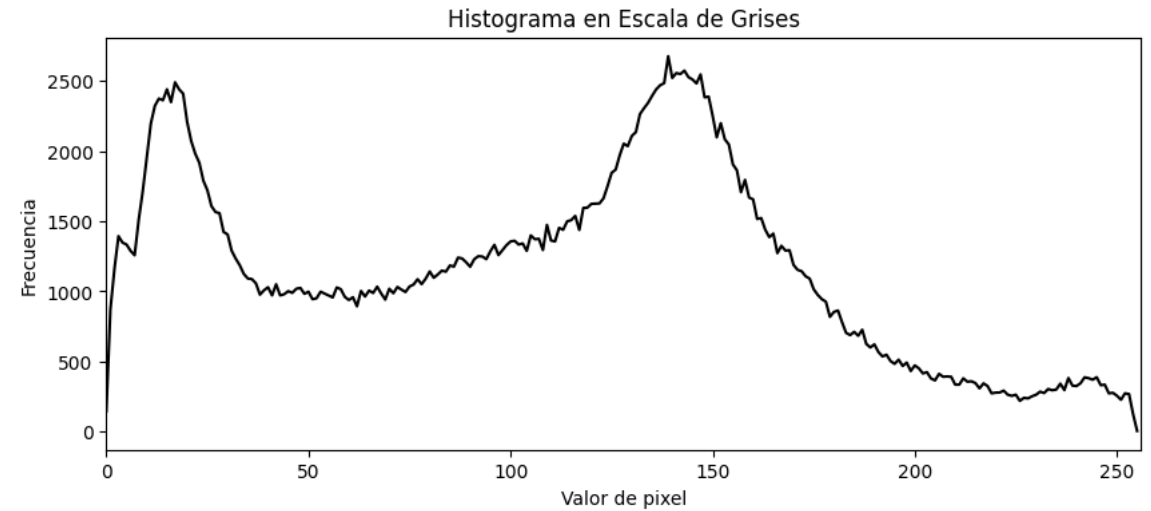
Histograma



Un histograma describe cómo se distribuyen los valores de una señal o dato, sin importar su posición exacta.



-
- El histograma **no se usa como gráfica**, sino como **vector de características**.



- No solo se puede aplicar a imágenes, también a audio y otro tipo de señales.

$$h = [146, 870, 1159, 1394, 1348, 1334, 1289, \dots]$$

Ventajas

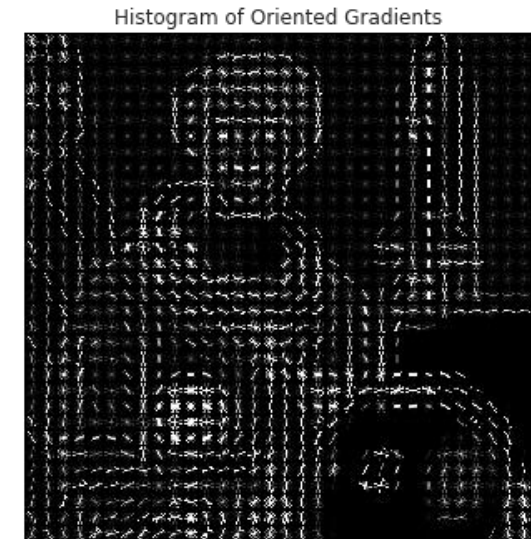
- Simples
- Robustos al ruido
- Invariantes a permutaciones/traslaciones
- Dimensión controlable
- Funcionan bien con pocos datos

Desventajas

- Pierden información espacial/temporal
- Sensibles al número de bins
- No capturan relaciones entre valores vecinos

- Para más adelante...

HOG



Para la próxima clase...

- Características Geométricas y Descriptores Clásicos

