

UNIVERSIDAD
NACIONAL
DE COLOMBIA

VISIÓN ARTIFICIAL

JOHN W. BRANCH

PROF. TITULAR

DEPARTAMENTO DE CIENCIAS DE COMPUTACIÓN Y DE LA DECISIÓN

DIRECTOR DEL GRUPO GIDIA

ALBERTO M. CEBALLOS / JAIRO A. RODRIGUEZ

ASISTENTE DE DOCENCIA / MONITOR

Nota: Este material se ha adaptado con base en el material de los profesores Domingo Mery (U. de Chile), María Patricia Trujillo (Univalle), Ginés García (U. de Murcia) y Nicolas Fernández (U. de Córdoba)

EN LA CLASE DE HOY ...

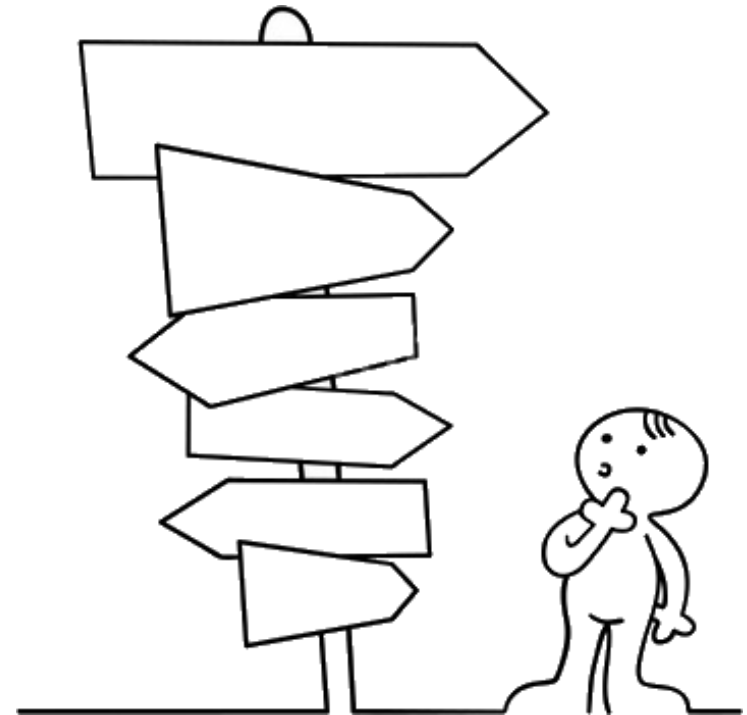
🦋 EXTRACCIÓN DE CARACTERÍSTICAS

🌀 Selección de Características

- 🌀 Introducción
- 🌀 El Problema de la Dimensionalidad
- 🌀 Métodos de Selección de Características

🌀 Reconocimiento de Patrones

- 🌀 Introducción
- 🌀 Métodos Supervisados y no Supervisados
- 🌀 K-NN
- 🌀 Clustering
- 🌀 Redes Neuronales
- 🌀 Máquinas de Vectores Soporte (SVM)

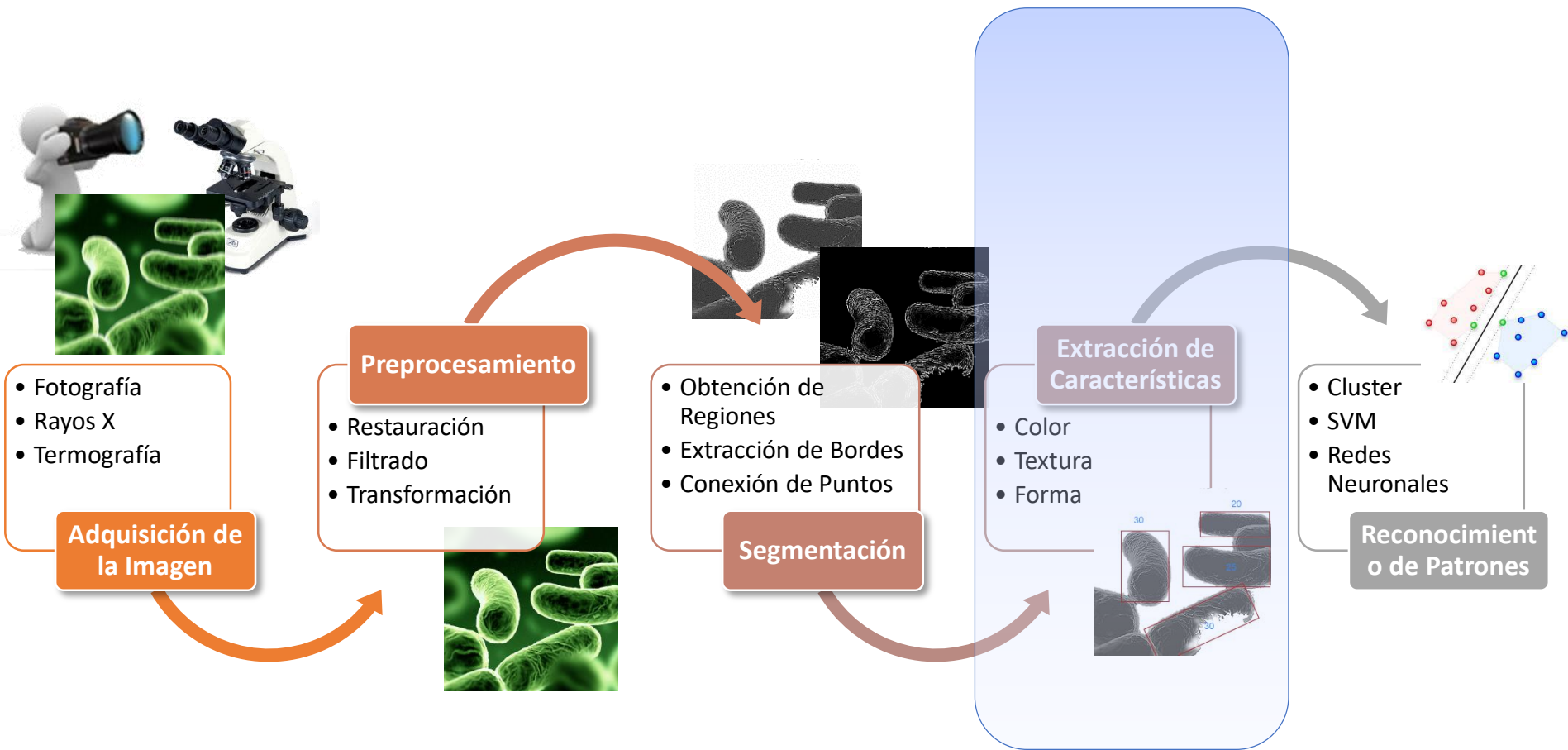


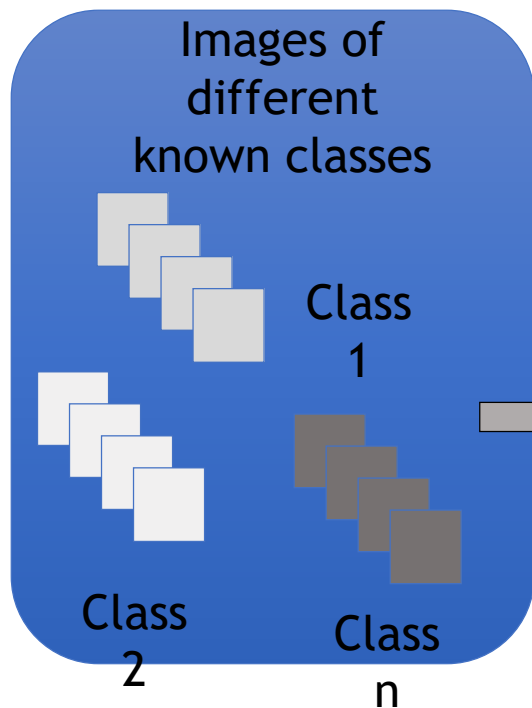
Extracción de Características



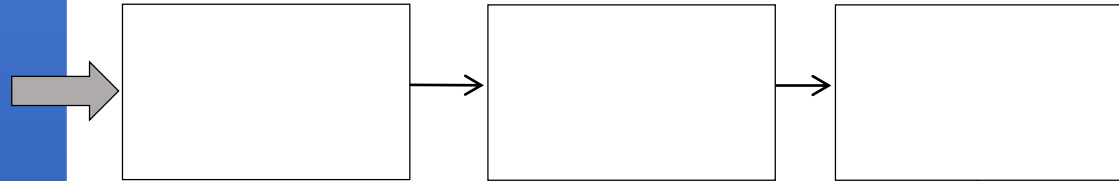
A MODO DE RESUMEN ...

ETAPAS DE UN SISTEMA DE VISIÓN ARTIFICIAL





FROM FEATURE EXTRACTION TO CLASSIFICATION



Stage 1:

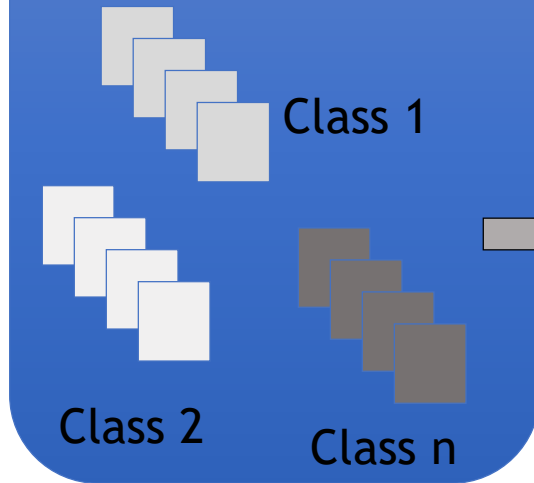
Stage 2:

Input :

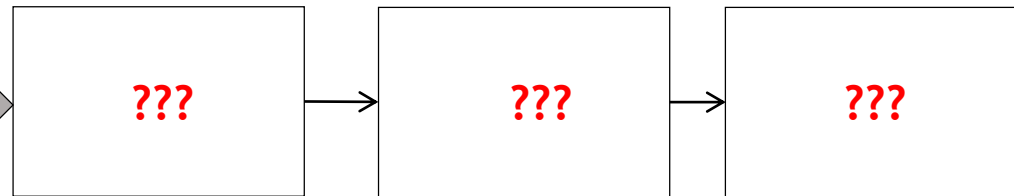


Output:

Images of different
known classes



FROM FEATURE EXTRACTION TO CLASSIFICATION



Stage 1: ???

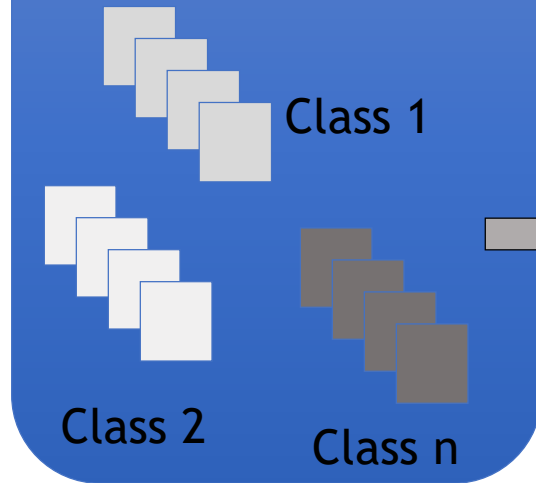
Stage 2: ???

Input : ???

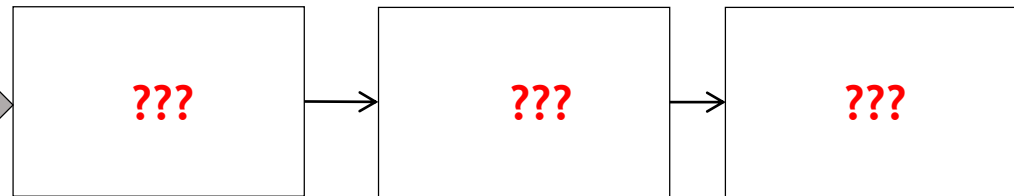


Output: ???

Images of different
known classes

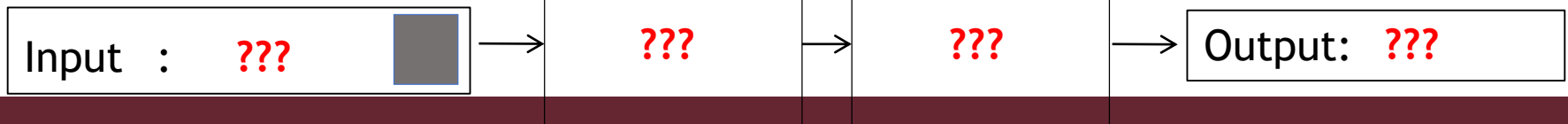


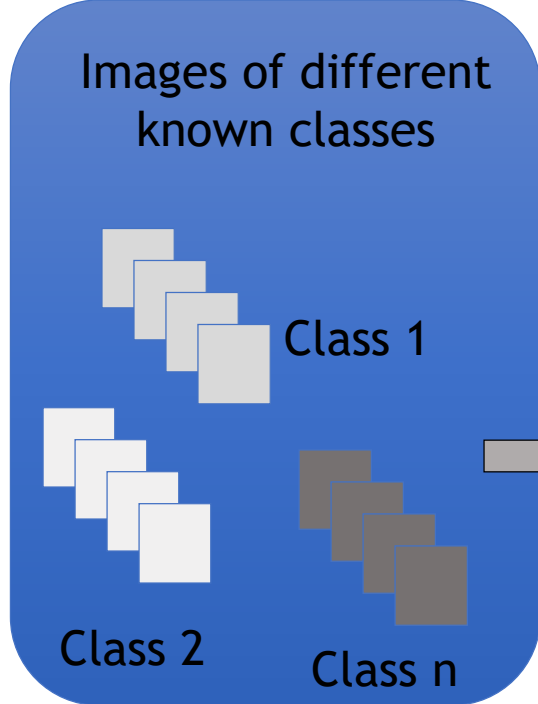
FROM FEATURE EXTRACTION TO CLASSIFICATION



Stage 1: **TRAINING**

Stage 2: **TESTING**





FROM FEATURE EXTRACTION TO CLASSIFICATION



Stage 1: **TRAINING**

Stage 2: **TESTING**



Images of different
known classes

Class 1

Class 2

Class n

FROM FEATURE EXTRACTION TO CLASSIFICATION

Feature
Extraction

Feature
Selection

Design of
Classifier

Stage 1: **TRAINING**

Stage 2: **TESTING**

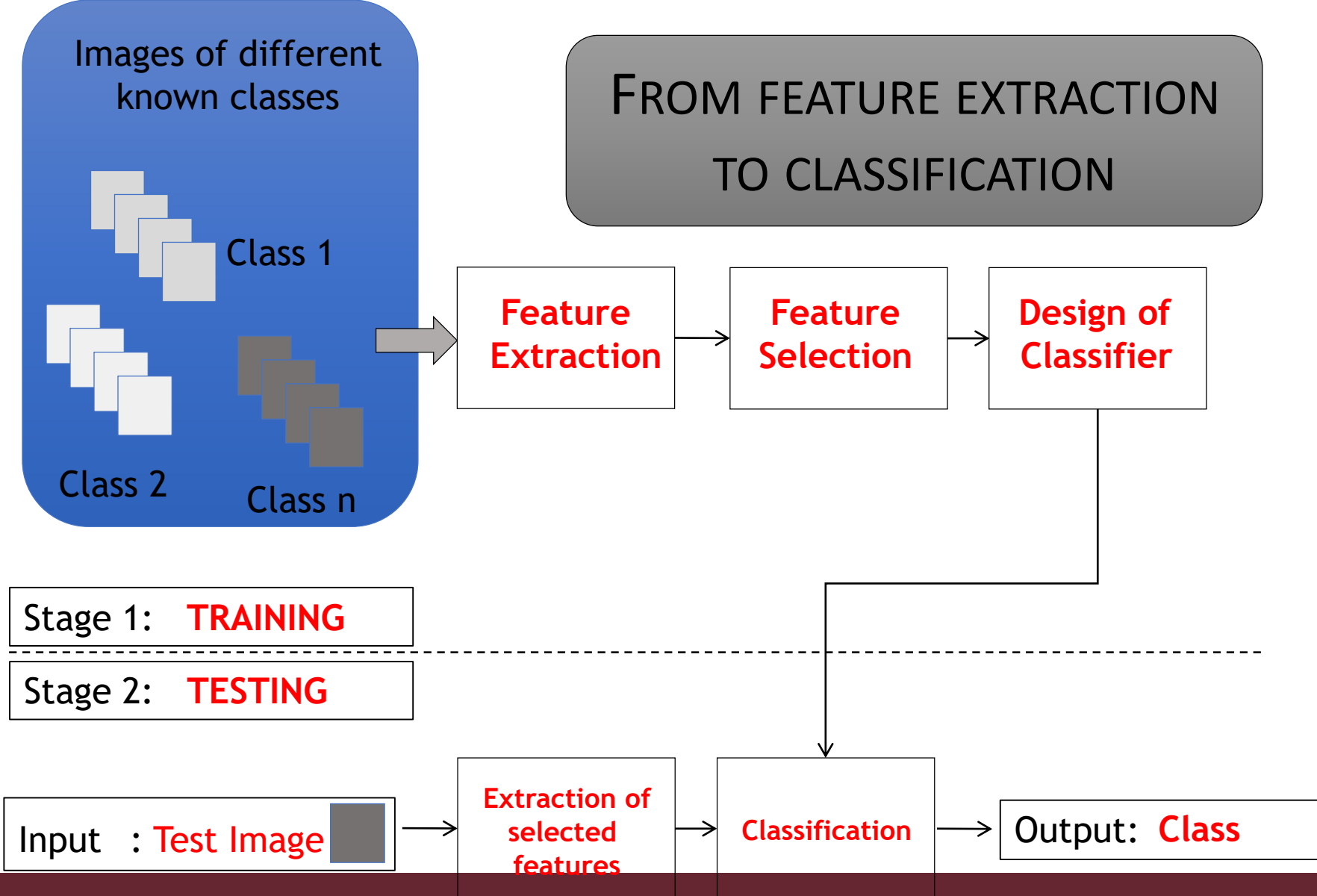
Input : **Test Image**

???

???

Output: **Class**

FROM FEATURE EXTRACTION TO CLASSIFICATION

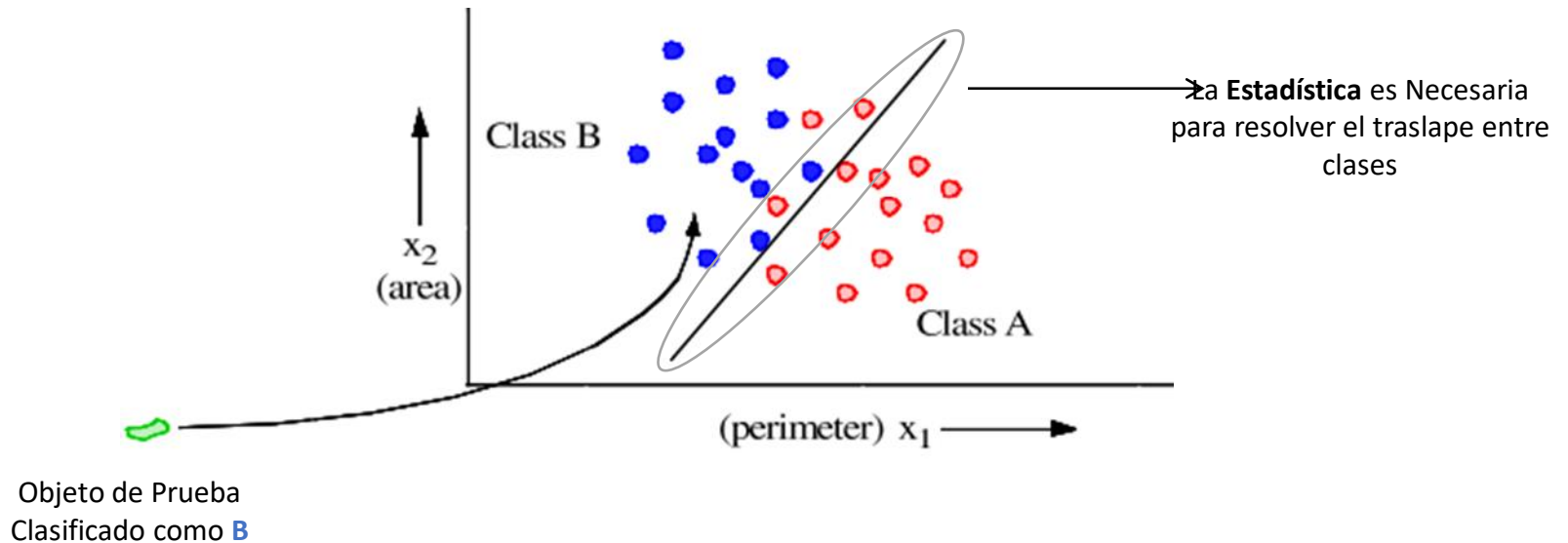
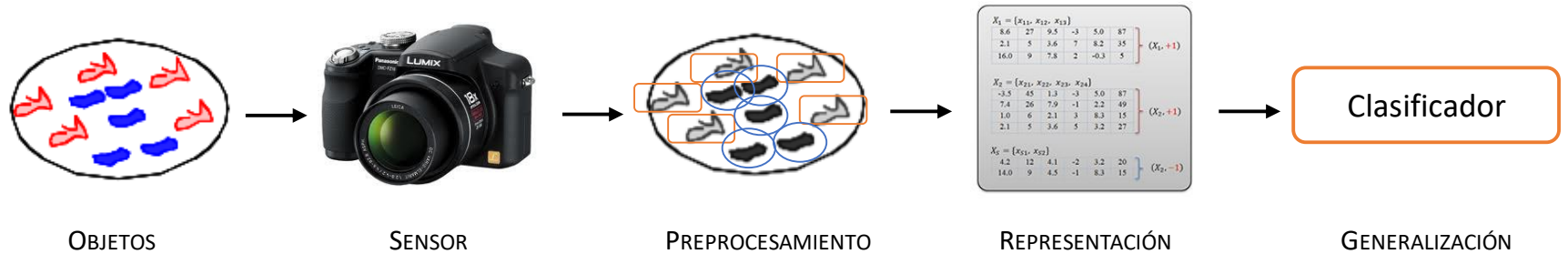


Selección de Características

INTRODUCCIÓN

Selección de Características

INTRODUCCIÓN



Selección de Características

INTRODUCCIÓN – REGLA DE BAYES

- El **Teorema de Bayes** que expresa la probabilidad a posteriori de un evento aleatorio A dado x en términos de la distribución de probabilidad condicional y la probabilidad marginal.

- Para un problema de 2 clases:

$$p(A|x) > p(B|x) \rightarrow A \text{ else } B$$

$$\frac{p(x|A) p(A)}{p(x)} > \frac{p(x|B) p(B)}{p(x)} \rightarrow A \text{ else } B$$

- $S(x) = p(x|A)P(A) - p(x|B)P(B) > 0 \rightarrow A \text{ else } B$

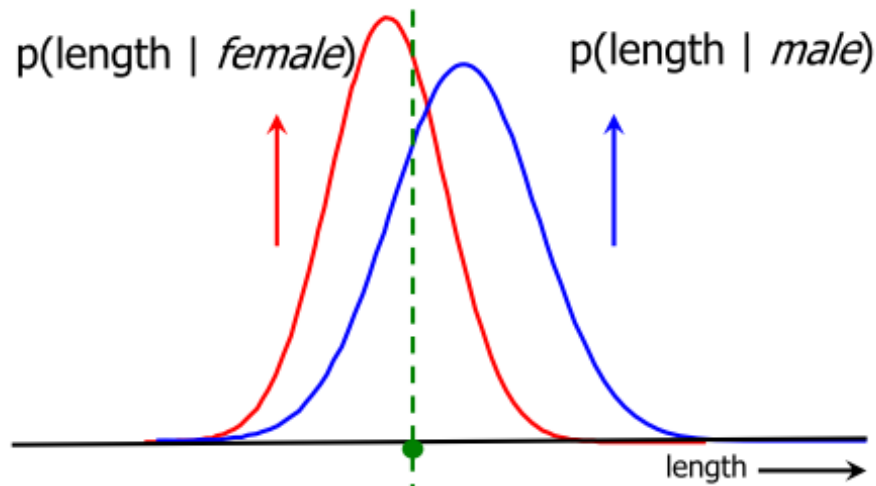
$$p(x|A) p(A) > p(x|B) p(B) \rightarrow A \text{ else } B$$

- Para múltiples clases : $\text{Clase}(x) = \arg \max_{\omega} (p(x|\omega)P(\omega))$

Selección de Características

🦋 INTRODUCCIÓN – REGLA DE BAYES

🎯 Cuál es el género de alguien con esta altura?



Bayes: $\begin{cases} p(female | length) = p(length | female) p(female) / p(length) \\ p(male | length) = p(length | male) p(male) / p(length) \end{cases}$

Selección de Características

INTRODUCCIÓN – REGLA DE BAYES

Regla de Clasificación de Bayes:

$$p(\textit{female} \mid \textit{length}) > p(\textit{male} \mid \textit{length}) \rightarrow \textit{female} \text{ else } \textit{male}$$

Bayes:

$$\frac{p(\textit{length} \mid \textit{female}) p(\textit{female})}{p(\textit{length})} > \frac{p(\textit{length} \mid \textit{male}) p(\textit{male})}{p(\textit{length})}$$

$$p(\textit{length} \mid \textit{female}) p(\textit{female}) > p(\textit{length} \mid \textit{male}) p(\textit{male}) \rightarrow \textit{female} \text{ else } \textit{male}$$

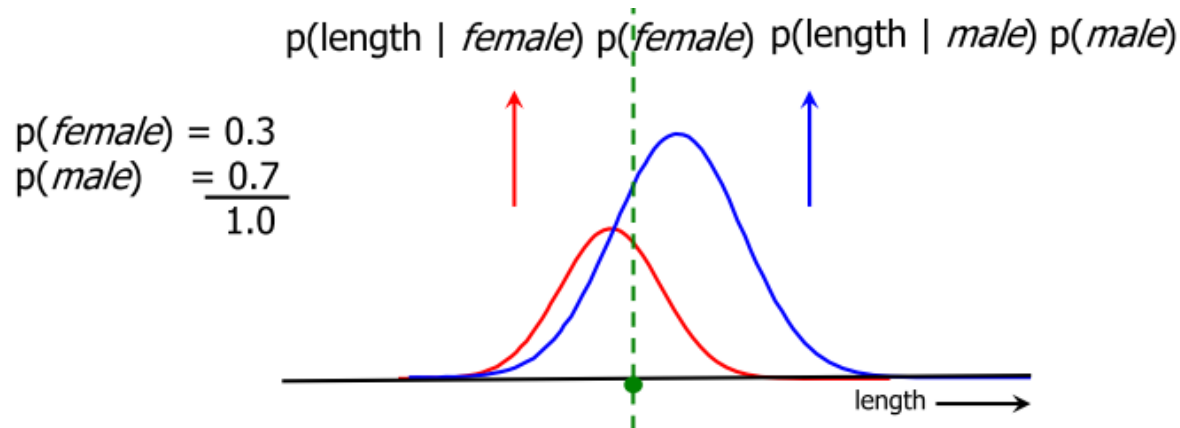
pdf estimated from training set

class prior probabilities
known, guessed or estimated

Selección de Características

🦋 INTRODUCCIÓN – REGLA DE BAYES

- 🌀 Cuál es el género de alguien con esta altura?



What is the gender of somebody with this length?

Bayes: $\begin{cases} p(\text{female} \mid \text{length}) = p(\text{length} \mid \text{female}) p(\text{female}) / p(\text{length}) \\ p(\text{male} \mid \text{length}) = p(\text{length} \mid \text{male}) p(\text{male}) / p(\text{length}) \end{cases}$

¿Es la altura una buena característica para diferenciar hombres de mujeres?

Selección de Características

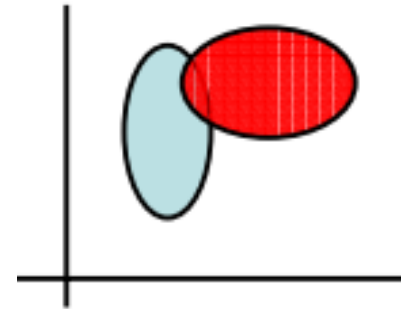
BUENAS REPRESENTACIONES

Selección de Características

🦋 BUENAS REPRESENTACIONES

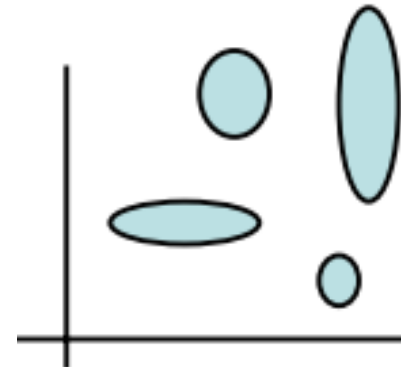
🌀 Específicas a las Clases

Diferentes clases deberían estar representados en diferentes posiciones en el espacio de representación.



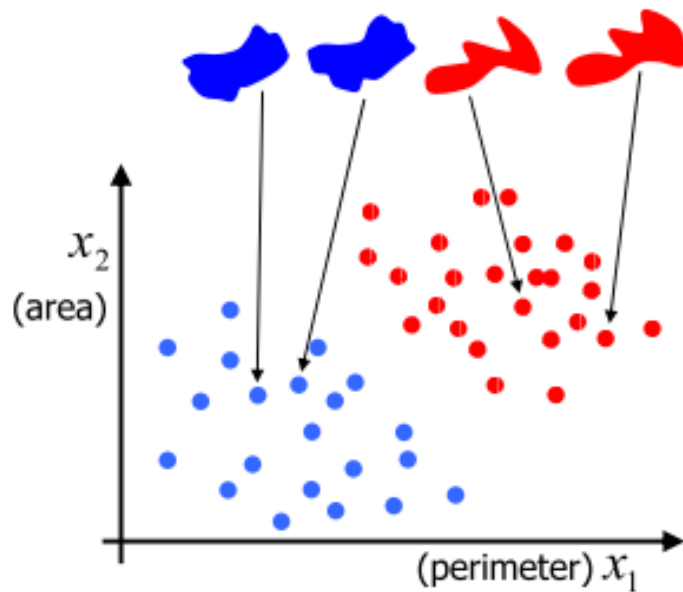
🌀 Deben ser Compactas

Cada clase debe estar representada en un pequeño conjunto de dominios finitos.



Selección de Características

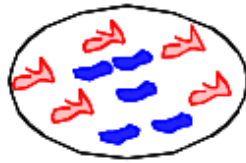
🦋 BUENAS REPRESENTACIONES



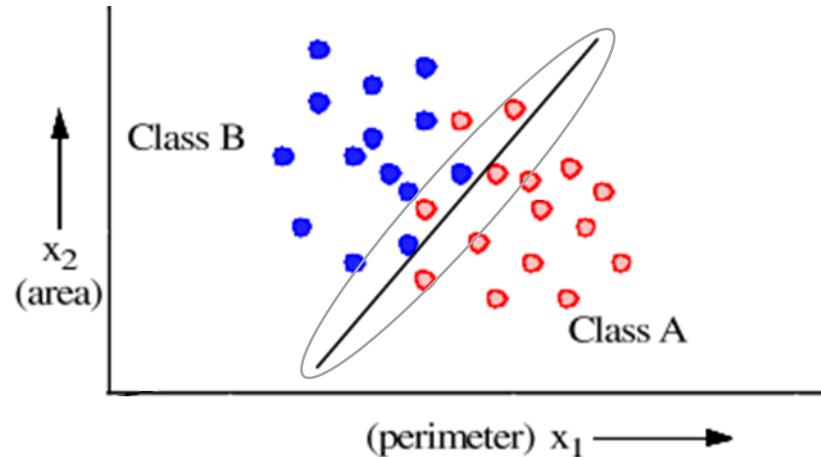
Objetos **SIMILARES** están cerca
y
Objetos **DISÍMILES** son distantes

Selección de Características

BUENAS REPRESENTACIONES



OBJETOS DE
ENTRENAMIENTO



Conocimiento del Problema:

→ Buenas Características

→ Clases Separables (o Casi Separables)

Falta de Conocimiento:

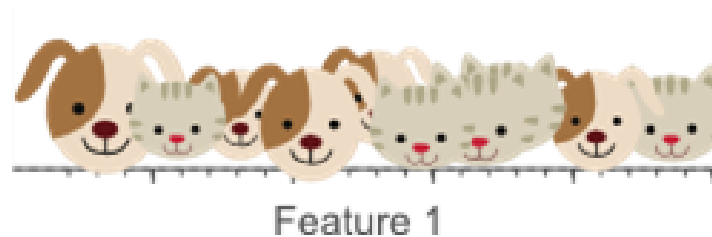
→ Malas Características (Muchas, además)

→ Clases difícilmente separables

Selección de Características

🦋 EL PROBLEMA DE LA DIMENSIONALIDAD

- 🌀 Clasificación de gatos y perros con 10 instancias ... usando un clasificador lineal
- 🌀 Podemos empezar por una sola característica, por ejemplo, el color promedio de 'rojo' en la imagen:



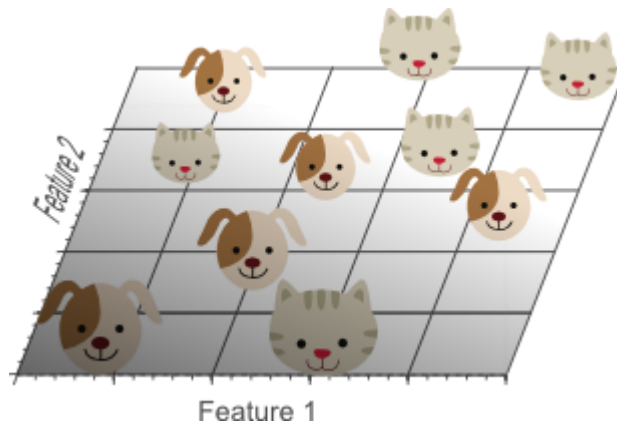
Una sola característica no resulta en una separación perfecta de nuestros datos de entrenamiento.

Tomado de Computer vision for dummies: <http://www.visiondummys.com/2014/04/curse-dimensionality-affect-classification/>

Selección de Características

🦋 EL PROBLEMA DE LA DIMENSIONALIDAD

- 🌀 Clasificación de gatos y perros con 10 instancias ...
- 🌀 Decidimos añadir otra característica, por ejemplo, el color promedio de "verde" en la imagen:



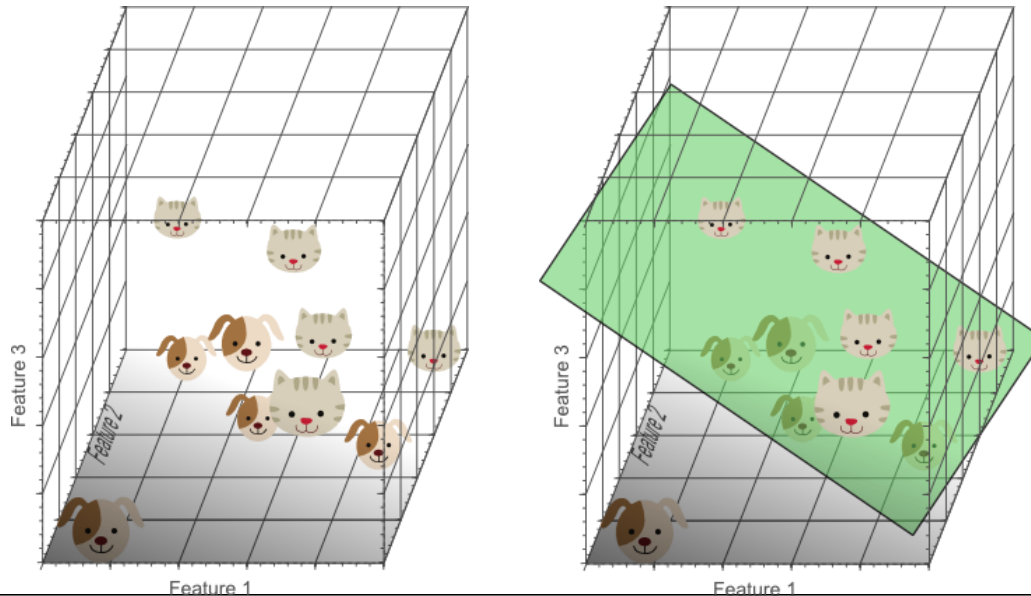
Aún agregando una segunda característica **el problema NO es linealmente separable**: No hay una sola línea que pueda separar todos los gatos de todos los perros en este ejemplo.

Selección de Características



EL PROBLEMA DE LA DIMENSIONALIDAD

- Clasificación de gatos y perros con 10 instancias ...
- Finalmente, decidimos añadir una característica más: el color promedio de “azul” en la imagen:

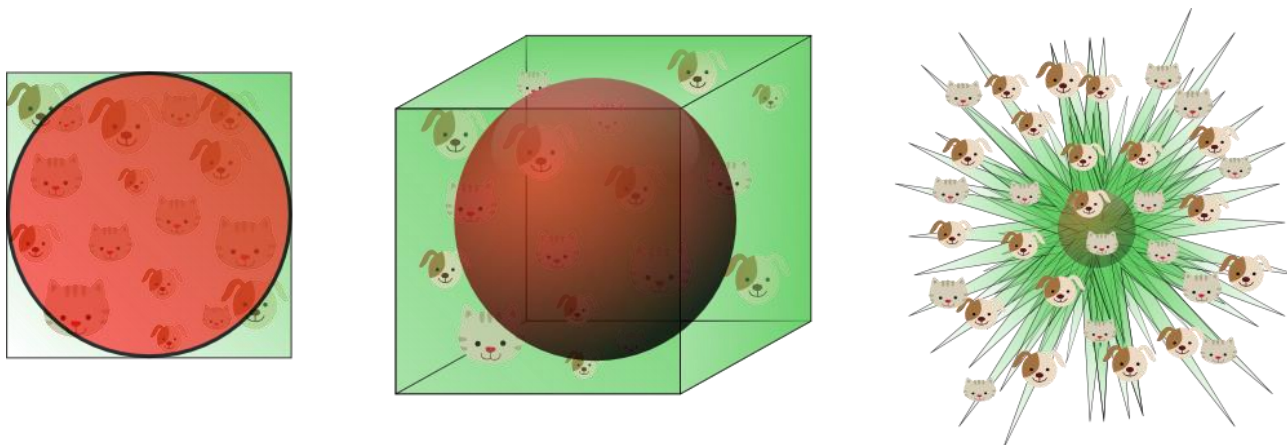


Agregar una tercera característica resulta en un problema separable: existe un plano que separa perfectamente perros de gatos

Selección de Características

🦋 EL PROBLEMA DE LA DIMENSIONALIDAD

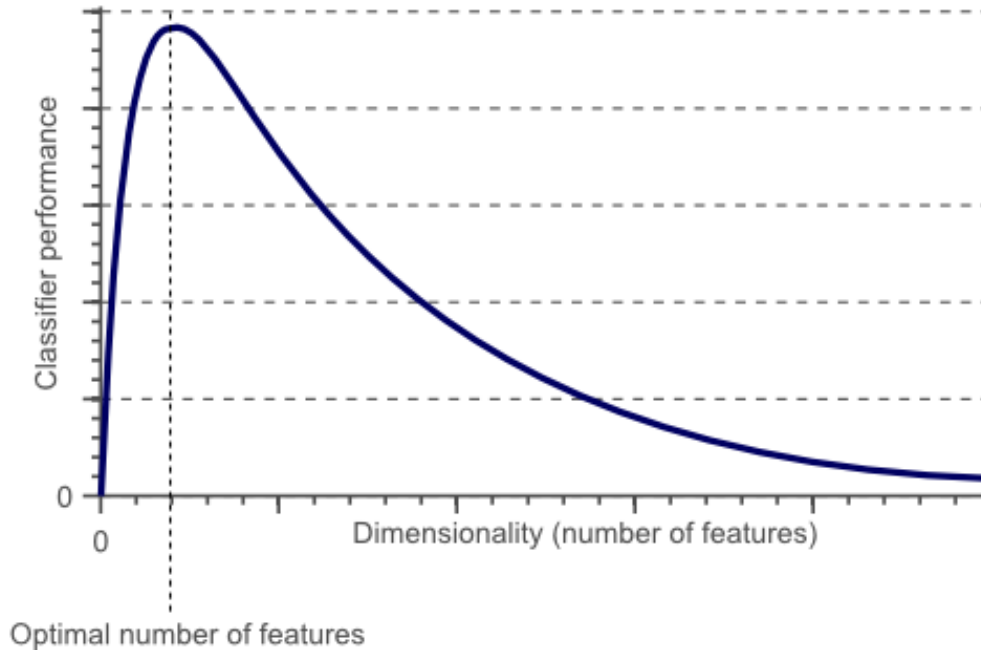
- 🌀 Clasificación de gatos y perros con 10 instancias ...
- 🌀 Sin embargo, tenga en cuenta cómo la densidad de las muestras de entrenamiento disminuye exponencialmente cuando aumentamos la dimensionalidad del problema.



Como la dimensionalidad aumenta, un mayor porcentaje de los datos se ubica en las esquinas del espacio: se dispersan los datos

Selección de Características

EL PROBLEMA DE LA DIMENSIONALIDAD



La **MALDICIÓN DE LA DIMENSIONALIDAD**: cuando aumenta la dimensionalidad, el volumen del espacio aumenta exponencialmente haciendo que los datos disponibles se vuelven dispersos

Selección de Características

¿POR QUÉ SELECCIONAR CARACTERÍSTICAS?

Selección de Características

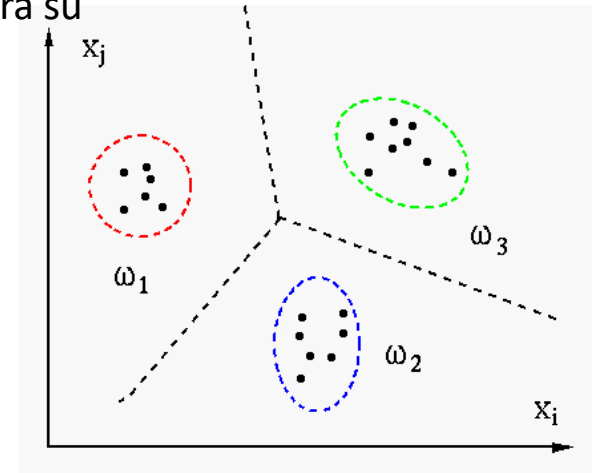
🚀 EL PORQUE ...

🎯 ¿Por qué reducir la dimensionalidad del problema?

- 🌀 **Precisión** → Un conjunto pequeño de características define un espacio menor
- 🌀 **Velocidad** → Menos características exigen menos mediciones
- 🌀 **Memoria** → Menos características menos memoria para su almacenamiento

🎯 ¿Cómo reducir el número de características utilizado?

- 🌀 Eliminar características “ruidosas”
- 🌀 Medición del costo



🦋 APROXIMACIONES A LA REDUCCIÓN DE LA DIMENSIONALIDAD

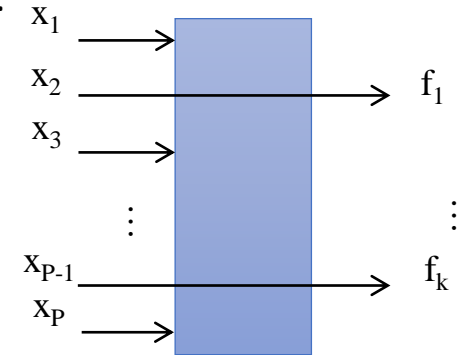
🌀 Selección de Características: **Seleccione** k características de P:

🌀 Ventajas

- ✓ Disminuye las mediciones
- ✓ Fácil interpretación

🌀 Desventajas

- ✓ Métodos costosos computacionalmente
- ✓ Son aproximaciones



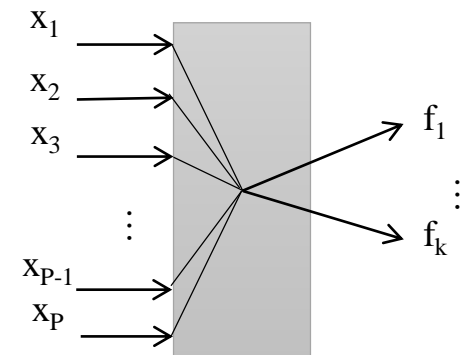
🌀 Transformación de Características: **Transforme** P características en k:

🌀 Ventajas

- ✓ Suelen ser menos costosas
- ✓ Pueden ser No-Lineales

🌀 Desventajas

- ✓ Necesita todas las características
- ✓ Sub-óptimas dependientes de un criterio



Selección de Características

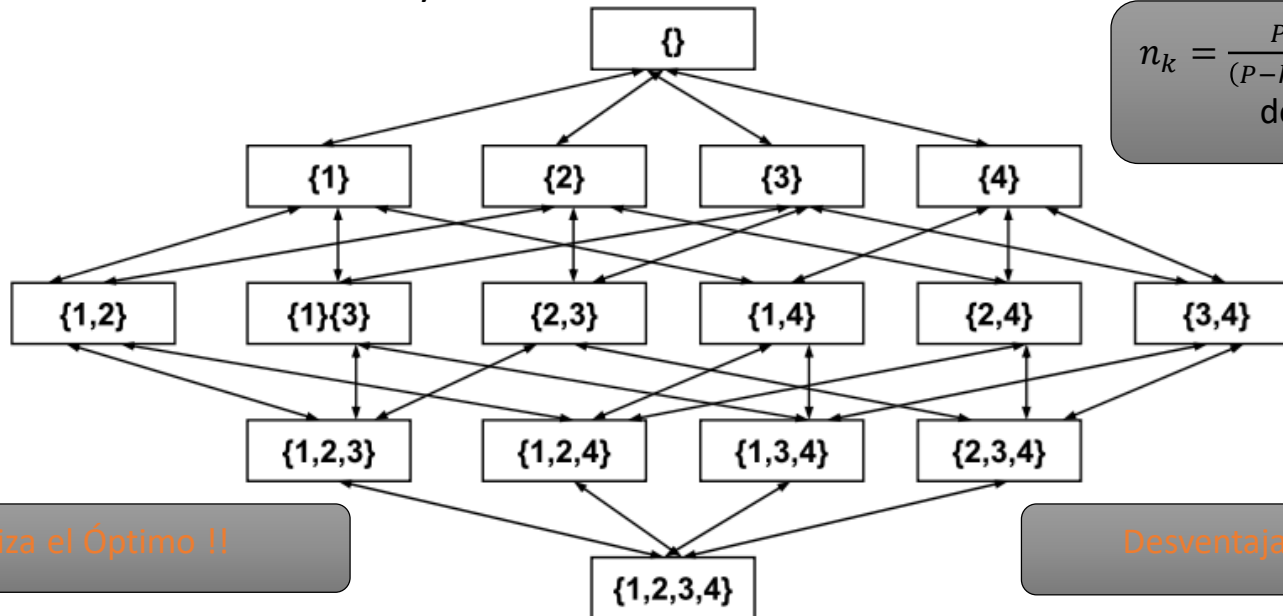
MÉTODOS DE SELECCIÓN

Selección de Características



MÉTODOS DE SELECCIÓN

- Estos métodos seleccionan un **subconjunto** de k de características optimizando un criterio determinado (o función objetivo).
- La solución trivial es **BÚSQUEDA EXHAUSTIVA**: evalúa **TODO**s los posibles subconjuntos de características. ¿Cuántos hay ? ...



$$n_k = \frac{P!}{(P-k)!k!} \text{ Subconjuntos de tamaño } k$$

Garantiza el Óptimo !!

Desventaja: Complejidad

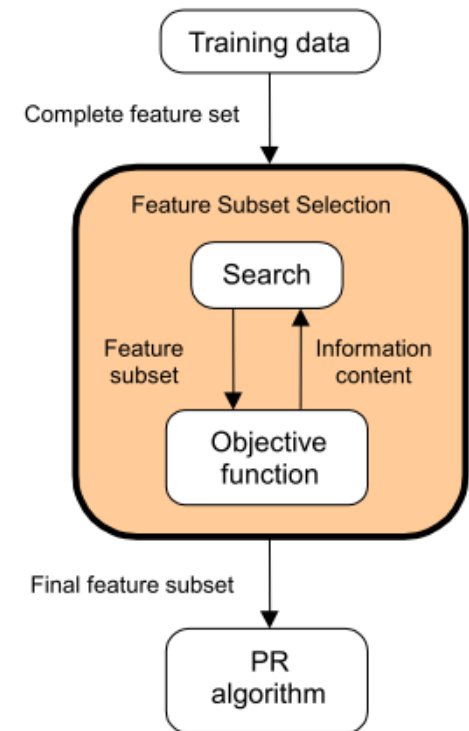
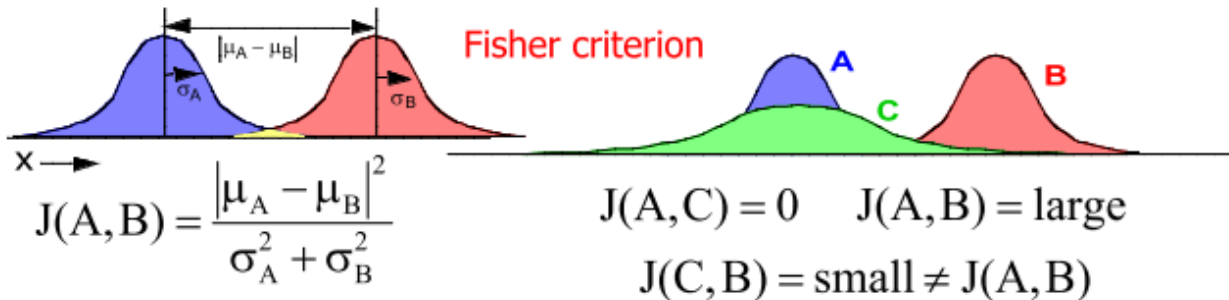
Selección de Características

🦋 MÉTODOS DE SELECCIÓN

🌀 **FUNCIÓN OBJETIVO:** evalúa subconjuntos candidatos y devuelve una medida de su "bondad".

🌀 Esta retroalimentación se utiliza por la estrategia de búsqueda para seleccionar nuevo candidatos.

🌀 Una función simple: El criterio de Fisher




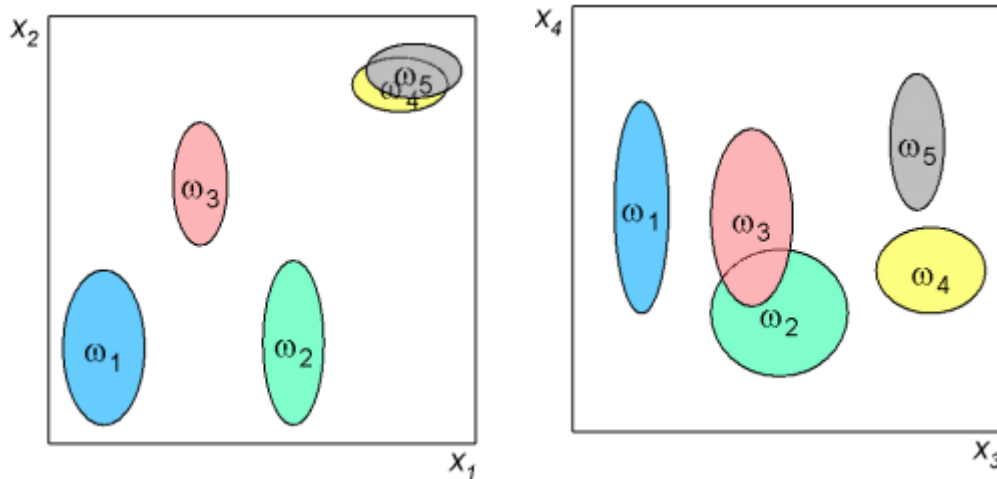
Selección de Características

MÉTODOS DE SELECCIÓN

 **SELECCIÓN SIMPLE:** Seleccionar las mejores k características individuales.

 Ventaja: Rápido de Calcular



 Desventaja: Combinando variables que individualmente son buenas, no siempre conduce a buenos resultados

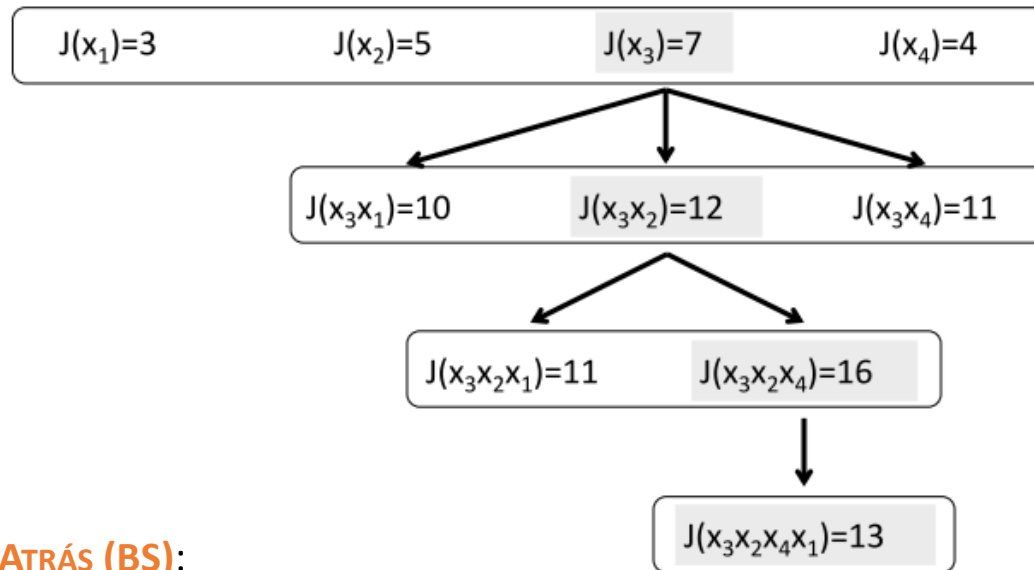


- Cualquier función buena haría un ranking como este: $J(x_1) > J(x_2) \cong J(x_3) > J(x_4)$
- Solución óptima: $\{x_1, x_4\}$
- Solución del algoritmo: $\{x_1, x_2\}$



MÉTODOS DE SELECCIÓN

SELECCIÓN HACIA ADELANTE (FS):

-  Comience con un conjunto vacío.
-  Adicione una característica a la vez tal que se maximice J



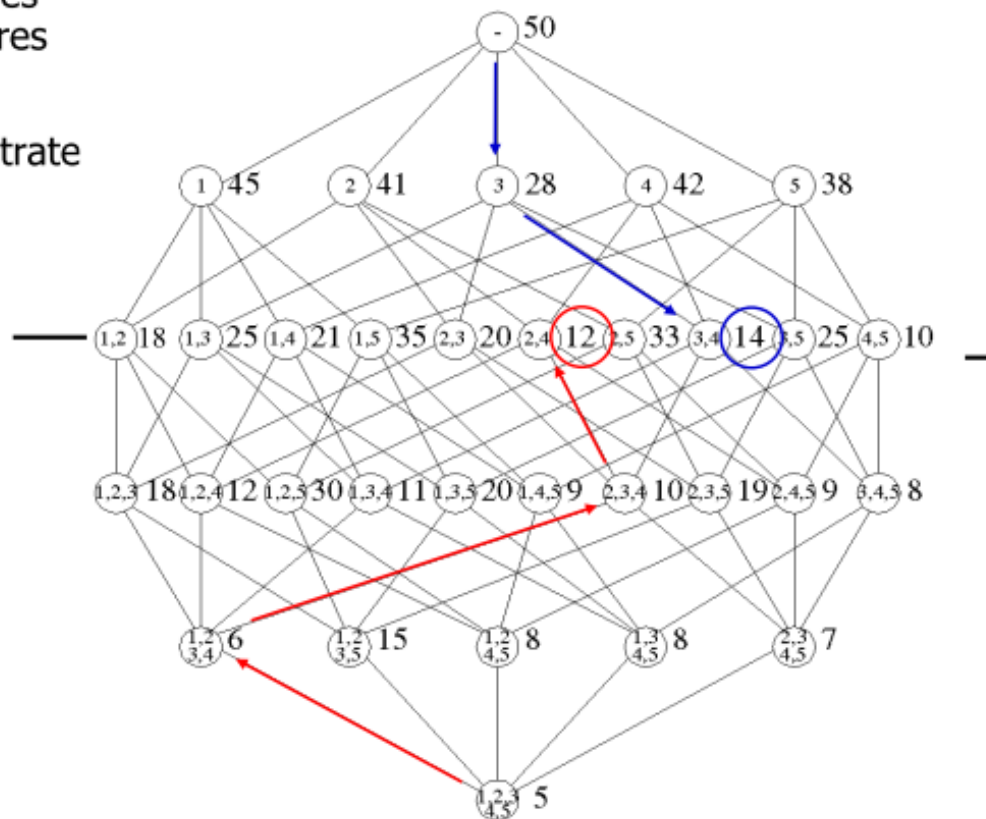
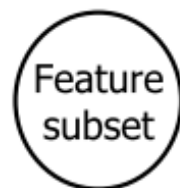
SELECCIÓN HACIA ATRÁS (BS):

-  Comience el conjunto de todas las características
-  Elimine una característica a la vez tal que se maximice J al eliminarla

Selección de Características




MÉTODOS DE SELECCIÓN

- Select $l = 2$ features out of $k = 5$ features
- Suboptimality illustrate
 - forward
 - backward



Selección de Características

MÉTODOS DE SELECCIÓN

-  **SELECCIÓN BRANCH & BOUND:** es una búsqueda hacia atrás con retroceso que garantiza encontrar el subconjunto óptimo de característica bajo el supuesto de monotonidad.
-  Inicia con el conjunto completo de características formado un árbol de evaluación.
-  Los nodos cuya función objetivo son menores que la mejor solución actual no se exploran por el supuesto de monotonidad el cual asegura de que los hijos de esos nodos no van a contener una solución mejor.



MÉTODOS DE SELECCIÓN



EN RESUMEN:

- El número de características normalmente se limita en el aprendizaje debido a los costos y a la precisión del clasificador (maldición de la dimensionalidad)
- Encontrar buenas características **no es trivial**. La selección de características puede mejorar el rendimiento y ayudar a la interpretación:
 - **Requisitos:** criterio y el algoritmo de búsqueda.
 - **Estimación de rendimiento:** por ejemplo, validación cruzada.



Pasos:

- Conocimiento especializado (por ejemplo, iniciar con 100 características)
- Ranking Individual (Seleccionar 40)
- Selección hacia adelante (Seleccionar 20)
- Selección hacia atrás usando branch & bound (10)

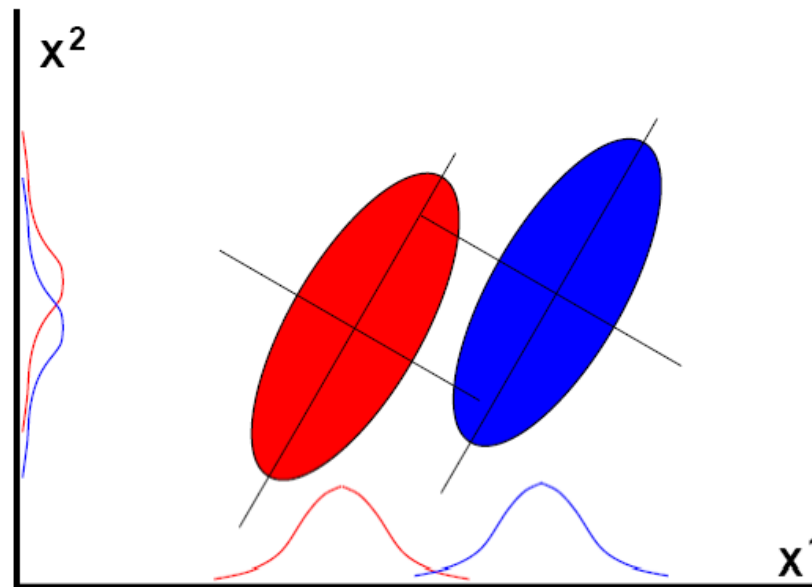
Selección de Características

MÉTODOS DE EXTRACCIÓN: PCA

Selección de Características


🦋 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

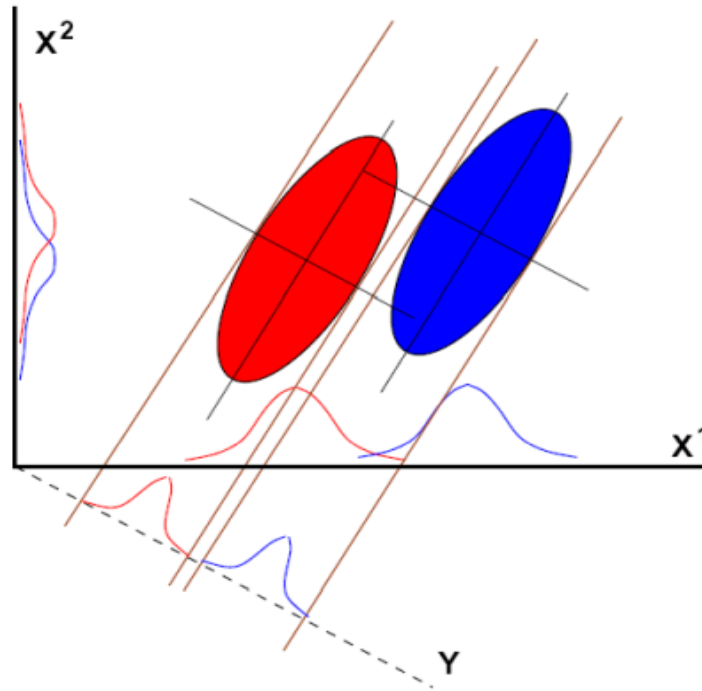
- 🌀 Hay casos en los que la selección de cualquiera de las variables originales no proporciona una buena solución.



Selección de Características

ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

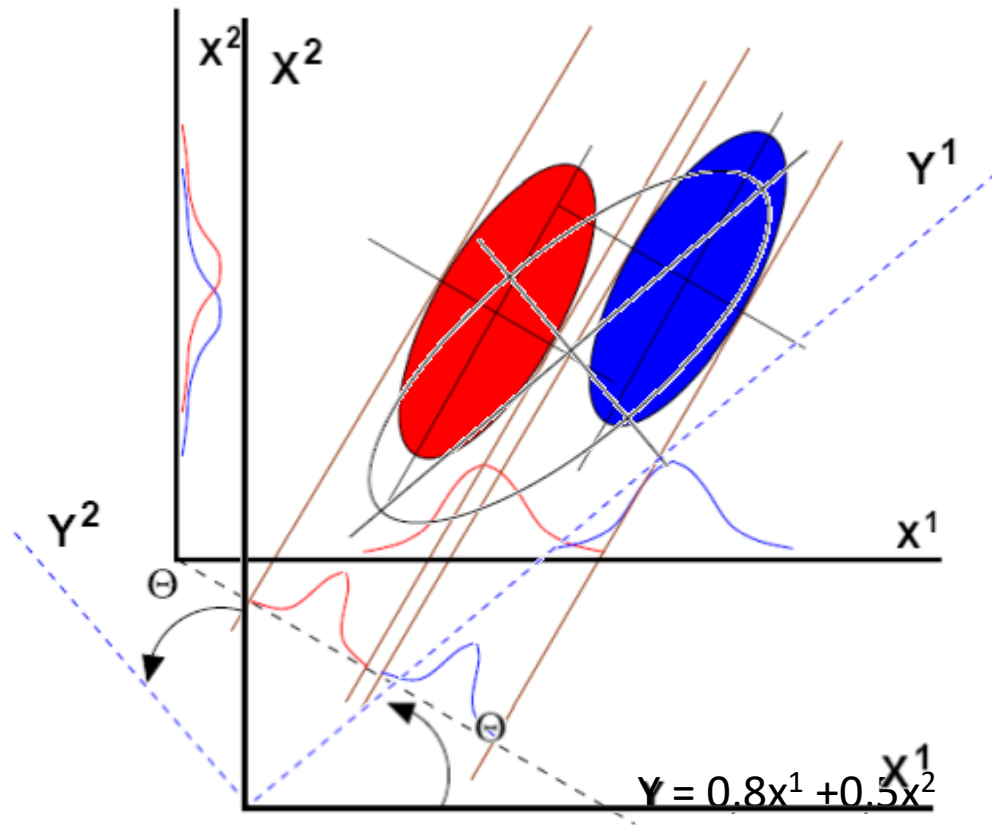
-  Si transformamos el espacio de manera lineal usando una nueva variable Y que tenga la forma $Y = 0,8x^1 + 0,5x^2$, se obtienen mejores resultados en la separación:





ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

- PCA es el método más popular de extracción de características
- PCA es una transformación lineal




Objetivo: Transformar el espacio de representación P en uno nuevo P' , en el que los datos estén no correlacionados, es decir, que su matriz de sea diagonal.





Es decir, se trata de encontrar un nuevo conjunto de ejes ortogonales en el que la varianza de los datos sea la máxima en alguna dirección.

Selección de Características

ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

 La relación entre los nuevos ejes y los antiguos consiste en:

$$\begin{aligned}Y_1 &= W_{11}X_1 + W_{12}X_2 \\ Y_2 &= W_{21}X_1 + W_{22}X_2\end{aligned}$$

-  La mayor parte de la información contenida en el espacio P puede retenerse únicamente en el primer eje principal Y_1 , lo que implica una selección de características ...
-  Los nuevos ejes se definen secuencialmente de manera que un nuevo eje se define como aquel que es perpendicular al seleccionado anteriormente y su dirección es la de la máxima varianza de entre todos los ejes posibles.

Selección de Características

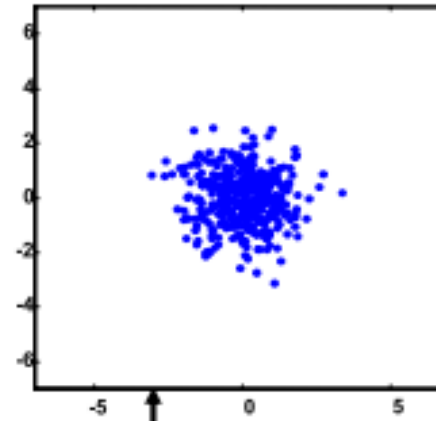
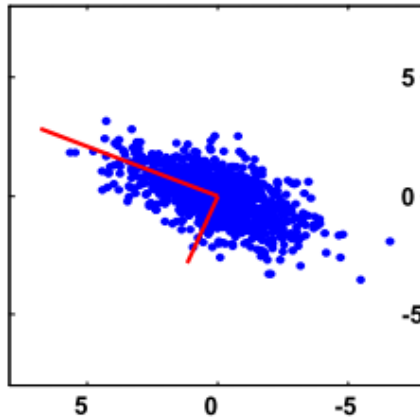


ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

Encontrar W_{ij} :

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

- Los ejes que definen P' son ortogonales
- Los datos en P' están no correlacionados
- Se debe minimizar el error de reconstrucción



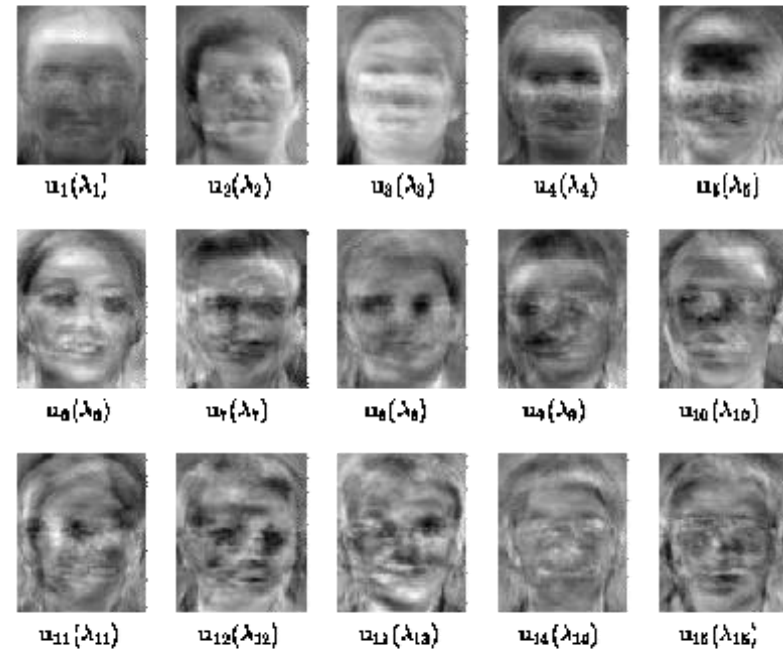
Los **componentes principales** son los primero k vectores propios de la transformación ...

Selección de Características

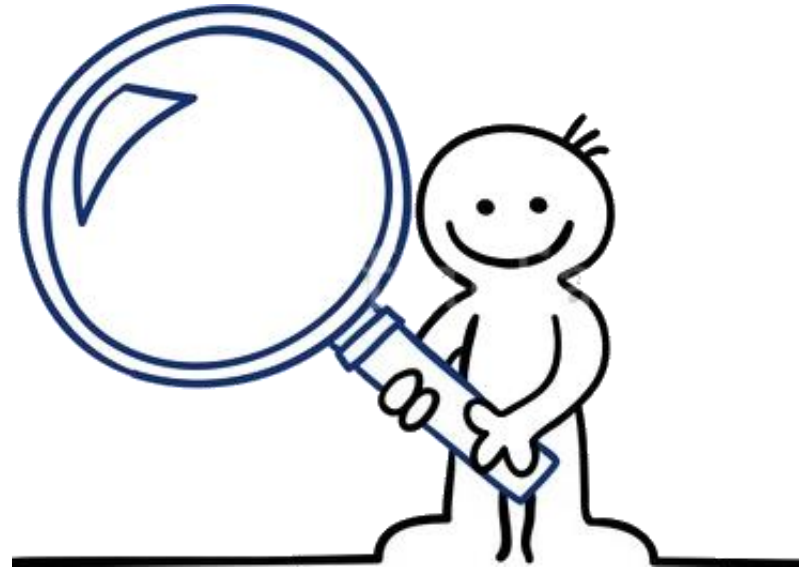
🦋 ANÁLISIS DE COMPONENTES PRINCIPALES (PCA)

🌀 Ejemplo con Imágenes para Reconocimiento de un Patrón en las imágenes

Eigenfaces for Recognition, Turk, M. & Pentland, A. ,
Journal of Cognitive Neuroscience, 3, 71-86, 1991.



RECONOCIMIENTO DE PATRONES

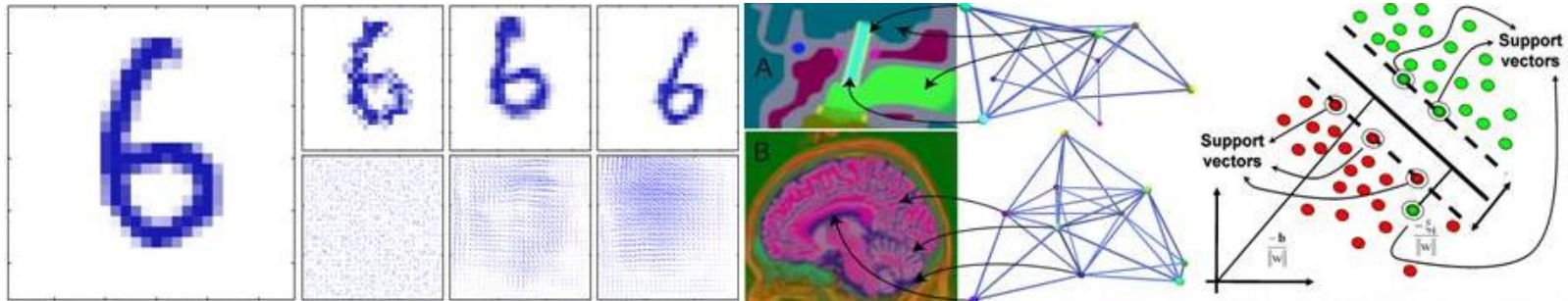


RECONOCIMIENTO DE PATRONES

Reconocimiento de Patrones

🦋 INTRODUCCIÓN

- 🌀 El *Reconocimiento de Patrones* es la última etapa dentro de un sistema de visión artificial, en la que a partir de las características encontradas, los posibles objetos se **CLASIFICAN** en dos o más clases.



- 🌀 **Clasificar** (o reconocer) significa, en este contexto, asociar a clases (o prototipos) una serie de elementos (u objetos). Esta asociación se realiza en base a las características o propiedades de los objetos.

Reconocimiento de Patrones



CONSIDERACIONES

Las características de las regiones u objetos segmentados se representan usando **vectores de características** normalizados.

Las características usadas para el reconocimiento deben ser cuidadosamente seleccionadas (p. ej. elección de características invariantes a transformaciones geométricas)

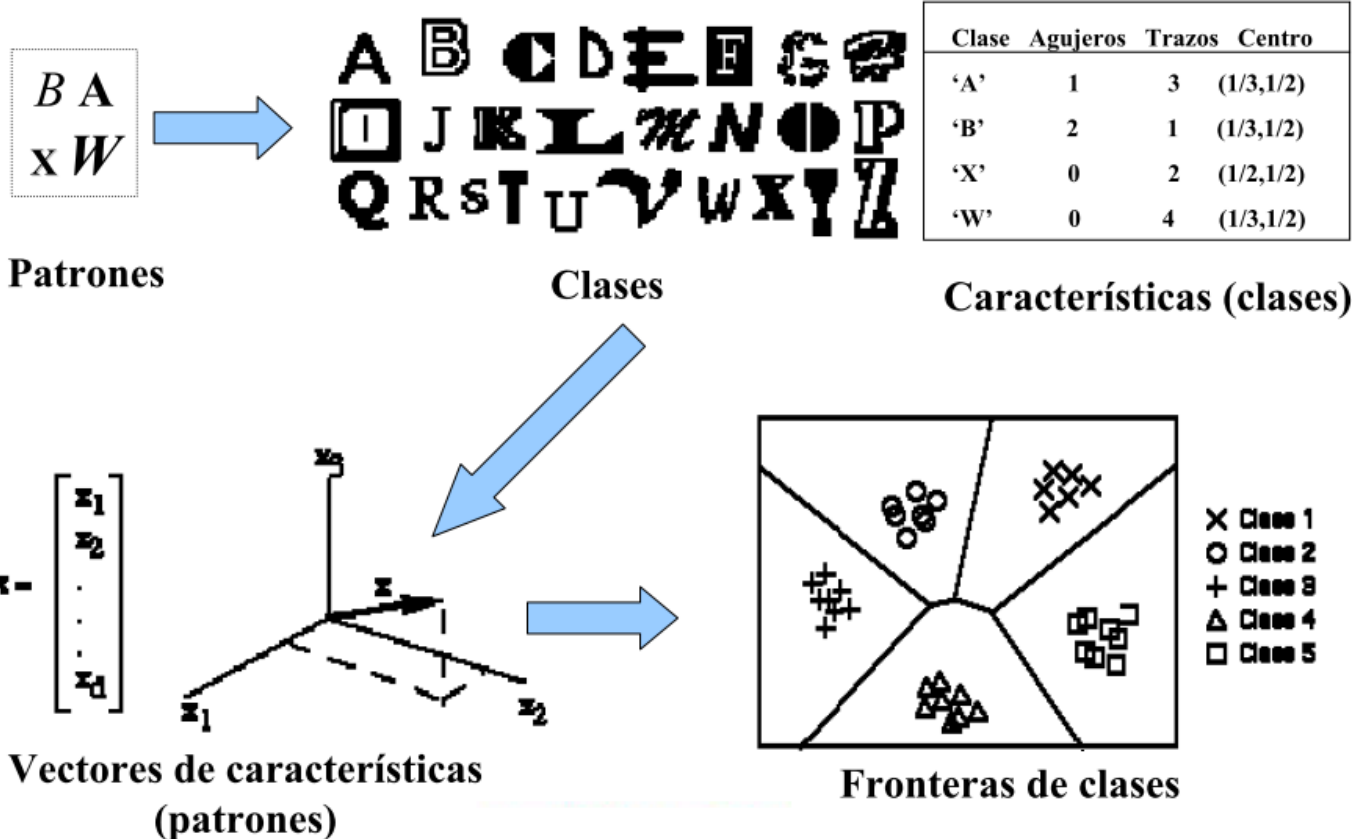
Reconocer o clasificar no son tareas fáciles: las clases pueden no estar correctamente definidas, la información sobre los objetos a clasificar puede ser incompleta.

La interpretación de imágenes (o escenas) requiere el uso de modelos y técnicas de **Inteligencia Artificial**

Métodos de clasificación diferentes → clasificaciones diferentes.

Reconocimiento de Patrones

🦋 EJEMPLO DE RECONOCIMIENTO DE CARACTERES



Reconocimiento de Patrones

🦋 **IMPORTANTE:** Si los descriptores elegidos son adecuados, objetos similares tendrán patrones próximos en el espacio de características.



Patrones que describen objetos de una misma clase, presentan características similares.

Patrones que describen objetos de diferentes clases presentan características diferenciadas.

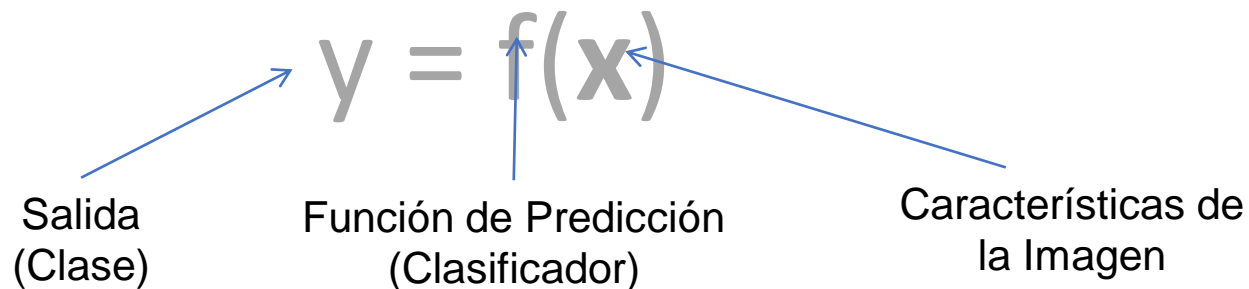
Reconocimiento de Patrones

🦋 MODELO GENERAL DE UN CLASIFICADOR

- 🌀 Aplicar una función de predicción en una representación de las características de la imagen para obtener el resultado deseado

$f(\text{manzana}) = \text{"manzana"}$

$f(\text{vaca}) = \text{"vaca"}$

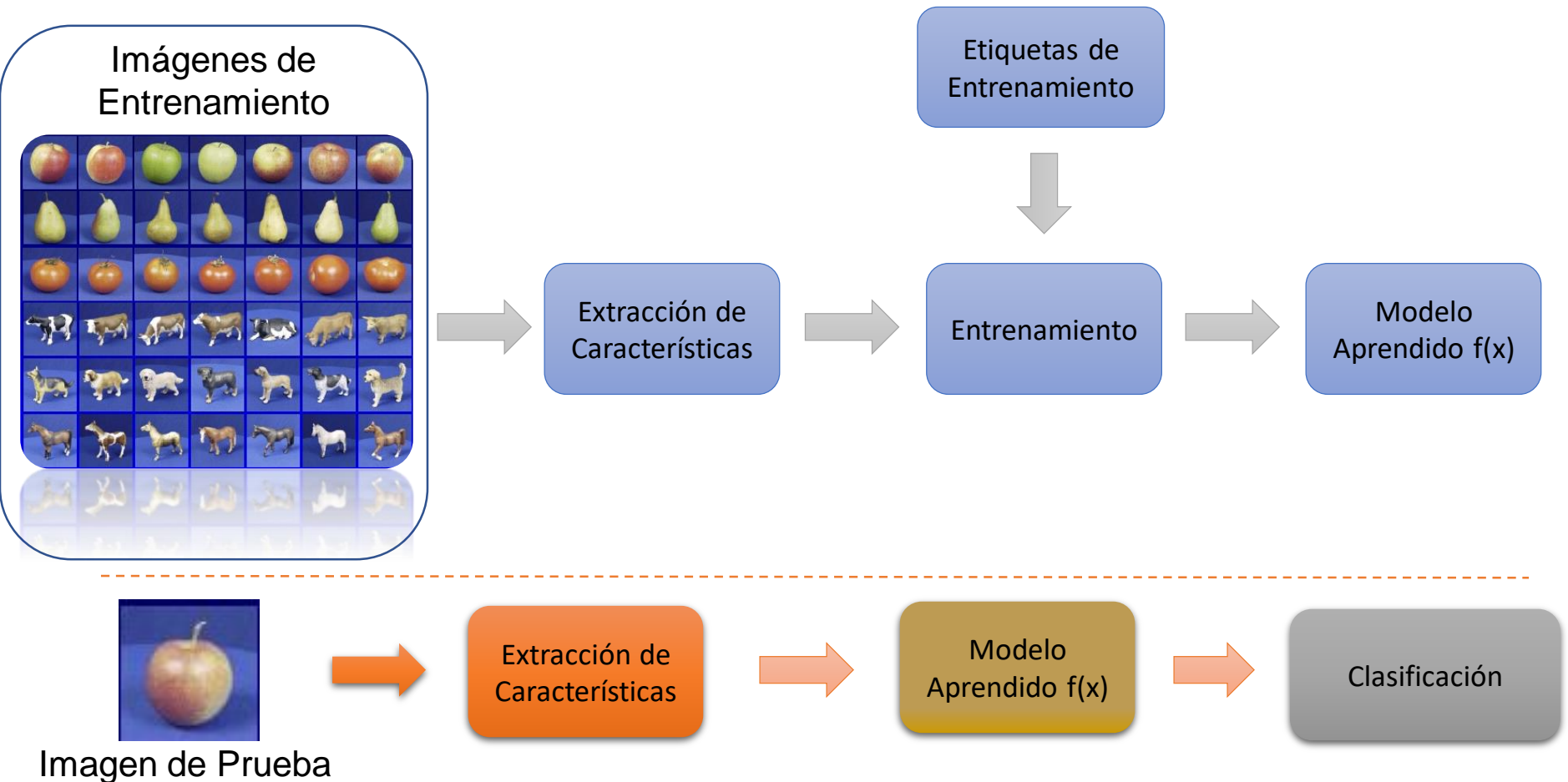


- 🌀 **Entrenamiento:** dado un conjunto de ejemplos $\{(x_1, y_1), \dots, (x_n, y_n)\}$, calcular la predicción de la función f , **REDUCIENDO AL MÍNIMO EL ERROR DE PREDICCIÓN** en el conjunto de entrenamiento

Reconocimiento de Patrones






MODELO GENERAL DE UN CLASIFICADOR

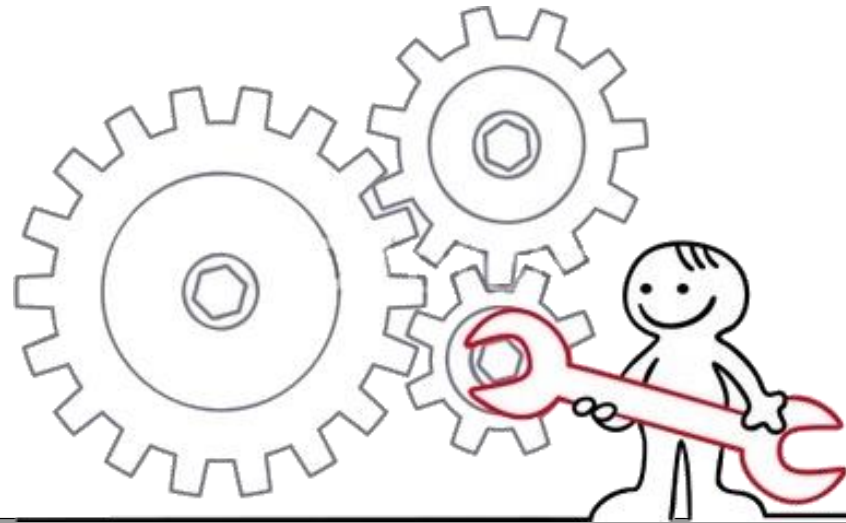


Reconocimiento de Patrones

TIPOS DE CLASIFICADORES

-  Atendiendo a la información que se proporciona en el proceso de construcción del clasificador se puede hablar de dos tipos de clasificadores: supervisados y No supervisados:
-  **Clasificadores NO Supervisados:** sin la necesidad de ningún supervisor externo, el clasificador determina las clases que representan los datos de entrenamiento.
-  **Clasificadores Supervisados:** el conjunto de entrenamiento es dividido por el maestro en las diferentes clases ya conocidas en las que se desea clasificar, así el clasificador aprende las características que definen cada clase.

RECONOCIMIENTO DE PATRONES



TÉCNICAS DE RECONOCIMIENTO DE PATRONES

Reconocimiento de Patrones

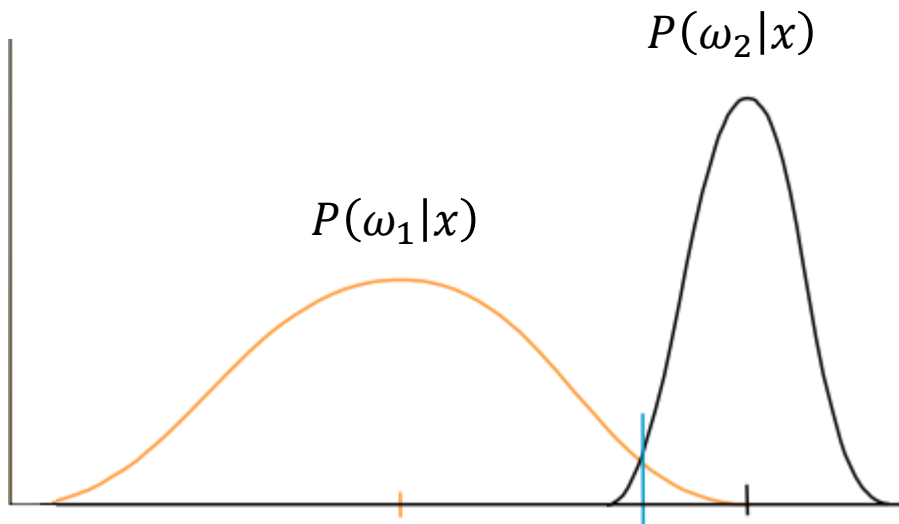


CLASIFICADORES SUPERVISADOS – TEOREMA DE BAYES



Usar la teoría de la probabilidad para clasificar el objeto en la clase que tenga mayor probabilidad posteriori

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{p(x)}$$



$P(\omega_i)$ = Probabilidad de que en la población haya un objeto de clase ω_i


$p(x|\omega_i)$ = Probabilidad de que en la clase ω_i se de un vector de características x

$P(\omega_i|x)$ = Probabilidad de que el objeto de vector de características x pertenezca a la clase ω_i

$$g(x) = \begin{cases} 1 & \text{si } P(\omega_1|x) > P(\omega_2|x) \\ 2 & \text{en otro caso} \end{cases}$$

Reconocimiento de Patrones

CLASIFICADORES SUPERVISADOS – TEOREMA DE BAYES

 Asumiendo que $p(x|\omega_i) \sim N(\mu_i, \Sigma_i)$ la función discriminante para la clase ω_i es:

$$g_i(x) = \ln(p(x|\omega_i)) + \ln(P(\omega_i))$$


Donde, $p(x|\omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp(-\frac{1}{2} (x - \mu_i)^t \Sigma_i^{-1} (x - \mu_i))$

 **Caso 1:** $\Sigma_i = \sigma^2 I$

LDA

 **Caso 2:** $\Sigma_i = \Sigma$

Clasificador con Distancia Mahalanobis al Cuadrado

 **Caso 3:** $\Sigma_i = \text{Arbitraria}$

QDA

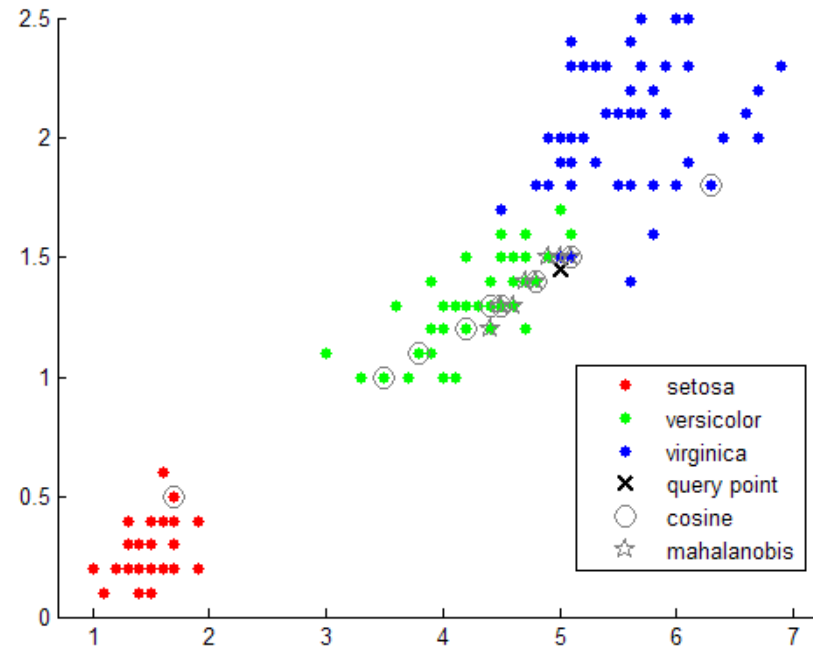
Reconocimiento de Patrones

CLASIFICADORES SUPERVISADOS – K VECINOS MÁS CERCANOS

- La idea básica del método considera la utilización de un conjunto de vecinos para etiquetar el nuevo objeto. Esta regla basa su operación en el supuesto de considerar a los patrones cercanos, como aquellos que tienen la mayor probabilidad de pertenecer a la misma clase.

Así el algoritmo asigna la etiqueta de clase que tengan la mayoría de los k – vecinos.

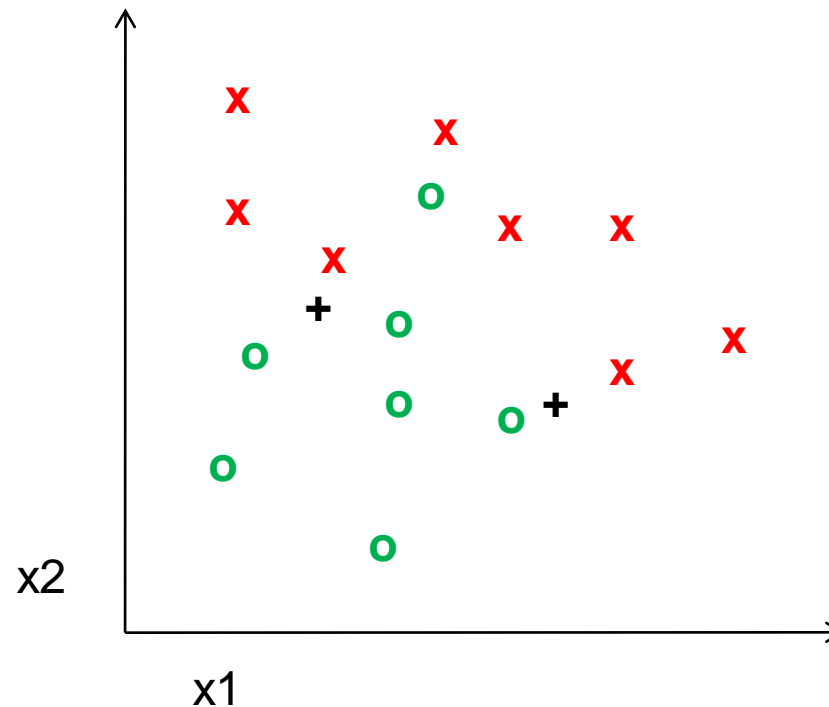
En Balu: `Bcl_knn`



Reconocimiento de Patrones



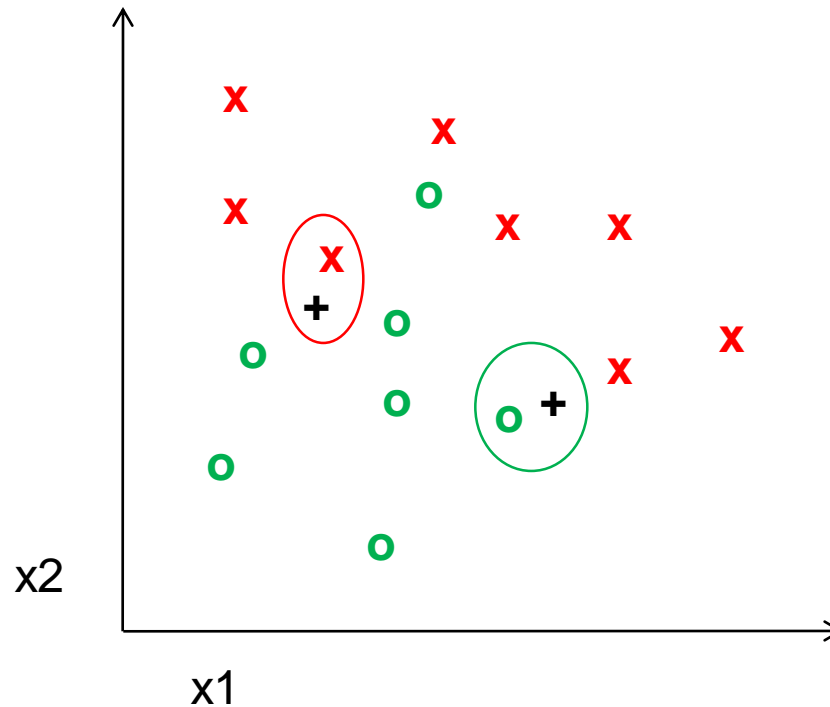
CLASIFICADORES SUPERVISADOS – K VECINOS MÁS CERCANOS



Reconocimiento de Patrones

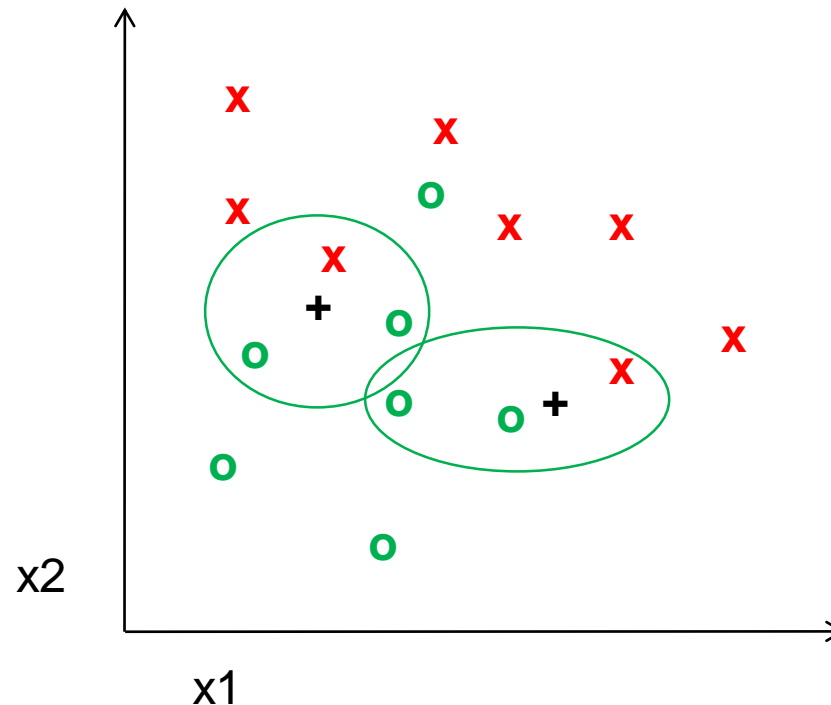


CLASIFICADORES SUPERVISADOS – 1 VECINO MÁS CERCANOS



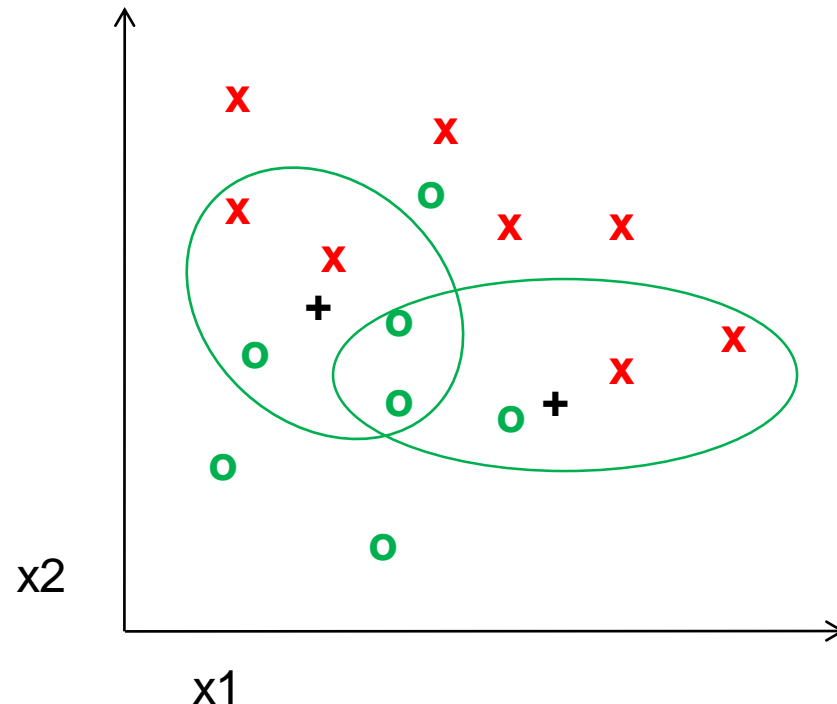
Reconocimiento de Patrones

🦋 CLASIFICADORES SUPERVISADOS – 3 VECINOS MÁS CERCANOS



Reconocimiento de Patrones

🦋 CLASIFICADORES SUPERVISADOS – 5 VECINOS MÁS CERCANOS

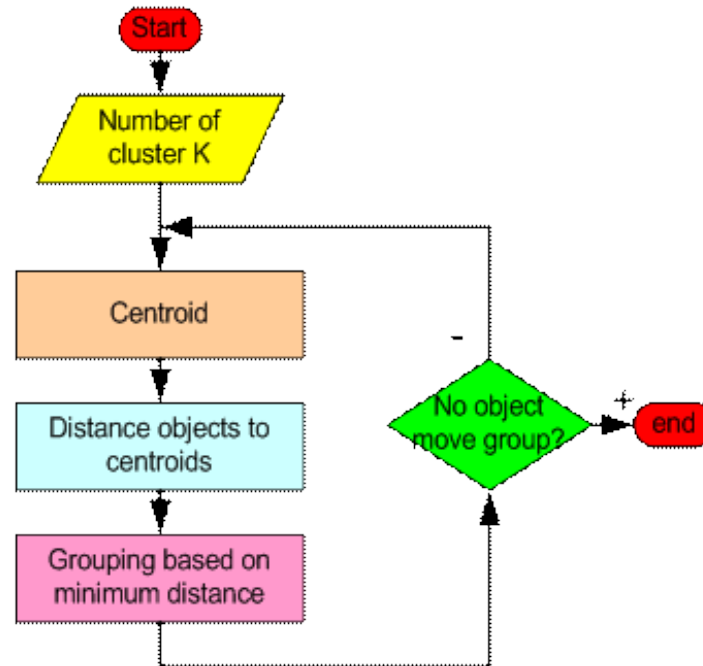


Reconocimiento de Patrones



CLASIFICADORES NO SUPERVISADOS - CLUSTERING

- Los algoritmos de **clustering** o agrupamiento intentan **dividir** el conjunto de datos de entrenamiento en **k grupos**, de acuerdo con un criterio de cercanía que se define en términos de una función de distancia, como la Euclidiana, la Manhattan o la de Mahalanobis.



Reconocimiento de Patrones



CLASIFICADORES NO SUPERVISADOS - CLUSTERING

- Cada una de las k clases se representa con un prototipo Z_k o centroide que es un vector d -dimensional:

$$Z_k = \frac{1}{n_k} \sum_{j=1}^{n_k} x_{kj}$$

Siendo: x_{kj} el j -ésimo vector de características (patrón) de la clase k .

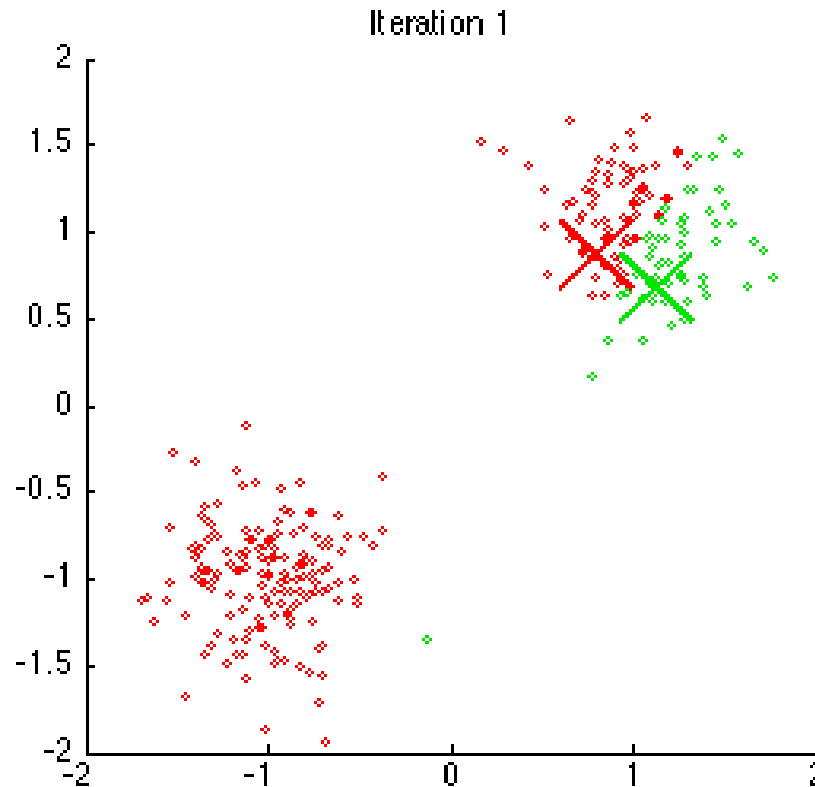
- La distancia euclídeana d_E de un nuevo patrón X a la clase C_k es:

$$d_E(X, Z_k) = \|X - Z_k\| = \sqrt{\sum_{i=1}^d (X_i - Z_{ki})^2}$$

La fórmula anterior es equivalente a evaluar la expresión de la función discriminante de cada clase $fd_k(X)$, siendo: $k \in 1..N$, para el patrón X y asignarlo a la clase C_k para la que $fd_k(X)$ sea máximo.

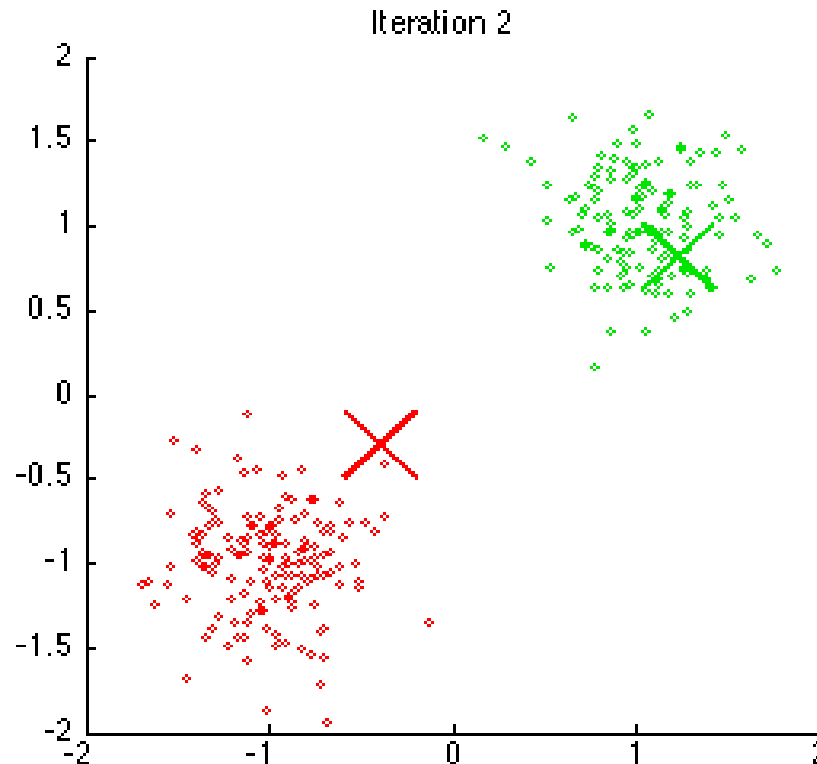
Reconocimiento de Patrones

CLASIFICADORES NO SUPERVISADOS - CLUSTERING



Reconocimiento de Patrones

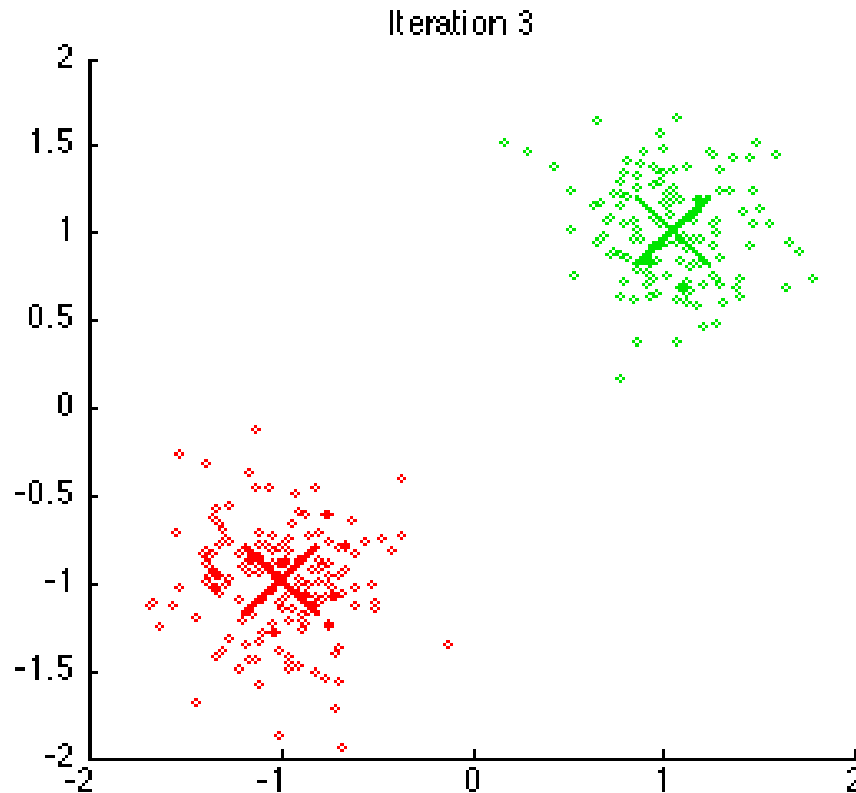
CLASIFICADORES NO SUPERVISADOS - CLUSTERING



Reconocimiento de Patrones



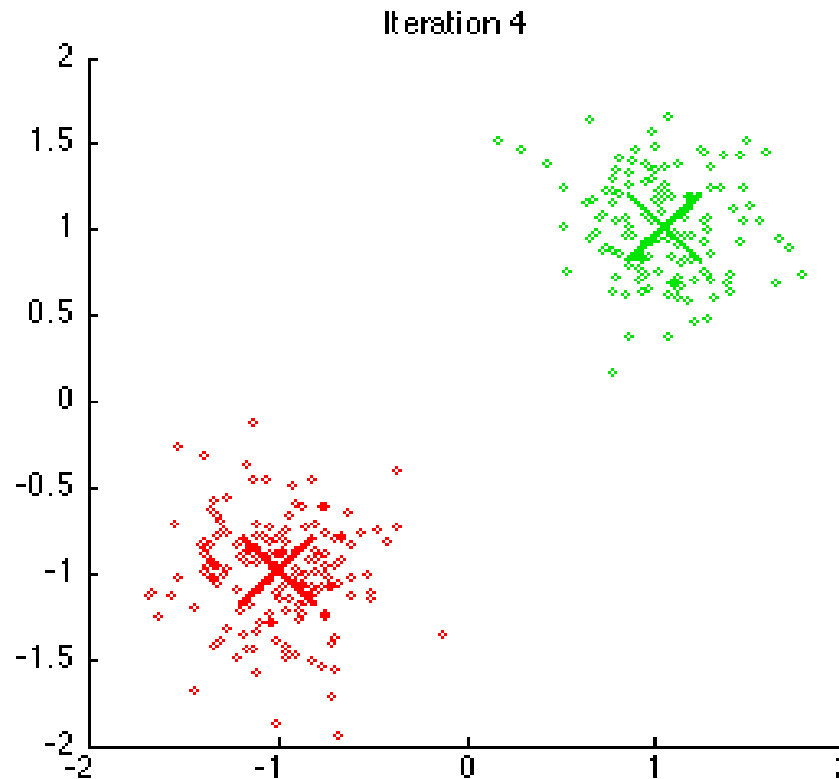
CLASIFICADORES NO SUPERVISADOS - CLUSTERING



Reconocimiento de Patrones



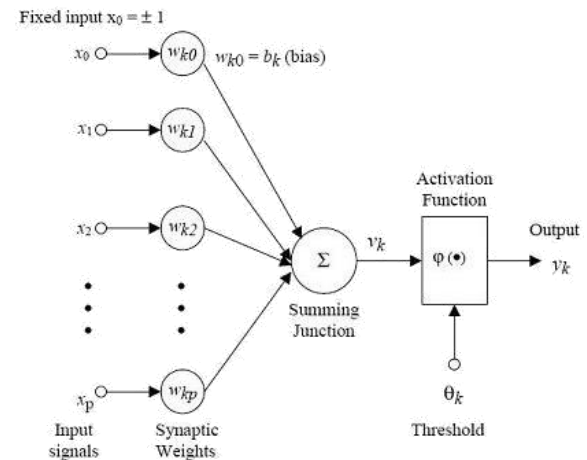
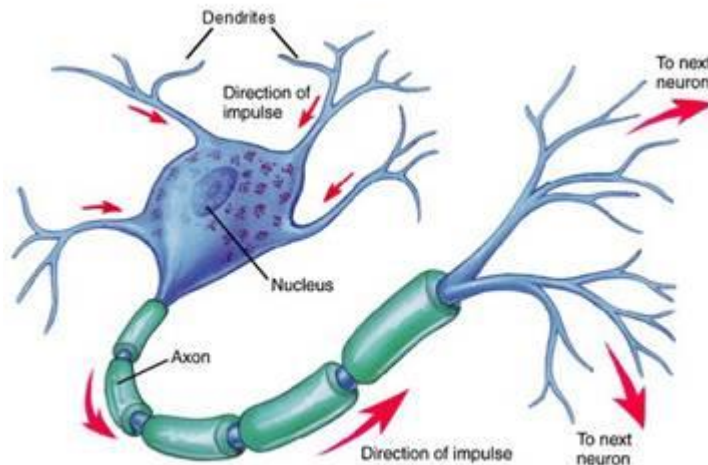
CLASIFICADORES NO SUPERVISADOS - CLUSTERING



Reconocimiento de Patrones

CLASIFICADORES SUPERVISADOS – REDES NEURONALES

- Por su capacidad de aprendizaje **las neuronas** de los organismos biológicos se han estudiado para su aplicación en sistemas de aprendizaje automático.
- Al igual que las neuronas biológicas están conectadas, las redes de neuronas artificiales están formadas por elementos sencillos de cómputo interconectados según diferentes



Neurona artificial

Reconocimiento de Patrones

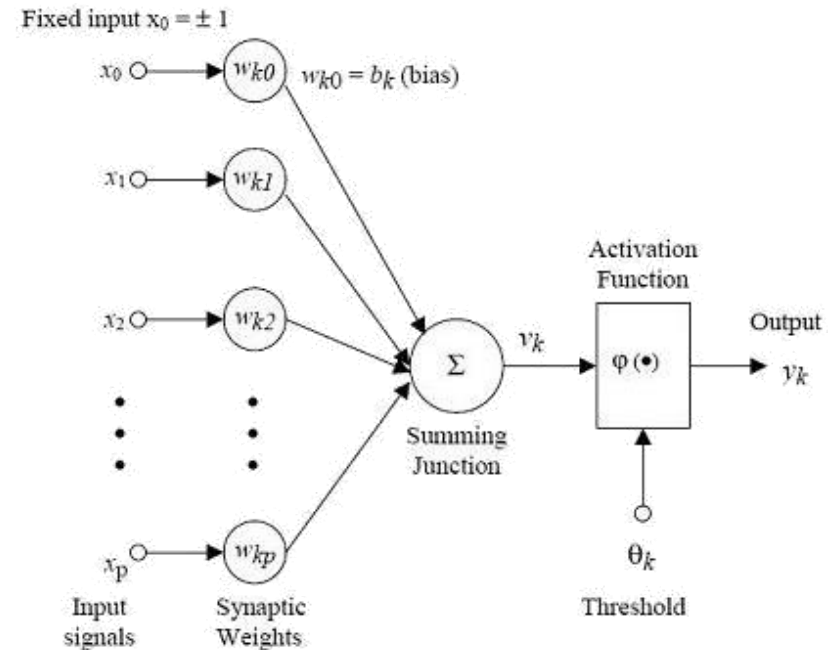
CLASIFICADORES SUPERVISADOS – REDES NEURONALES

- El **Perceptrón**, en su forma básica, consiste en una neurona que es capaz de aprender una función discriminante lineal v_k , que permite dividir a dos conjuntos de entrenamiento linealmente separables. Su respuesta consiste en una suma ponderada de sus entradas que representa la ecuación de un hiperplano en el espacio p-dimensional :

$$v_k = \sum_{j=1}^p w_{kj} x_j$$

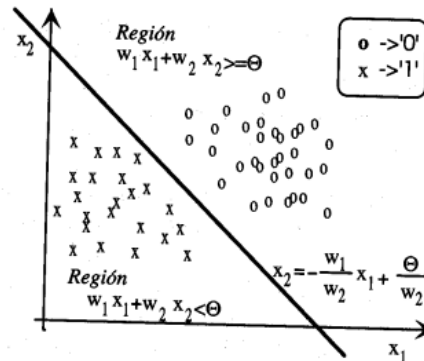
- A la salida se aplica una función de activación $\phi(v)$ (escalón, sigmoide, etc) que indica si se activa o no la neurona.

$$\phi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \quad \phi(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)}$$

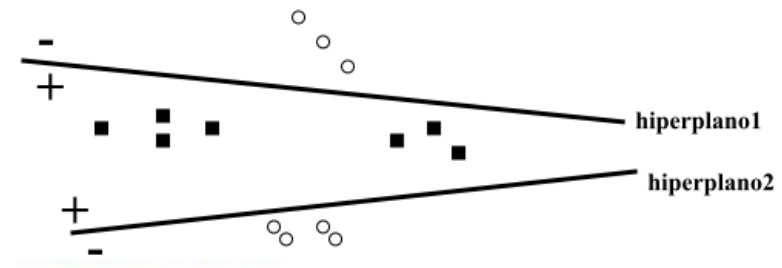
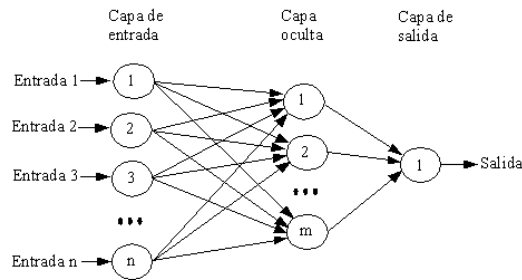




- Perceptrón: Separación de dos clases (regiones) con un perceptrón:

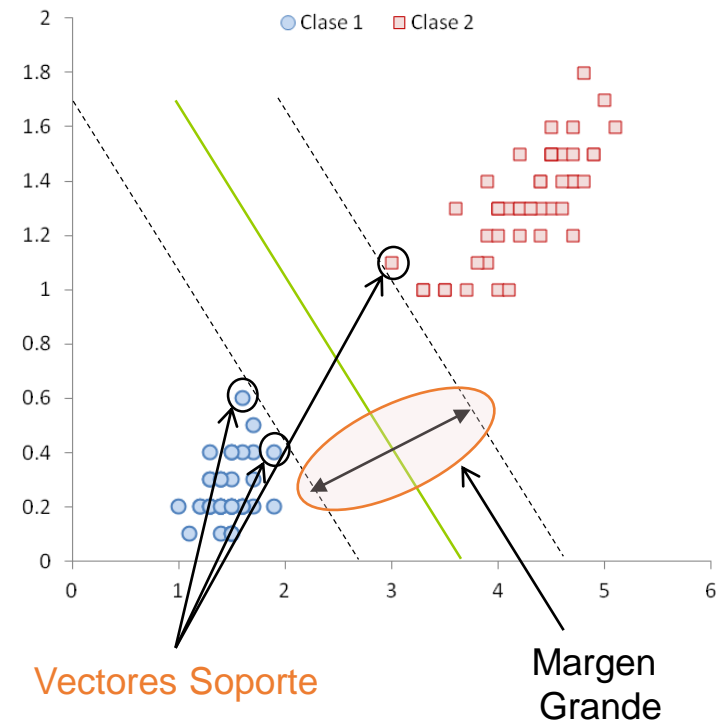
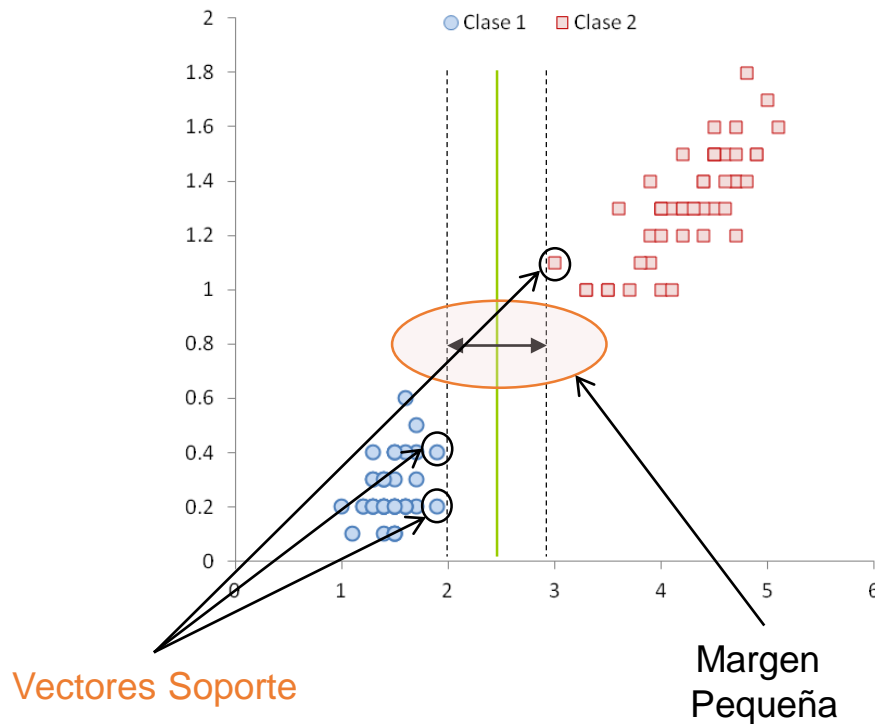


- Perceptrón de dos capas (multicapa) y ejemplo de frontera de decisión realizable con esta red:




Una vez una vez entrenada la red con un conjunto de patrones de entrenamiento, ésta es capaz de resolver el problema para patrones desconocidos.

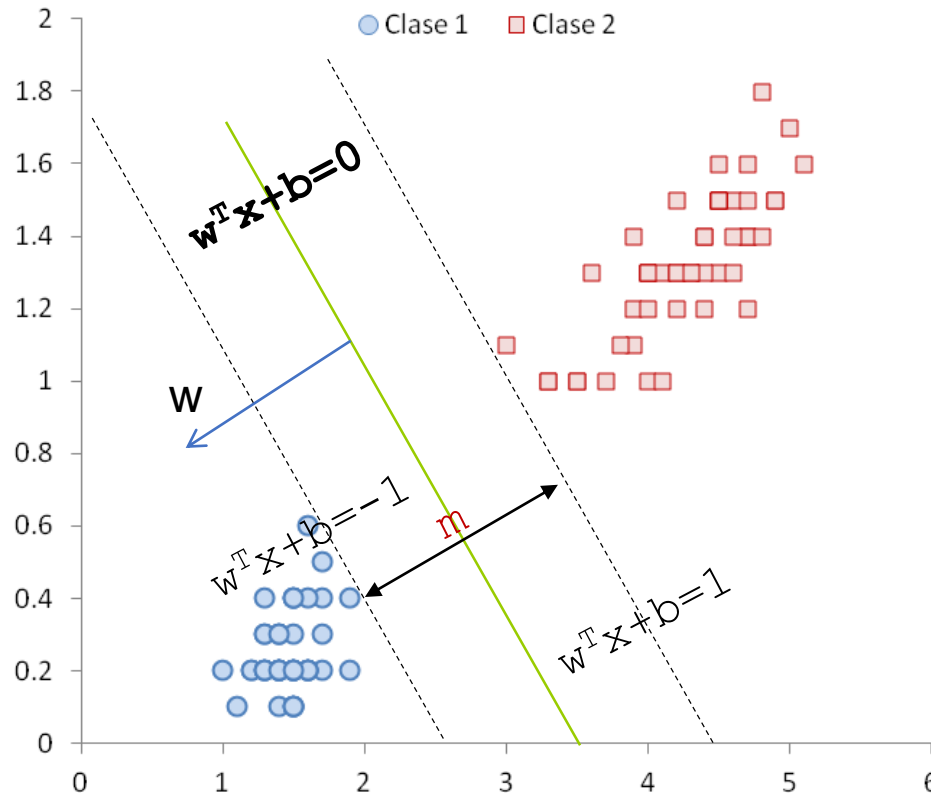
- Las **SVM** son un tipo de clasificadores de patrones basados en técnicas estadísticas de aprendizaje y están a la cabeza de los métodos de clasificación por permitir construir fronteras de decisión flexibles, y su buena capacidad de generalización.



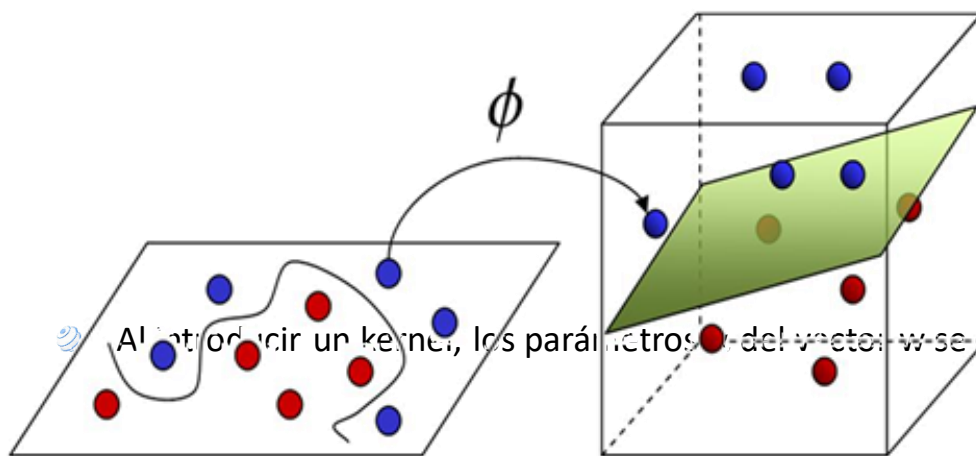
Reconocimiento de Patrones

CLASIFICADORES SUPERVISADOS – MÁQUINAS DE VECTORES SOPORTE

-  Clasificación Lineal: Las SVM generan un hiperplano que separa el espacio en dos o más regiones, una para cada clase.



- La **Clasificación NO Lineal** con una SVM realiza una transformación del espacio de entrada a otro de dimensión más alta, en el que los datos son separables linealmente.



Lineal: $K(x_i, x_j) = x_i \cdot x_j$

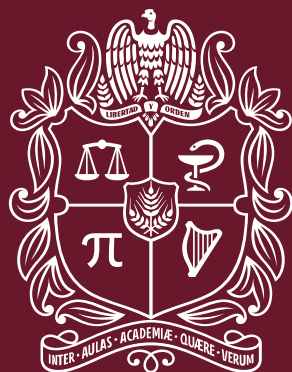
Polinómico: $K(x_i, x_j) = (x_i \cdot x_j + 1)^d$

Gausiano: $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$

$$\tilde{L}(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j)$$

Preguntas





UNIVERSIDAD
NACIONAL
DE COLOMBIA