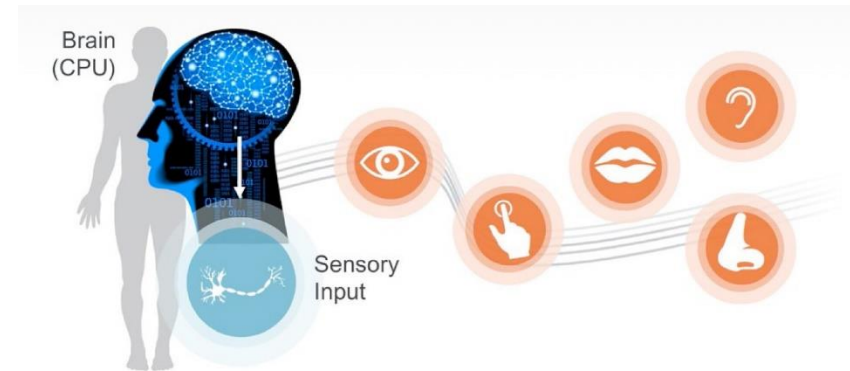


■ Percepción Humana

- A través de la evolución, el hombre ha desarrollado sistemas sofisticado para captar el ambiente que lo rodea y tomar acciones según lo captado.
- Algunas sistemas de R.P. son:
 - Reconocimiento de rostros
 - Entendimiento del habla
 - Reconocimiento de gestos
 - Lectura de caracteres
 - Reconocimiento de formas
 - Identificación de alimentos en buen estado



■ Percepción Humana – Percepción de Máquina

- Los algoritmos de reconocimiento de patrones están fuertemente influenciados por el conocimiento sobre cómo los patrones son modelados y reconocidos en los sistemas naturales.
- La investigación sobre percepción de máquina también ayuda a comprender y apreciar los sistemas de reconocimiento de patrones en la naturaleza.
- Existen muchas técnicas totalmente numéricas que no tienen relación con los sistemas naturales.

■ Reconocimiento de Patrones

- Se centra en el estudio de como las máquinas pueden:
 - Percibir el ambiente
 - Aprender a diferenciar los patrones de interés
 - Tomar decisiones acertadas y razonables sobre las categorías de los patrones.



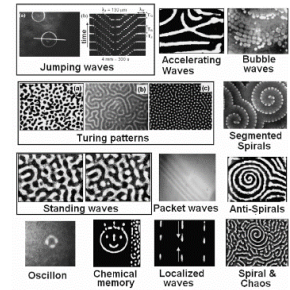
■ Reconocimiento de Patrones

■ Qué es un Patrón?

- Es una abstracción de la representación de un objeto físico, representado por un conjunto de medidas.

■ Qué tipos de patrones existen?

- Tantos como objetos/detalles existan
- Visuales, temporales, sonoros, lógicos, frecuenciales, ...



■ Reconocimiento de Patrones

■ Qué es una *Clase* de Patrones?

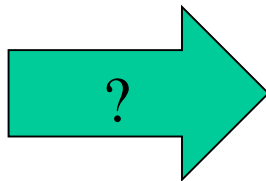
- Conjunto de patrones que comparten atributos comunes.
- Colección de objetos similares, no necesariamente idénticos.



■ Reconocimiento de Patrones

■ Qué es una *Clase* de Patrones? ...

- Sirven para asignar un objeto dado a una clase preestablecida durante el proceso de reconocimiento.



■ Reconocimiento de Patrones

■ Técnicas de Procesamiento de Patrones

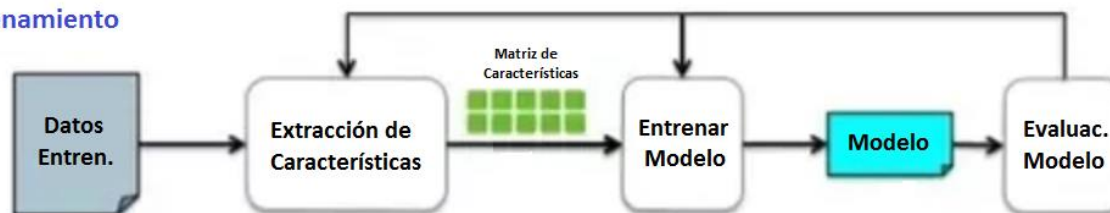
- No existe una técnica unificadora debido a la gran variedad de problemas, formas de representar los patrones y métodos de reconocerlos.
- Existen algunas técnicas estándar de amplia aplicación
 - R.P. Estadístico (incluido el fuzzy)
 - R.P. Sintáctico o structural
 - R.P. Basado en conocimiento

■ Reconocimiento de Patrones

■ Etapas Básicas del Proceso de R.P.

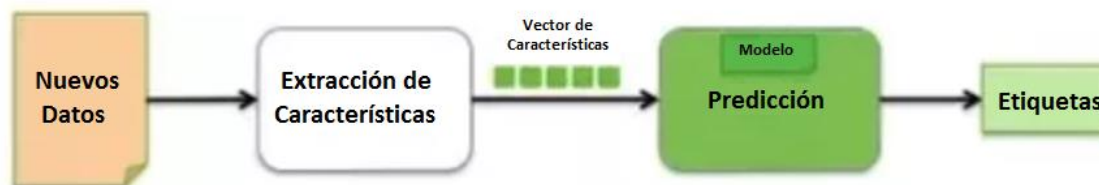
■ Entrenamiento/Aprendizaje

Entrenamiento



■ Detección/Clasificación

Clasificación



■ Ejemplo

- Una planta despachadora de pescado desea automatizar el proceso de ordenar los pescados que entran al empacador según la especie.
- El proyecto piloto desea separar róbalo y salmón con sensores ópticos



Sea bas - Róbalo



Salmon - Salmón

■ Desarrollo

- **Plataforma:** banda transportadora.



■ Desarrollo

■ Sensor

- Mediante imágenes de muestra capturadas por una cámara se identifican diferencias entre el róbalo y salmón.

■ Características

- Rasgos de los tipos de pescados que pueden utilizarse para diferenciarlos y emplearse en el clasificador.
- Longitud, brillo, ancho, número y forma de aletas, posición de la boca, ...

■ Ruido

- Variaciones de las imágenes por iluminación, posición de los pescados, estática del sistema electrónico, ..

■ Desarrollo

■ Modelo

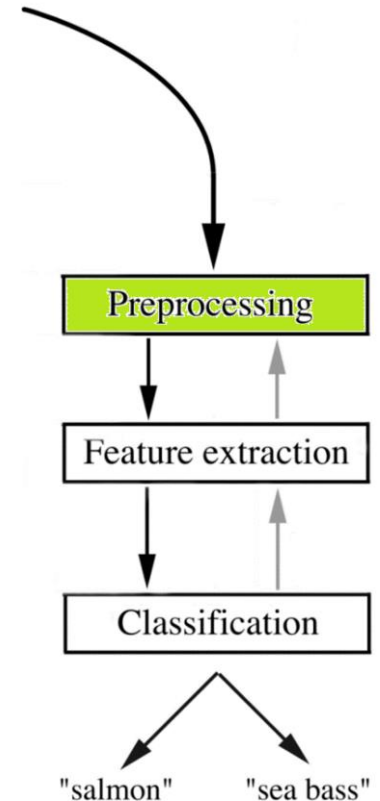
- Las diferencias entre la población de salmón y róbalo puede describirse en forma matemática mediante modelos.

■ Meta

- Establecer una hipótesis sobre la clase de modelos, procesar los datos obtenidos para eliminar ruido y escoger el modelo que corresponda mejor.

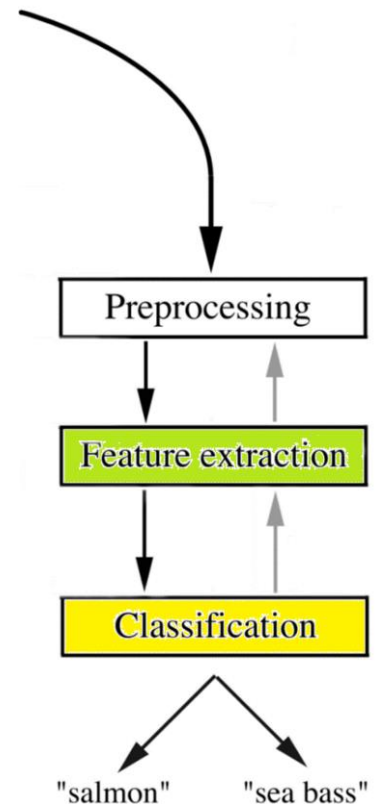
■ Desarrollo ...

- **Pre-procesamiento:** Las imágenes deben adecuarse previamente para simplificar las operaciones subsiguientes sin pérdida de información relevante.
 - Ej. Las imágenes pueden segmentarse para separar los pescados y eliminar el fondo.



■ Desarrollo ...

- **Extracción de Características:** reducción de datos y medición de ciertas características o propiedades de cada pescado.
- **Clasificador:** evalúa la evidencia presentada (características) y toma una decisión sobre la especie.



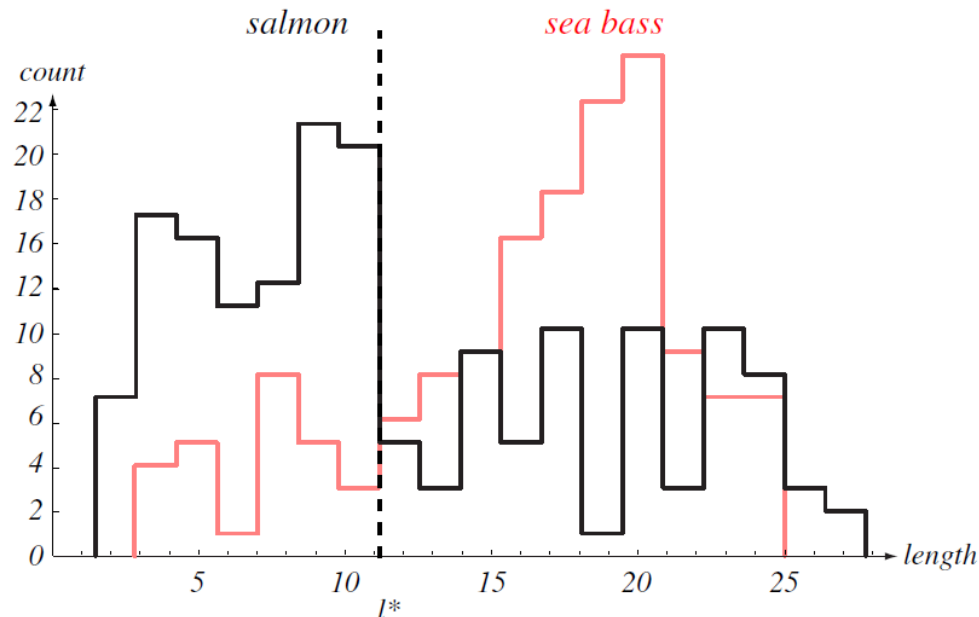
■ Desarrollo ...

■ Muestras de Entrenamiento

- **Modelo tentativo 1:** El róbalo tiene una longitud típica y es mayor que la del salmón
- **Característica:** longitud l
- **Clasificador:** si $l > T_l$ la especie es róbalo
- **Umbral:** T_l puede obtenerse de un conjunto de *muestras de entrenamiento* de las diferentes especies al medir la longitud y analizar los resultados.

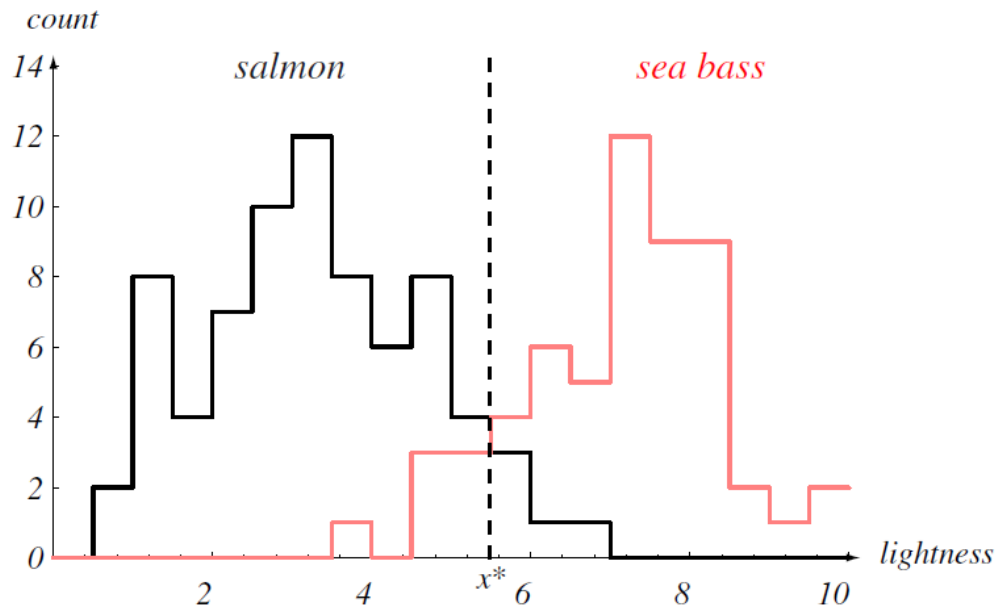
■ Muestras de entrenamiento ...

- Las mediciones pueden analizarse a través de histogramas
- Una sola medición no permite encontrar un valor de l^* óptimo que separe las dos especies de pescado.



■ Muestras de entrenamiento ...

- Se requiere adicionar una nueva medida : iluminación x
- El histograma muestra que las clases están mejor separadas.



- **Costo:** existe un costo asociado a la decisión de asignar un pescado a un especie incorrecta.
 - **Costo igual:** se asume que las consecuencias de las asignaciones tienen igual costo
 - Es simple
 - Generalmente no es conveniente.
 - **Costo diferente:** se asignan costos en función de la penalidad que pueda causar la decisión.
 - Empacar piezas de róbalo dentro de los paquetes de salmón tienen un impacto negativo en los consumidores, pues el salmón es más costoso que el róbalo.

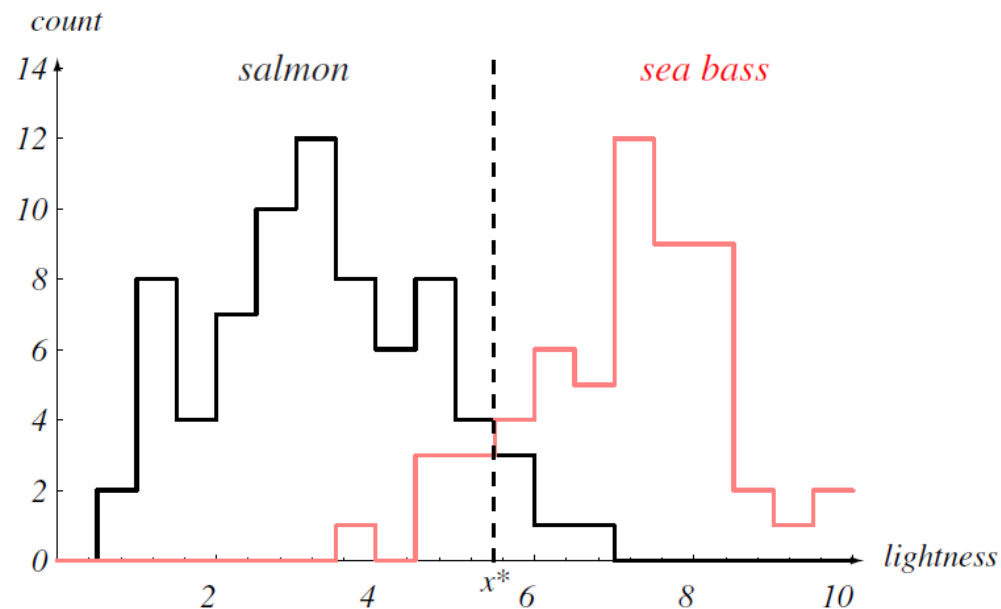
■ Teoría de la Decisión:

- Busca establecer una regla de decisión que minimice el costo de la asignación.
- Subcampo de alta importancia en la Clasificación de Patrones.
- Es posible que conociendo el umbral óptimo y el costo asociado no se tenga una solución satisfactoria.
 - Se puede desplazar la frontera de decisión ajustando el umbral x^* para reducir el número de róbalo clasificados como salmón.

■ Teoría de la Decisión...

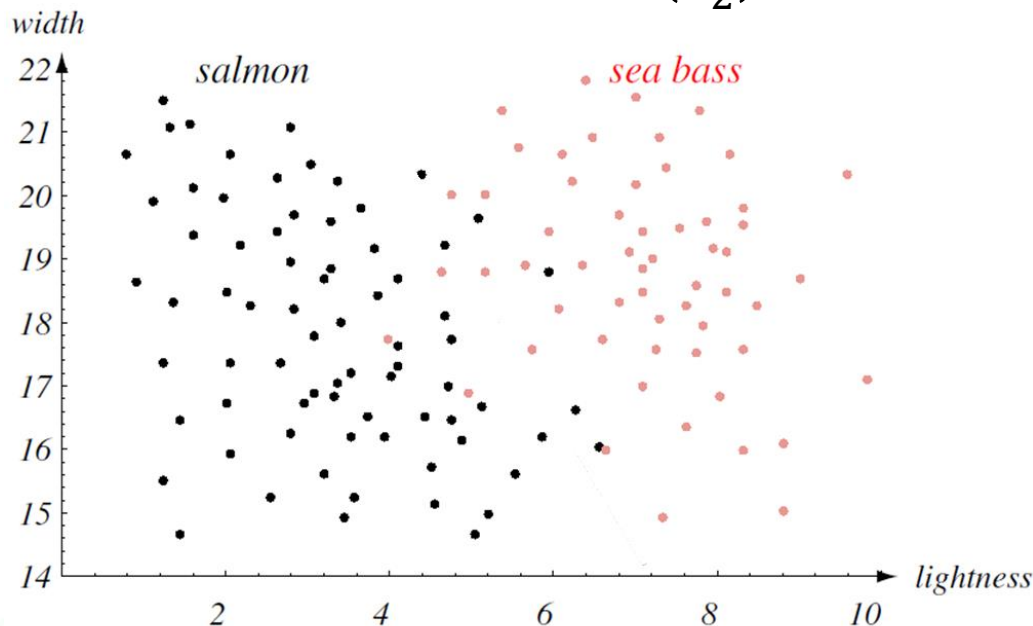
■ Solución no satisfactoria!:

- Puede beneficiar al comprador pero no al vendedor.



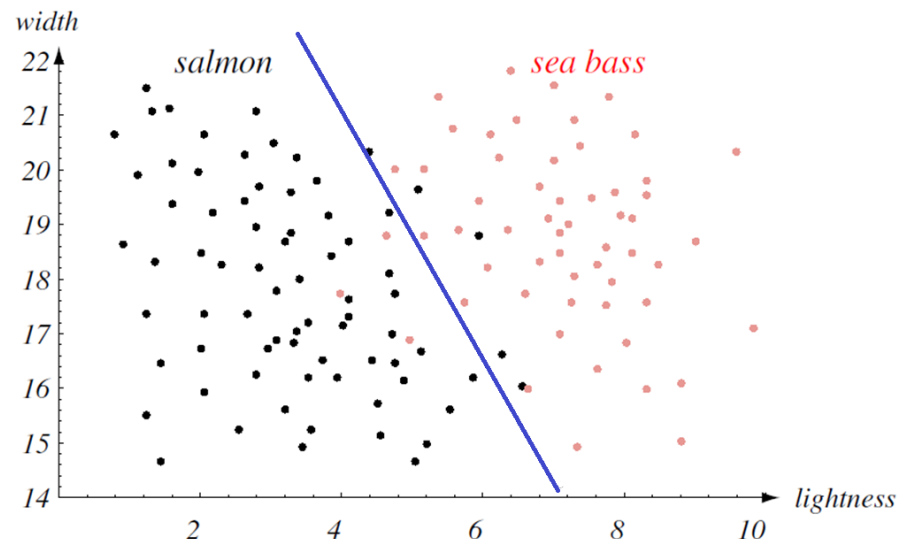
■ Teoría de la Decisión ...:

- Cuando el desempeño de la clasificación no es satisfactoria se puede recurrir a espacios de más de una dimensión.
 - De la observación: el róbalo es más ancho que el salmón.
 - Vector de características: $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $x_1 \approx$ brillo, $x_2 \approx$ ancho



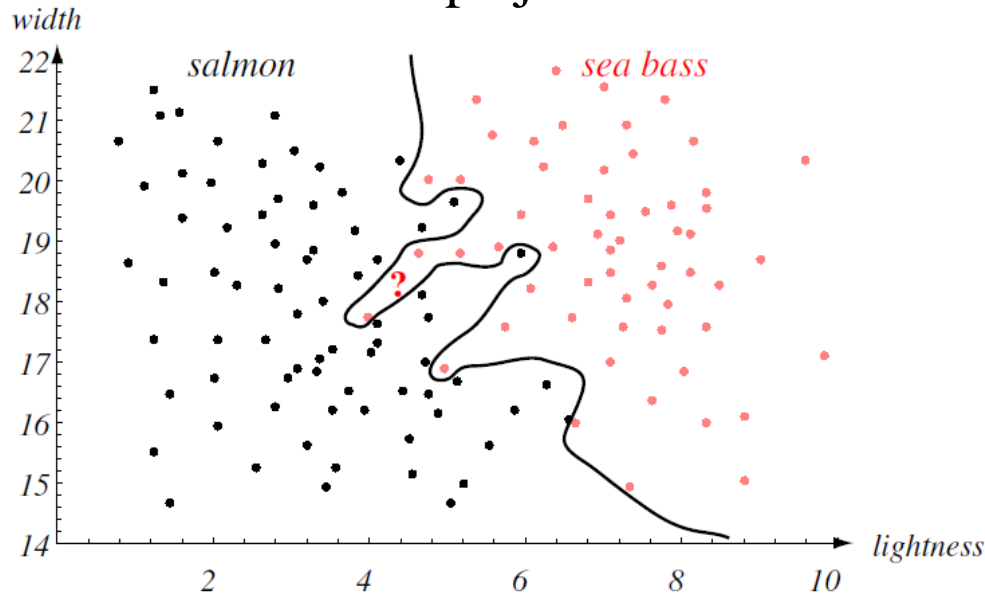
■ Frontera de decisión

- Particionar el espacio de características en regiones.
 - Regla de decisión para el ejemplo de pescados
 - Región sobre la frontera : salmón
 - Región debajo de la frontera: salmón
 - Menor error de clasificación (con dos características)



■ Frontera de decisión:

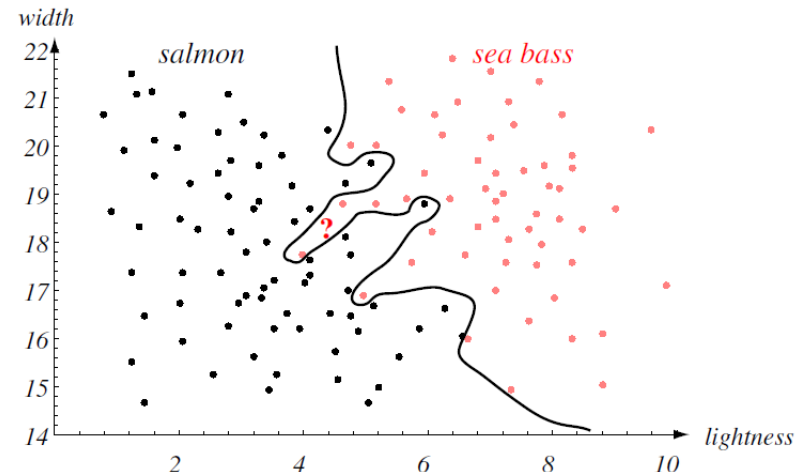
- No siempre es posible o conveniente aumentar el número de características.
 - Maldición de la dimensionalidad.
- Puede aumentarse la complejidad del modelo para dar origen a una frontera más compleja.



■ Desarrollo ...

■ Generalización

- Una frontera de decisión compleja puede *aparentar* ser la solución perfecta para los datos iniciales.
- Puede presentar bajo desempeño con nuevos patrones: solución ajustada para los patrones existentes (mala generalización)
- Se buscan fronteras simples que clasifiquen bien nuevos patrones (buena generalización)



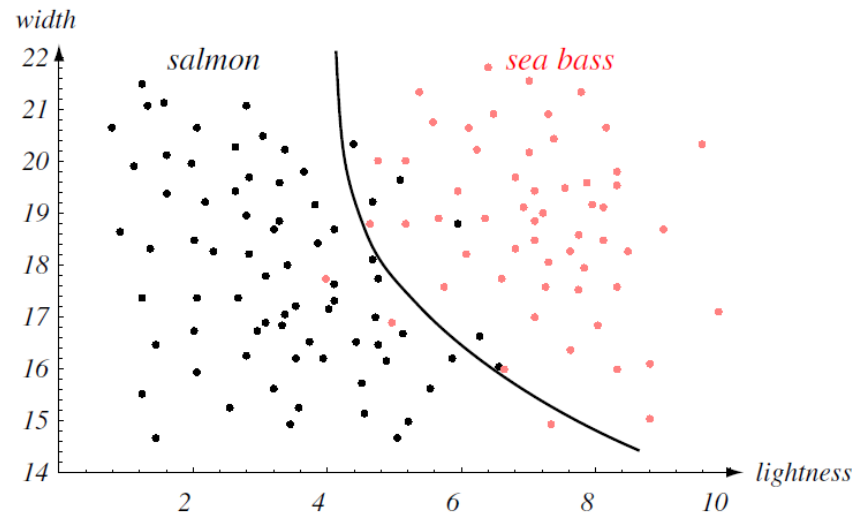
■ Desarrollo ...

■ Generalización ...

- Aumentando el número muestras de entrenamiento se puede mejorar la estimación de las características y la frontera de decisión.
 - En muchas aplicaciones la cantidad de muestras es muy limitado.
 - Si se aumentan las muestra y persiste una frontera de decisión complicada, no se mejora la generalización.

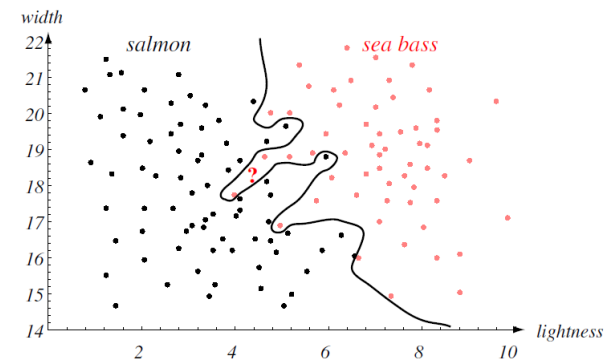
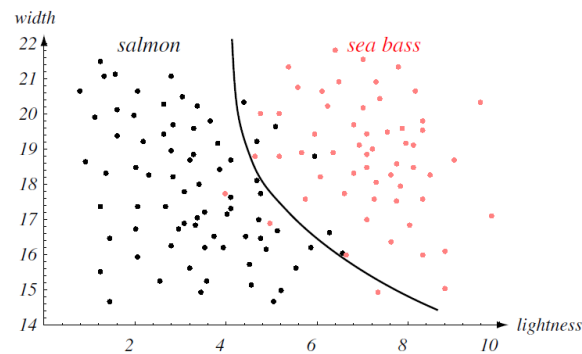
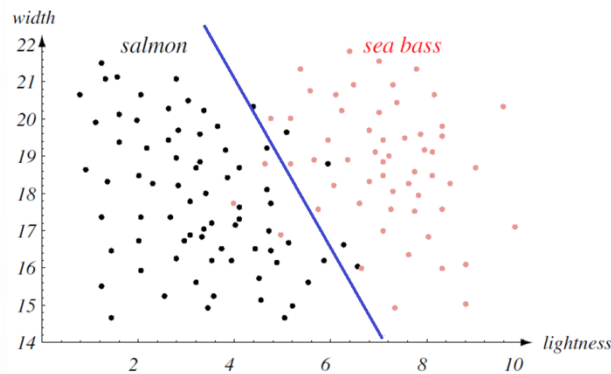
■ Generalización ...

- Buena generalización puede obtenerse con una frontera de decisión más simple:
 - con menor desempeño para las muestras de entrenamiento y mejor respuesta ante nuevos patrones.
- Modelos complejos pueden conducir a clasificadores con baja precisión.



■ Generalización ...

- Problemas de Reconocimiento estadístico de patrones:
 - Cómo determinar automáticamente cuál de las fronteras de decisión es mejor (simple, compleja)?
 - Cómo predecir que tan bien generalizará ante nuevos patrones?



■ Generalización ...

- Para un mismo objeto (patrón) de entrada es posible tener **diferentes funciones** de costo o **fronteras de decisión**.
 - Ejemplo de los pescados:
 - Si se requiere venta de huevos de pescado:
 - *Tarea: separar hembras y machos.*
 - Si se requiere comida de gatos:
 - *Tarea: separar pescados dañados y buenos.*
- Sistema Artificiales de RP de **propósito general** son muy difíciles de implementar.
 - Diferentes tareas de decisión requieren diferentes características y fronteras de decisión.

■ Generalización ...

- El problema de clasificación puede entenderse como el proceso de recobrar el modelo que genera el patrón.
- **Se busca** un clasificador que :
 - Requiera bajo número de características
 - Presente fronteras de decisión simples
 - Requiera un entrenamiento fácil
 - Presente alto porcentaje de generalización
 - Sea rápido y fácil de implementar

■ Generalización ...

■ Análisis por síntesis:

- Cuando se tiene *insuficientes* datos de entrenamiento debe incorporarse conocimiento sobre el dominio del problema.
- El *conocimiento ideal*: contar con el modelo que genera los patrones!!!
- Sistemas de R.P. deben incorporar *algún conocimiento* sobre el método de *producción de los patrones* o su *uso funcional* para asegurar una buena representación del problema.
 - Ejemplo: En un sistema OCR se conoce previamente:
 - Con trazos se construyen los caracteres.
 - Los trazos pueden recuperarse de una imagen digital
 - A partir de los trazos se puede identificar un caracter

■ Campos Relacionados

■ Procesamiento de Señales e Imágenes

- Manipulación matemática de una señal de una o varias dimensiones para extraer información, modificarla o mejorarla.
- Operaciones en el propio dominio o transformado discreto (tiempo, espacio frecuencia discreta, u otro dominio).



Image (a)

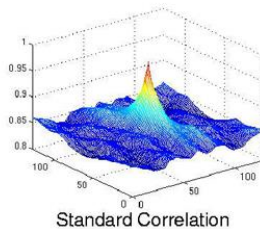
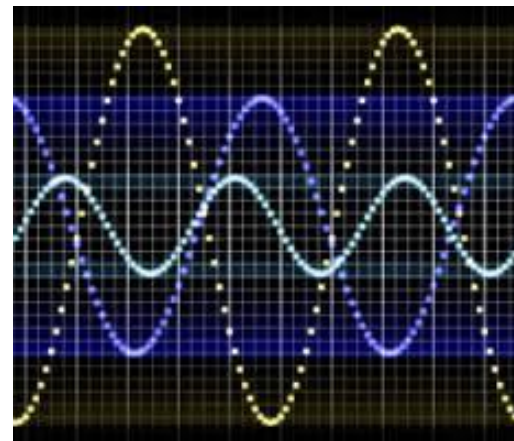
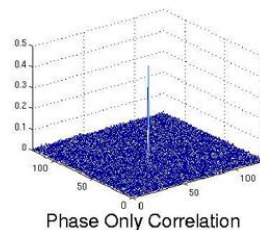


Image (b)



■ Campos Relacionados...

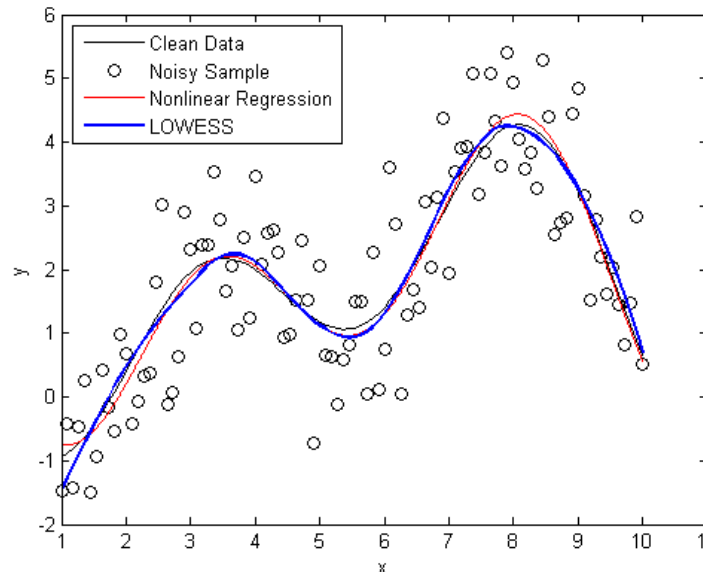
■ Memoria Asociativa

- Almacenamiento y recuperación de información a partir del conocimiento parcial de su contenido, por asociación con otras informaciones sin saber su localización de almacenamiento.
- En actos de memoria asociativa, el sistema toma un patrón y emite otro patrón que es representativo de un grupo general de patrones.
- La información del patrón de salida es menor que la de entrada.

■ Campos Relacionados...

■ Regresión

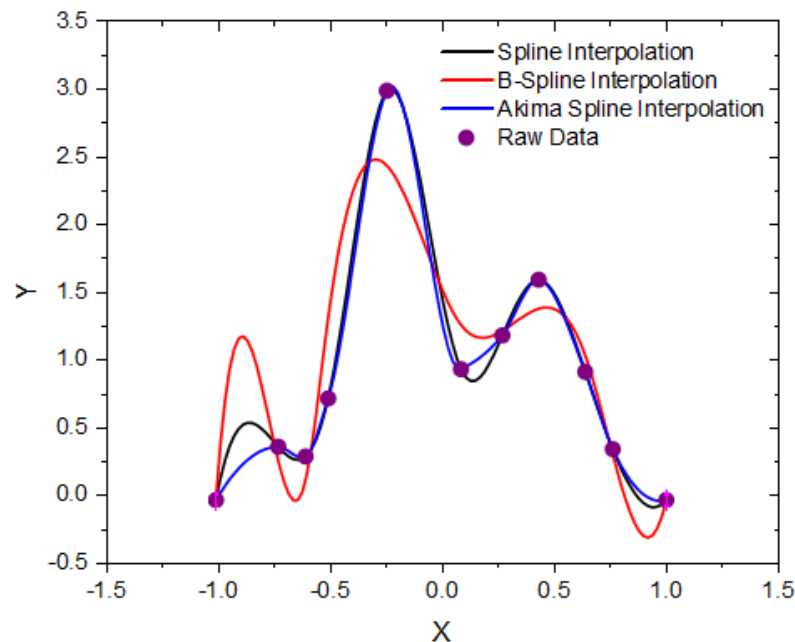
- Proceso para encontrar una descripción funcional de los datos, con la finalidad de predecir valores para nuevas entradas.
- Ampliamente utilizado para la predicción y previsión.



■ Campos Relacionados...

■ Interpolación

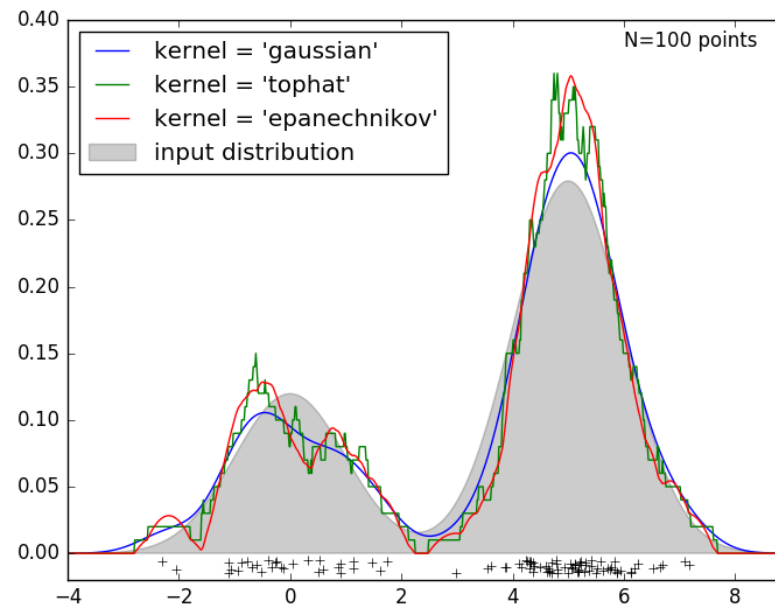
- Encontrar una función para determinar los valores intermedios a partir del conocimiento de un conjunto de datos en un rango de la señal.



■ Campos Relacionados...

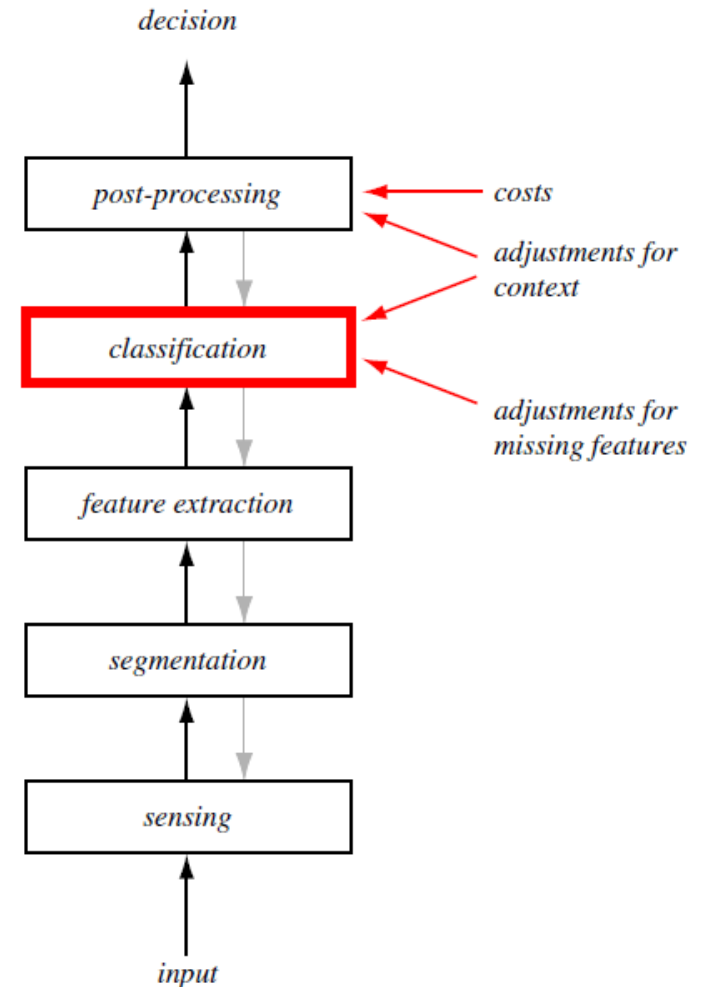
■ Estimación de la densidad

- Obtención de una función que asigna a cada suceso definido sobre la variable aleatoria, la probabilidad de que dicho suceso ocurra.



■ Sistema de Reconocimiento de Patrones

- *Sensado*
- *Segmentación y Agrupamiento*
- *Extracción de características*
- *Clasificación*
- *Post- Procesamiento*



■ Sensado

- Con frecuencia la entrada de un sistema de Reconocimiento de Patrones es un transductor o los datos generados por éste. (Eje. Cámara, micrófono, electrodos, etc.)
- La dificultad del problema está asociada con las características y limitaciones del transductor (ancho de banda, resolución, sensibilidad, distorsión, SNR, latencia, etc.)



■ Segmentación y Agrupamiento

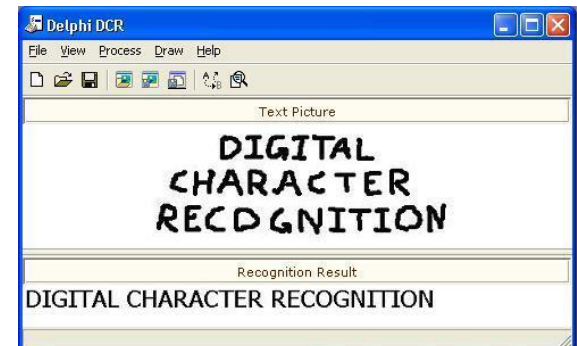
■ Segmentación

- Establecer las fronteras que delimitan un objeto de interés



■ Agrupamiento

- Reconocer como un todo las diversas partes que componen un objeto.



■ Extracción de características

■ Caracterizar un objeto mediante medidas:

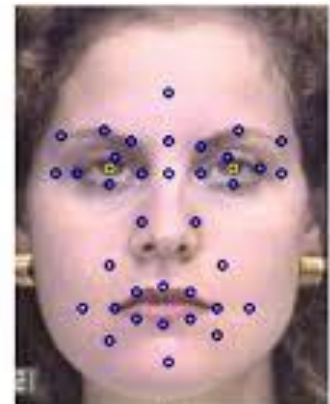
- Muy similares para objetos de la misma categoría.
- Muy diferentes para objetos de distintas categorías.

Multi-State Models for Geometric Feature Extraction



(a)

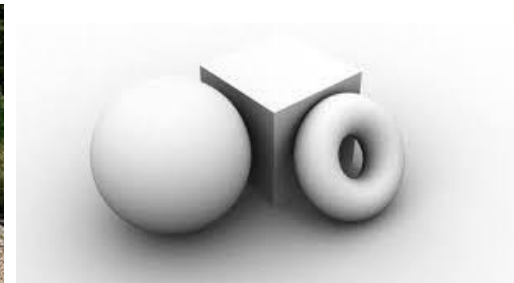
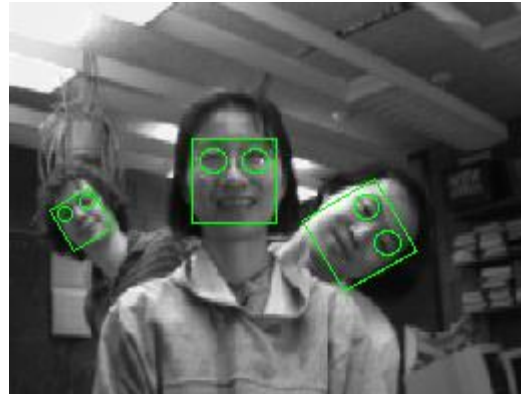
- One State for Brow and Cheek*
- Two States for Eye*
- Two States for Furrows*
- Three States for Lip*



(b)

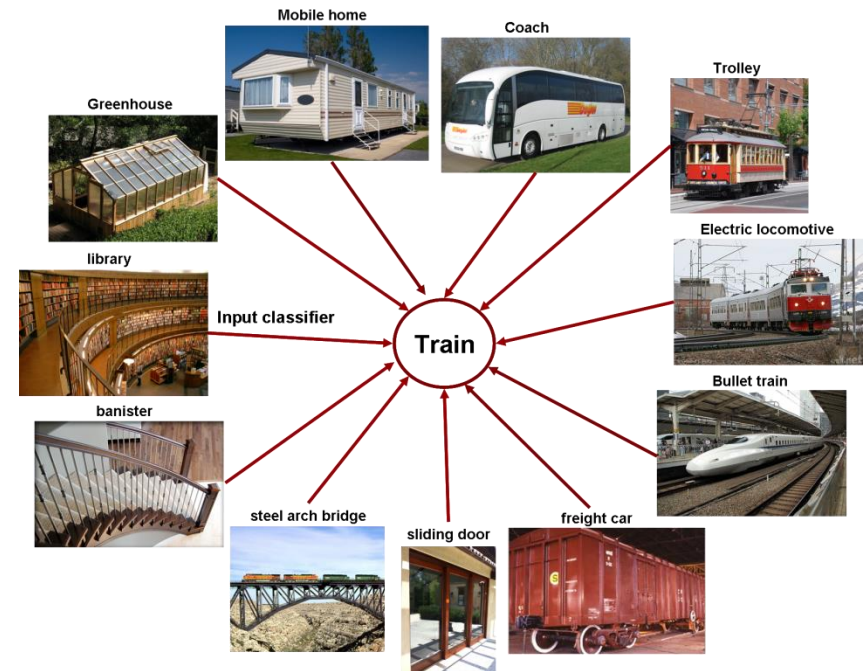
■ Características Invariantes

- Traslación
- Rotación
- Escala
- Oclusión
- Distorsión proyectiva
- Tasa de ocurrencia
- Deformación



■ Selección de características

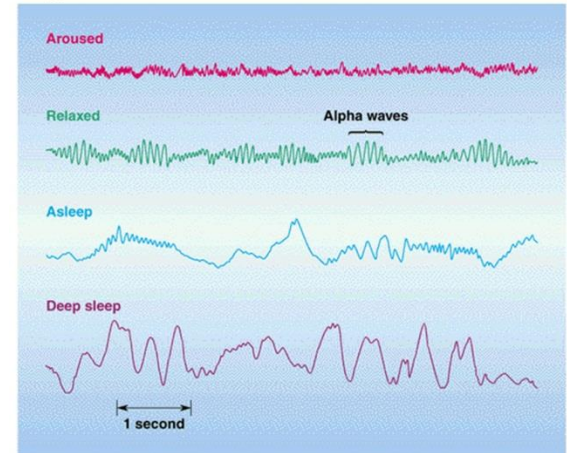
- Escoger las características más relevantes o discriminantes a partir de un conjunto.
- Combinar o transformar características.



■ Clasificación

- Tiene como función asignar una categoría a un objeto a partir de las características.
- Frecuentemente no se obtiene un resultado de clasificación perfecto.
- Dificultad:
 - Variabilidad de las características en una misma categoría respecto a la diferencia entre diferentes clases.
 - Características perdidas (por oclusión, deformación, etc.)

Typical EEGs



EEG Pattern	Frequency (cps)	Typical States of Consciousness
Beta	13 - 24	Normal walking thought, alert problem solving
Alpha	8 - 12	Deep relaxation, blank mind, meditation
Theta	4 - 7	Light sleep
Delta	Less than 4	Deep sleep

■ Post-Procesamiento

- Los clasificadores se utilizan para recomendar un acción posteriormente.
- Cada acción posterior tiene un costo asociado.
- El post-procesador utiliza la salida del clasificador para decidir la acción que finalmente se realizará.



Download from:
DesignTime.com

© 2008
All Rights Reserved

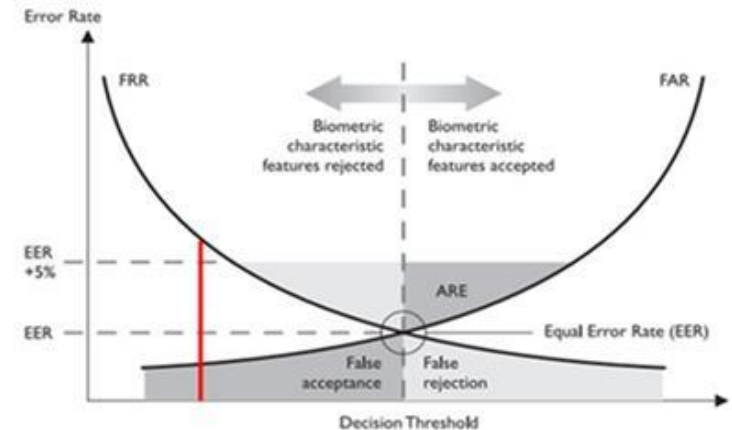
■ Post-Procesamiento

■ Tasa de error

- Porcentaje de nuevos patrones que son asignados a clases erradas.
- Medida simple para determinar el desempeño del clasificador

■ Riesgo

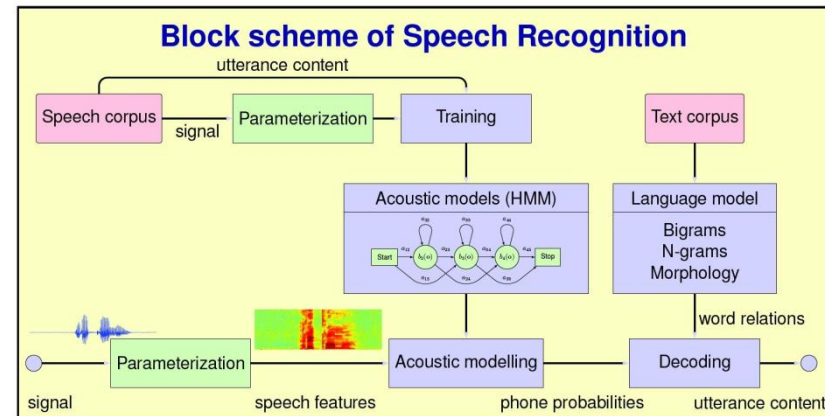
- Mínimo costo total esperado



■ Post-Procesamiento

■ Contexto

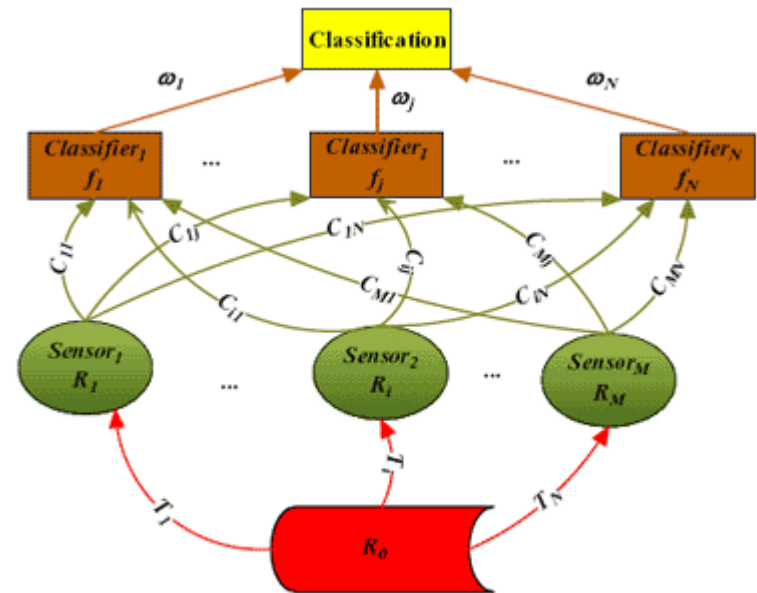
- Información dependiente del problema diferente al patrón mismo.
 - Ejemplo: frecuencia de vocales en reconocimiento de palabras.
- Sirve para mejorar el desempeño del clasificador.



■ Post-Procesamiento

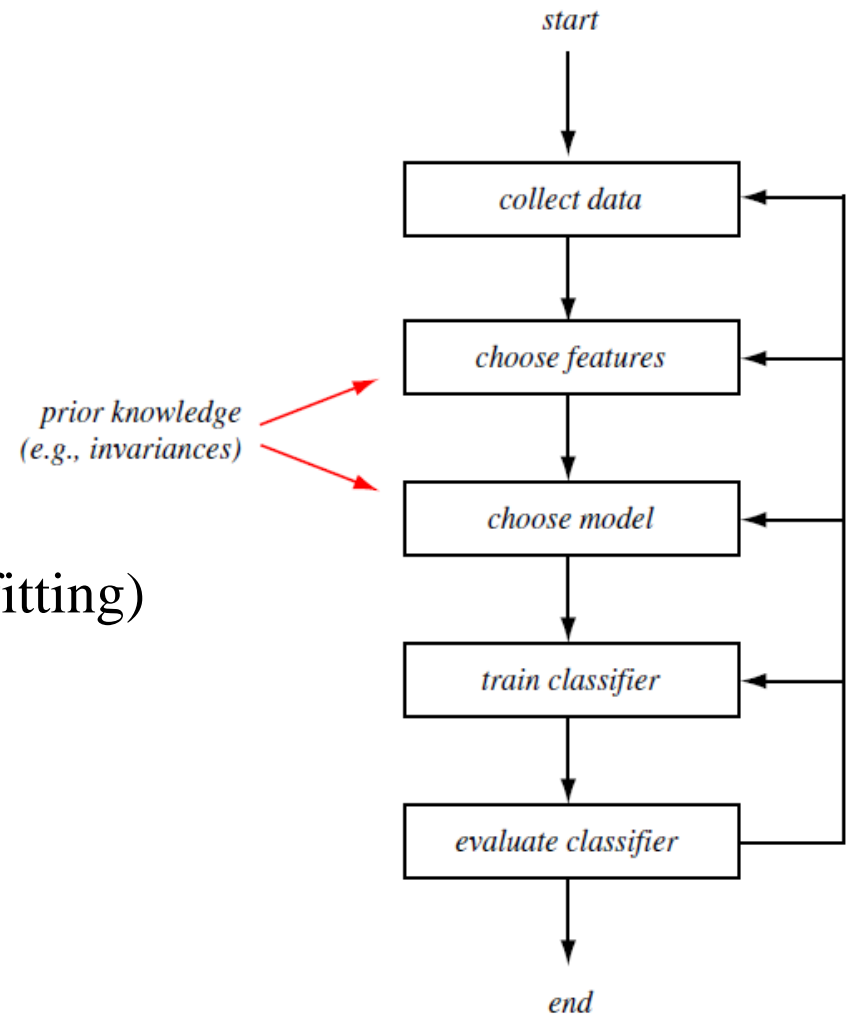
■ Clasificadores Múltiples

- Uso de diversos clasificadores para mejorar el desempeño global de la clasificación.
- Los clasificadores se especializan en diferentes aspectos del patrón de entrada.
- Se requiere analizar el conjunto de salidas para tomar la mejor decisión.



■ Ciclo de Diseño

- Recolección de datos
- Selección de características
- Selección del modelo
- Entrenamiento
- Evaluación
 - Sobre-entrenamiento (Overfitting)
- Complejidad computacional

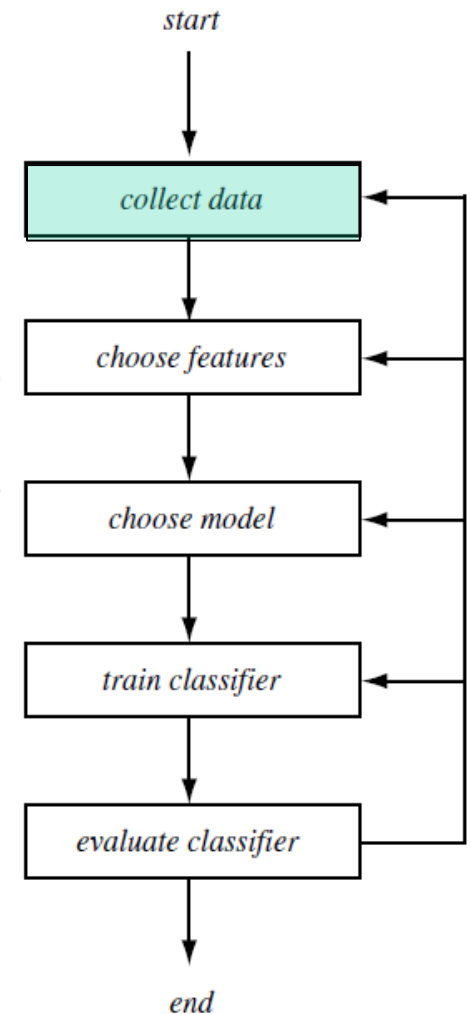


■ Ciclo de Diseño

■ Recolección de datos

- Puede requerir gran parte del costo del desarrollo del sistema de reconocimiento de patrones.
- Puede realizarse un estudio de viabilidad preliminar con un pequeño data-set.
- Un buen desempeño del sistema se obtiene con data sets grandes.
- Cómo establecer si el tamaño del data set es adecuado y representativo?

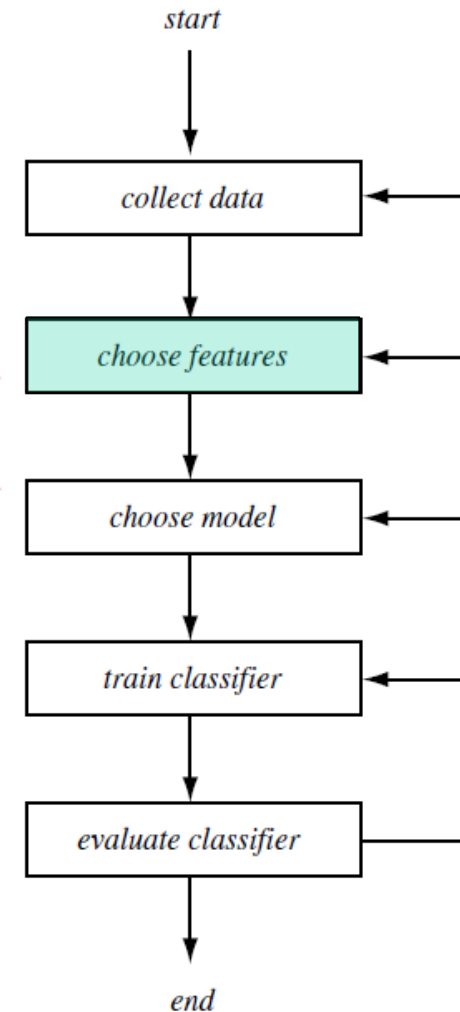
*prior knowledge
(e.g., invariances)*



■ Ciclo de Diseño

■ Selección de características

- Escoger rasgos diferenciadores es esencial y depende del dominio del problema.
- El conocimiento *a priori* es fundamental en la selección, pero requiere mayor experiencia.
- Las características deben ser *simples* de extraer, invariantes a transformaciones irrelevantes al problema, insensibles a ruido, y discriminantes entre categorías.
- Cómo combinar información *a priori* y datos empíricos para encontrar características relevantes y efectivas?

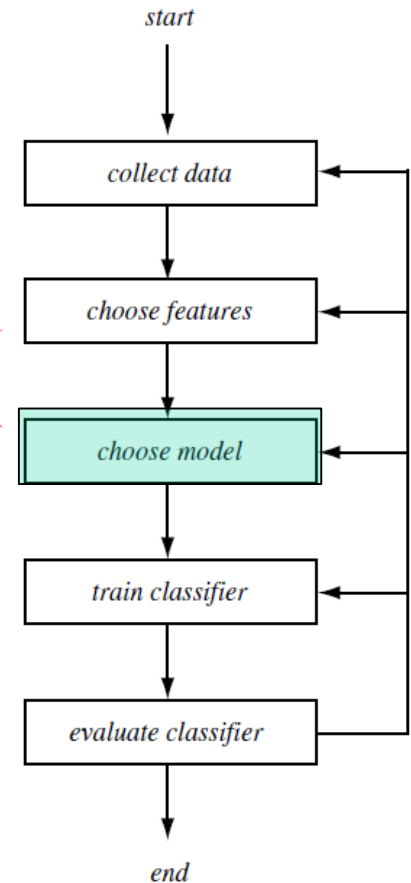


■ Ciclo de Diseño

■ Selección del modelo

- Cómo determinar cuánto difiere el modelo seleccionado del modelo subyacente generador de los patrones?
- Cuándo es necesario un nuevo modelo?
- Cuánto puede mejorar el desempeño?
- Puede desecharse una clase del modelo e introducir otra?

prior knowledge
(e.g., invariances)

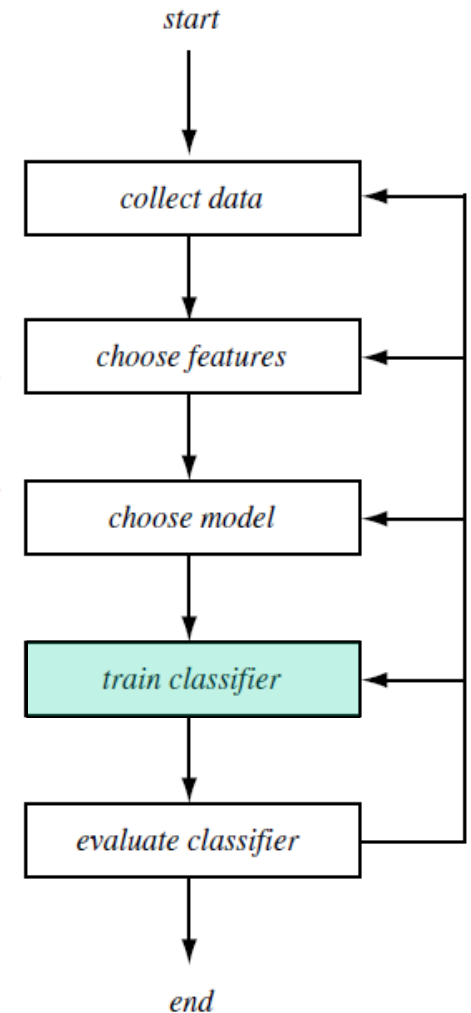


■ Ciclo de Diseño

■ Entrenamiento

- Proceso de ajustar los parámetros del clasificador a partir de los datos.
- No existe un método universal de entrenamiento.
- Existen muchos procedimientos de entrenamiento de clasificadores y selección del modelo.
- Los métodos más efectivos involucran aprendizaje a partir de patrones ejemplo.

*prior knowledge
(e.g., invariances)*

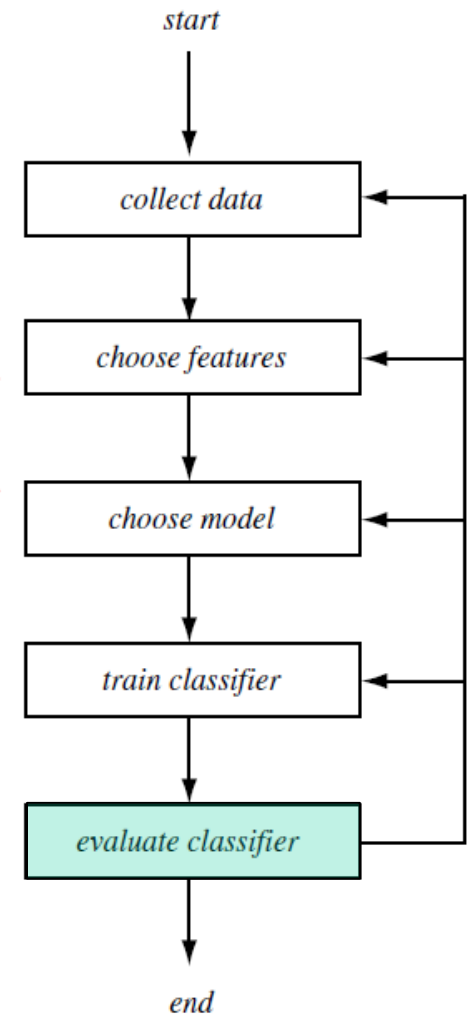


■ Ciclo de Diseño

■ Evaluación

- Necesario para medir el desempeño del sistema e identificar la necesidad de mejoras en sus componentes.
- Excesiva complejidad en el sistema puede conducir a una perfecta clasificación de las muestras de entrenamiento (Overfitting)
- Cómo encontrar la mejor complejidad de un clasificador?

*prior knowledge
(e.g., invariances)*

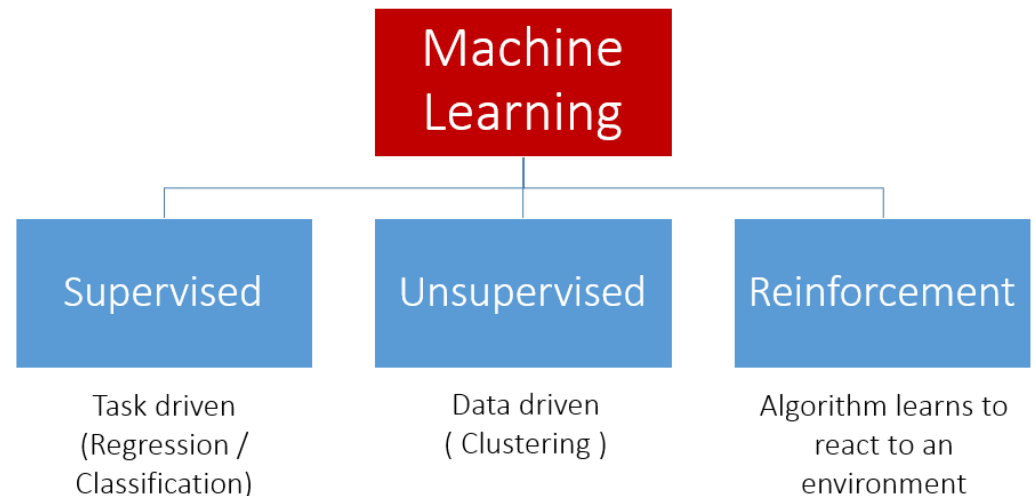


■ Aprendizaje

■ Tipos de aprendizaje

- Supervisado
- No-supervisado
- Con reforzamiento

Types of Machine Learning



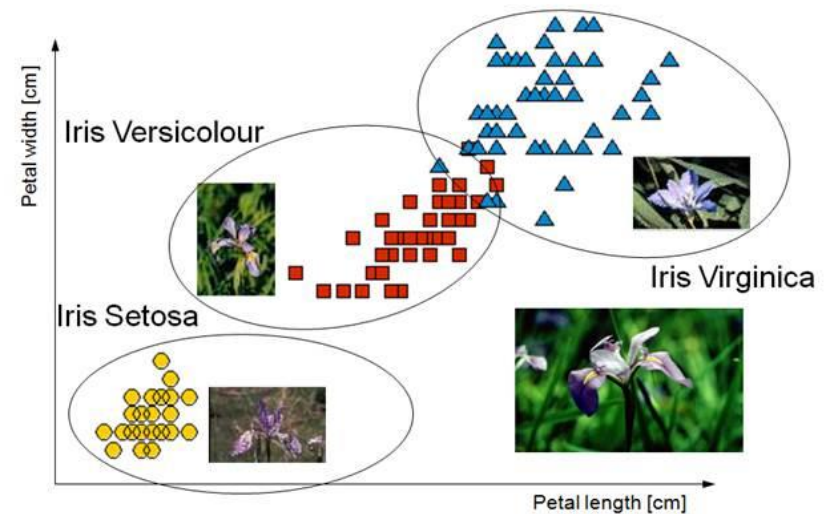
■ Cómo una máquina aprende el modelo a partir de los datos?

- Mediante procedimientos de adaptación de un modelo a partir de datos.
- Utilización de información a partir de muestras de entrenamiento para diseñar un clasificador.
- Algoritmo para ajustar los parámetros de un modelo para reducir una medida de error a partir de un conjunto de datos de entrenamiento.



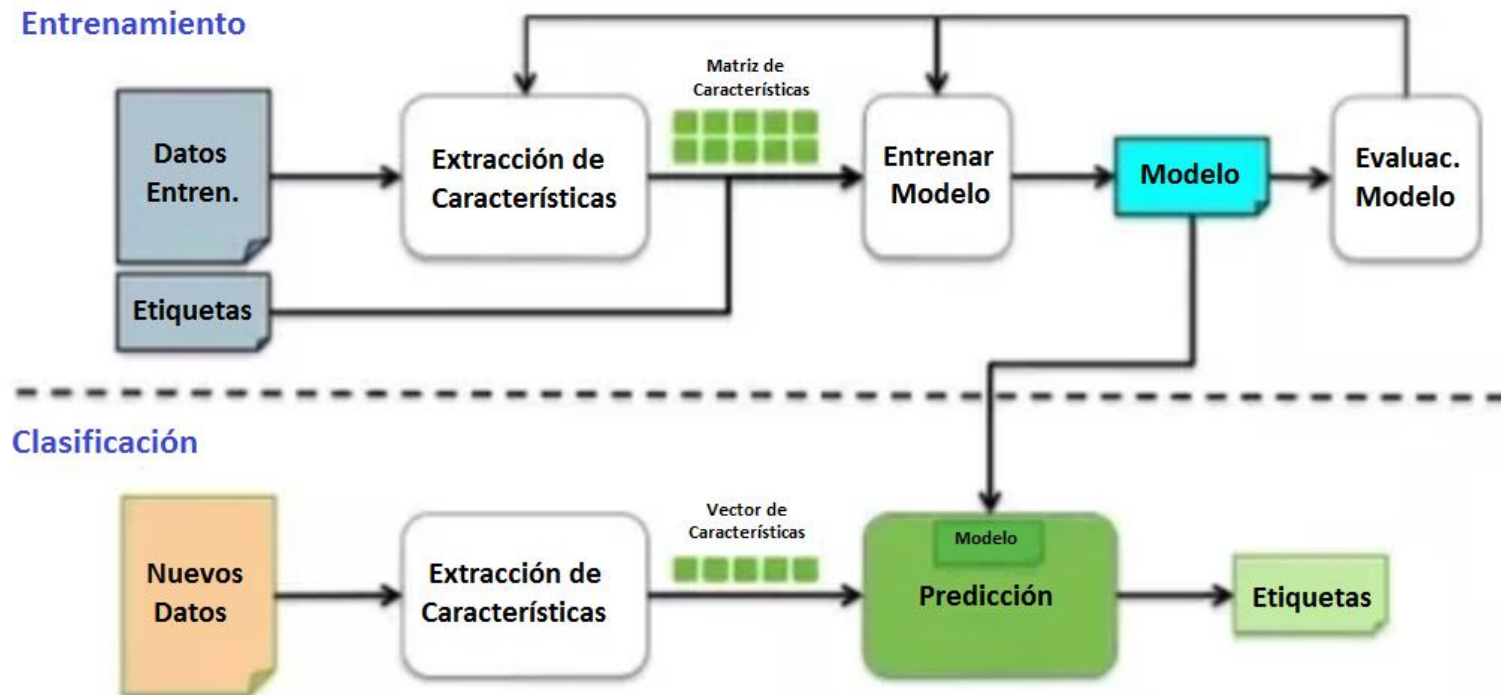
■ Aprendizaje Supervisado

- Se cuenta con información sobre la membresía (etiqueta) o costo de cada patrón del conjunto de entrenamiento.
- El data-set de entrenamiento contiene información sobre la clase a la cual pertenece cada patrón.
- Los algoritmos de aprendizaje utilizan el error entre la salida estimada y la salida real para ajustar los parámetros del modelo.



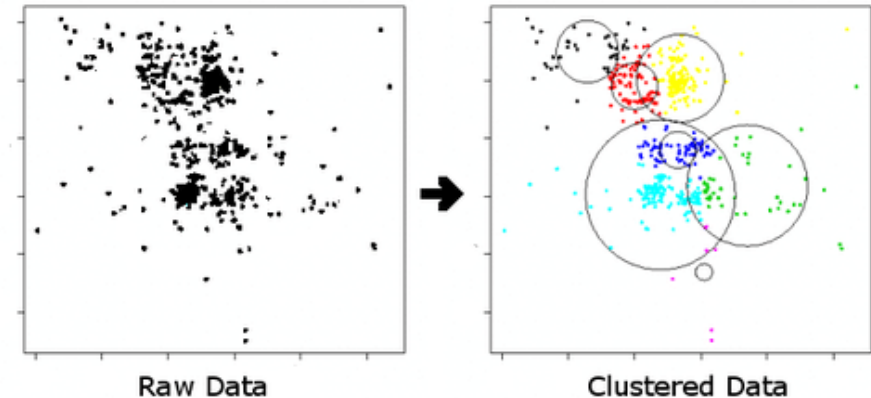
http://www.ebtic.org/source/ebtics_view_ml_big_data/image001.jpg

■ Aprendizaje Supervisado



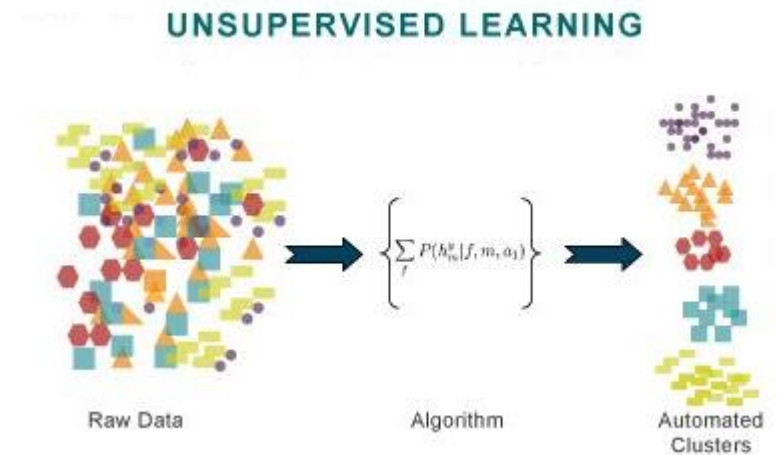
■ Aprendizaje **NO** Supervisado

- El data-set de entrenamiento **NO** contiene información sobre la clase a la cual pertenece cada patrón.
- Los algoritmos de aprendizaje forman grupos *naturales* (clusters) a partir de los datos de entrenamiento basados en criterios de similitud.

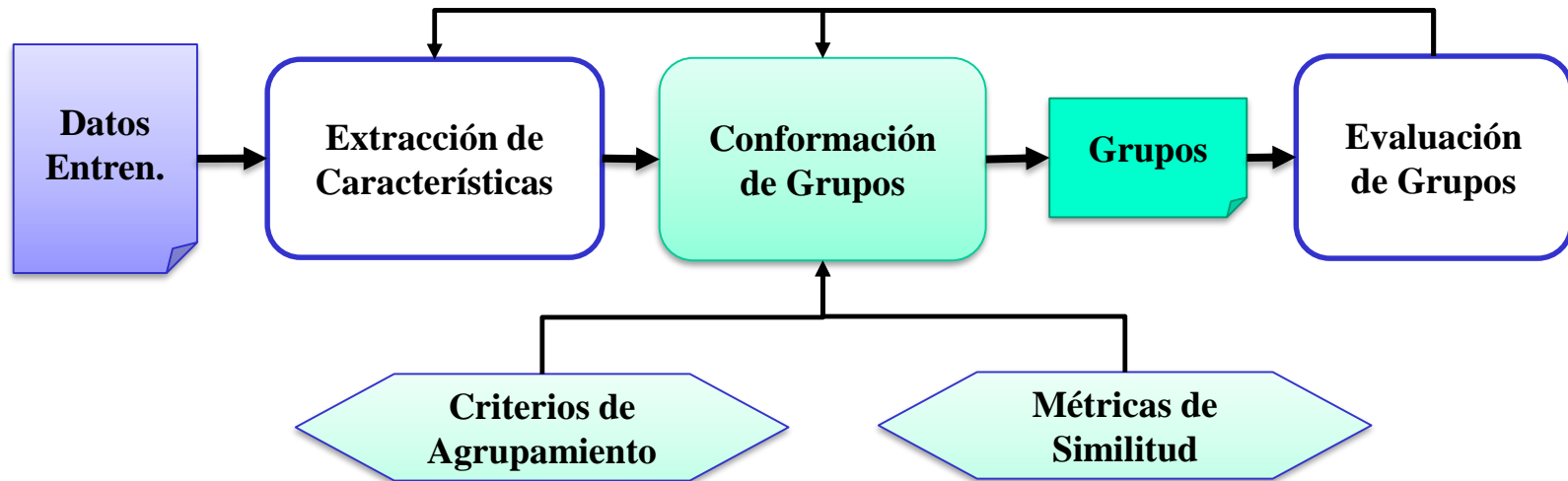


■ Aprendizaje NO Supervisado

- La información de los clusters se utilizan para ajustar los parámetros del modelo.
- No requiere de una “enseñanza”
- Diferentes métodos conducen a diferentes grupos .

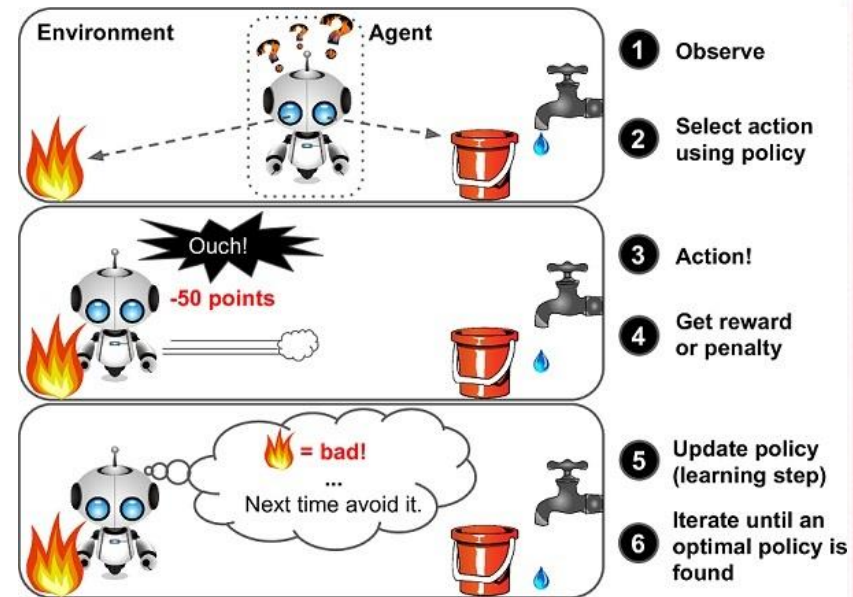


■ Aprendizaje **NO** Supervisado



■ Aprendizaje por Reforzamiento

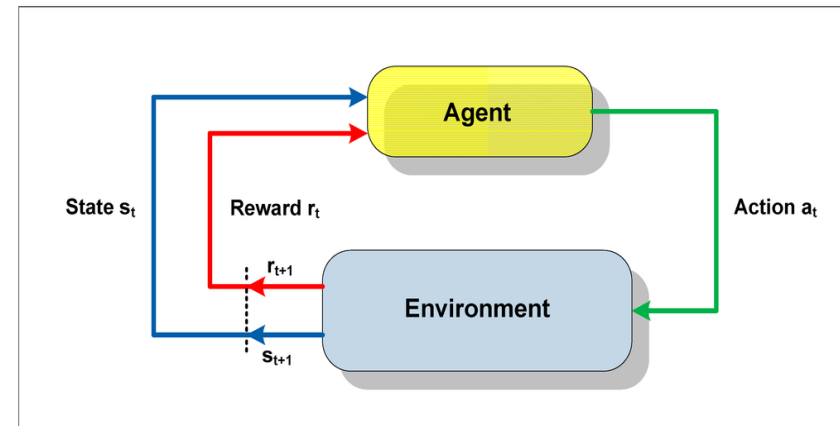
- No se cuenta con información de la clase de pertenencia de los patrones de entrenamiento (estados).
- Se cuenta con una evaluación (recompensa) sobre si la decisión (acción) es correcta o incorrecta.
- Busca maximizar las recompensas al asociar los estados con las acciones.



<http://generation-ai.com/tags/artificial-intelligence/>

■ Aprendizaje Por Reforzamiento

- Se consideran agentes software autónomos en un ambiente, que deben tomar decisiones en función de su estado actual.
- El ambiente otorga recompensas (positiva o negativa) al agente en función de su decisión.
- Se busca maximizar las recompensas positivas.
- No se da información sobre la clase ni de la magnitud del error obtenido.





PSI Percepción y Sistemas Inteligentes



Universidad del Valle

Facultad de Ingeniería
Escuela de Ingeniería Eléctrica y Electrónica

humberto.loaiza@correounivalle.edu.co

