

Clase I: ¿Qué es Machine Learning?



- Introducción al Machine Learning,
- 1er cuatrimestre 2019
- Ana Georgina Flesia
- <http://www.famaf.unc.edu.ar/~flesia>
- georgina.flesia@gmail.com
- Oficina 370

Acerca de este curso

2

- Programa
<http://www.famaf.unc.edu.ar/wpcontent/uploads/2019/02/Flesia.pdf>
- Cuando: Miércoles y viernes de 14 a 16 hs.
- Donde: Aula TIC frente a vinculación Tecnológica.
- Horario de consulta: Jueves 14 a 16 hs. Oficina 370.
- Moodle
 - <http://www.famaf.proed.unc.edu.ar/course/view.php?id=470>

Bibliografía

3

- Mitchell, T. **Machine Learning**, McGraw Hill, 1997.
- Marsland, S. **Machine Learning: an algorithmic perspective**. CRC Press 2009
- Raschka, S., & Mirjalili, V. **Python Machine Learning**, 2nd Ed. Birmingham, UK: Packt Publishing. 2018.
- Murphy K. **Machine Learning: a probabilistic perspective**. MIT Press 2012.
- Duda, R., Hart, P., Stork, D. **Pattern Classification**, Wiley 2002.
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). **The Elements of Statistical Learning**. Springer series in statistics.

Pantallazo general

4

- Objetivo general del área
- Como construir programas de computación que mejoren
 - automáticamente
 - con nuevas experiencias

Pantallazo general

5

- Como construir programas de computación que mejoren automáticamente con nuevas experiencias
- Tom Mitchell, Carnegie Mellon Professor:
 - Un programa se dice que aprende de la **experiencia E** con respecto a alguna clase de **tareas T** y medida de **desempeño P** si su desempeño en la tarea **T**, medida por **P**, mejora con la experiencia **E**.

Tareas y algoritmos

6

- Clase de tareas **T**
 - Reconocimiento de patrones
 - Creación de relaciones
 - Toma de decisiones
- La medida **P** y el tipo de experiencias **E** dividen los algoritmos en clases

Ejemplo: Reconocer dígitos escritos a mano

7

- Un programa se dice que aprende de la **experiencia E** con respecto a alguna clase de **tareas T** y medida de **desempeño P** si su desempeño en la tarea **T**, medida por **P**, mejora con la experiencia **E**.

Elements of Statistical Learning (2nd Ed.) ©Hastie, Tibshirani & Friedman 2009 Chap. 1



FIGURE 1.2. Examples of handwritten digits from U.S. postal envelopes.

Hastie, Tibshirani and Friedman

8

Elements of Statistical Learning (2nd Ed.) ©Hastie, Tibshirani & Friedman 2009 Chap 1



FIGURE 1.2. *Examples of handwritten digits from U.S. postal envelopes.*

Ejemplo

9

- **Tarea T**: clasificar dígitos a partir de imágenes
- **Medida de desempeño P**: porcentaje de dígitos clasificados correctamente
- **Experiencia de entrenamiento E**: data base de dígitos con clasificaciones anotadas, MNIST

Elements of Statistical Learning (2nd Ed.) ©Hastie, Tibshirani & Friedman 2009 Chap 1



FIGURE 1.2. Examples of handwritten digits from U.S. postal envelopes.

Clases de algoritmos de aprendizaje

10

- Aprendizaje supervisado
- Descubrimiento del conocimiento
- Aprendizaje reforzado
- Aprendizaje evolutivo

Sebastian Rashka 2018

11

Supervised Learning

- Labeled data
- Direct feedback
- Predict outcome/future

Unsupervised Learning

- No labels/targets
- No feedback
- Find hidden structure in data

Reinforcement Learning

- Decision process
- Reward system
- Learn series of actions

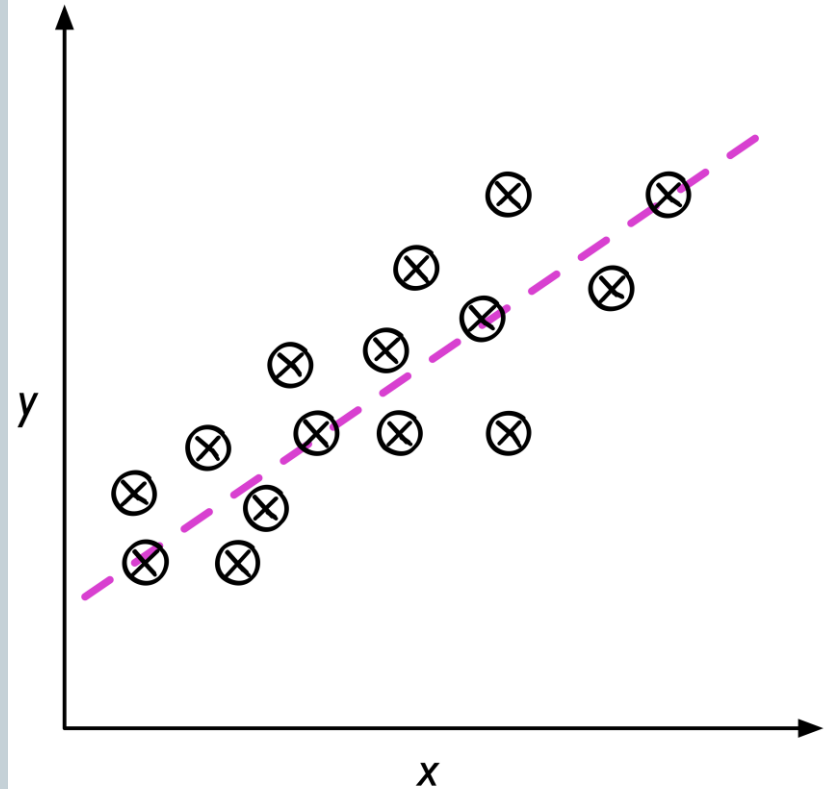
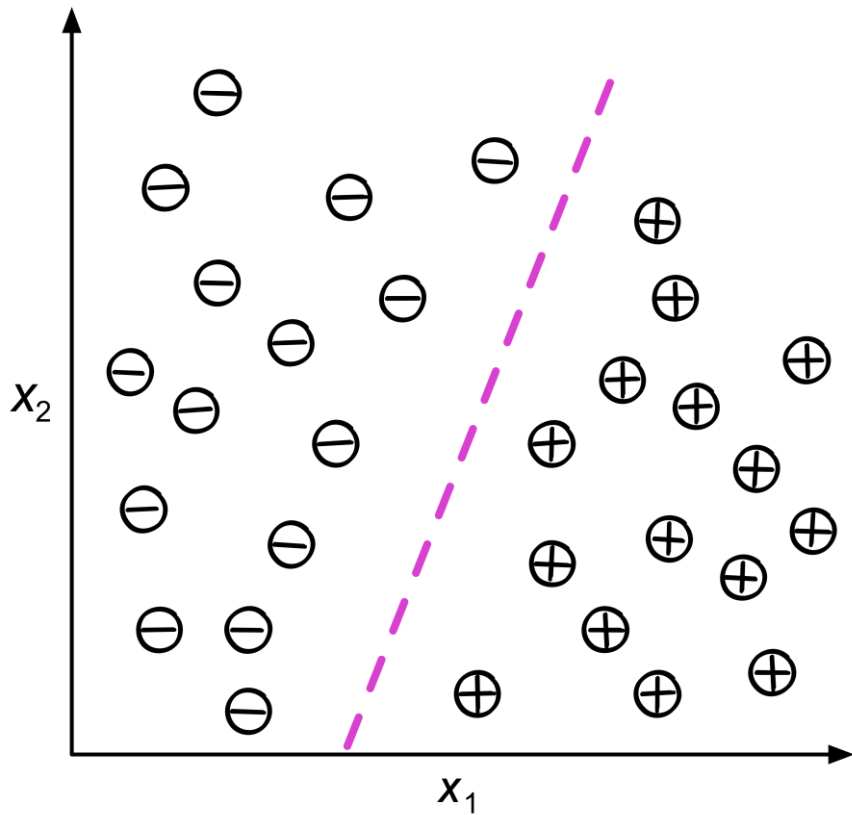
Aprendizaje Supervizado

12

- Aprendizaje supervisado
 - Un conjunto de muestras de entrenamiento con las respuestas correctas
 - basado en ese conjunto de entrenamiento, el algoritmo generalice automáticamente, prediciendo la etiqueta de un nuevo dato
- Esto se suele llamar también aprendizaje a partir de ejemplos.

Aprendizaje Supervizado

13



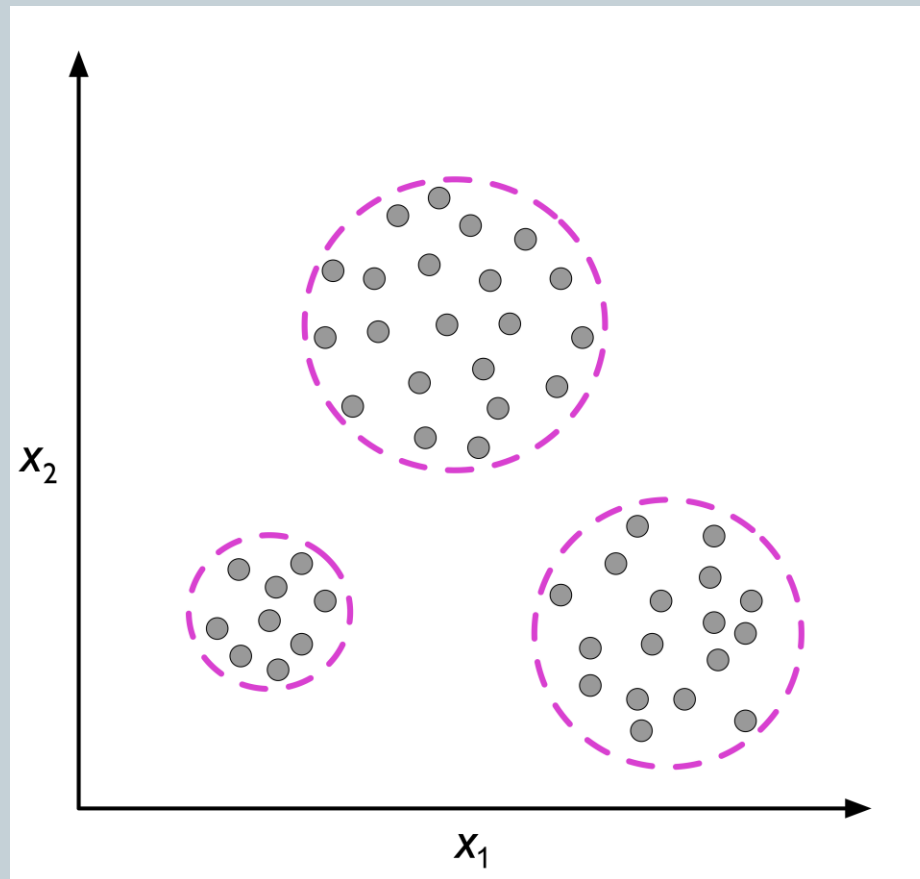
Aprendizaje no supervisado

14

- Aprendizaje no supervisado:
 - No se proveen respuestas correctas
 - El algoritmo trata de identificar similitudes entre los datos ingresados para agrupar juntos los datos con algo en común
- Se suele llamar Descubrimiento del Conocimiento

Aprendizaje no supervisado

15



Aprendizaje reforzado

16

- Se encuentra entre el aprendizaje supervisado y el no supervisado
- Se proveen respuestas parciales, como que el algoritmo erra pero no en qué se equivoca.
- Suele llamarse aprendizaje con un crítico porque este monitorea las respuestas pero no dice como mejorar.

Aprendizaje reforzado

17

Carnegie Mellon's Never Ending Language Learner

Uses tweeter to obtain feedback on new beliefs.

<https://www.cs.cmu.edu/link/campus-what-does-computer-believe>

Aprendizaje evolutivo

18

- Evolución biológica puede ser vista como un proceso:
 - organismos biológicos se adaptan para mejorar
 - su tasa de supervivencia
 - y su probabilidad de tener descendencia en su entorno.
- Modelos genéticos basados en medidas de aptitud y diferentes conceptos de mutación generan los llamados algoritmos genéticos.



Reconocimiento de Patrones

Reconocimiento de patrones: Duda, Hart, Stork

20

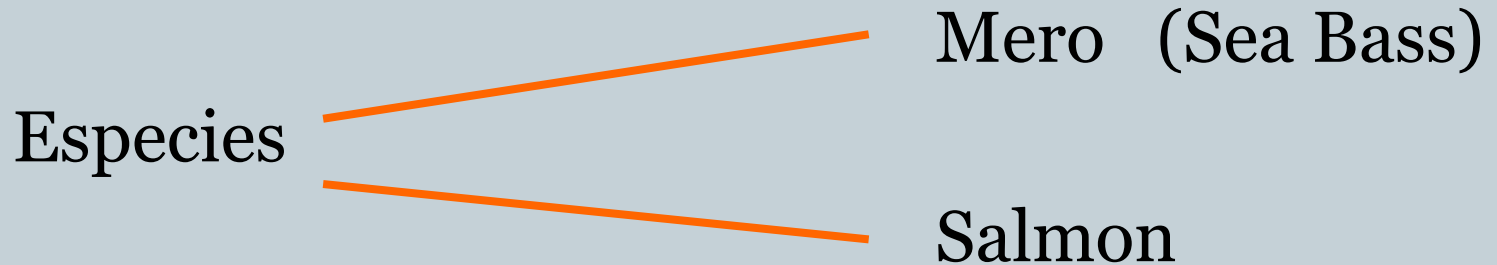
Construir una máquina capaz de reconocer patrones que mejore con la experiencias recibidas:

- Reconocimiento de frases habladas
- Identificación de huellas digitales
- Reconocimiento óptico de escritura
- Identificación de secuencias de ADN

Ejemplo Duda, Hart, Stork

21

- “Separar el pescado entrante en una cinta de transporte, de acuerdo con la especie, usando sensores ópticos”



Análisis del Problema

22

- Montar una cámara y tomar algunas imágenes de muestra para extraer características
 - ✦ Largo
 - ✦ Tono (color)
 - ✦ Tamaño
 - ✦ Numero y forma de las aletas
 - ✦ Posición de la boca, etc...
- Este es el conjunto de las características sugeridas para explorar en el clasificador

Preprocesamiento

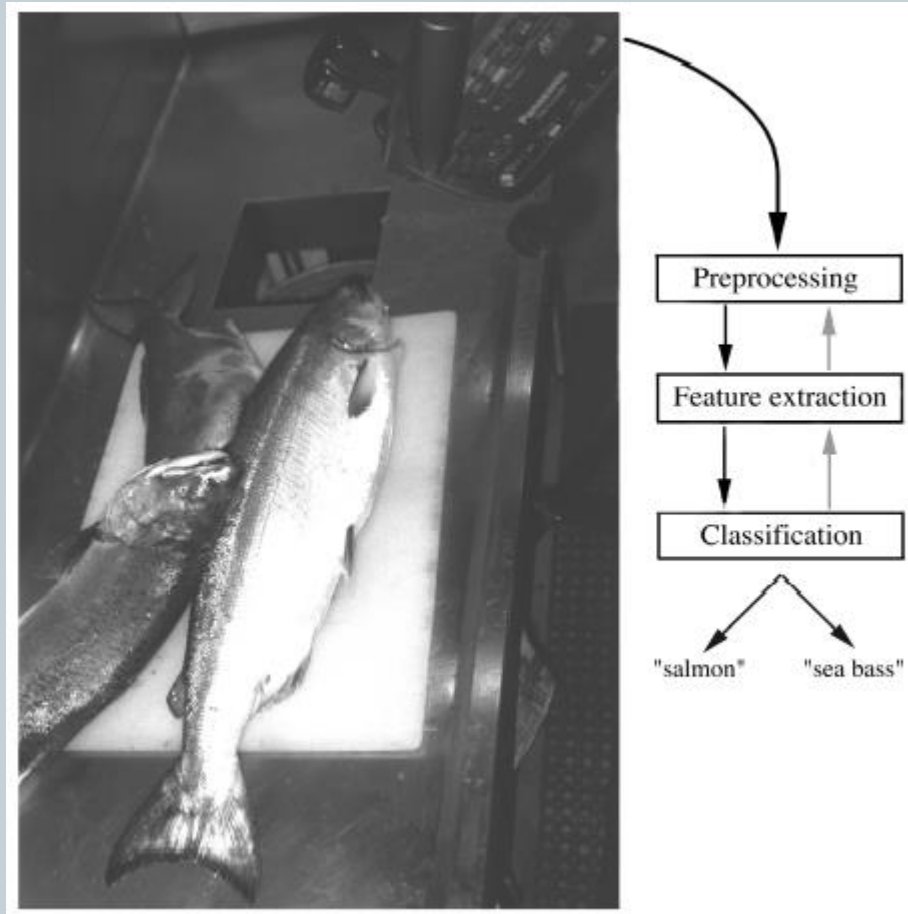
23

- Usar segmentación para aislar los peces uno de otro y con respecto al fondo
- Enviar información de cada pez al extractor de características

Extracción de características

24

- Extractor de características es un algoritmo que toma la imagen segmentada y calcula las características a enviar al clasificador
 - Largo
 - Ancho
 - Tamaño escamas
 - Tono

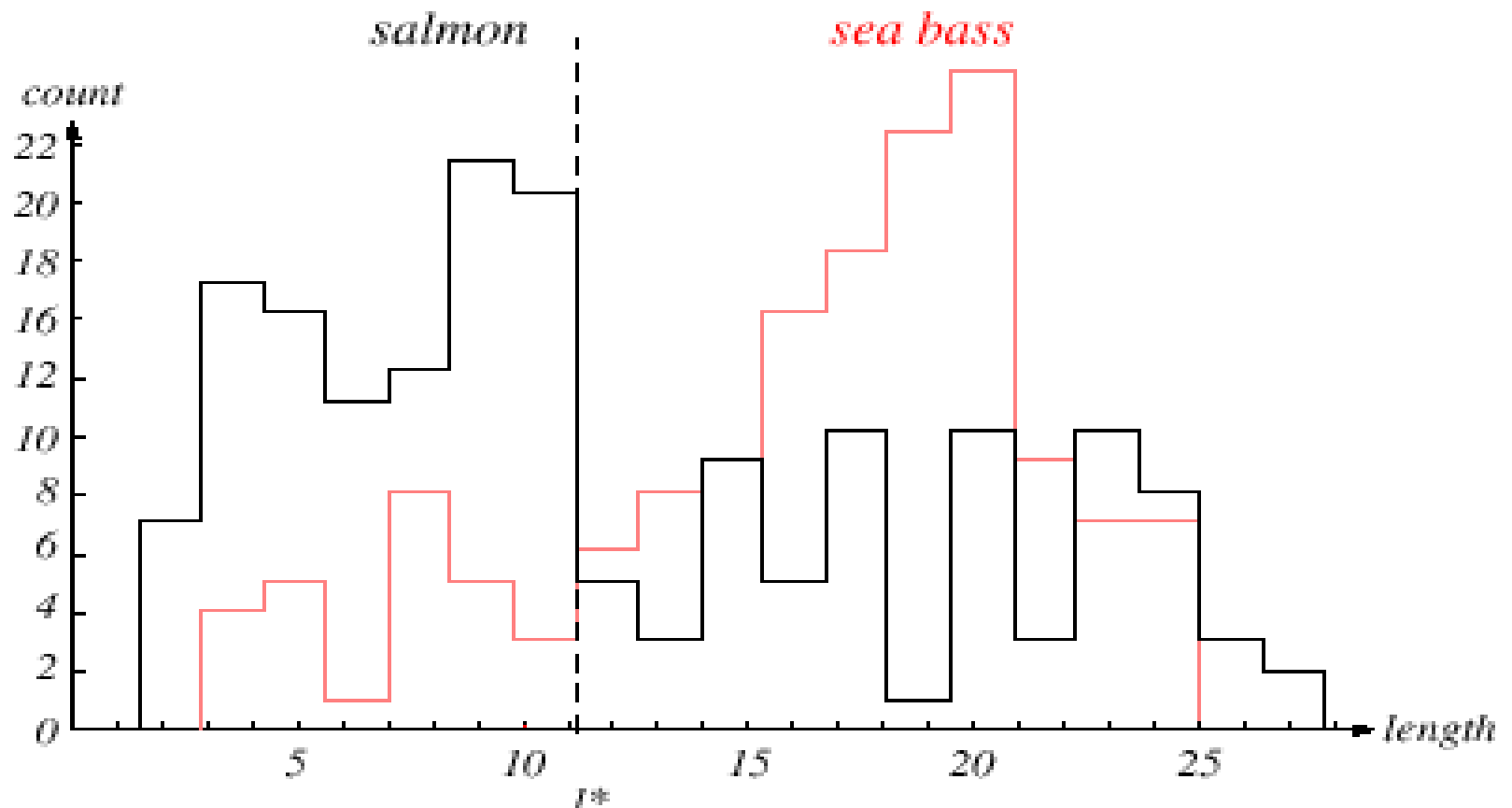


Clasificación

26

- Decidir el conjunto de características a usar
 - ✦ Conjunto “suficiente”
 - ✦ Conjunto “mínimo”
- Seleccionar el largo del pez como una posible característica para discriminación
- Clasificación lineal,

$\text{Largo}(x) < I^*$  $x = \text{salmon}$



En promedio son distintos, pero no es suficiente.....

El **largo** es una característica muy pobre!

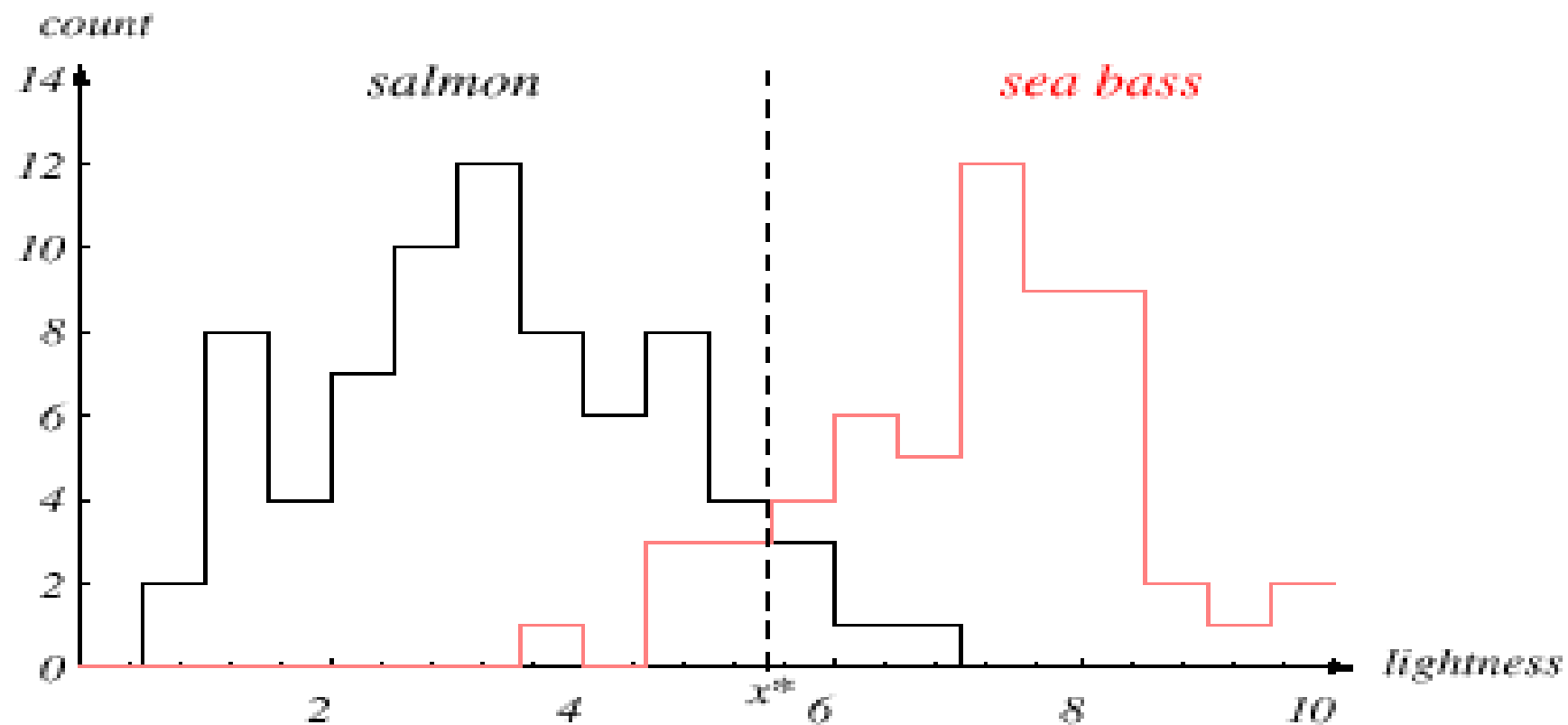
28

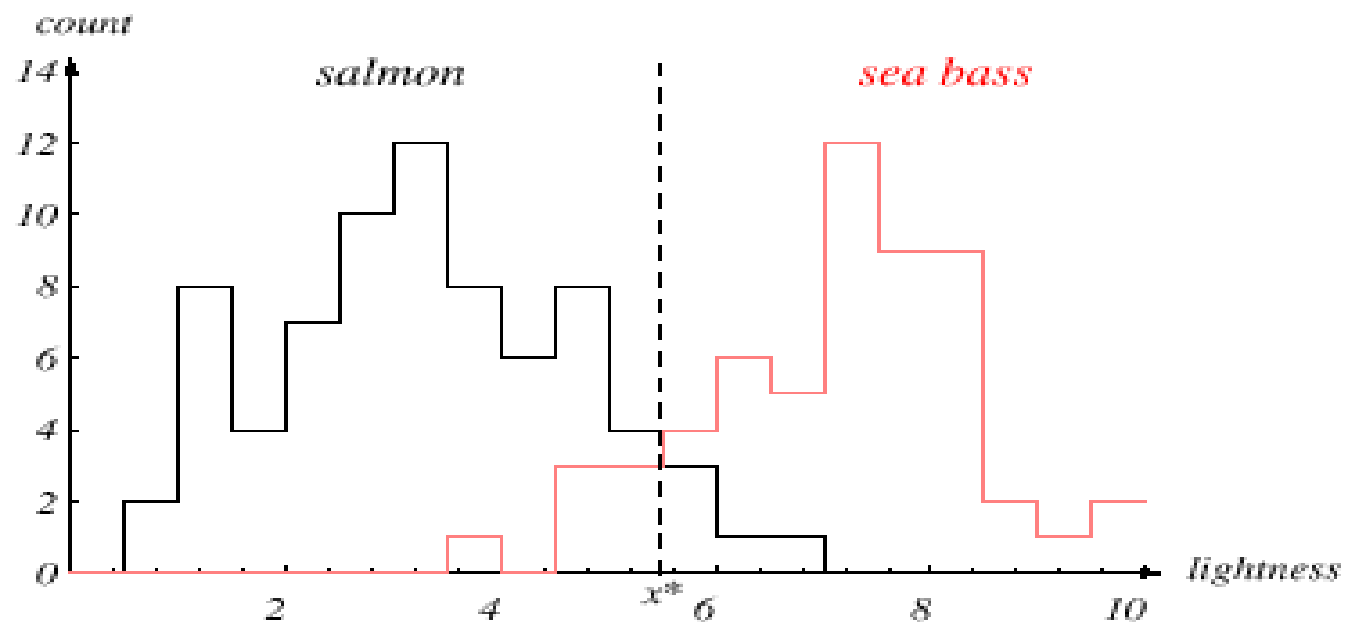
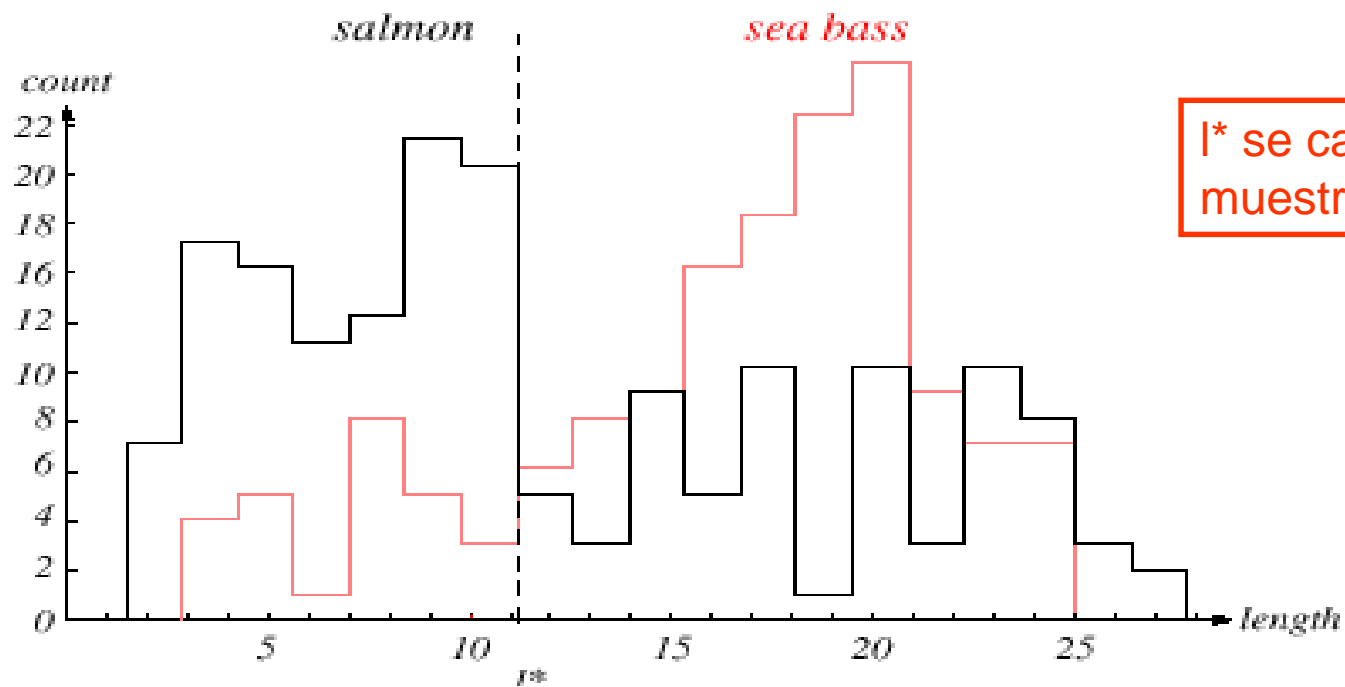
- Otras características
 - ✦ Largo
 - ✦ Tono (color)
 - ✦ Tamaño
 - ✦ Numero y forma de las aletas
 - ✦ Posición de la boca, etc...
- **Tono de las escamas** posible característica.
- Clasificación lineal,

$\text{Tono}(x) < x^*$



$x = \text{salmón}$

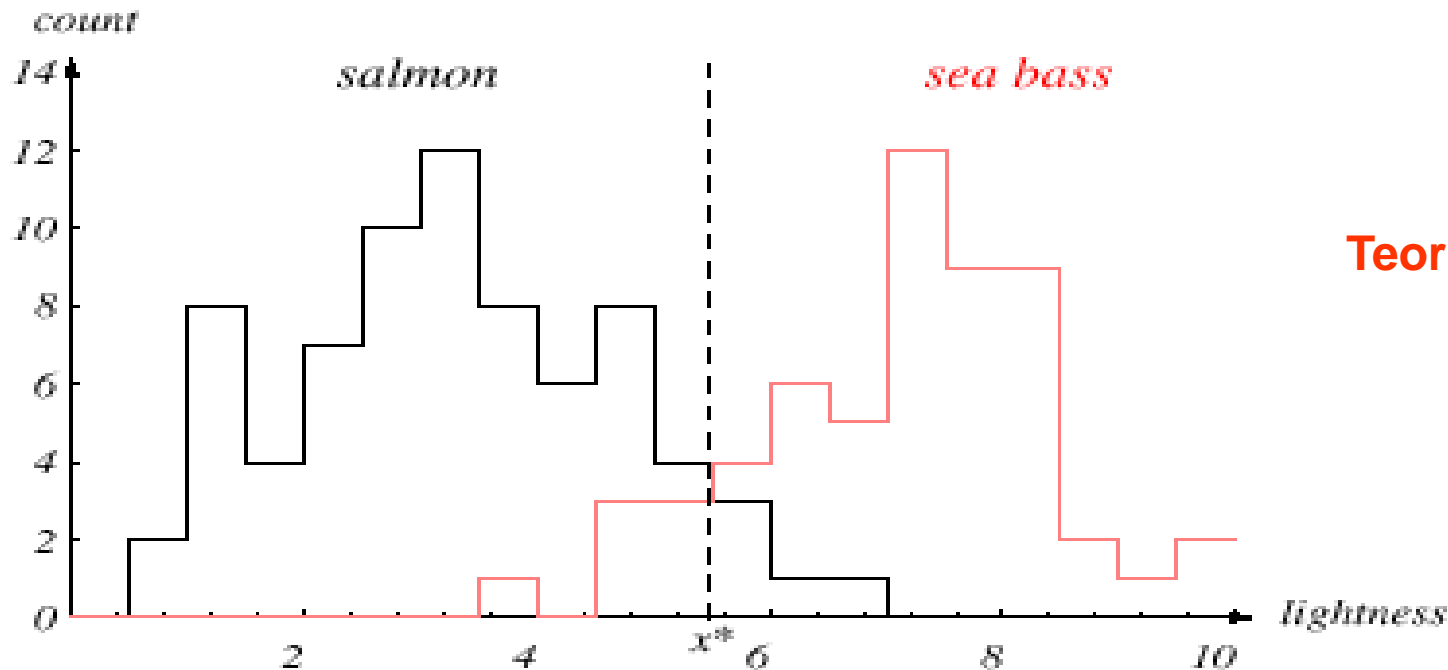




Umbral de decisión x^* y Costo de clasificación

31

- Mover nuestro borde de decisión hacia valores menores del Tono para minimizar el costo (reducir el numero de Meros que son clasificados como Salmones !)

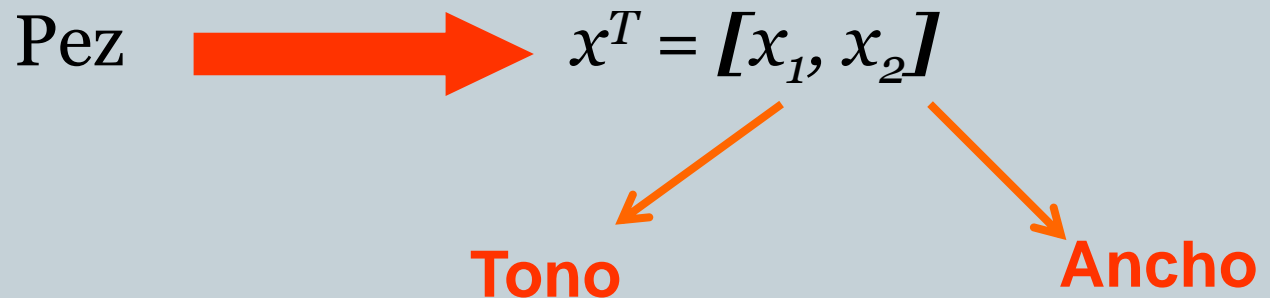


Teoría de la decisión

Todavía hay algunos salmones mal clasificados.

32

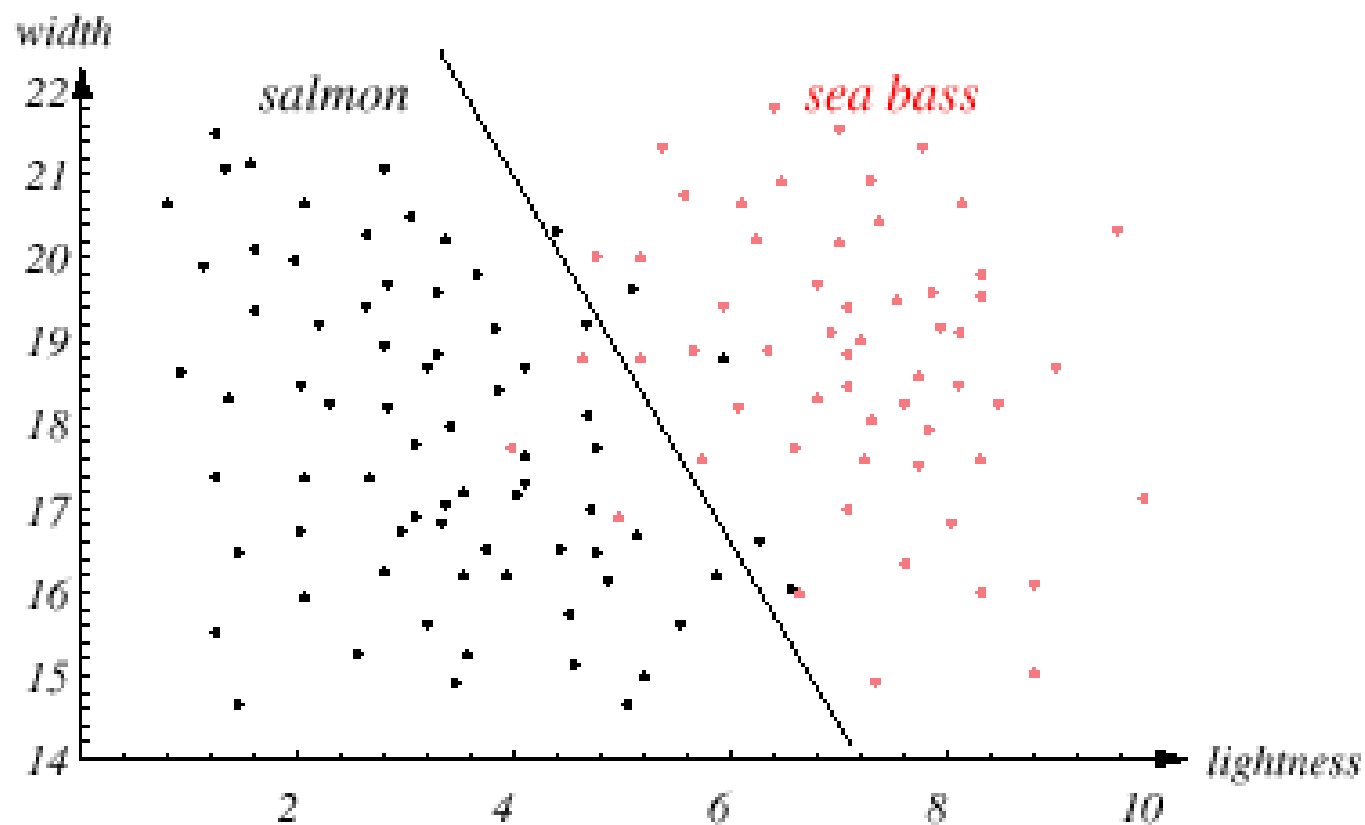
- Adoptar el Tono y agregar el ancho de las escamas del pez



Clasificador: Decisión Lineal,
 $x_2 < a x_1 + b,$
clasifica (x_1, x_2) en salmón

Clasificador lineal

33



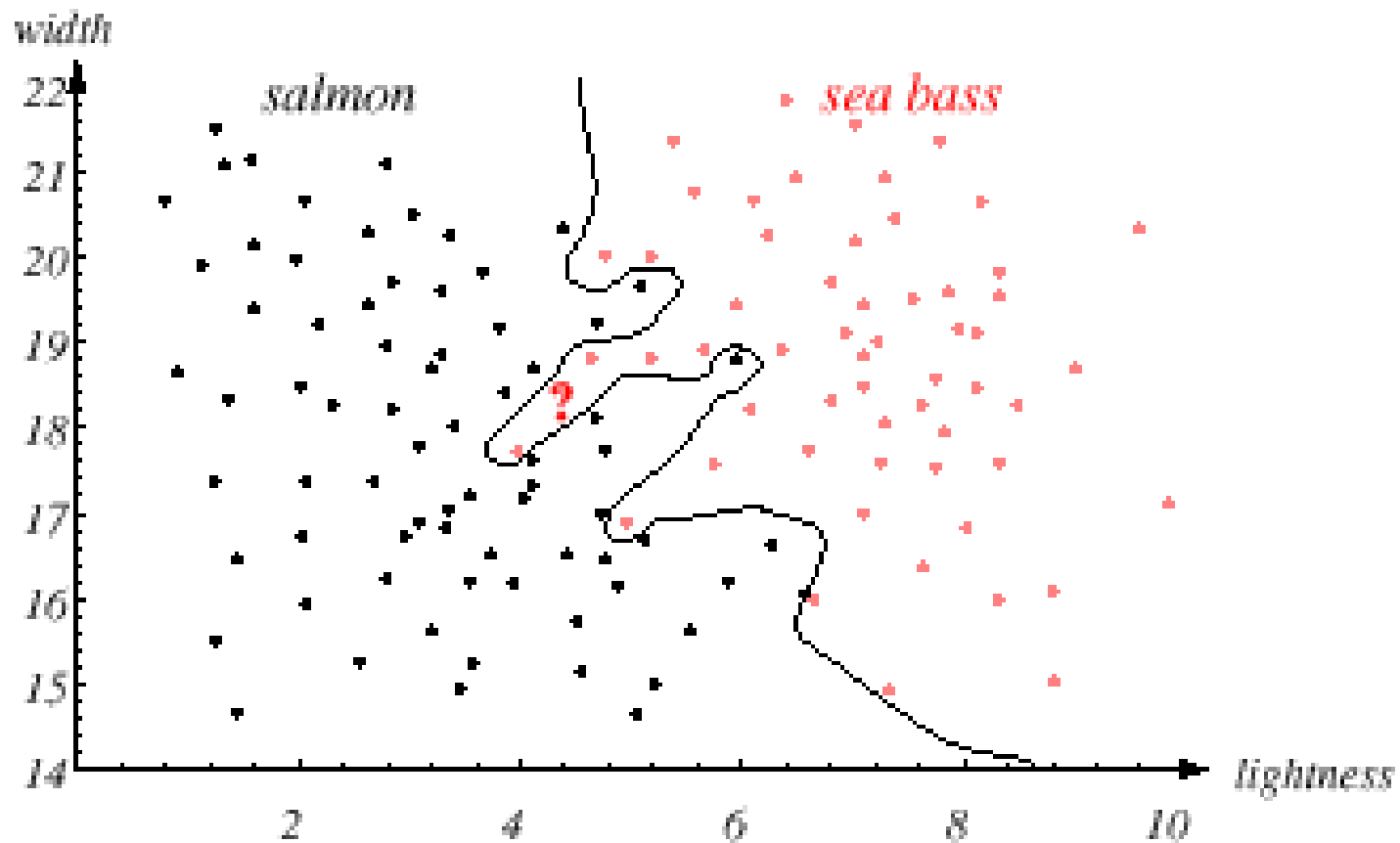
Otras opciones

34

- Podríamos agregar otras características que sean no correlacionadas con las que ya tenemos.
- Hay que tomar precauciones para no reducir el desempeño agregando características “ruidosas”.
- Idealmente, el mejor borde de decisión debería ser el que provee optimo desempeño como en la figura siguiente :

Clasificador óptimo

35



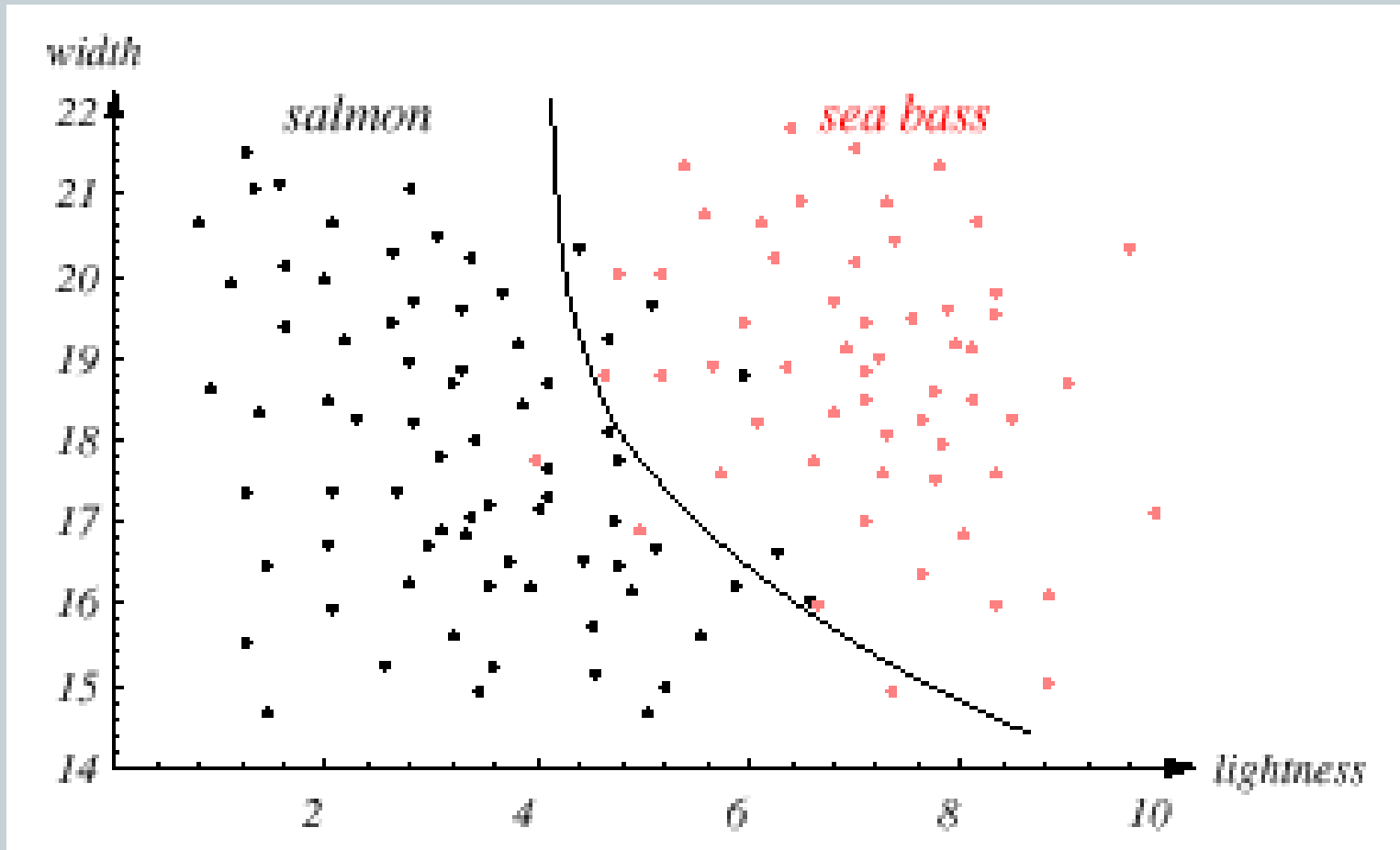
Generalización

36

Sin embargo, nuestra satisfacción es prematura porque el objetivo central es designar un clasificador que clasifique correctamente un nuevo elemento



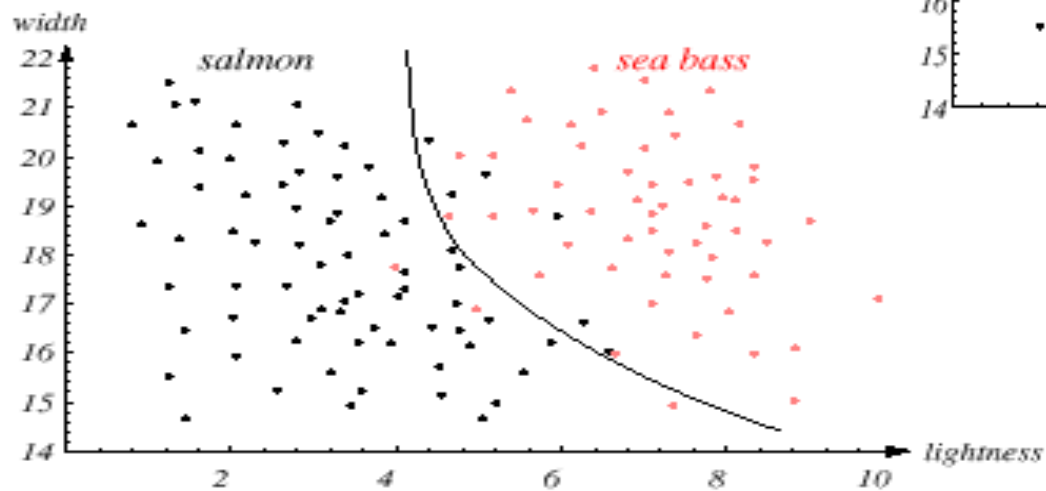
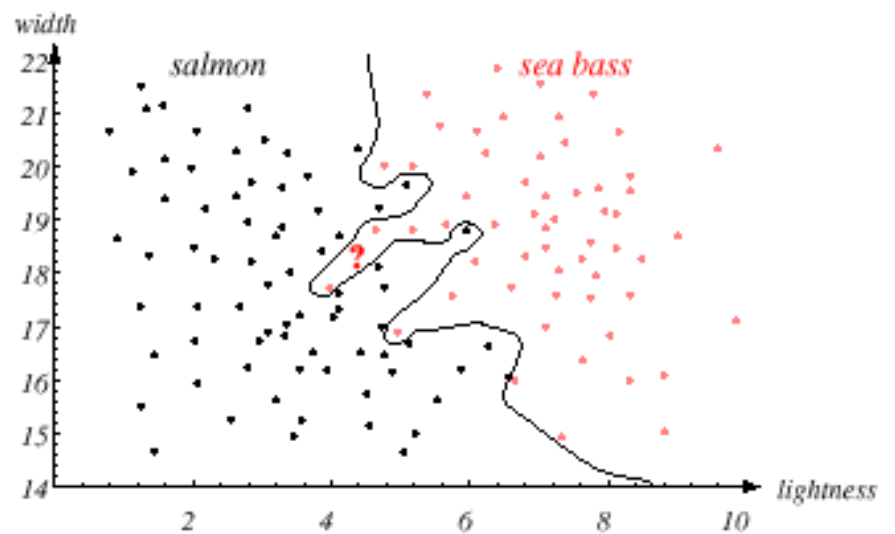
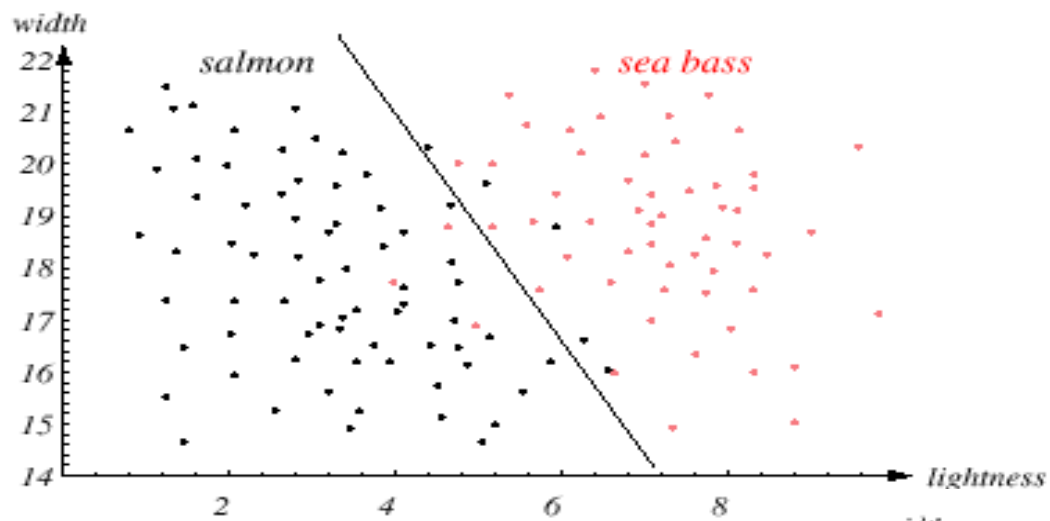
Se debe poder generalizar!



Objetivos del curso

38

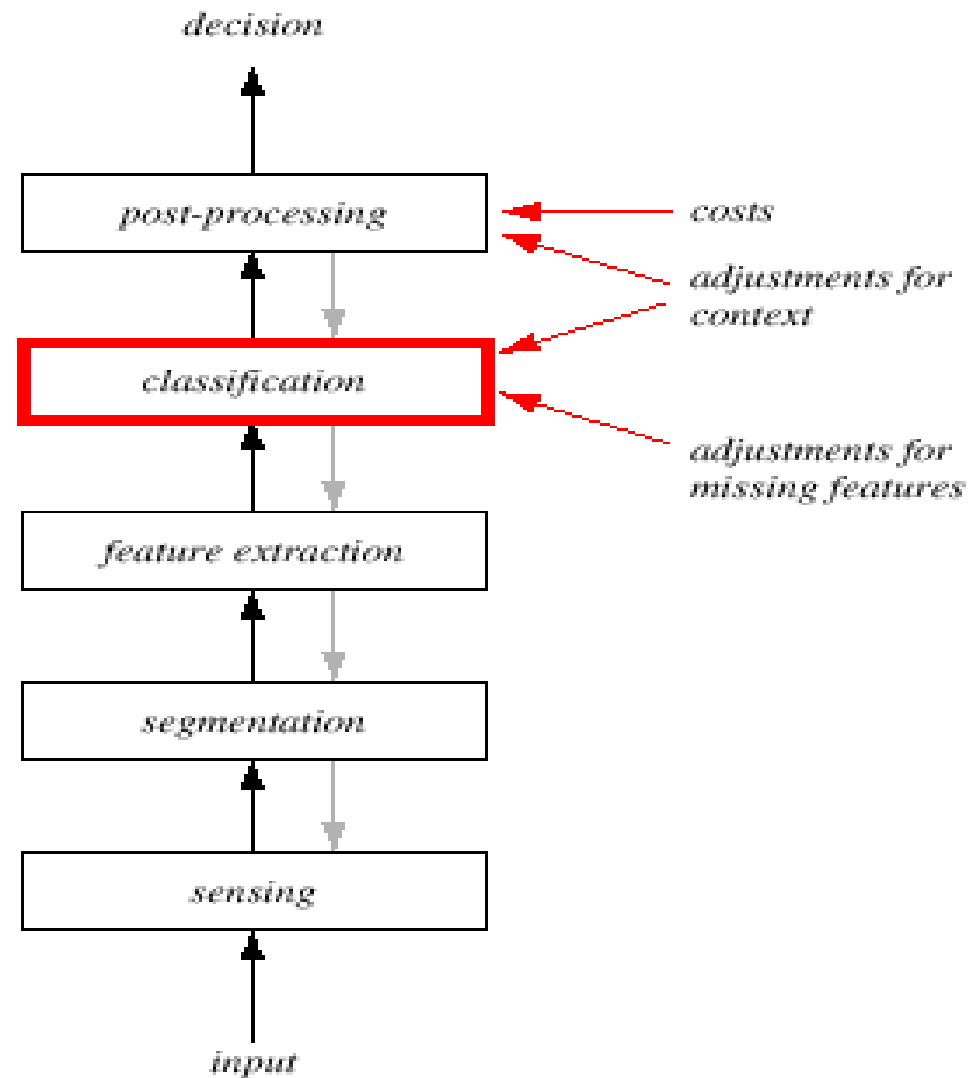
- Lograr realizar en un grupo de datos dado distintos tipos de clasificaciones
 - decidiendo las características a utilizar
 - planteando las ecuaciones
 - implementando en Python dichas ecuaciones
 - generando mapas visuales de clasificación
 - calculando tasas de error aparente y teóricas



Consideraciones adicionales de un Sistema de reconocimiento de patrones

40

- **Captación**
 - Medio (cámara o micrófono) para captar los datos
 - Características del medio :
 - ✦ ancho de banda,
 - ✦ resolución,
 - ✦ sensibilidad a la distorsión
- **Segmentación y agrupamiento de muestras**
 - Patrones deberían estar separados y no solaparse
 - Selección muestras de entrenamiento



Módulo de clasificación

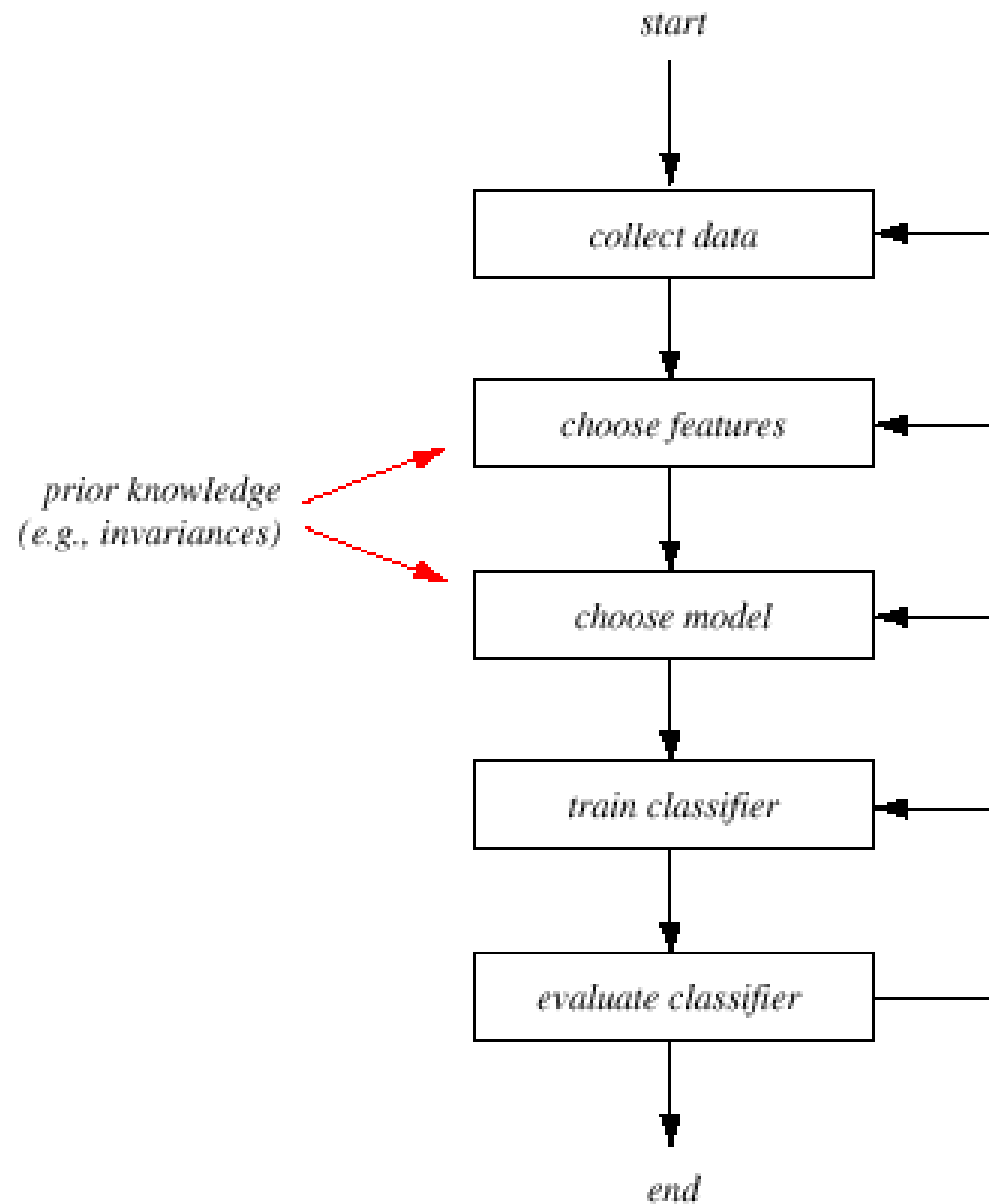
42

- Extracción de características
 - Características discriminativas
 - Características invariantes con respecto a la traslación, rotación y escala.
- Clasificación
 - Regla que usa el vector de características provisto para asignar el objeto a una categoría
- Post Procesamiento
 - Explotar la información de **contexto** del objeto para mejorar el desempeño

El Ciclo de Diseño

43

- Recolección de datos
- Elección de características
- Elección de Modelo
- Entrenamiento
- Evaluación
- Reajuste
 - complejidad computacional.
 - errores aceptables e inaceptables



Recolección de Datos

45

- Como sabemos cuando hemos recolectado una **colección adecuada** de ejemplos
 - en tamaño
 - representatividad,
 - para entrenar
 - y testear el sistema?

Elección de características

46

- **Dependen**
 - problema de clasificación,
 - del sistema de visión que genera los datos,
 - del costo de obtención de los datos.
- **Deben ser**
 - simples de extraer,
 - invariantes bajo transformaciones de la imagen
 - inmunes o resistentes al ruido de captación de la imagen.

Elección de Modelo Matemático

47

- Lineal
- Cuadrático
- Funcional
- Árbol
- Red neuronal
- Máquina de soporte vectorial
- etc

Entrenamiento

48

- Usamos datos para determinar los clasificadores.
- Muchos procedimientos diferentes para entrenar los clasificadores y elegir los modelos

Evaluación

49

- Medidas de desempeño bajo un modelo
 - Tasas de error
 - Costos
- Medidas de desempeño al cambiar de un grupo de características a otro

Complejidad computacional

50

- Cual es el balance entre computación fácil y desempeño?
- Como escala un algoritmo en función del numero de características, patrones y categorías?

Adaptación y Aprendizaje

51

- Aprendizaje supervisado
 - Un oráculo provee las etiquetas de categoría y costo de cada patrón en el conjunto de entrenamiento
- Aprendizaje no supervisado
 - El sistema forma grupos naturales de los patrones ingresados .

Clasificación y Segmentación

52

- Etiquetar toda una imagen como una clase
 - ✦ Imagen de interiores o de exteriores;
 - ✦ Imagen de Salmon o Imagen de Mero

Clasificación

- Dividir una imagen en regiones de diferente tipo.
 - ✦ Diferentes tipos de tejido en una imagen medica,
 - ✦ Diferentes ejemplares de peces en la cinta de transporte.

Segmentación

Clasificación en Imágenes

53

- Nombres distintos porque son problemas distintos
- Técnicas parecidas, podemos usar las mismas técnicas en ambos problemas.

Clasificación

54



Clase 1



Clase 2



Clase 3

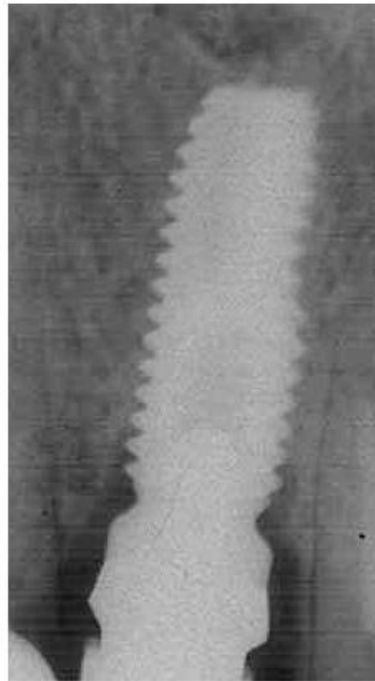


?

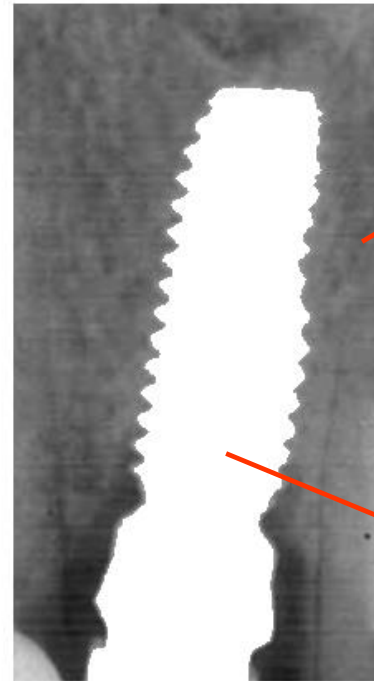
Segmentación

55

Original



Segmentada

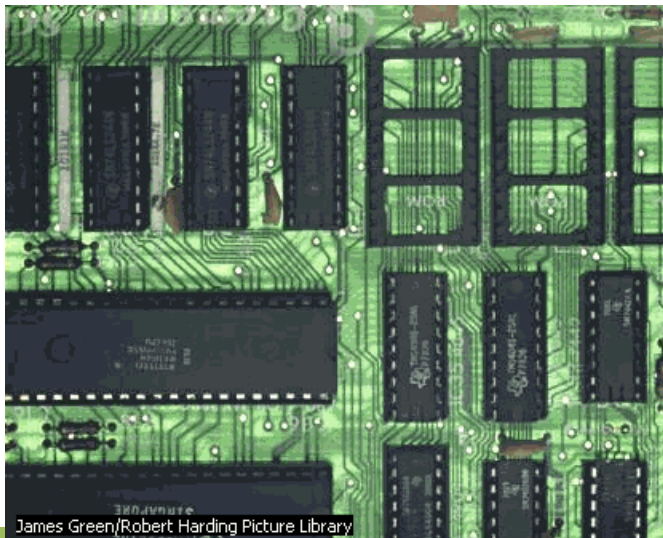
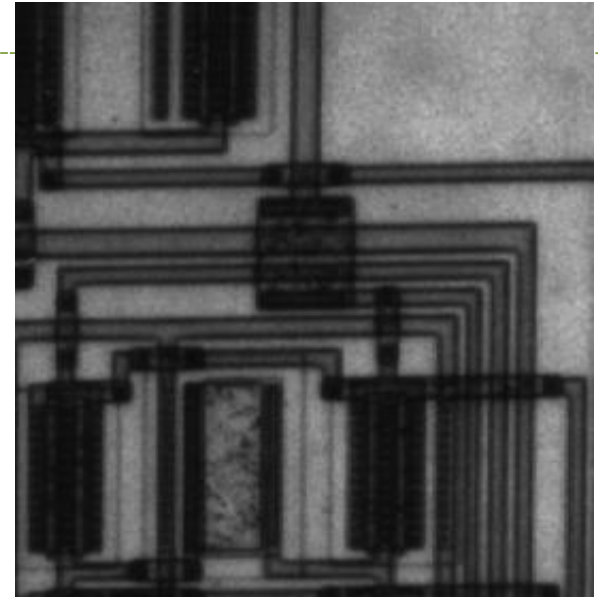
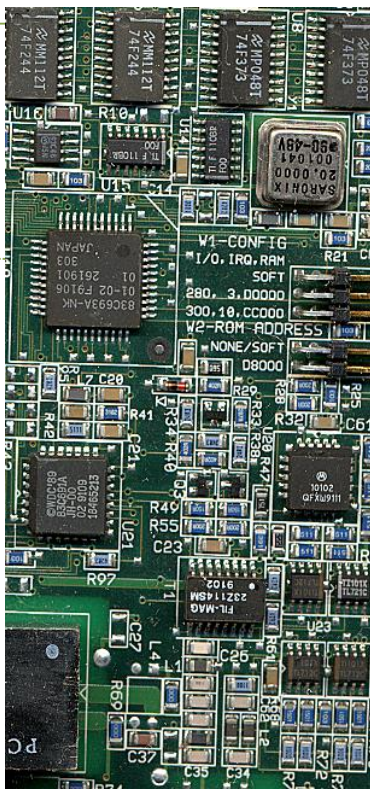


fondo

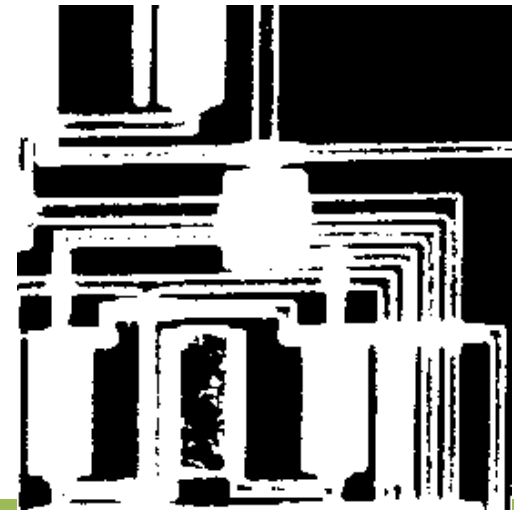
perno

Ejemplos

56



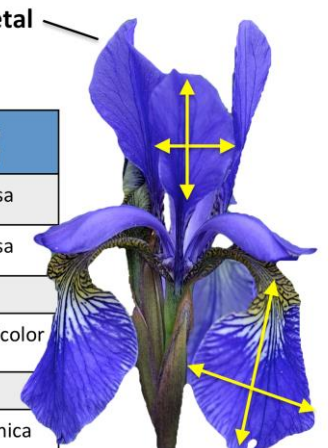
James Green/Robert Harding Picture Library



Ejercicios

57

- Estudiar el problema de clasificación de flores Iris
 - Características al recolectar los datos
 - Número de clases hipotéticas



The diagram shows a blue Iris flower. Yellow arrows indicate measurements: two vertical arrows on the upper petals labeled 'Petal', and two diagonal arrows on the lower petals labeled 'Sepal'.

Samples (instances, observations)					
	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
...					
50	6.4	3.5	4.5	1.2	Versicolor
...					
150	5.9	3.0	5.0	1.8	Virginica

Features
(attributes, measurements, dimensions)

Class labels
(targets)

- Estudiar cargado de datos y muestra de datos en pantalla según características elegidas.