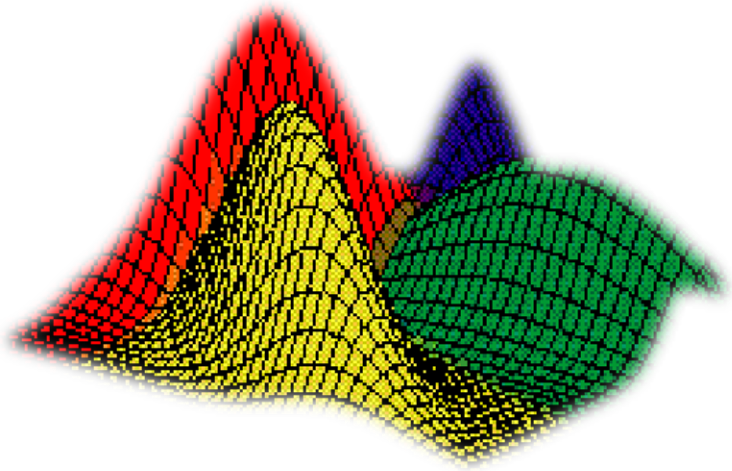
CSE 473
Pattern Recognition



Lecturer:

Dr. Md. Monirul Islam

Course Outline

- Introduction to Pattern Recognition
- Bayesian Classification and its variants
- Linear Classifiers: Perceptron Algorithms and its Variants, Linear SVM
- Non-Linear Classifiers: Multilayer Perceptrons, Non-Linear Support Vector Machines
- Context Dependent Classification
- Template Matching
- Syntactic Pattern Recognition: Grammar and Graph based Pattern Recognition
- Unsupervised Classification: Clustering Algorithms

Course Outcome

- have in-depth knowledge and understanding of classical and state-of-the-art pattern recognition algorithms
- identify and compare pros and cons of different pattern recognition techniques
- analyze real world pattern recognition problems and apply appropriate algorithm(s) to formulate solutions
- design and implement core pattern recognition techniques and
- develop/engineer new techniques for solutions of real world problems

Assessment

- Class Tests: 20%
- Attendance:10 %
- Term final: 70%

Text Books

- Pattern Recognition
  - S. Theodoridis & K. Koutrumbas
- Pattern Classification
  - R. Duda *et al*.
- Pattern Recognition *Statistical, Structural and Neural Approaches*
  - R. Shalkoff
- Introduction to Data Mining
  - Tan, Steinbach, Kumar

Schedule for Class Tests

* As per central routine

Pattern Recognition: What is it?

Perhaps one of the
oldest intelligent arts
of living beings

Pattern Recognition: What is it?

Perhaps one of the
oldest intelligent arts
of living beings

Pattern Recognition: What is it?

Perhaps one of the
oldest intelligent arts
of living beings

What Does It Do?

- Build a machine that can recognize patterns.

- The task: Assign unknown objects – patterns – into the correct class. This is known as classification.

What Does It Do?

- Areas:
  - Machine vision
  - Character recognition (OCR)
  - Computer aided diagnosis
  - Speech recognition
  - Face recognition
  - Bioinformatics
    - Image Data Base retrieval
    - Data mining
    - Biometrics
      - Fingerprint identification
      - Iris Recognition
    - DNA sequence identification

Representation of patterns

- Features:
  - measurable quantities from the patterns
  - determines the classification task
- Feature vectors: A number of features

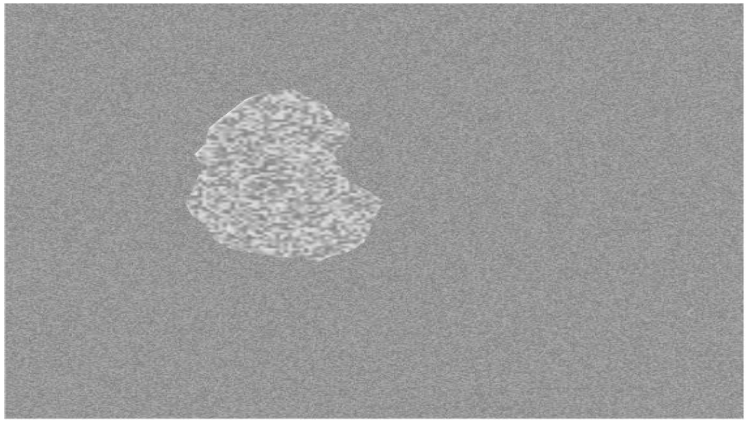
$$x_1, \ldots, x_l,$$

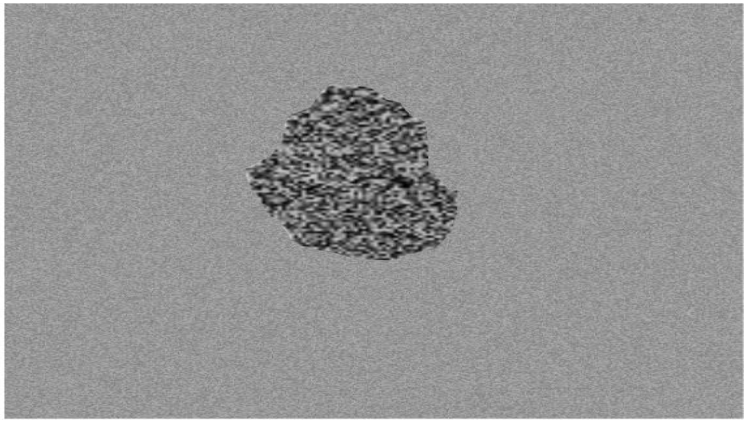
constitute the feature vector

$$\underline{x} = [x_1, \ldots, x_l]^T \in R^l$$

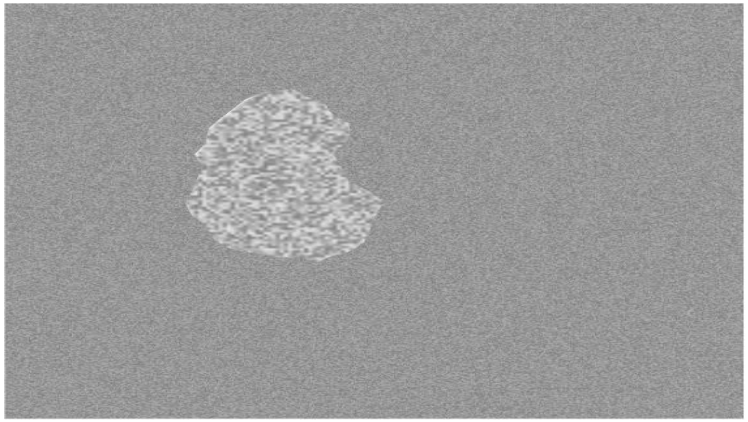
Feature vectors are treated as random vectors.

Example 1:

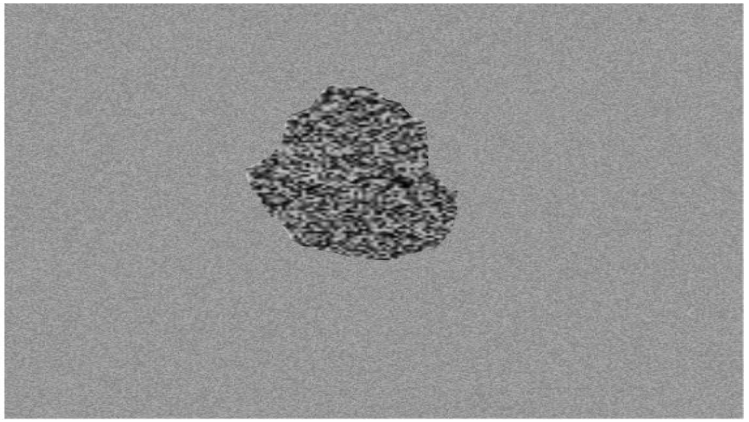




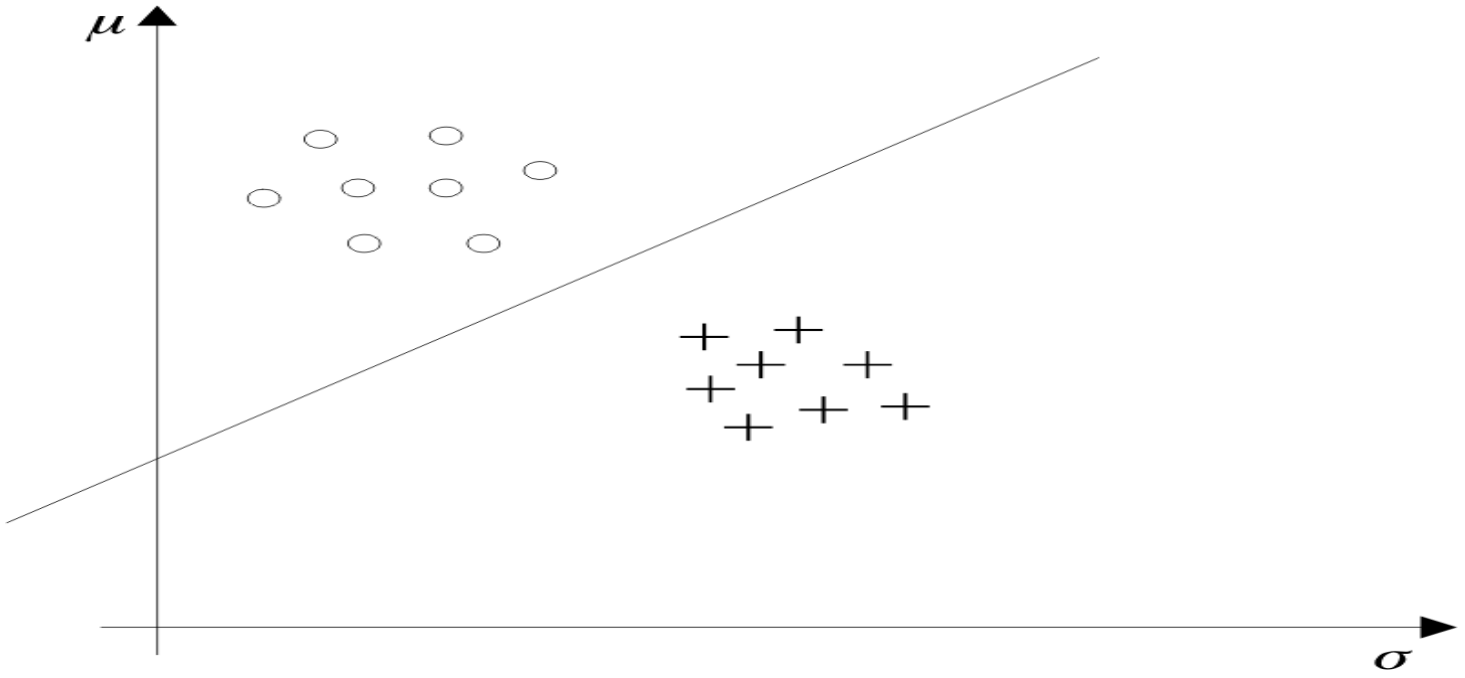
Example 1:



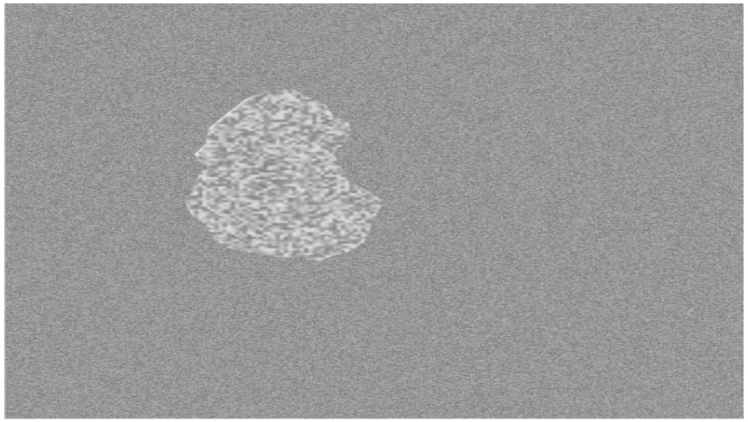
(a)



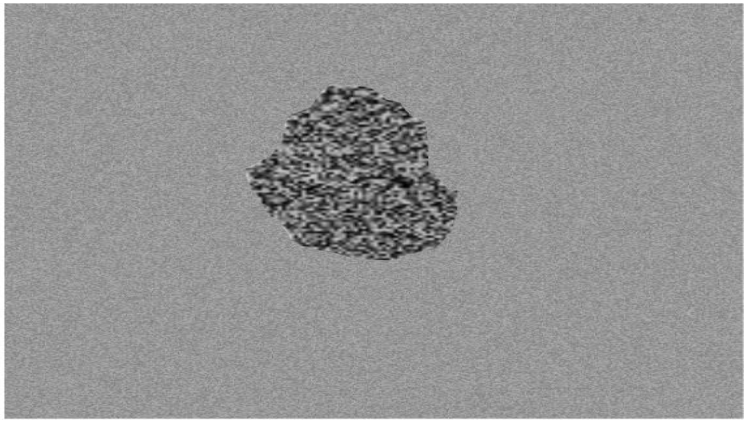
(b)



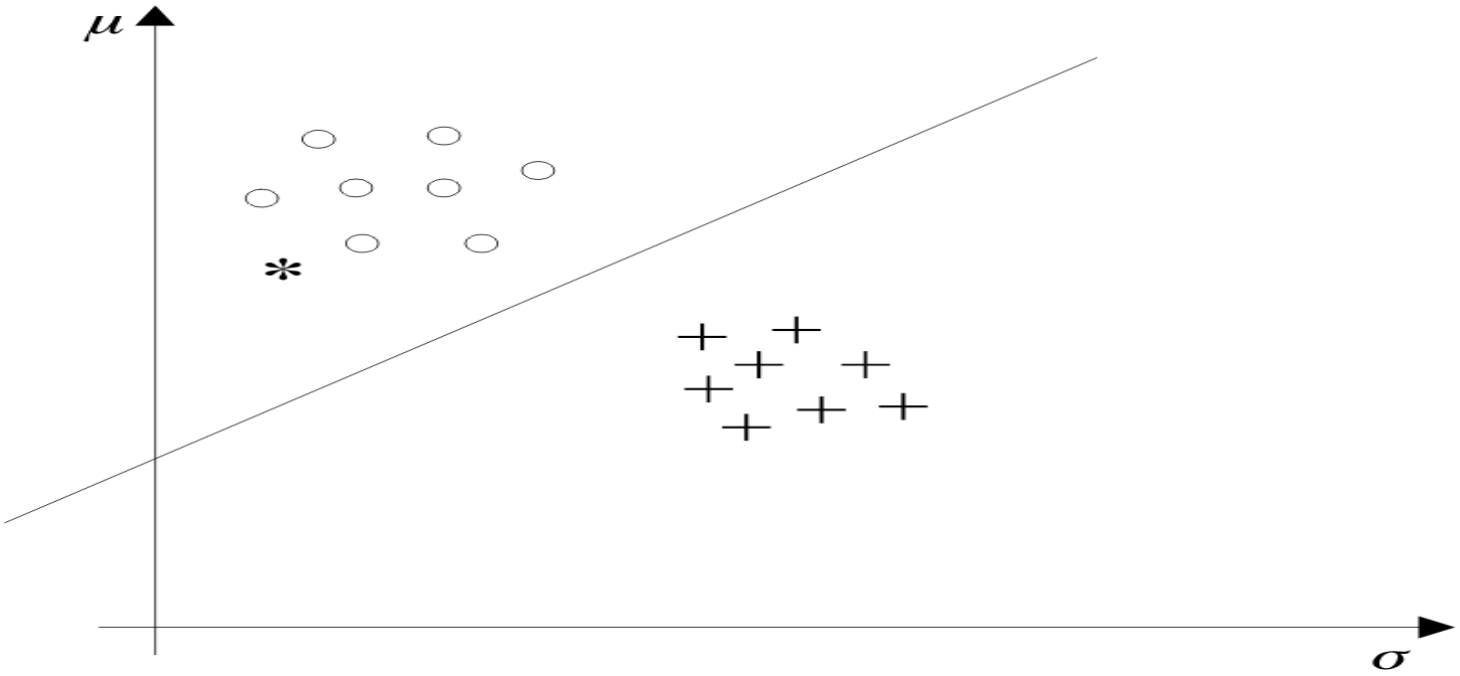
Example 1:



(a)



(b)

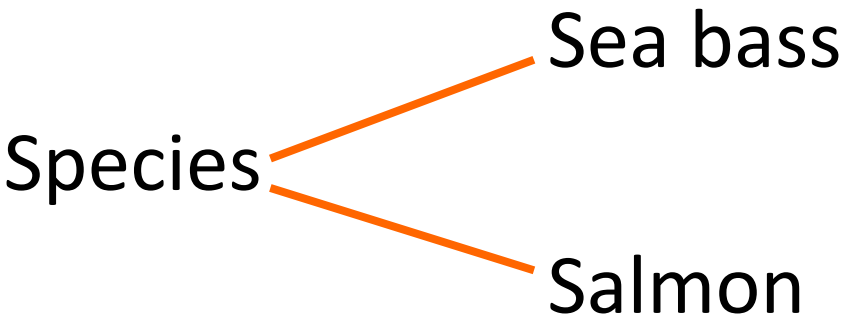


Issues in Pattern Recognition

- How are features generated?
- What is the best number of features?
- How are they used to design a classifier?
- How good is the classifier?

Example 2

- “Sorting incoming Fish on a conveyor according to species using optical sensing”

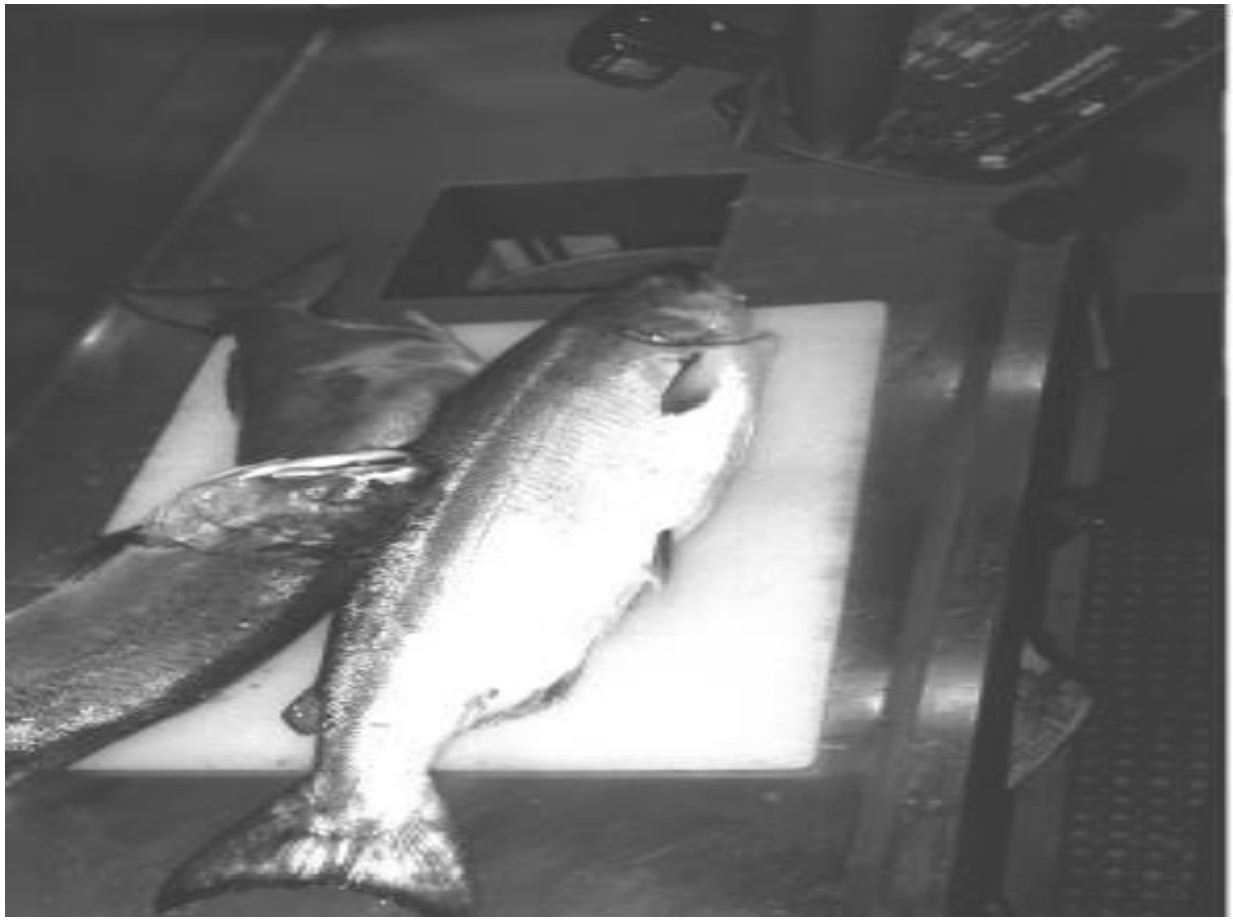


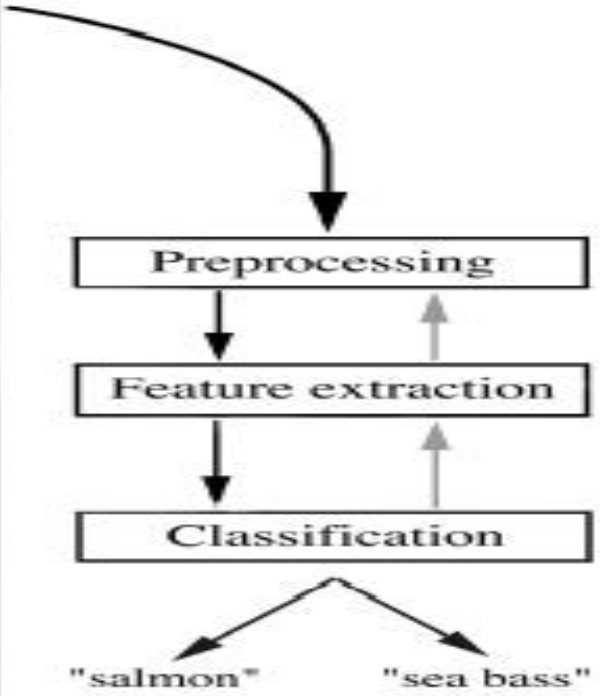
- Problem Analysis

  – Set up a camera and take some sample images to extract features

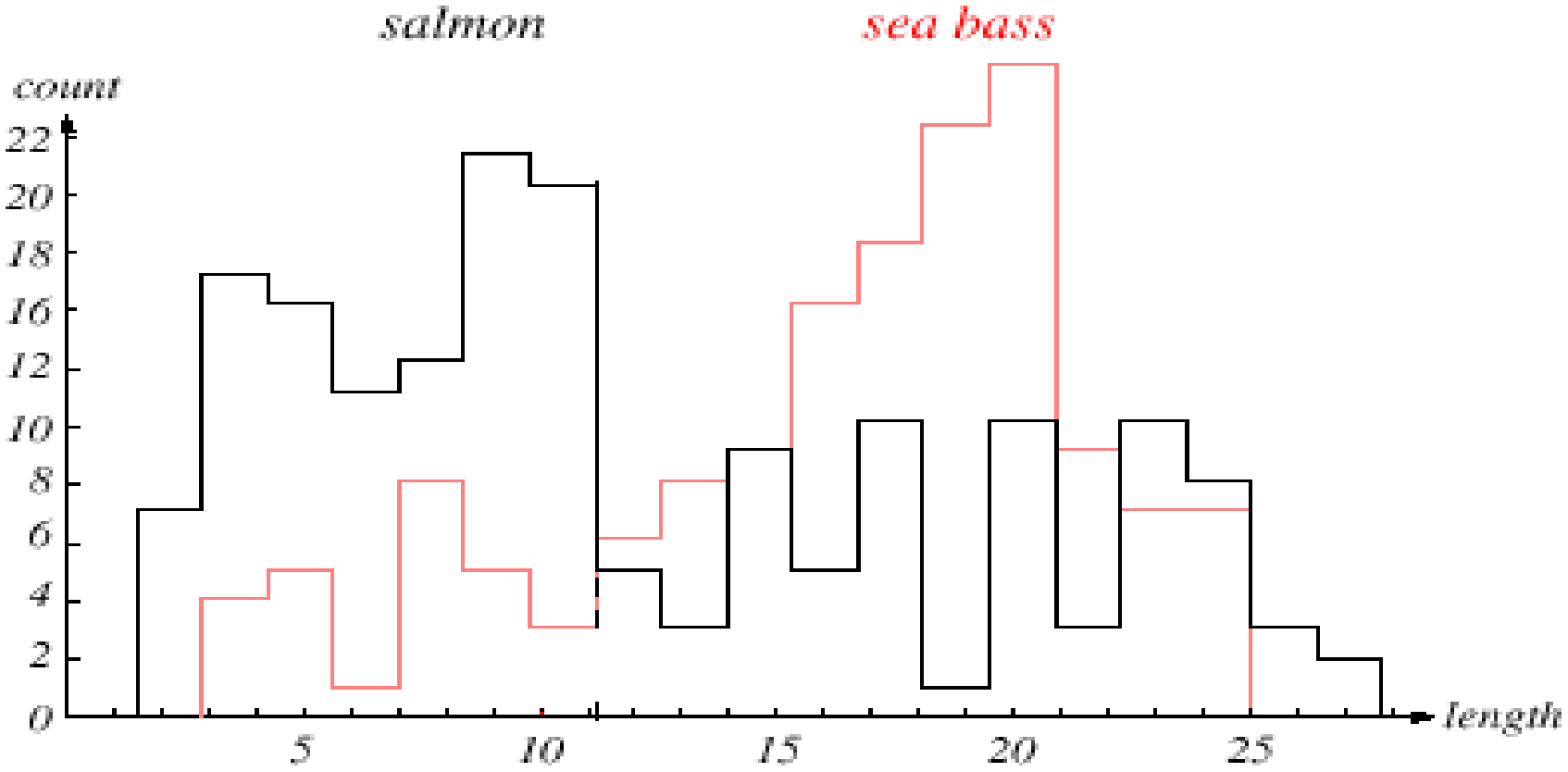
    - Length
    - Lightness
    - Width
    - Number and shape of fins
    - Position of the mouth, etc…

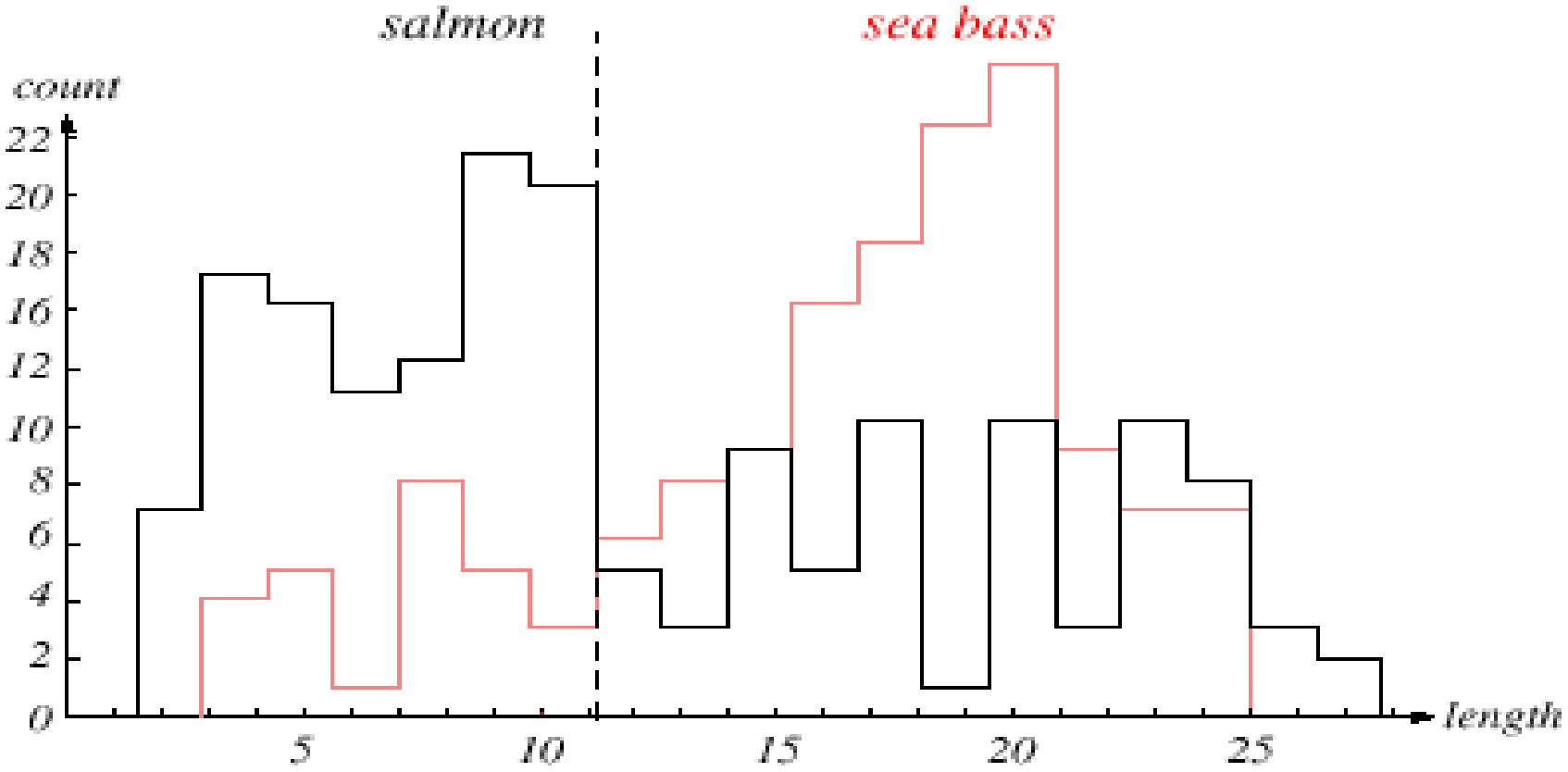
- Preprocessing
  - isolate fishes from one another and from the background
- Feature Extraction
  - send isolated fish image to feature extractor
  - it reduces the data, too
- Classification
  - pass the features to a classifier

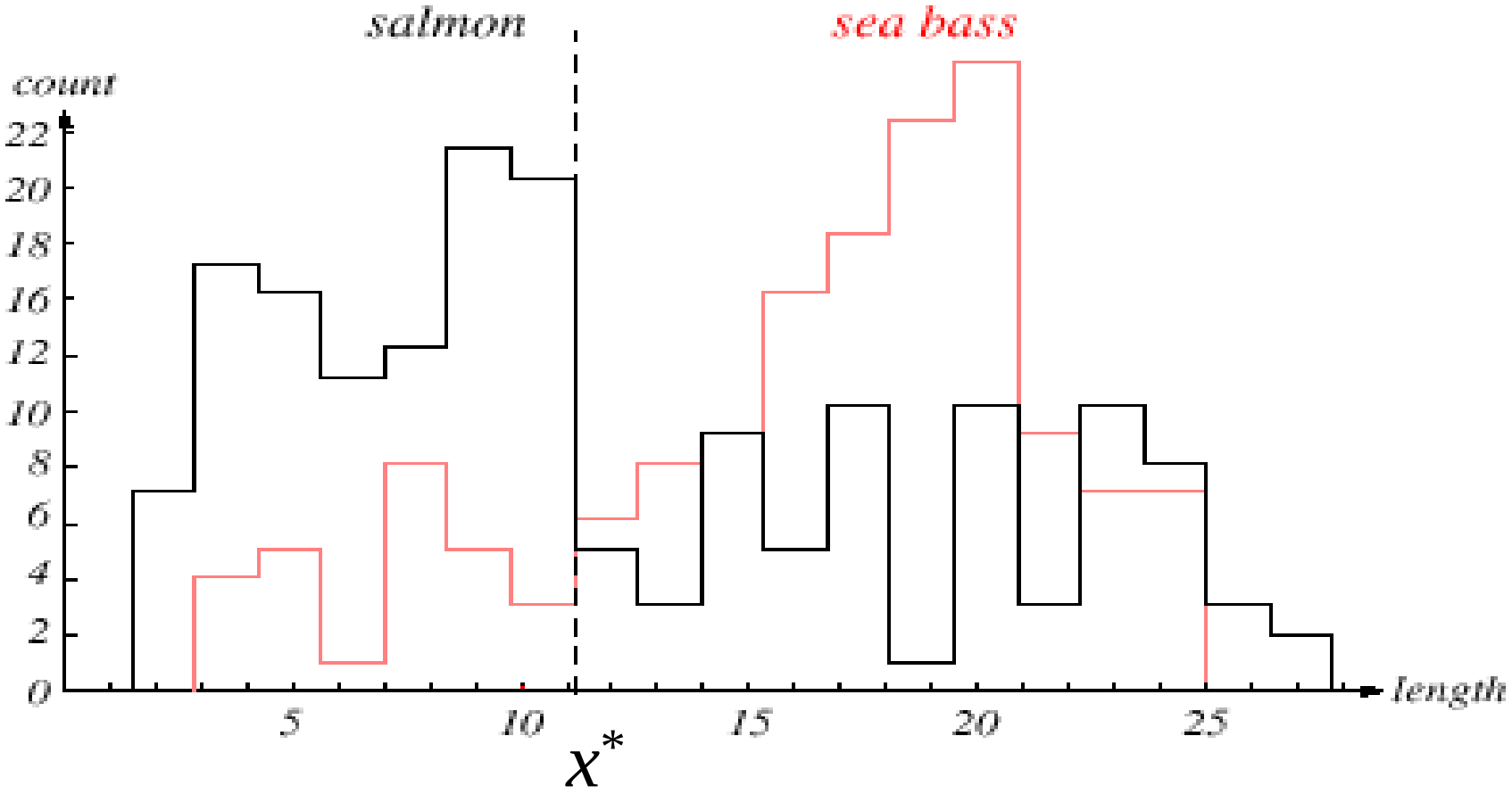



Preprocessing
Feature extraction
Classification
"salmon"
"sea bass"

- Classification
  - Select the length of the fish as a possible feature for discrimination

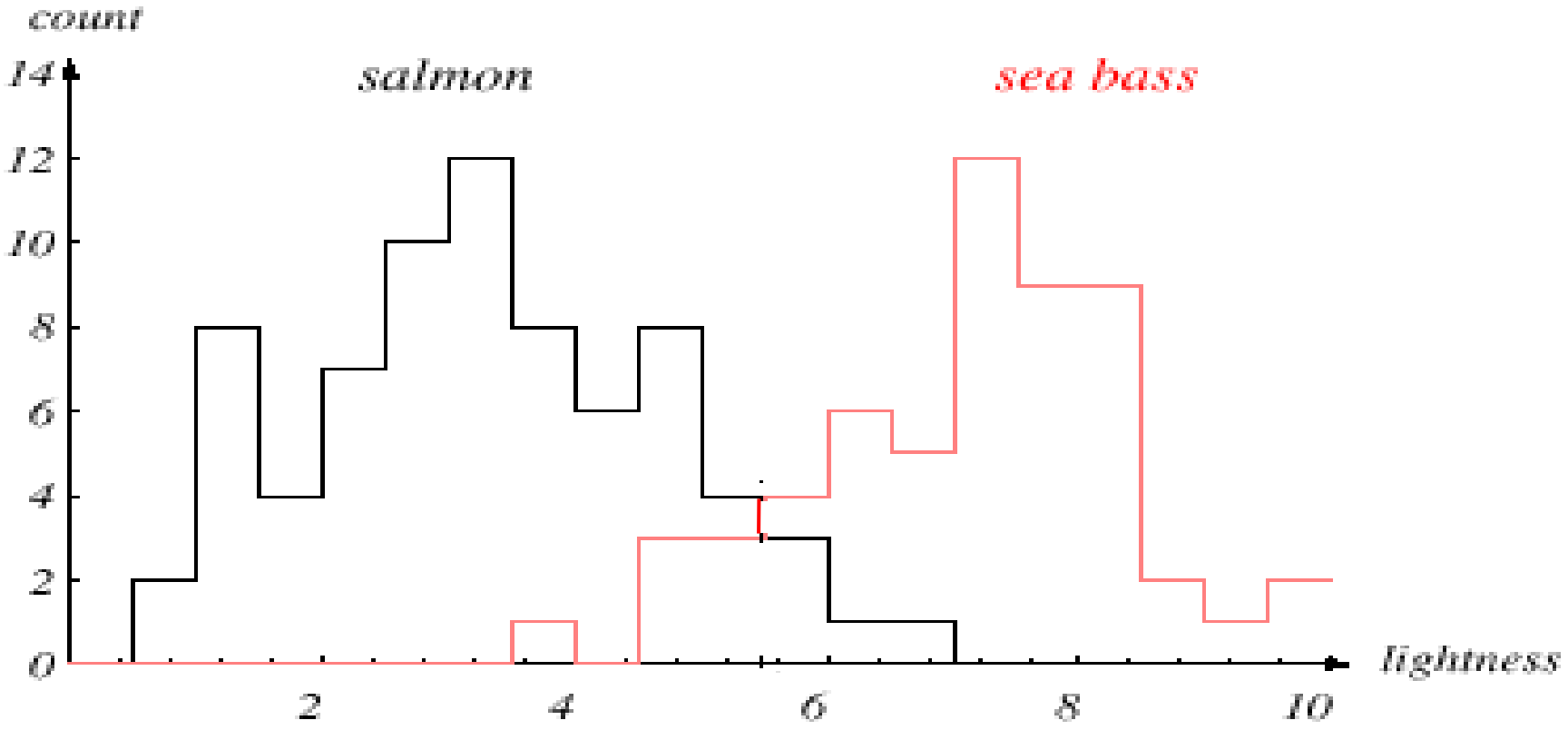

salmon
sea bass
count
22
20
18
16
12
10
8
6
4
2
0
5
10
15
20
25
length

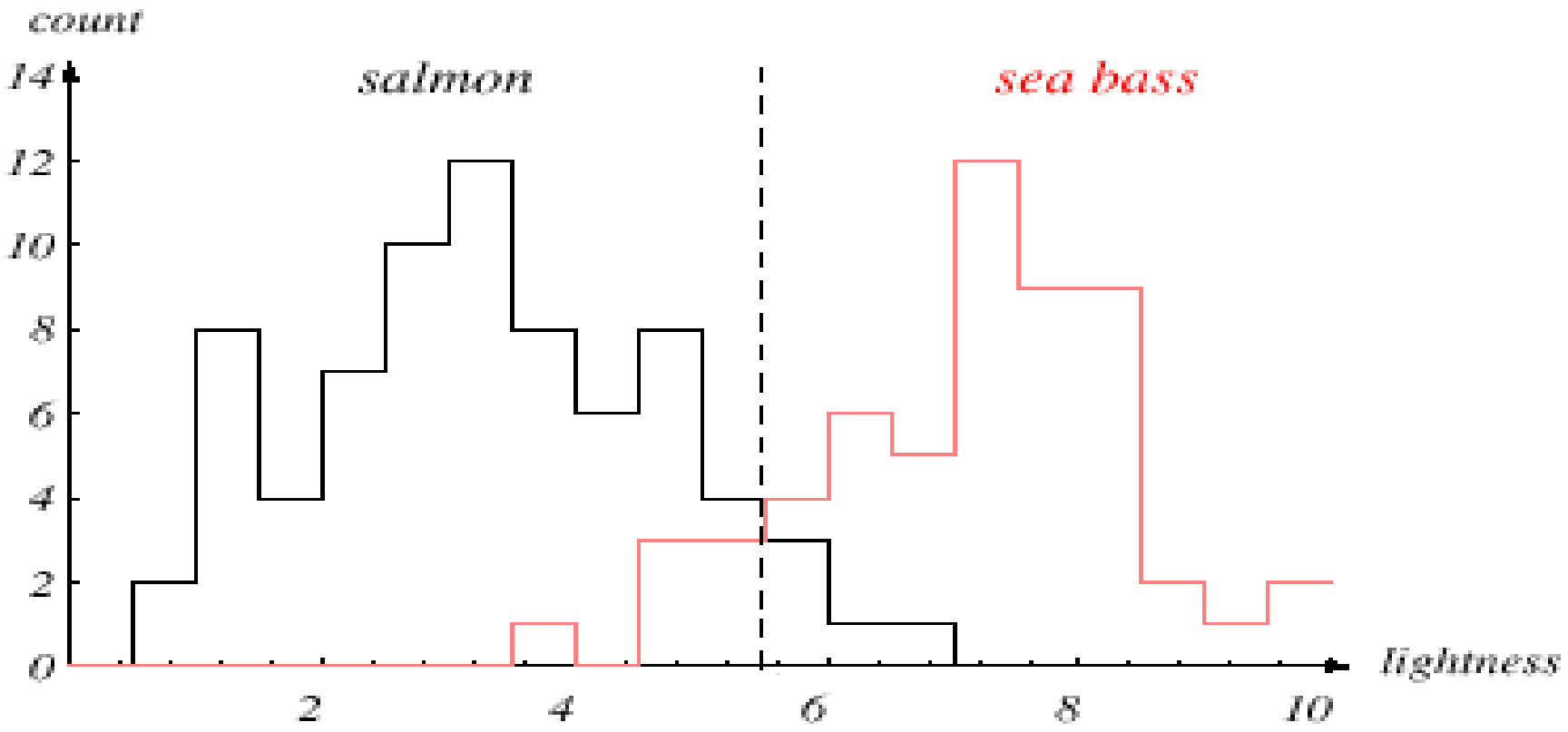

salmon
sea bass
count
22
20
18
16
12
10
8
6
4
2
0
5
10
15
20
25
length

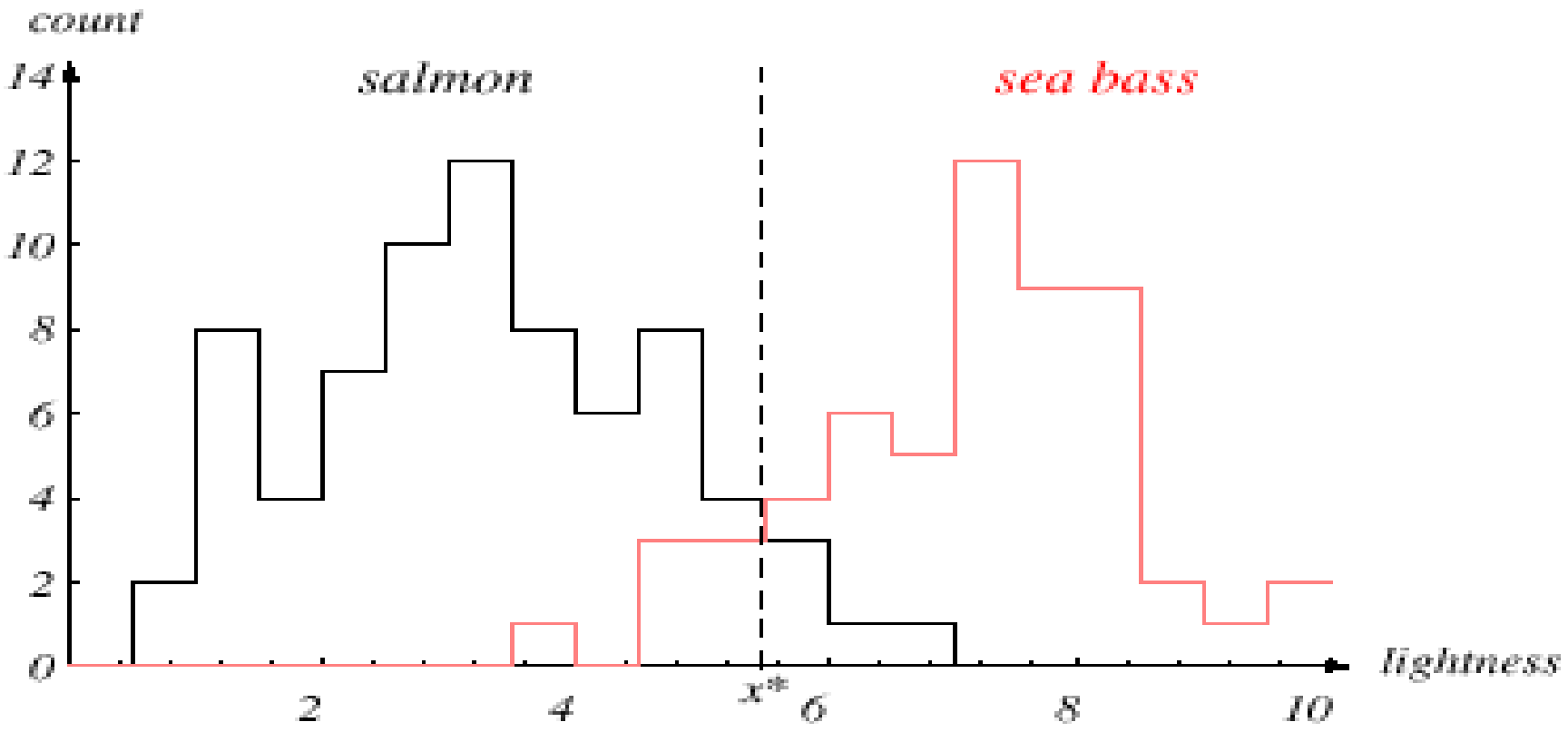

salmon
sea bass
count
22
20
18
16
12
10
8
6
4
2
0
5
10
15
20
25
length
x*

The length is a poor feature alone!

Select the lightness as a possible feature.

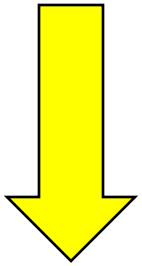

count
salmon
sea bass
14
12
10
8
6
4
2
0
2
4
6
8
10
lightness


count
salmon
sea bass
14
12
10
8
6
4
2
0
2
4
6
8
10
lightness


count
salmon
sea bass
14
12
10
8
6
4
2
0
2
4
x*
6
8
10
lightness

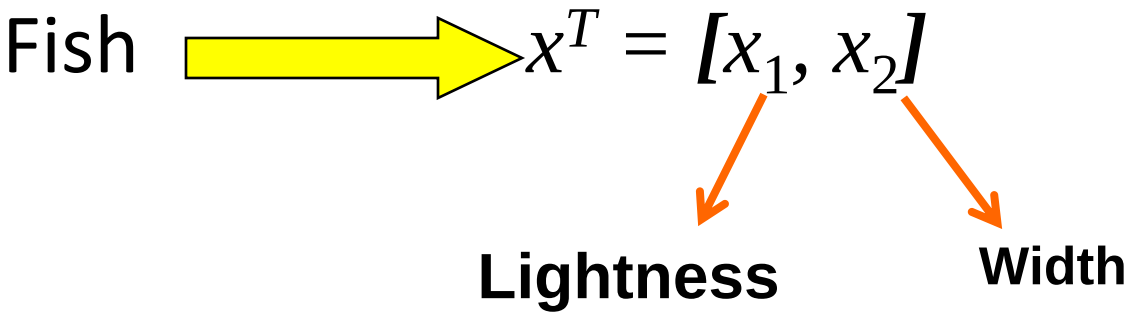
- Decision boundary and cost relationship

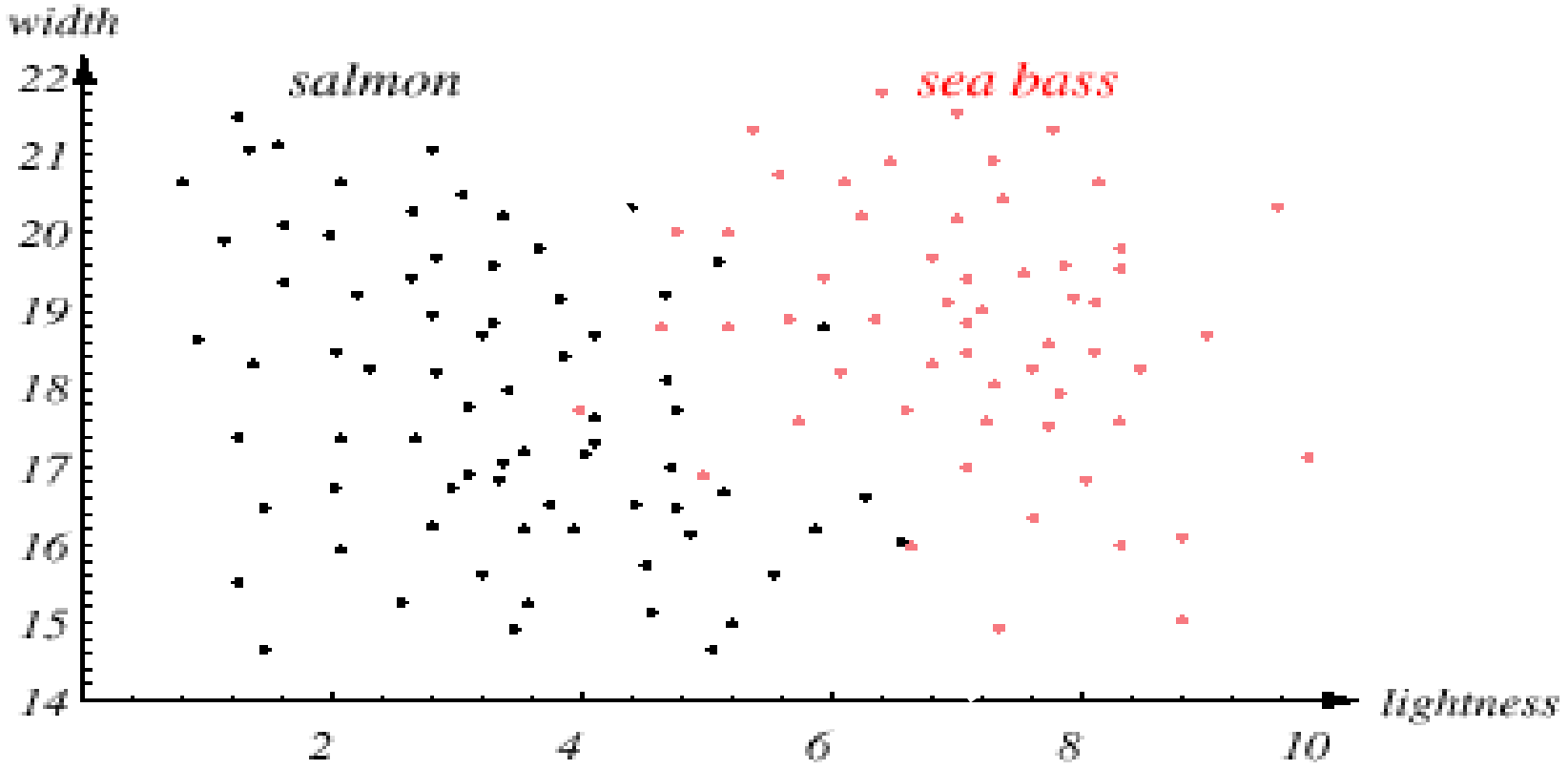
  – Move decision boundary toward smaller values of lightness in order to minimize the cost (reduce the number of sea bass that are classified as salmon!)

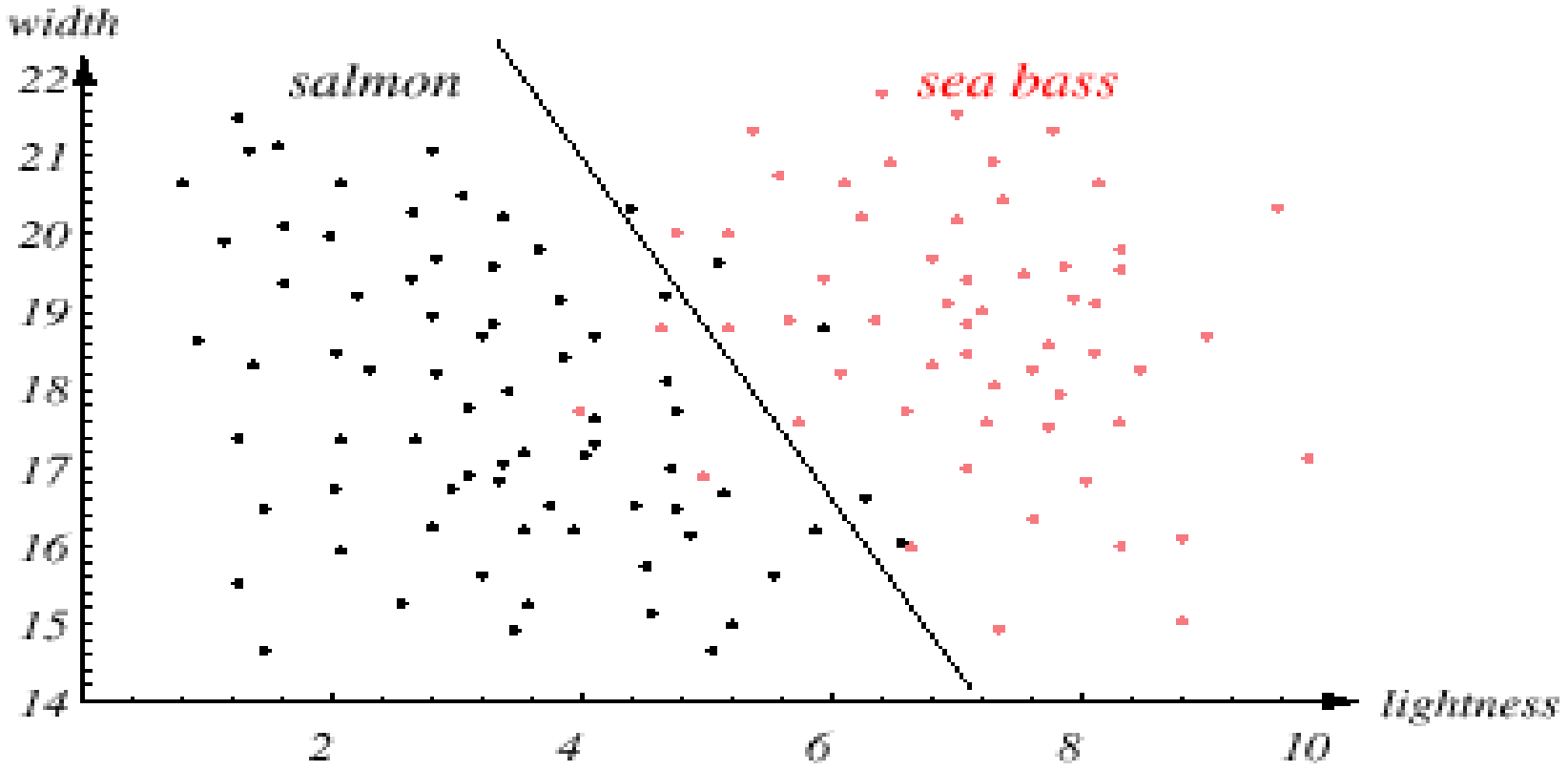


Task of decision theory

- Adopt the lightness and add the width of the fish

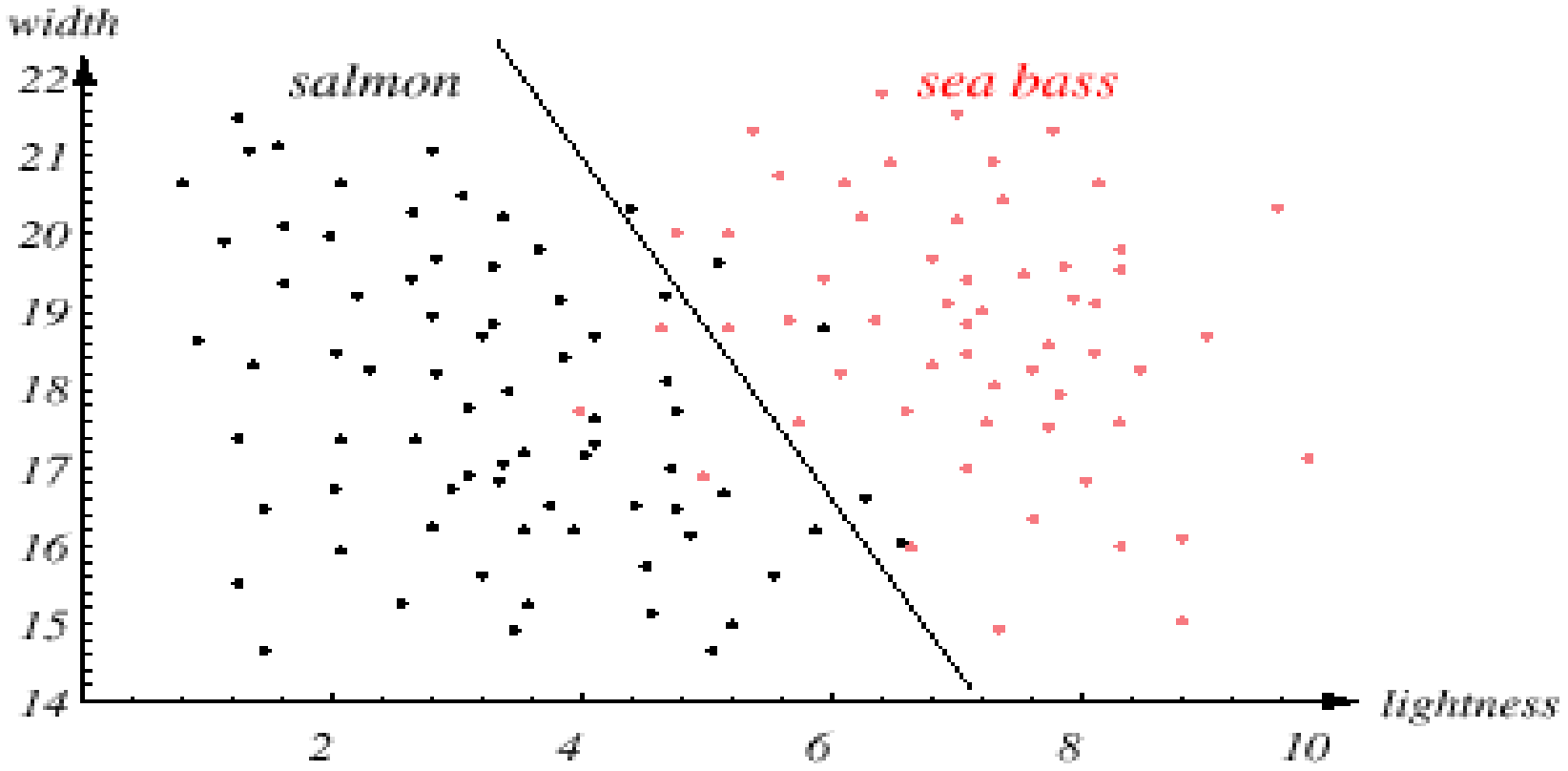



width
salmon
sea bass
22
21
20
19
18
17
16
15
14
2
4
6
8
10
lightness

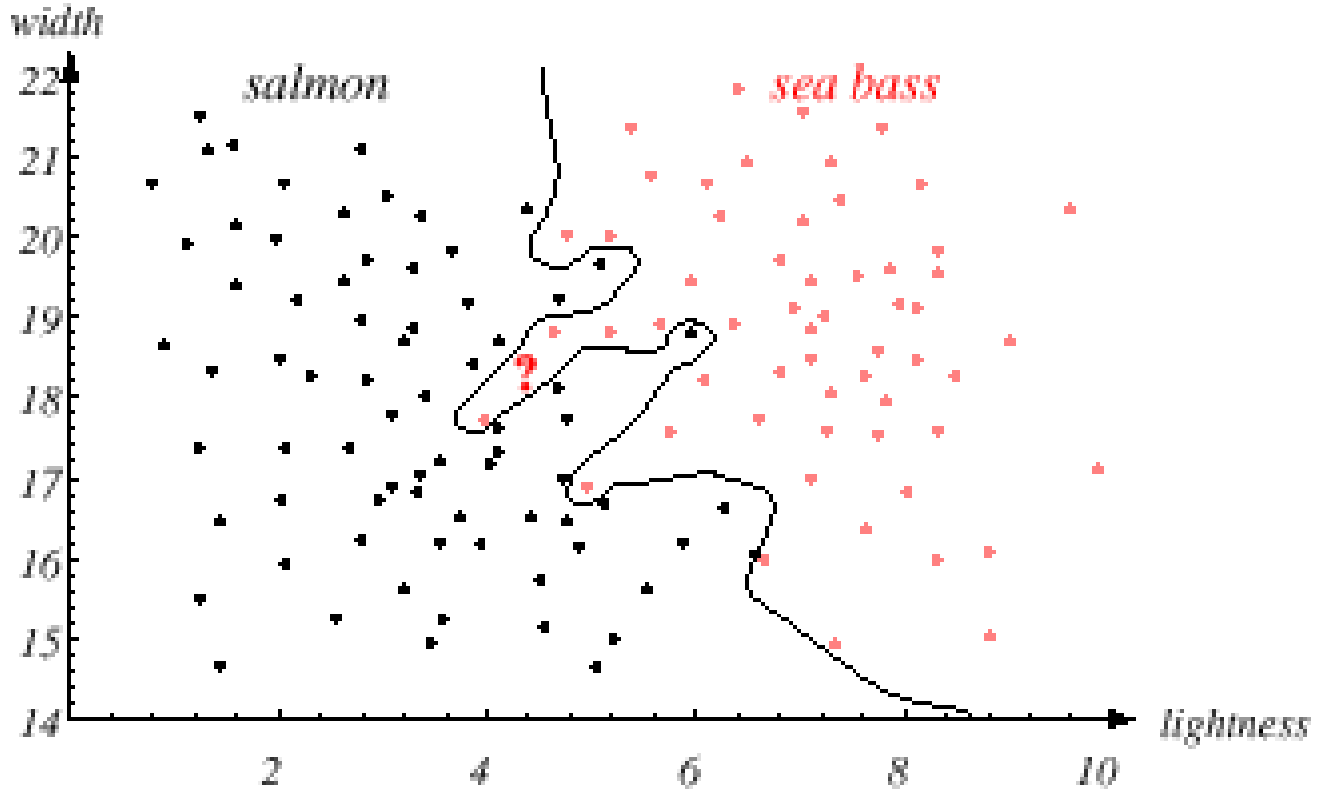

width
22
21
20
19
18
17
16
15
14
salmon
sea bass
2
4
6
8
10
lightness

- adding correlated feature does not improve anything and is thus redundant

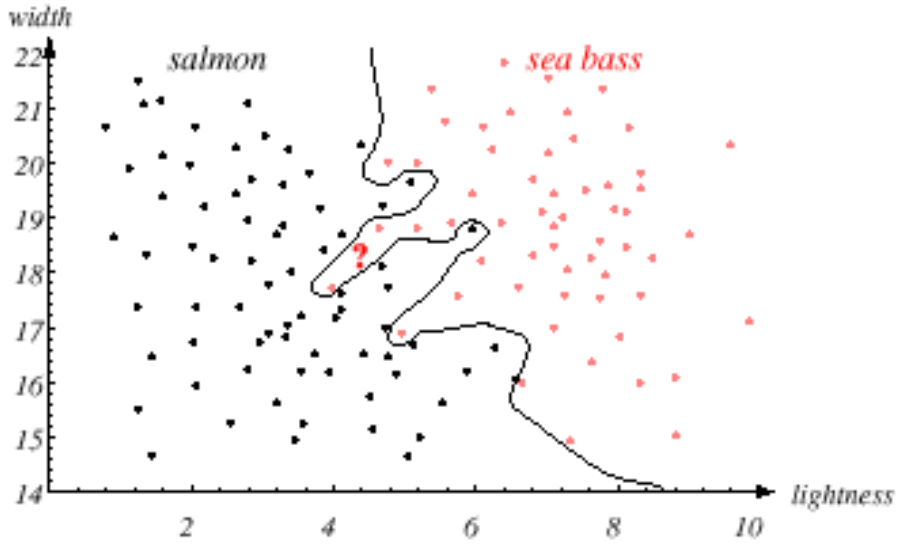
- too many features may lead to *curse of dimensionality*



still there are some misclassifications

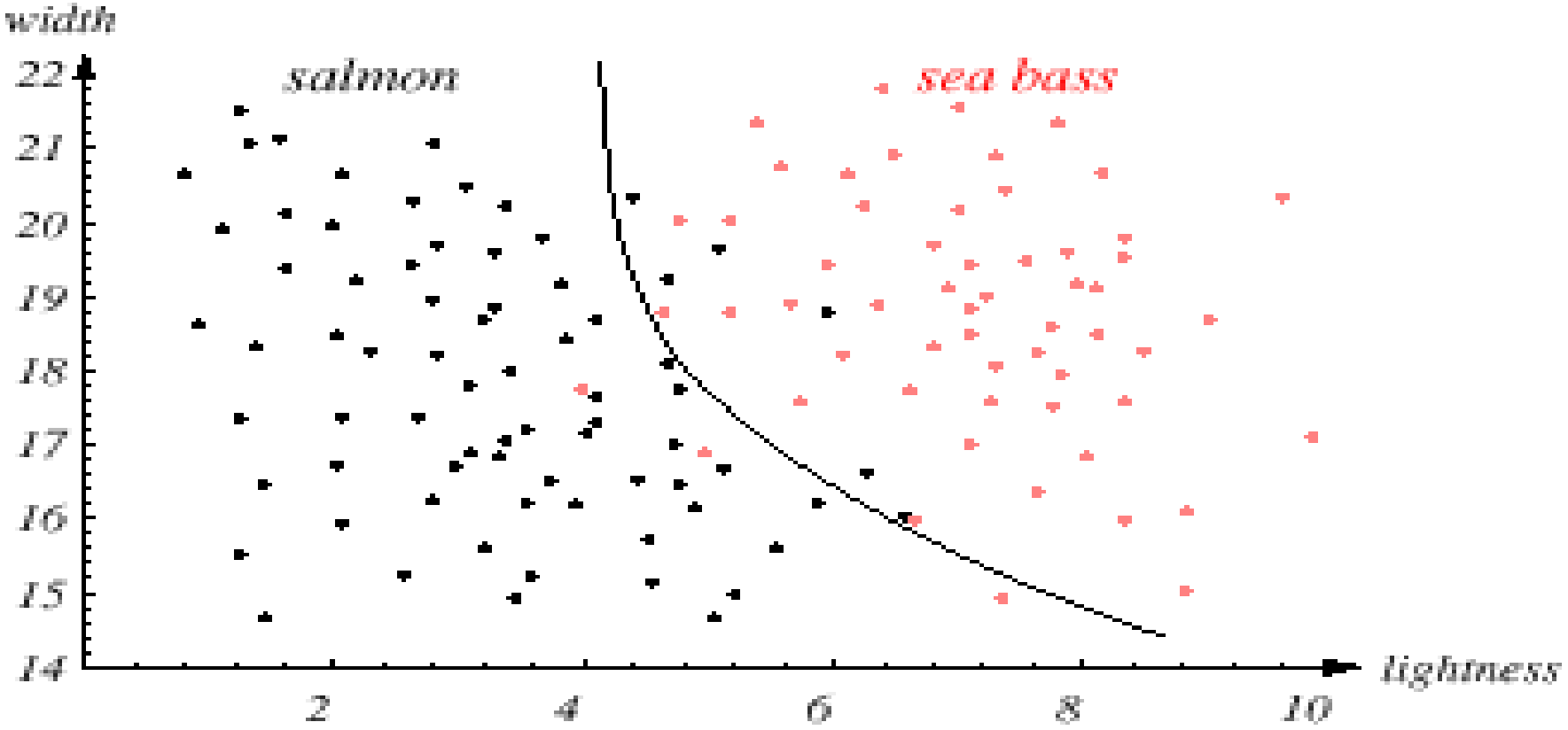


perhaps the best one, but too complex decision boundary



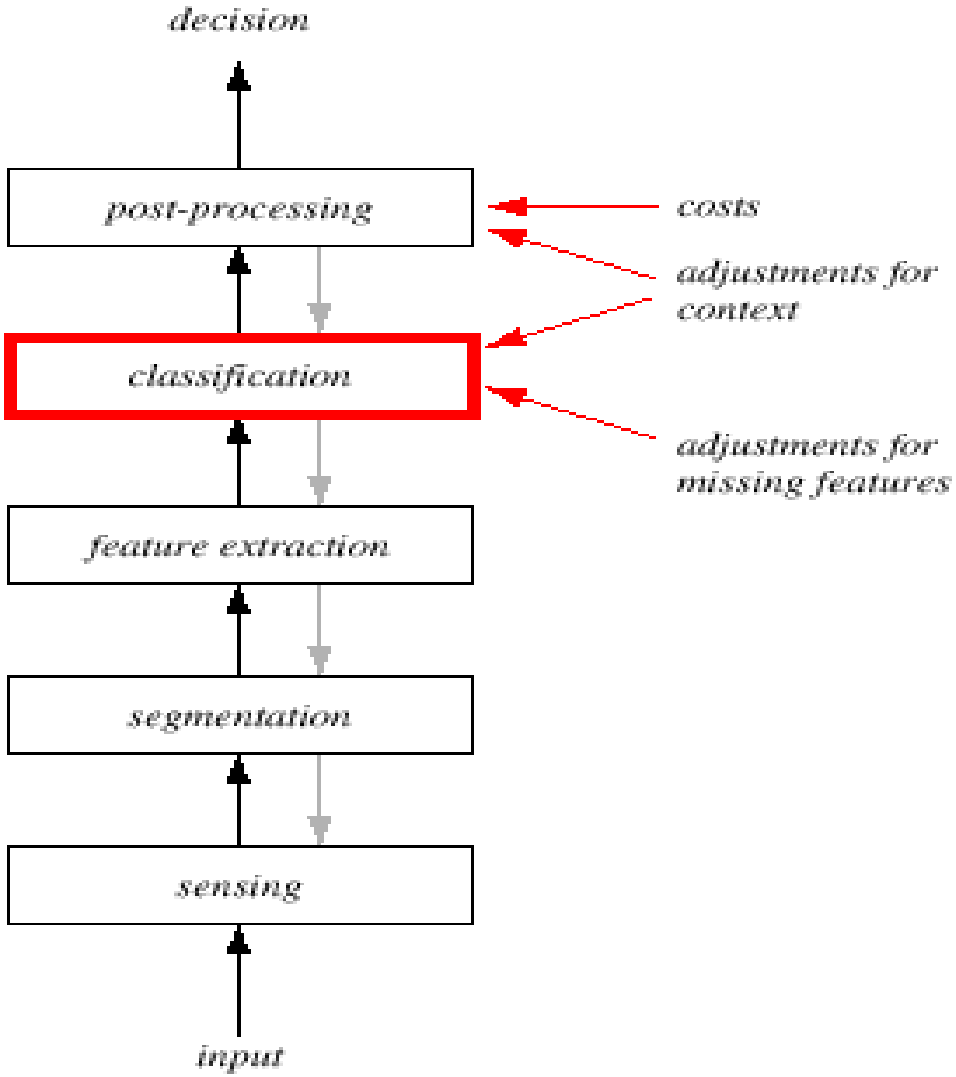
- satisfaction is premature
  - cause: aim of a classifier is to correctly classify unknown input

Issue of generalization!



A compromise between training and testing

Pattern Recognition System



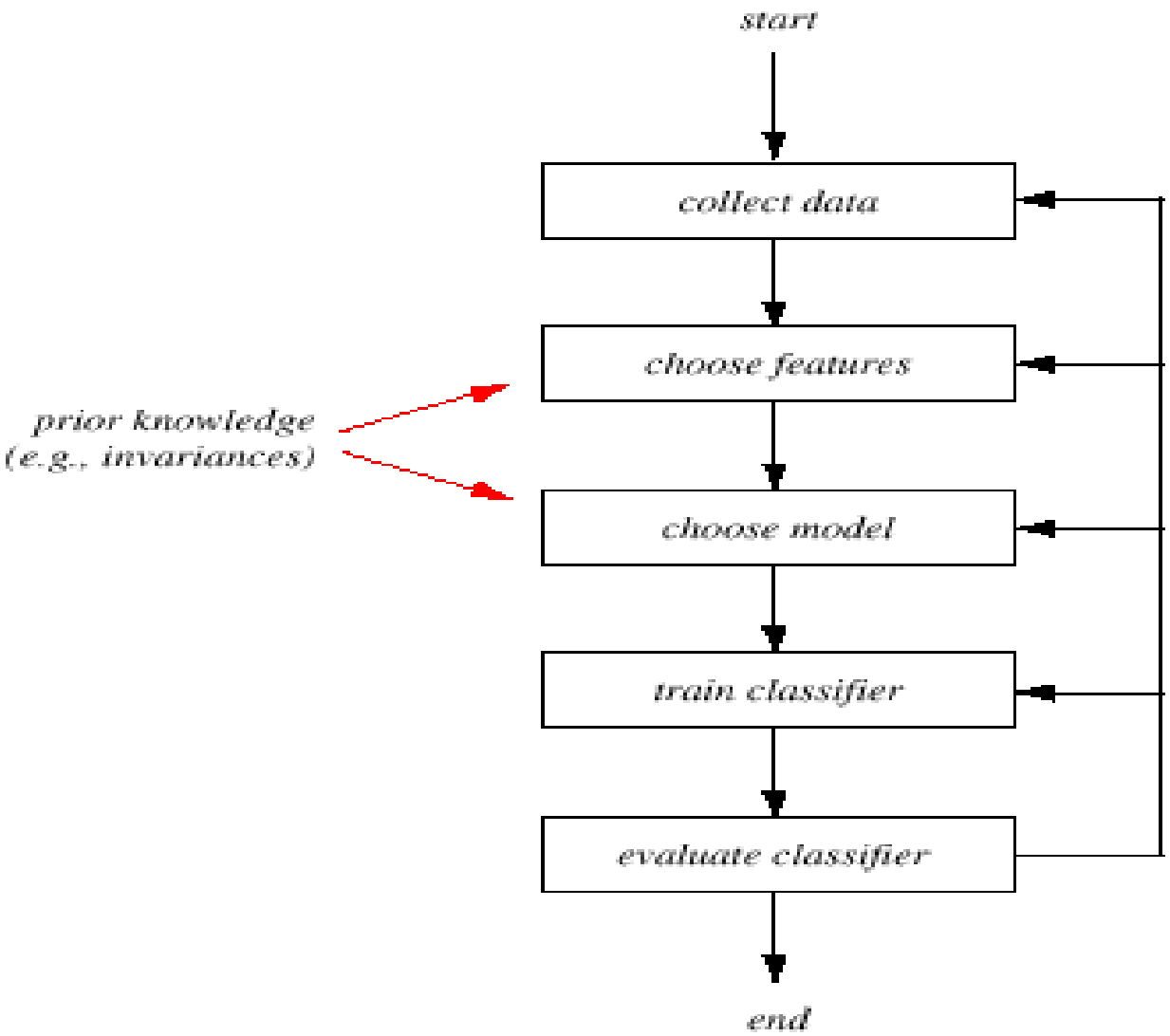
Pattern Recognition System

- Sensing
  - Use of a transducer (camera or microphone)
  - PR system depends on the bandwidth, the resolution sensitivity distortion of the transducer
- Segmentation and grouping
  - Patterns should be well separated and should not overlap

- Feature extraction
  - Discriminative features
  - Invariant features with respect to translation, rotation and scale.
- Classification
  - Use a feature vector provided by a feature extractor to assign the object to a category
- Post Processing
  - error rate
  - risk
  - use context

The Design Cycle

- Data collection
- Feature Choice
- Model Choice
- Training
- Evaluation
- Computational Complexity


start
collect data
choose features
prior knowledge
(e.g., invariances)
choose model
train classifier
evaluate classifier
end

- Data Collection

  - How do we know when we have collected an adequately large and representative set of examples for training and testing the system?

- Feature Choice

  – Depends on the characteristics of the problem domain.

  – Requirement

    - simple to extract
    - invariant to irrelevant transformation
    - insensitive to noise.

- Model Choice
  - too many classification models?
  - which one is best?

- Training
  - Use data to determine the classifier. Many different procedures for training classifiers and choosing models

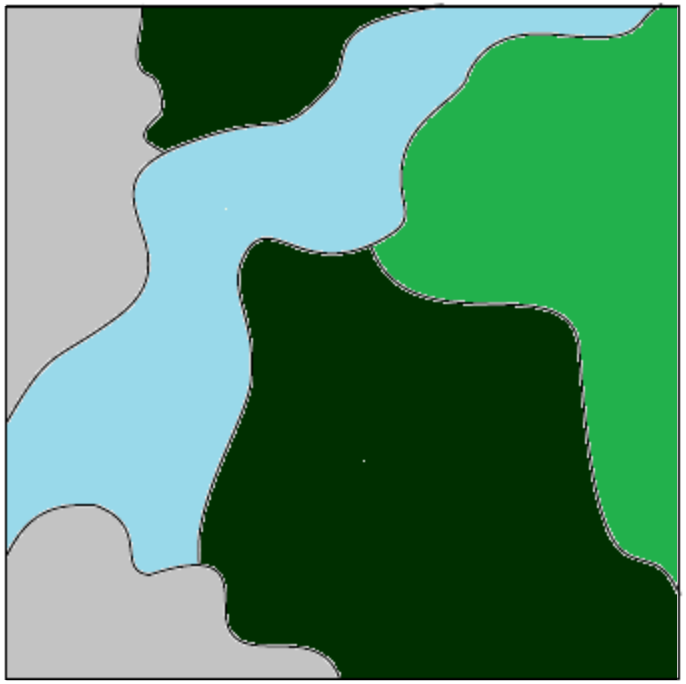
- Evaluation
  - Measure the error rate (or performance) and switch from one set of features to another

- Computational Complexity
  - What is the trade-off between computational ease and performance?

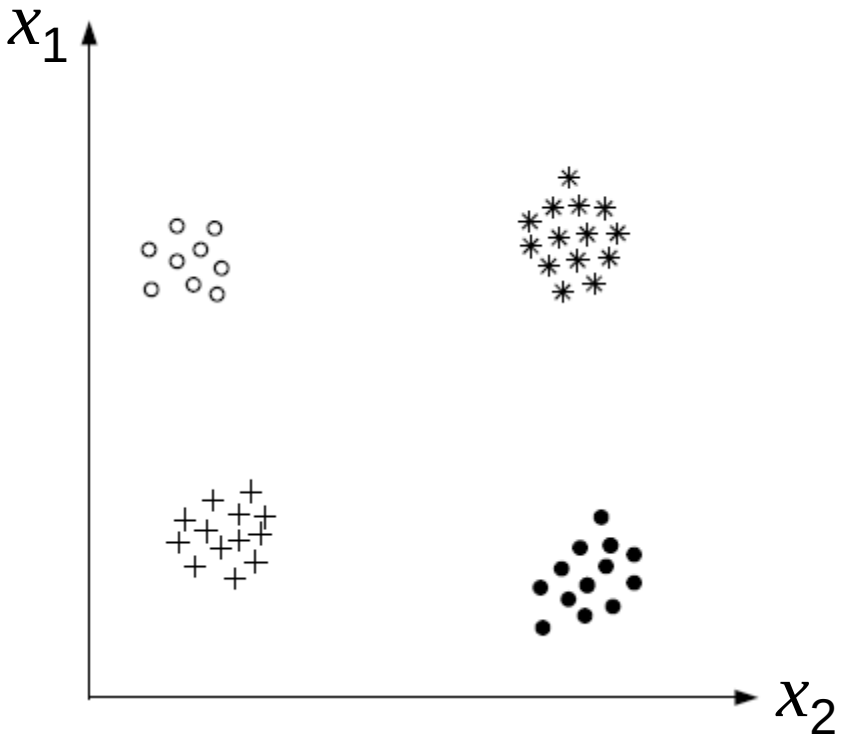
Supervised vs. Unsupervised Learning

- Supervised learning
  - A teacher provides a category label or cost for each pattern in the training set
- Unsupervised learning
  - The system forms clusters or “natural groupings” of the input patterns

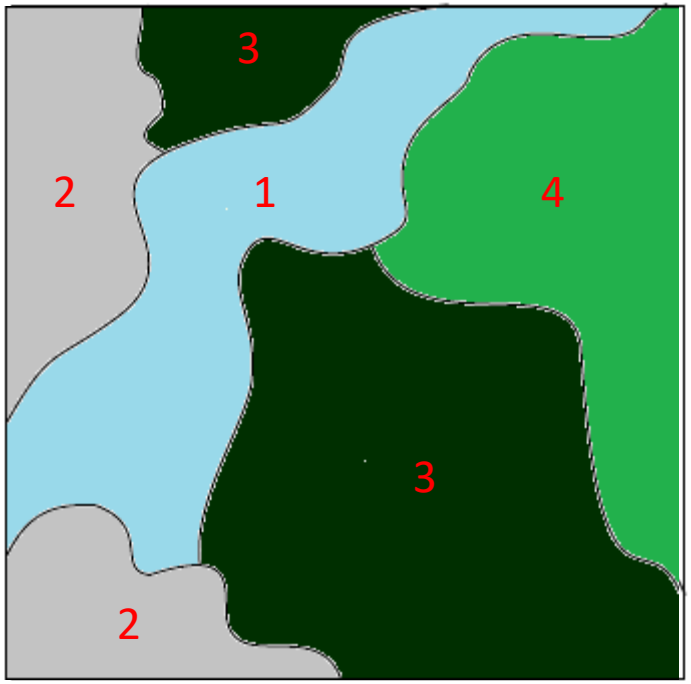
Unsupervised Learning

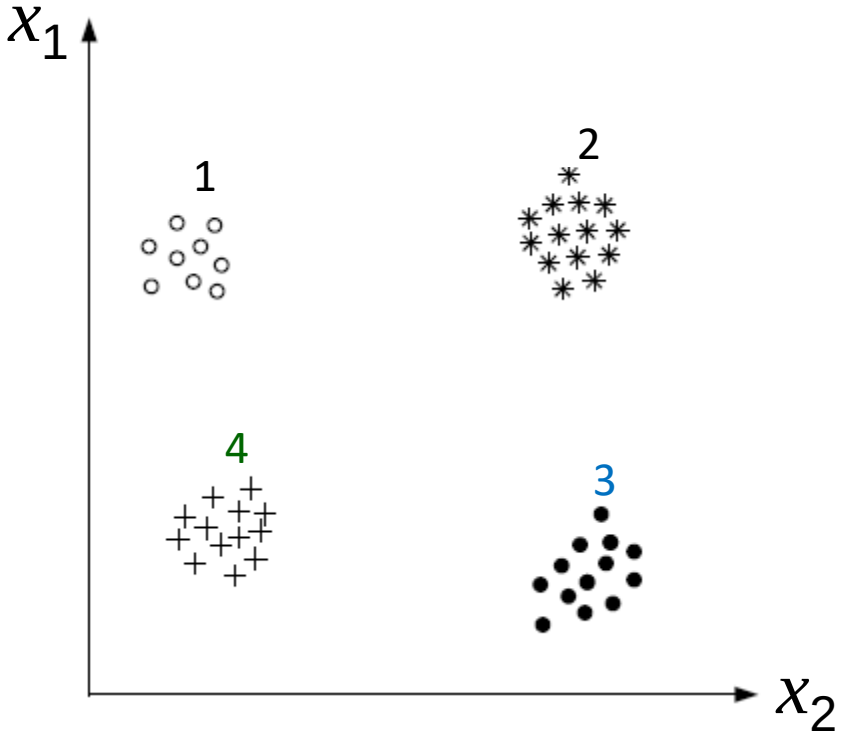


Unsupervised Learning



Unsupervised Learning





Unsupervised Learning