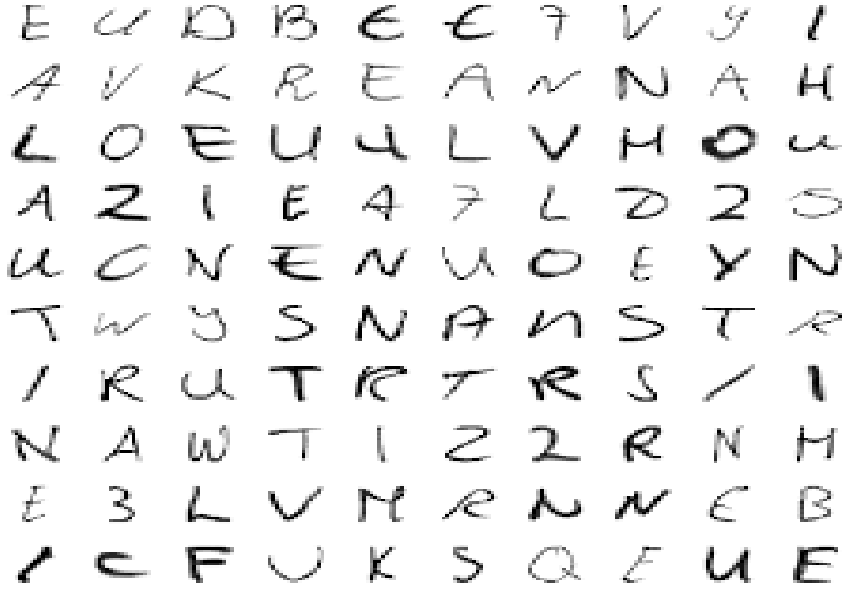
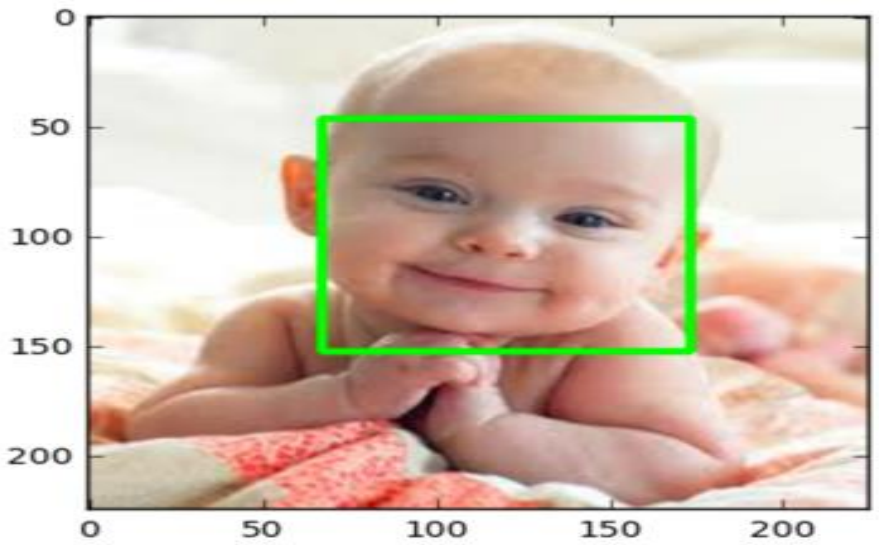
CSE 467
Pattern Recognition

Pre Requisite:
CSE 351: Artificial Intelligence

Pattern Recognition





Courtesy: https://www.google.com/

Pattern Recognition: With Example

Scenario:

Suppose that a fish packing plant wants to automate the process of sorting incoming fish on a conveyor belt according to species. As a pilot project it is decided to try to separate sea bass from salmon using optical sensing.

Pattern Recognition: With Example

Collect Data

- take some sample images
- some **physical differences** between the two types of fish — length, lightness, width, number and shape of fins, position of the mouth, and so on
- these suggest **features** to explore for use in our classifier

Pattern Recognition: With Example

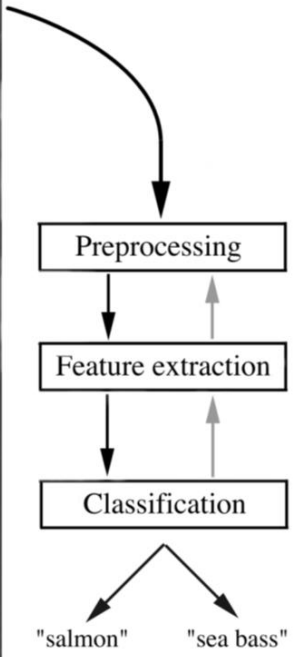
Model

- there truly are differences between the population of sea bass and that of salmon, we model view them as having different models — different descriptions, which are **typically mathematical in form**.
- process the sensed data to eliminate noise
- for any sensed pattern choose the model that corresponds best

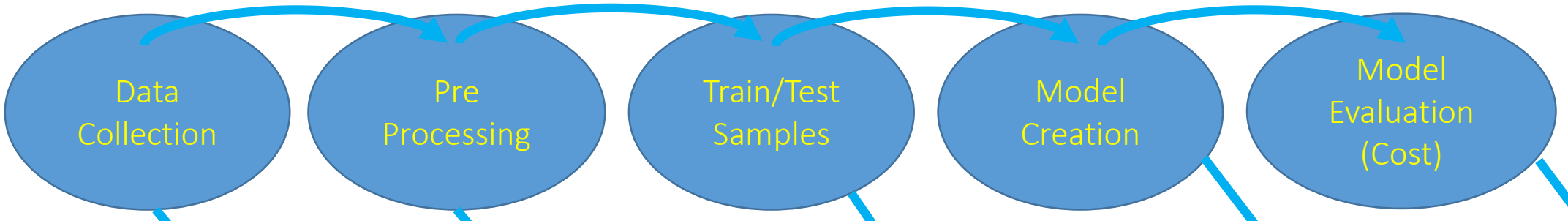
Pattern Recognition: With Example

Steps

- Camera captures an image of the fish
- Signals are preprocessed to simplify subsequent operations without loosing relevant information
- Segmentation
- Feature selection or extraction
- Choose training and testing samples
- Creating model
- Evaluate Model (cost)
- Decision



Pattern Recognition: Learning Steps or Cycle



- Image through camera
- Statistical Survey
- Transaction Information
- Web data
- Speech etc.

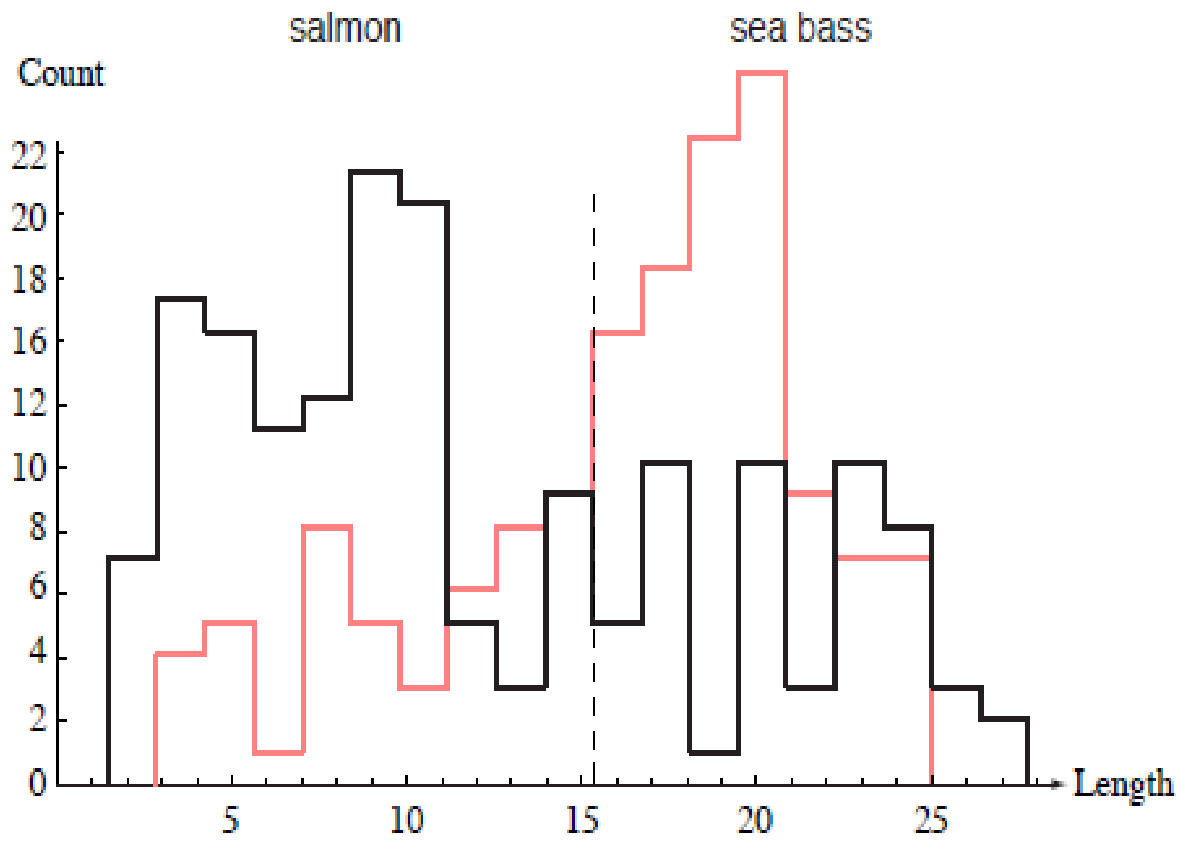
- Segmentation
- Noise Removing
- Feature Extraction

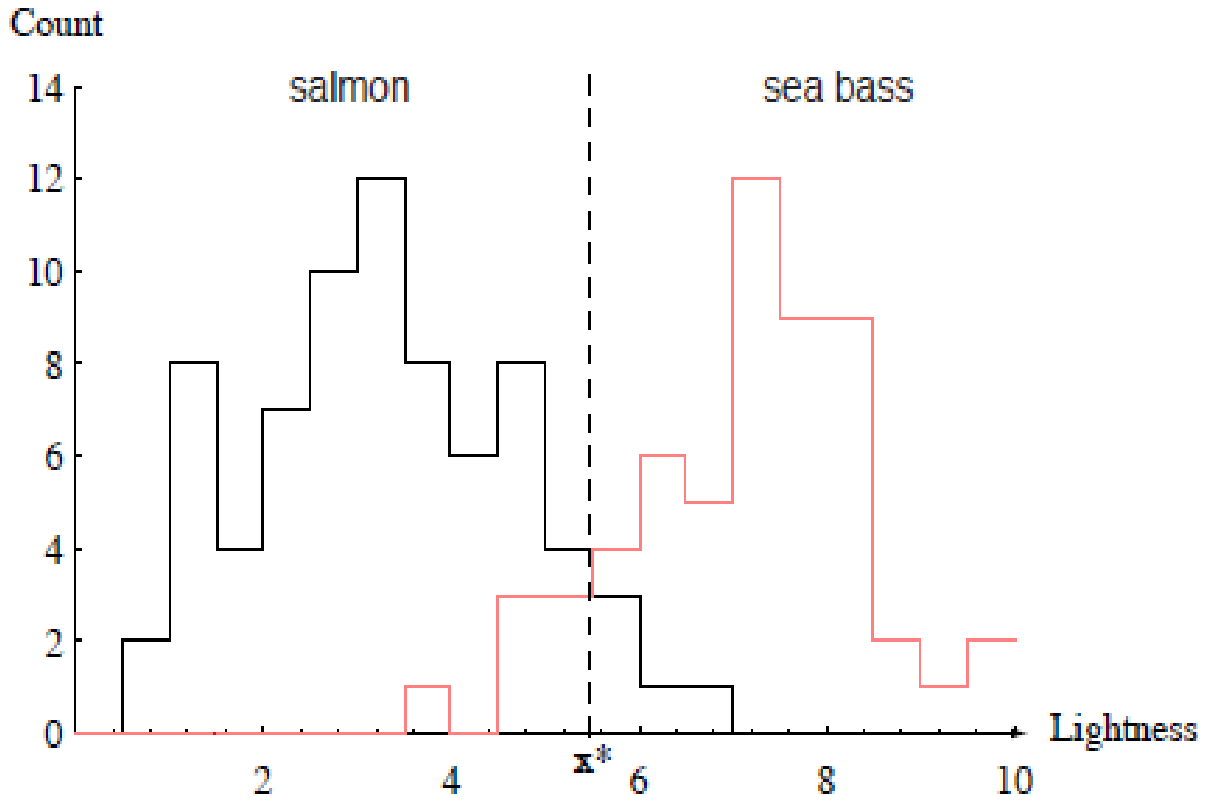
- 80-20 rules
- K-fold cross validation
- Manual Selection

- Classification
- Cluster
- Regression

Justify the Decision Boundary with different metrics such as accuracy, sensitivity, RMSE etc

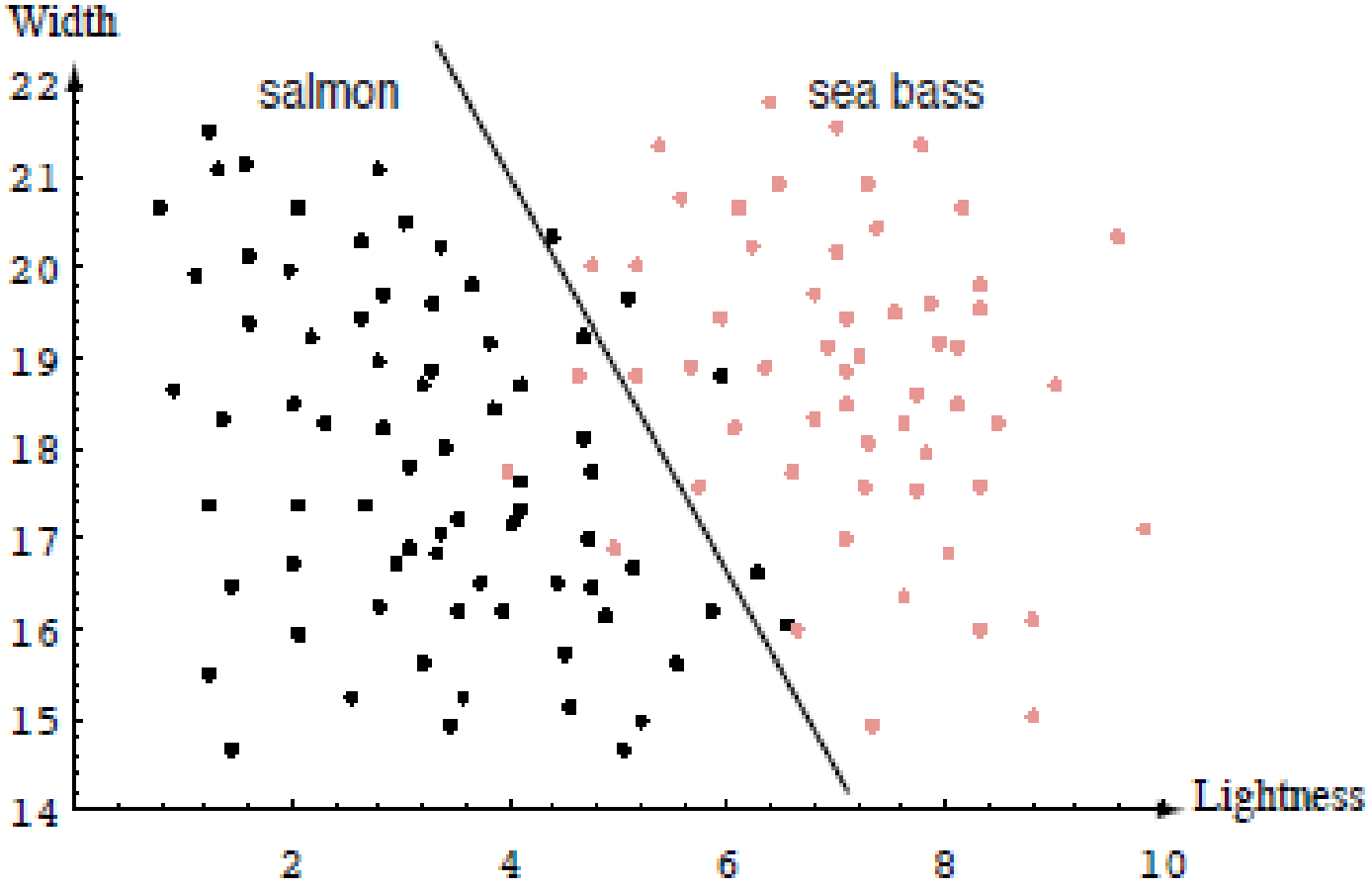
Decision Boundary With Single feature





Showing poor performance: Cost increase

Decision Boundary With Two features



Feature Vector:

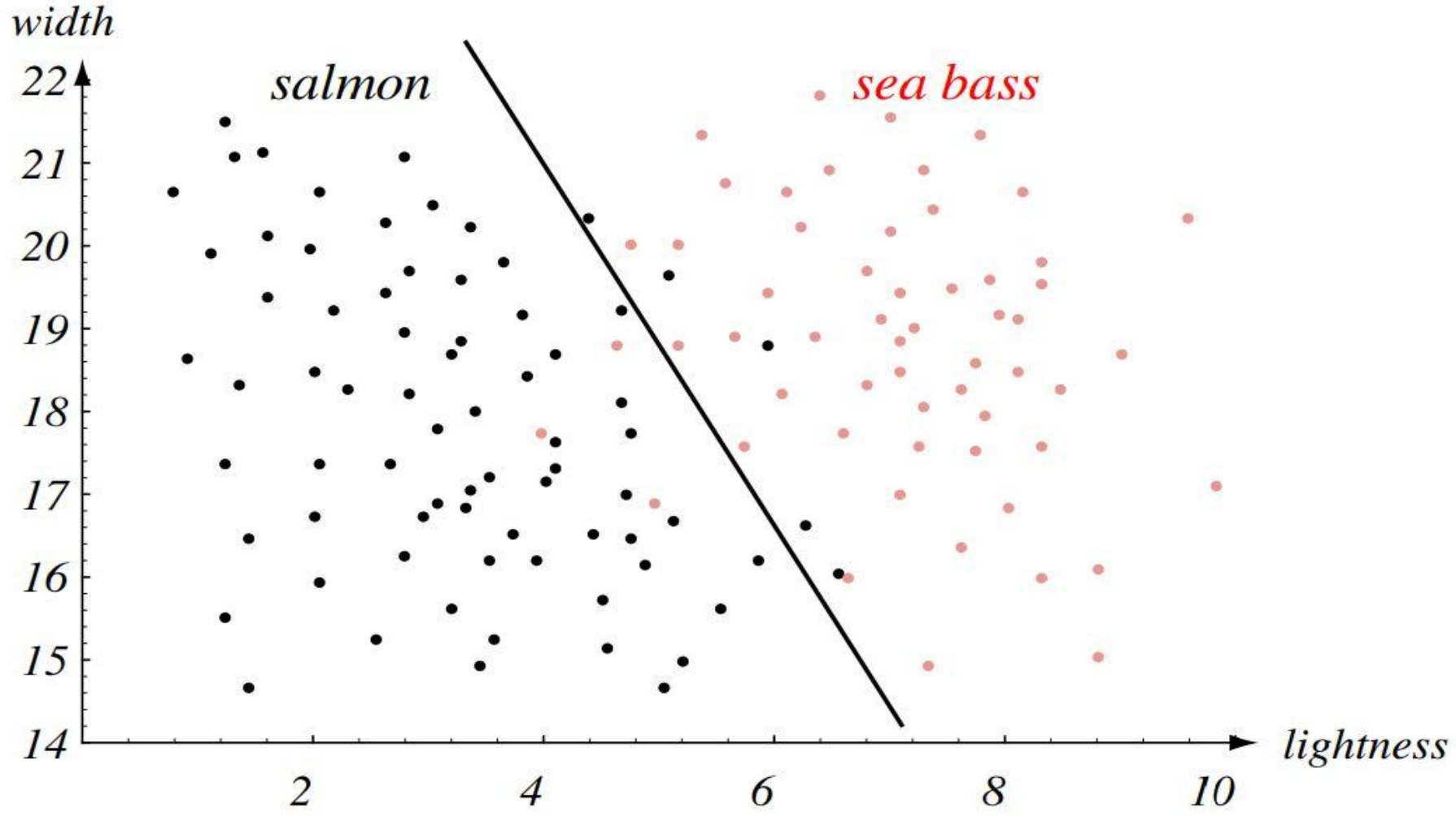
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- Still Poor Performance

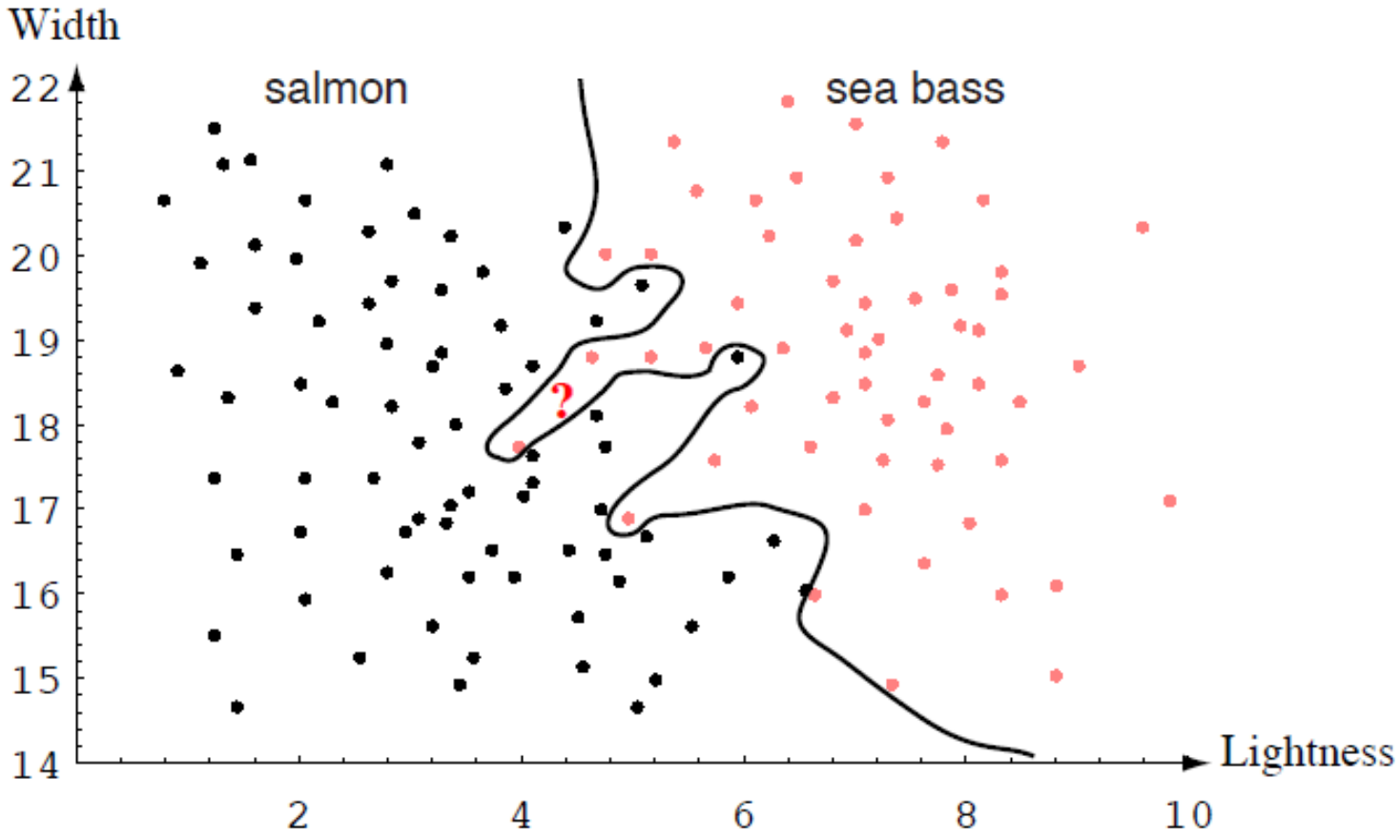
Curse of Dimensionality

- ❑ The two features obviously separate the classes much better than one alone.
- ❑ This suggests adding a third feature. And a fourth feature. And so on.
- ❑ Key questions
  - ✓ How many features are required?
  - ✓ Is there a point where we have too many features?
  - ✓ How do we know beforehand which features will work best?

Decision Boundary : Underfitting Problem [High Bias]

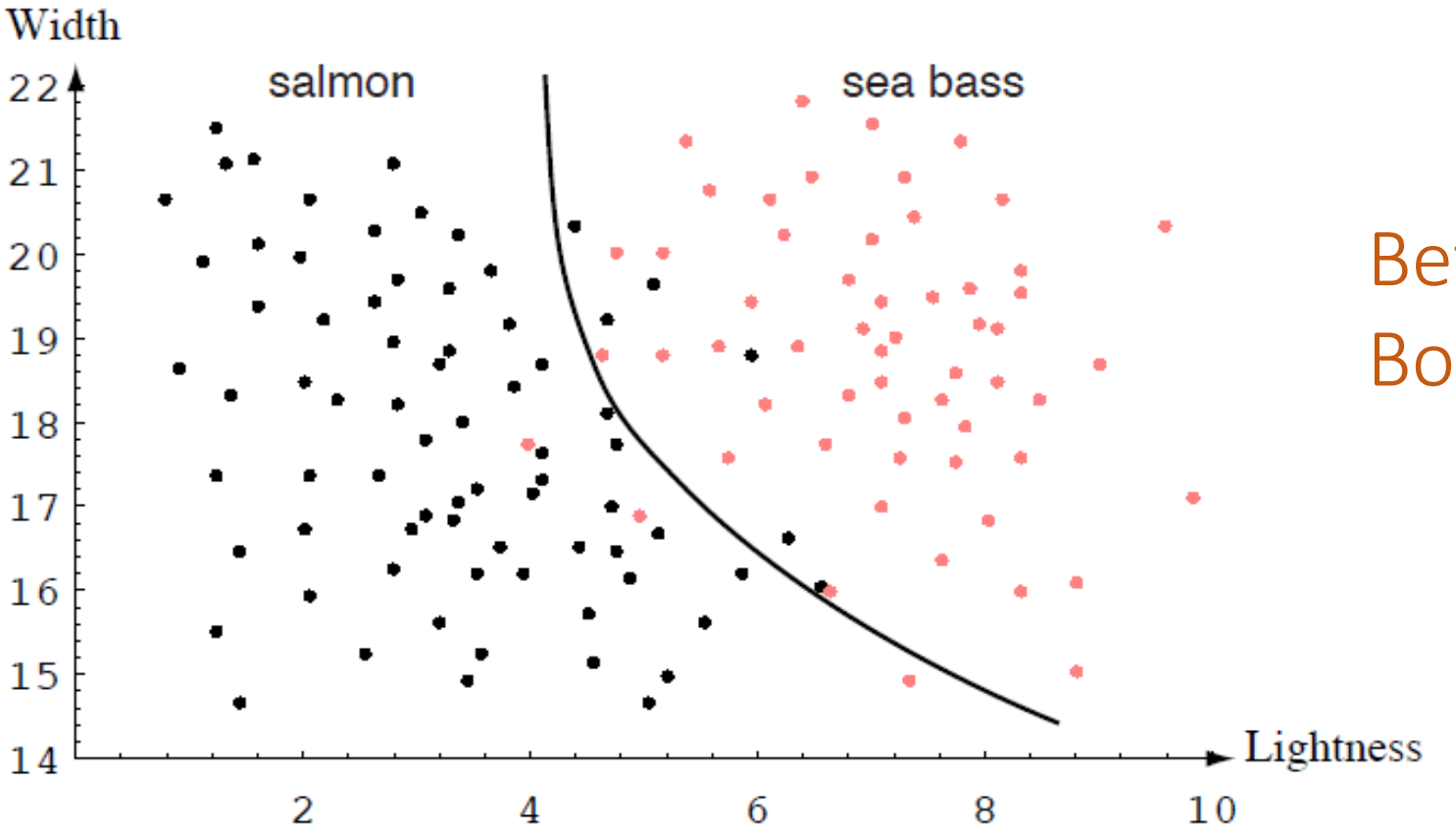


Decision Boundary : Overfitting Problem [High Variance]



?

Decision Boundary : Generalization [Low Bias ,Low Variance]



Better Boundary

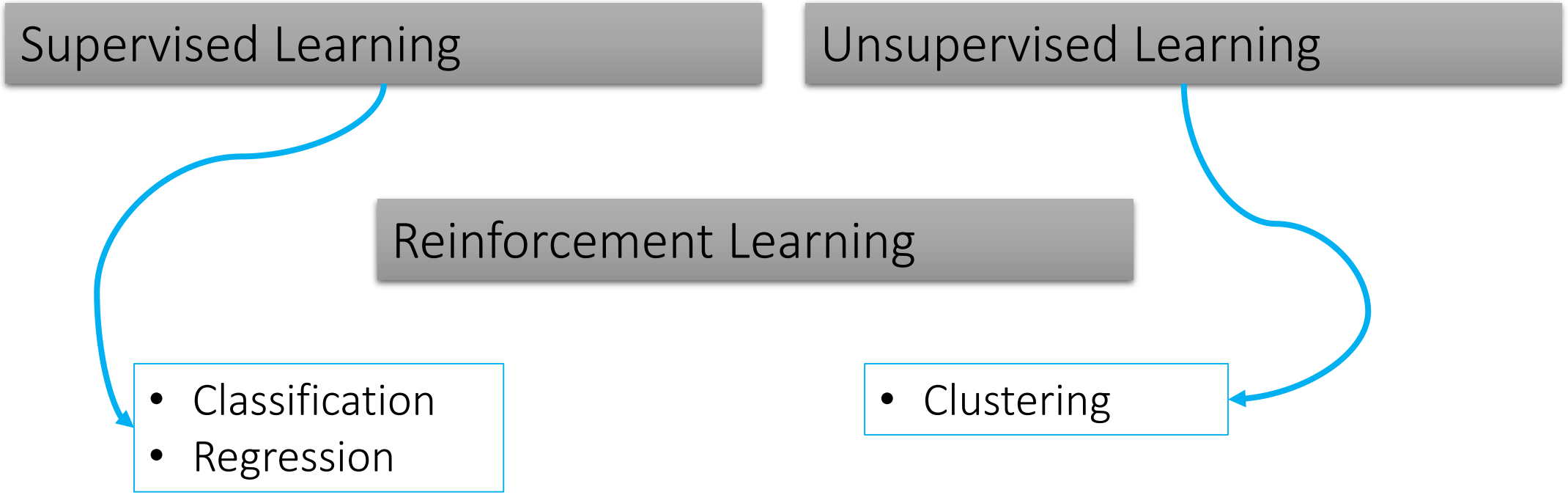
Pattern Recognition

Pattern is everything around in this digital world. A pattern can either be seen physically or it can be observed mathematically by applying algorithms.

Pattern recognition — the act of taking in raw data and taking an action based on the “category” of the pattern.

Examples: automated speech recognition, fingerprint identification, optical character recognition, DNA sequence identification and much more.

Learning and Adaptation



Some Terminologies

Model Selection

Noise

Feature Extraction

Overfitting

Missing Features

Computational Complexity

Segmentation

Context

Prior Knowledge

Mereology

Evidence Pooling

Costs and Risks

Thank You