

Pattern Recognition Systems

Dr. Shuang LIANG

School of Software Engineering
Tongji University
Fall, 2012

Today's Topics

- An example
- Pattern recognition systems
- The design cycle

Today's Topics

- *An example*
- Pattern recognition systems
- The design cycle

An Example

- Problem
 - Sorting incoming fish on a conveyor belt according to species
 - Assume that we have only two kinds of fish
 - Sea bass
 - salmon



Decision Process

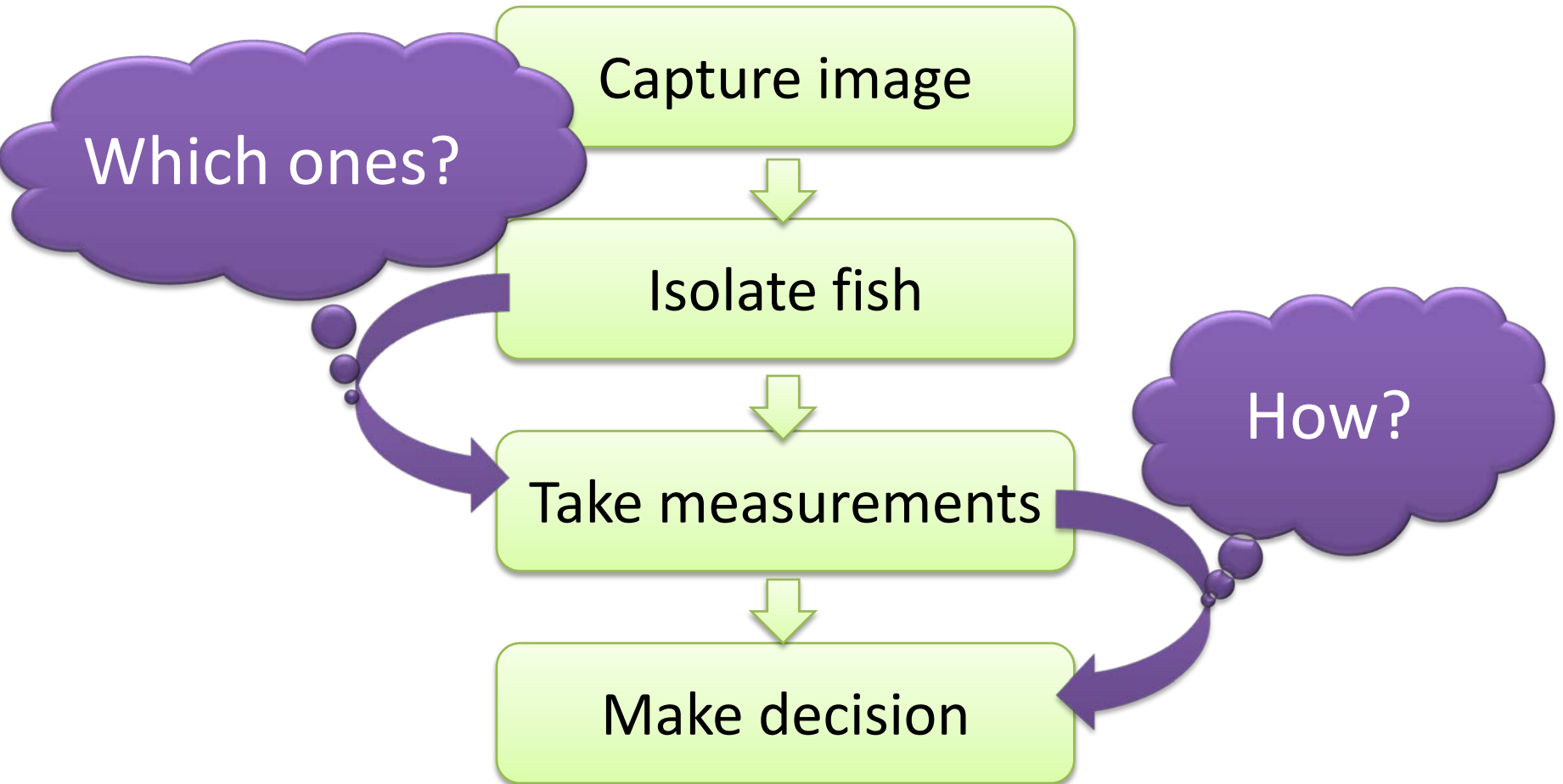
- Q 1: What kind of information can distinguish one species from the other?
 - Length
 - Width
 - Weight
 - Number and shape of fins
 - Tail shape
 -

Decision Process

- Q 2: What can cause problems during sensing?
 - Lighting conditions
 - Position of fish on the conveyor belt
 - Camera noise
 -

Decision Process

- Q 3: What are the steps in the process?

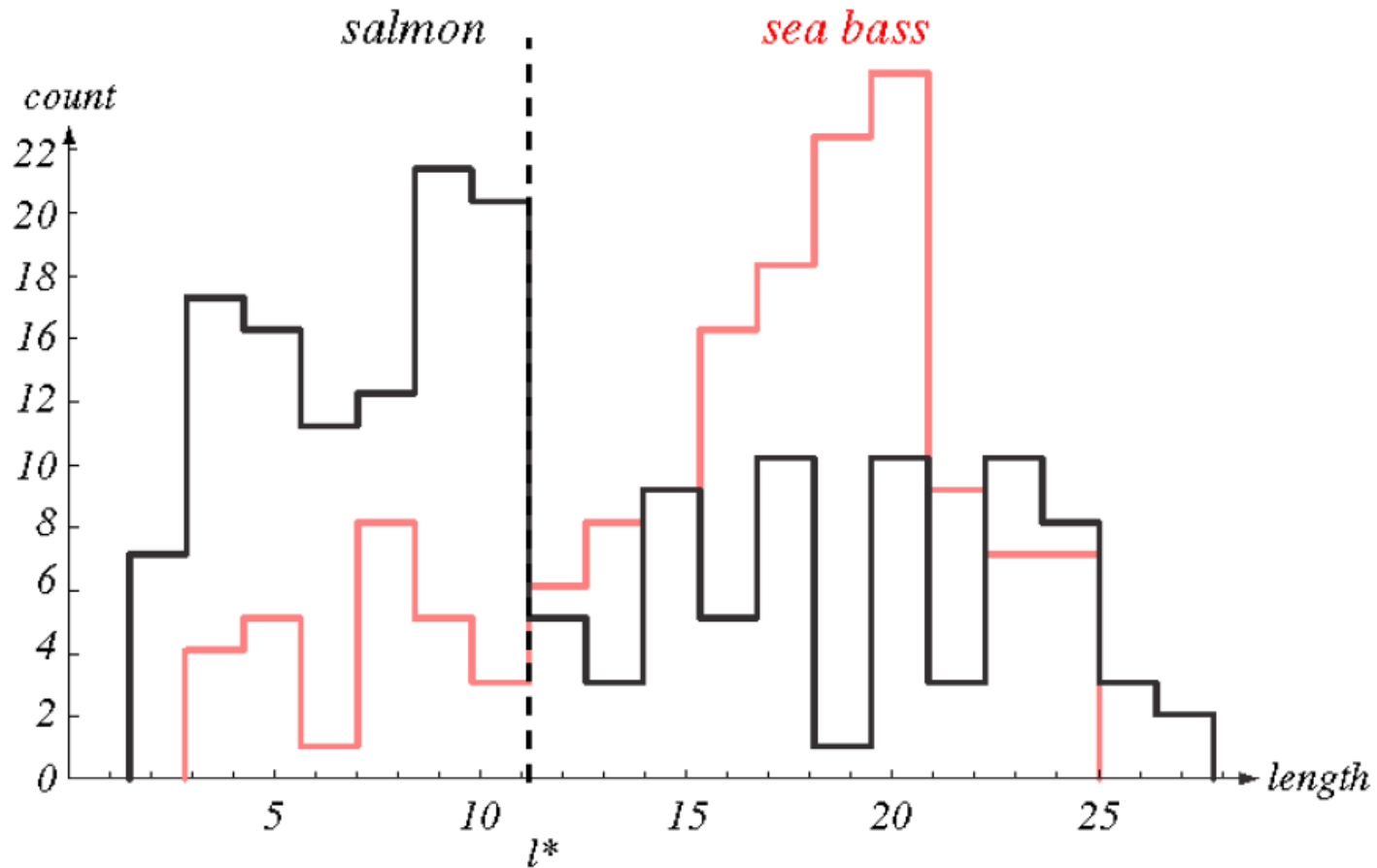


Selecting Features

- Assume a fisherman told us that a sea bass is generally longer than a salmon
- We can use length as a *feature* and decide between sea bass and salmon according to a threshold on length
- How can we choose this threshold?



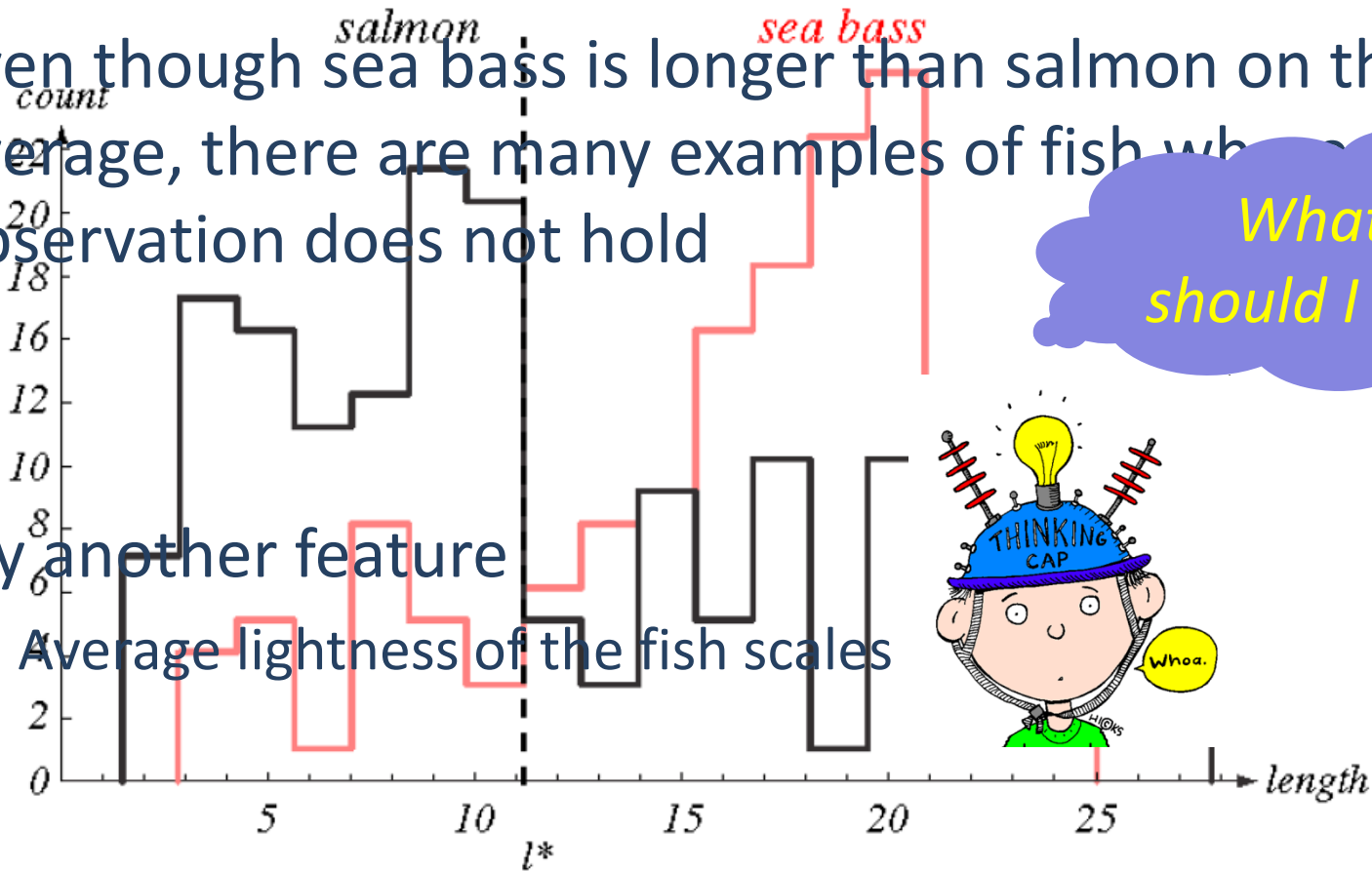
Selecting Feature (cont.)



- *Histograms* of the length feature for two types of fish in *training samples*.
- How can we choose the threshold l^* to make a reliable decision?

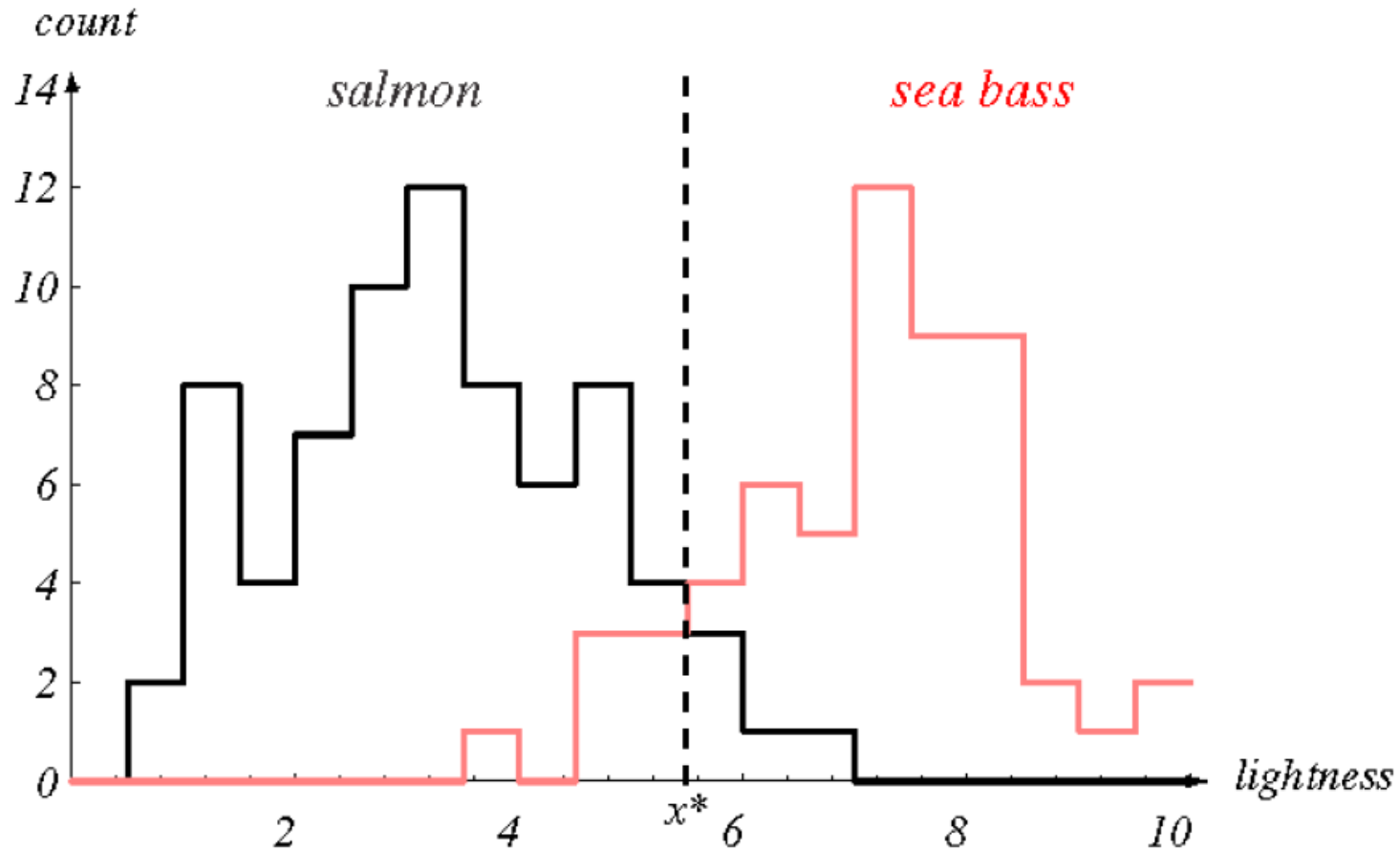
Selecting Features (cont.)

- Even though sea bass is longer than salmon on the average, there are many examples of fish whose observation does not hold



- Try another feature
 - Average lightness of the fish scales

Selecting Features (cont.)



- Histograms of the lightness feature for two types of fish in training samples
- It looks easier to choose the threshold x^* but we still cannot make a perfect decision

Cost of Error

- We should also consider *costs of different errors* we make in our decisions
- For example, If the fish packing company knows that
 - Customers who buy salmon will object vigorously if they see sea bass in their cans
 - Customers who buy sea bass will not be unhappy if they occasionally see some expensive salmon in their cans
- How does this knowledge affect our decision?

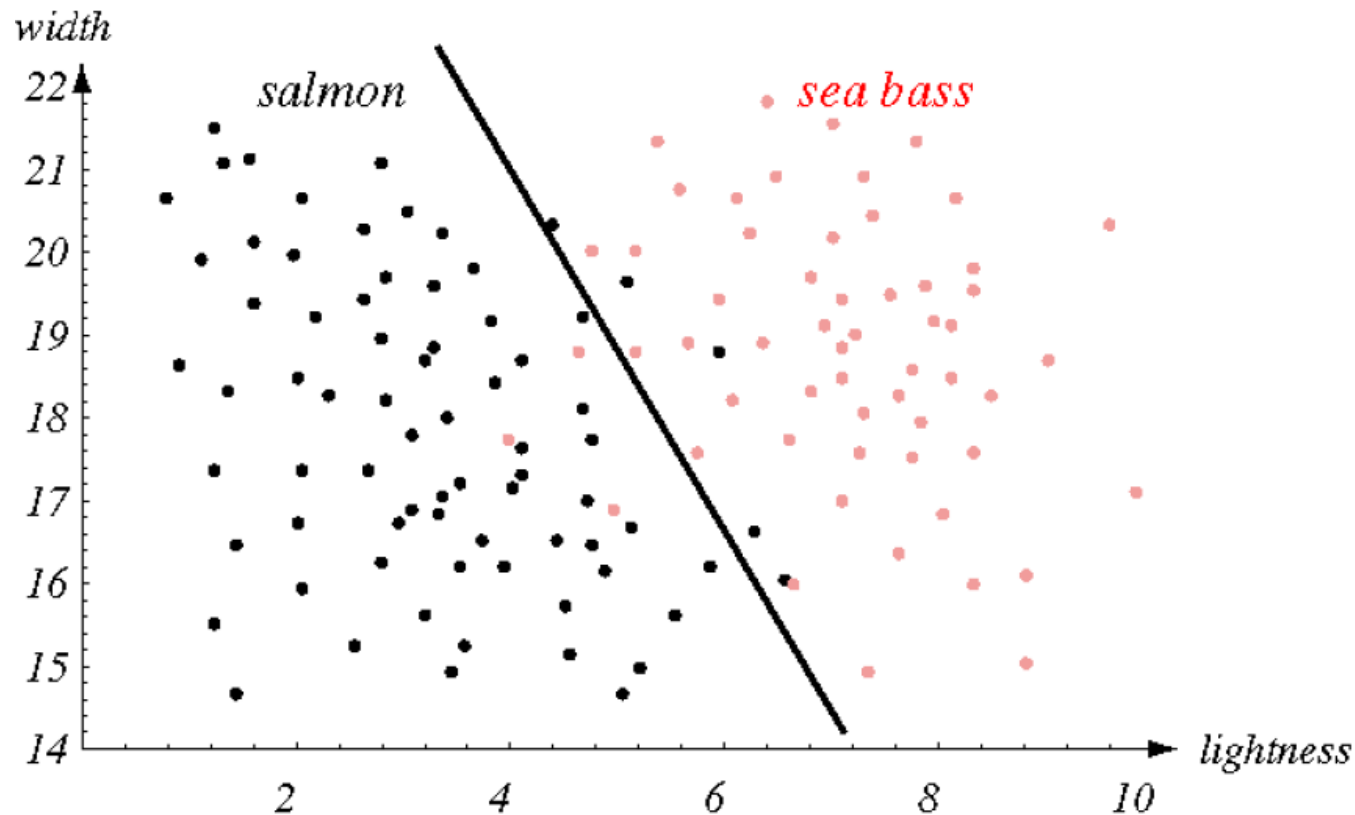
Multiple Features

- Assume we also observed that sea bass are typically wider than salmon
- We can use two features in our decision
 - Lightness: x_1
 - Width: x_2
- Each fish image is now represented as a point (*feature vector*)

$$x = \begin{cases} x_1 \\ x_2 \end{cases}$$

In a two-dimensional *feature space*

Multiple Features (cont.)



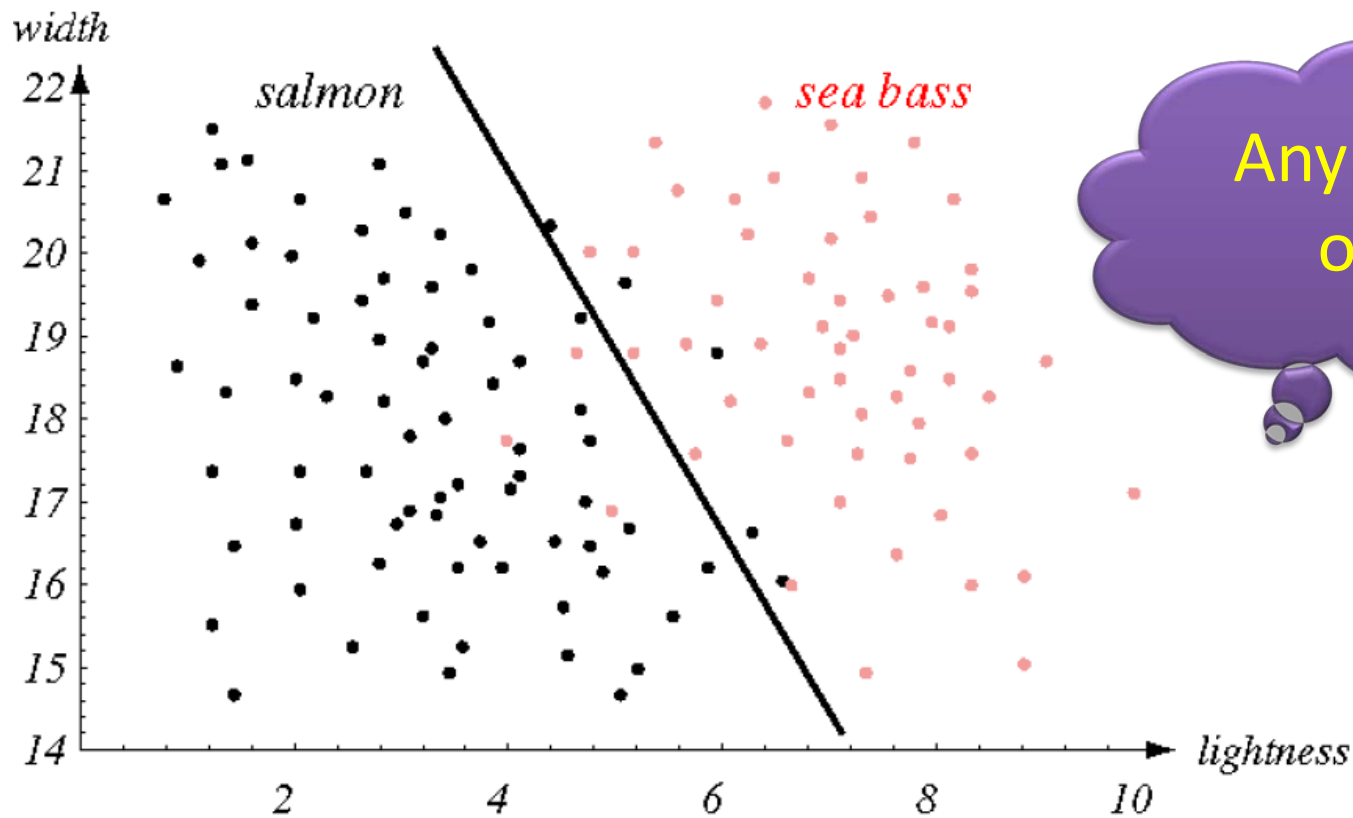
- *Scatter plot* of lightness and width features for training samples.
- We can draw a *decision boundary* to divide the feature space into two regions
- Does it look better than using only lightness?

Multiple Features (cont.)

- Does adding more features always improve the results?
 - Avoid unreliable features
 - Be careful about correlations with existing features
 - Be careful about measure costs
 - Be careful about noise in the measurements
- Is there some *curse* for working in very high dimensions?

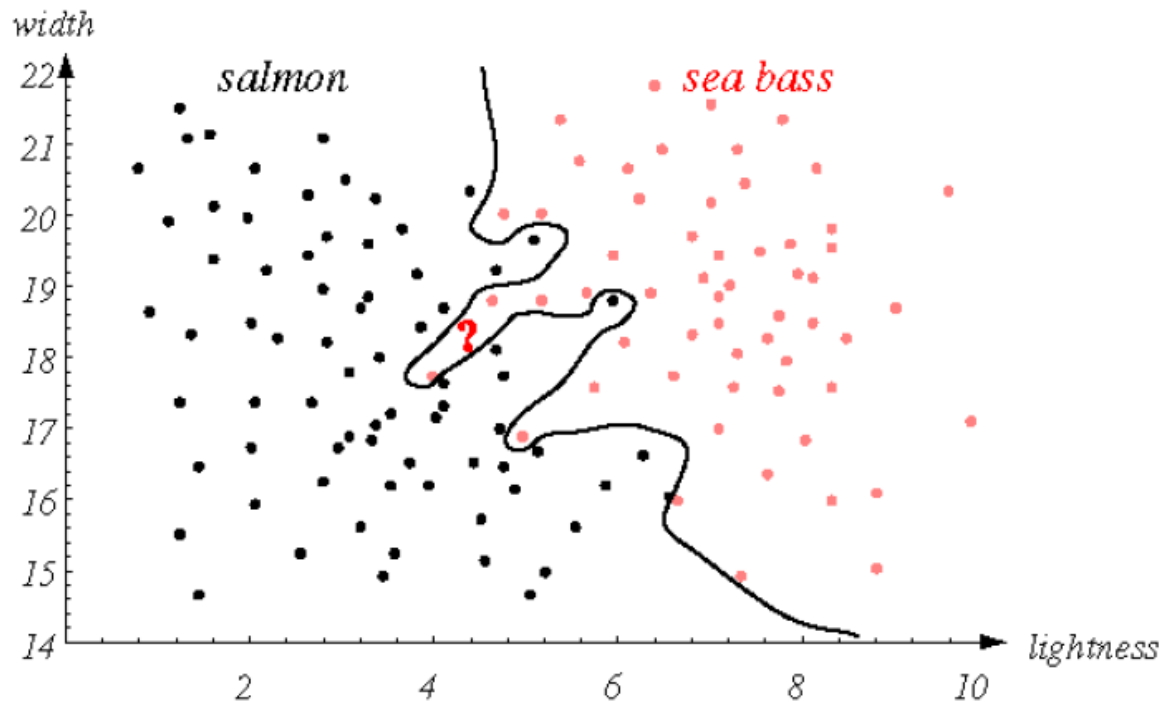
Decision Boundaries

- Can we do better with another decision rule?



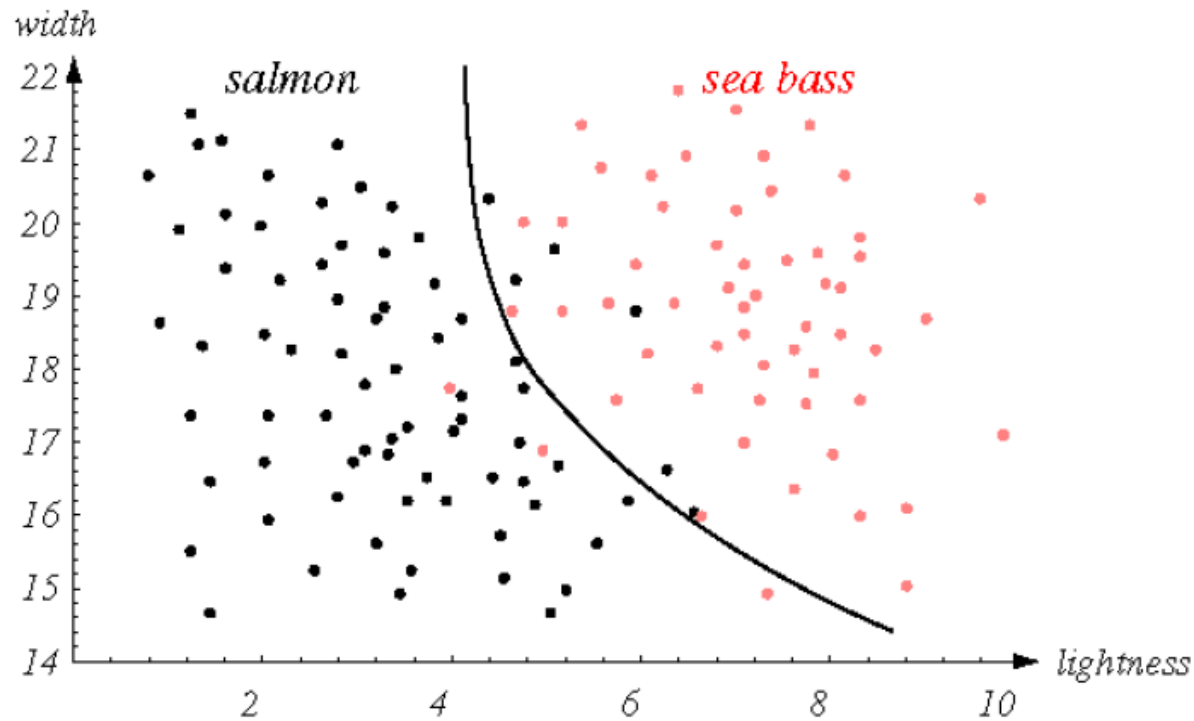
Decision Boundaries (cont.)

- More complex models result in more complex boundaries



Decision Boundaries (cont.)

- Different criteria lead to different decision boundaries

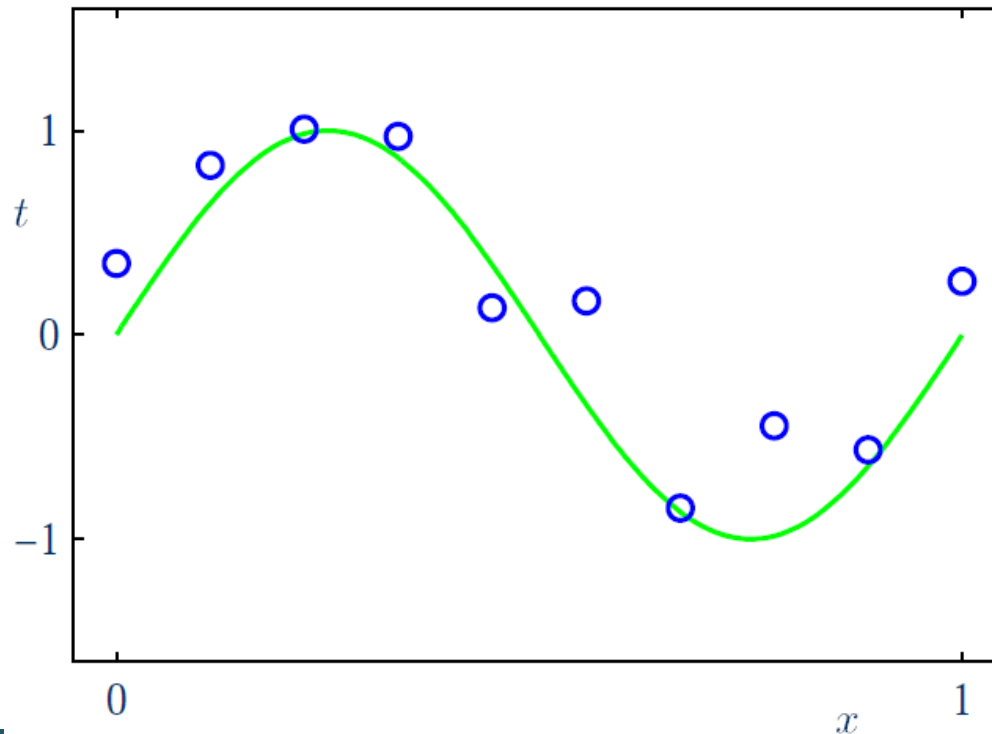


Decision Boundaries (cont.)

- Two aspects of concerns
 - We may distinguish training samples perfectly but how can we predict how well we can *generalize* to unknown samples
 - How can we manage the *tradeoff* between complexity of decision rules and their performance to unknown samples?

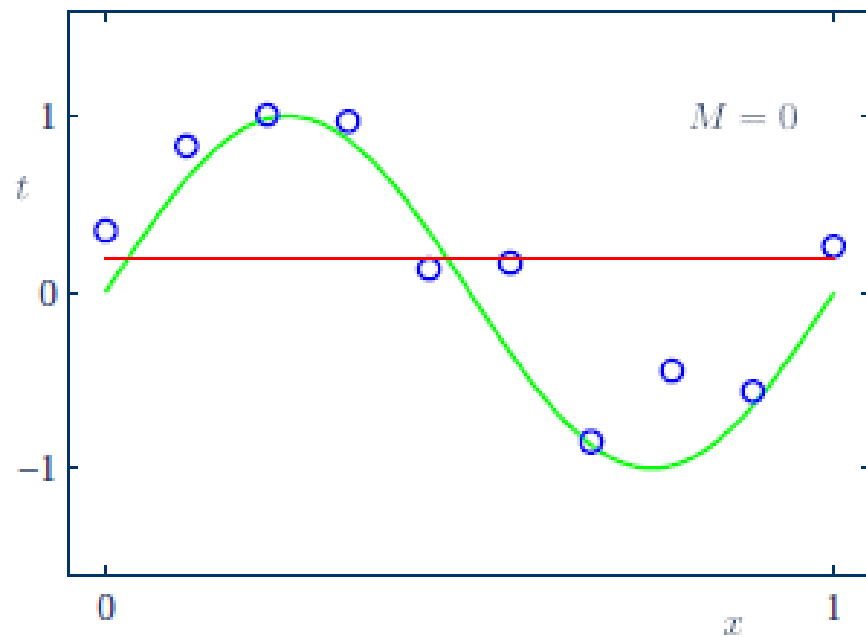
More on Complexity

- Regression example
 - Plot of 10 sample points for the input variable x along with the corresponding target variable t .
 - Green curve is the true function that generated the data

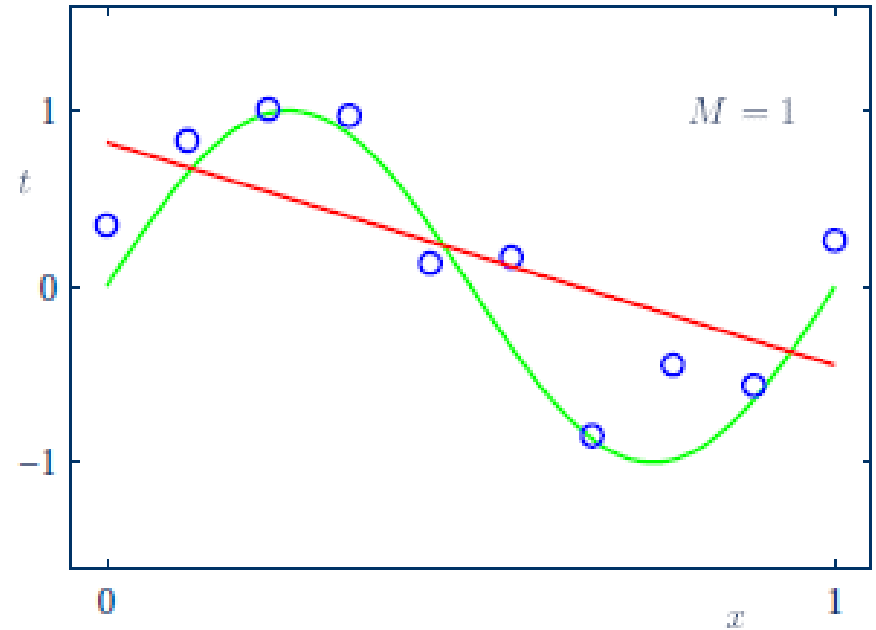


More on Complexity (cont.)

- Polynomial curve fitting
 - Plots of polynomials having various orders shown as red curved, fitted to the set of 10 sample points



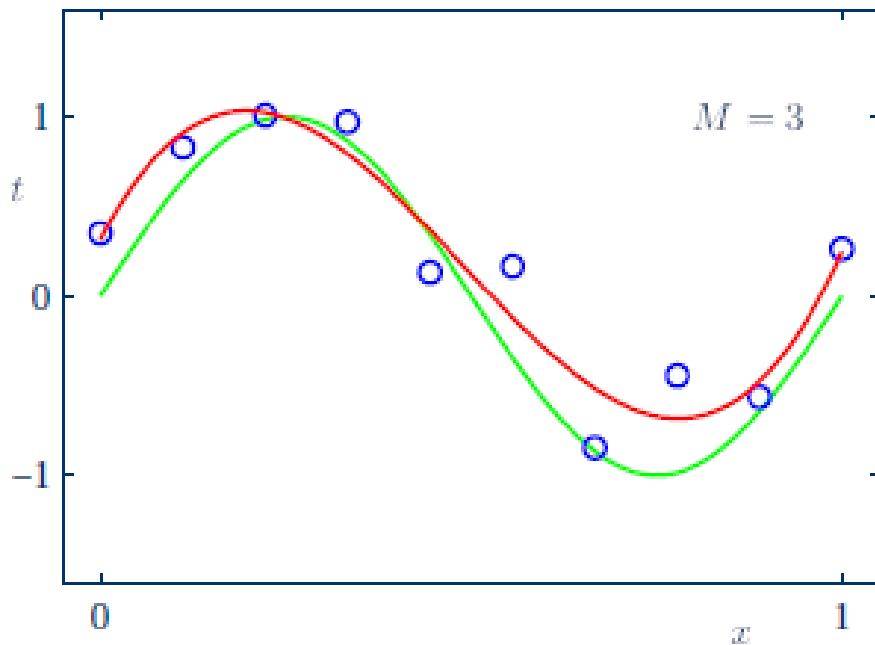
0'th order polynomial



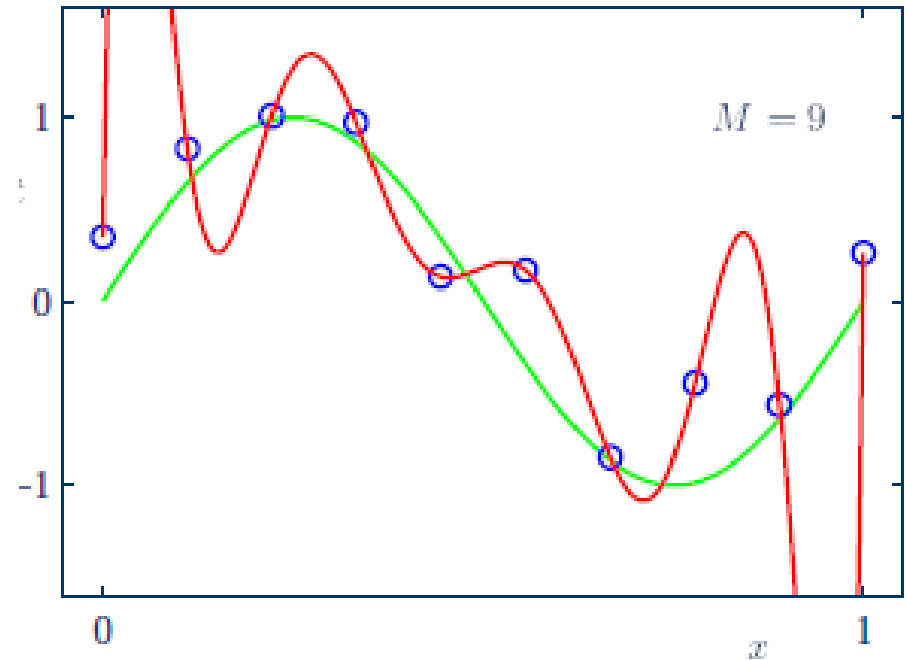
1'st order polynomial

More on Complexity (cont.)

- Polynomial curve fitting
 - Plots of polynomials having various orders shown as red curved, fitted to the set of 10 sample points

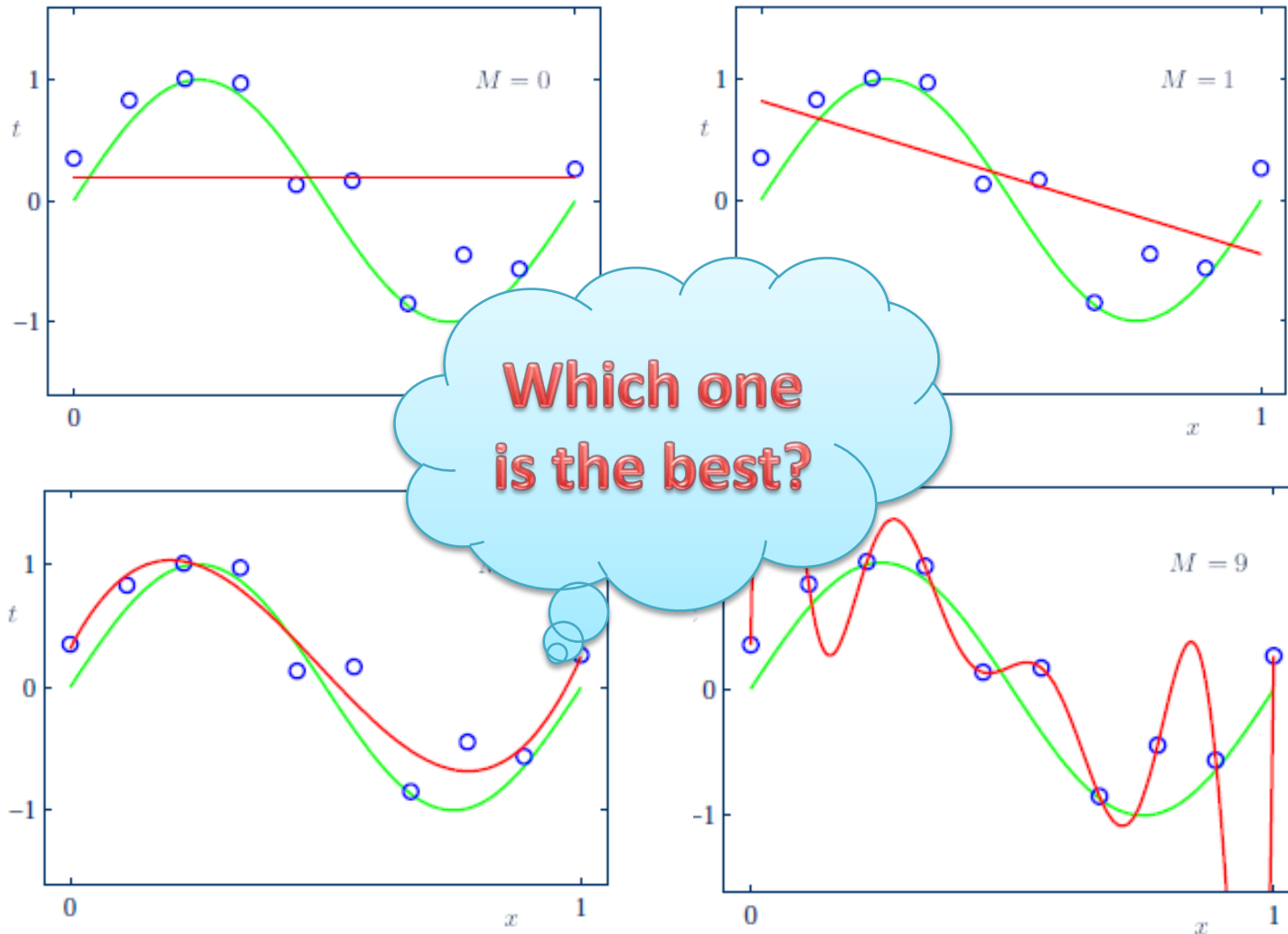


3'rd order polynomial



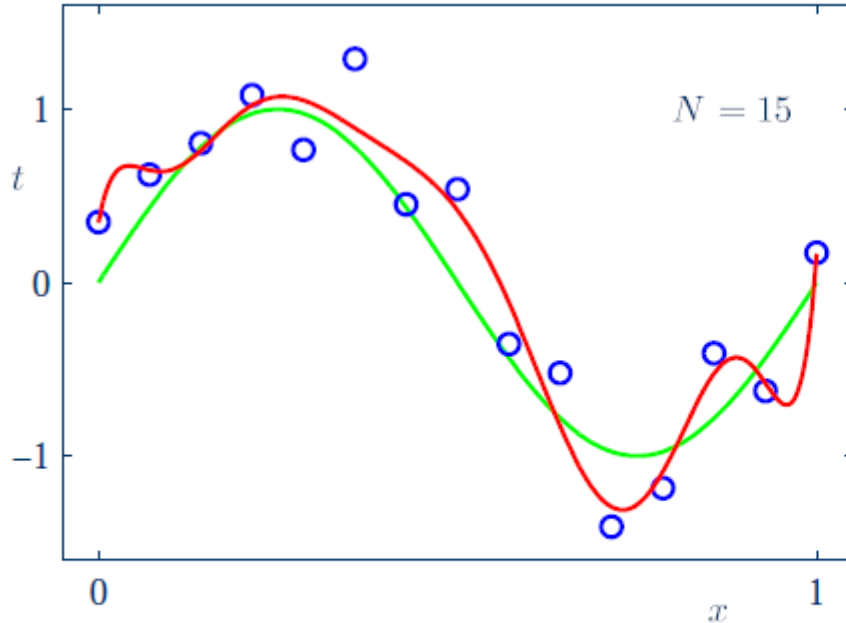
9'th order polynomial

More on Complexity (cont.)

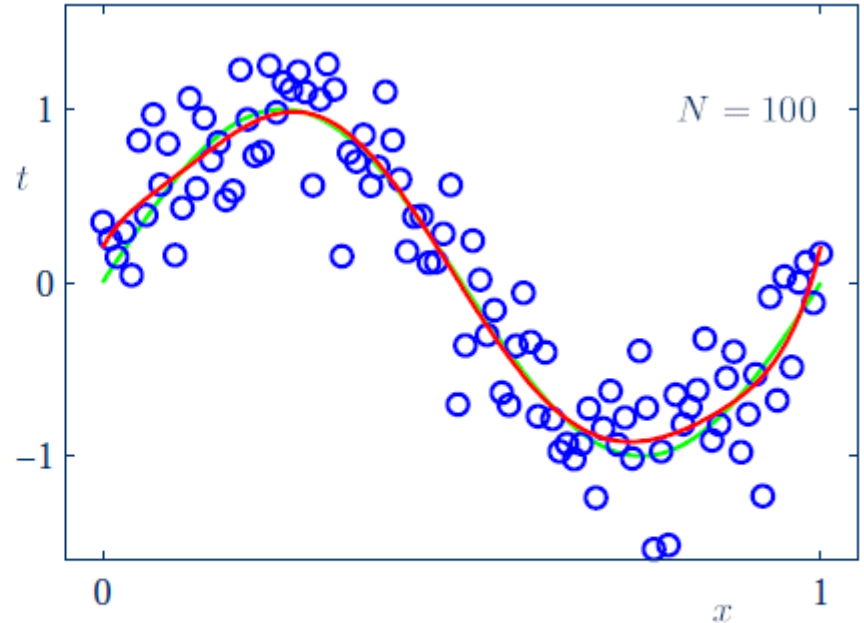


More on Complexity (cont.)

- Polynomial curve fitting
 - Plots of 9'th order polynomials fitted to 15 and 100 sample points



(a) 15 sample points



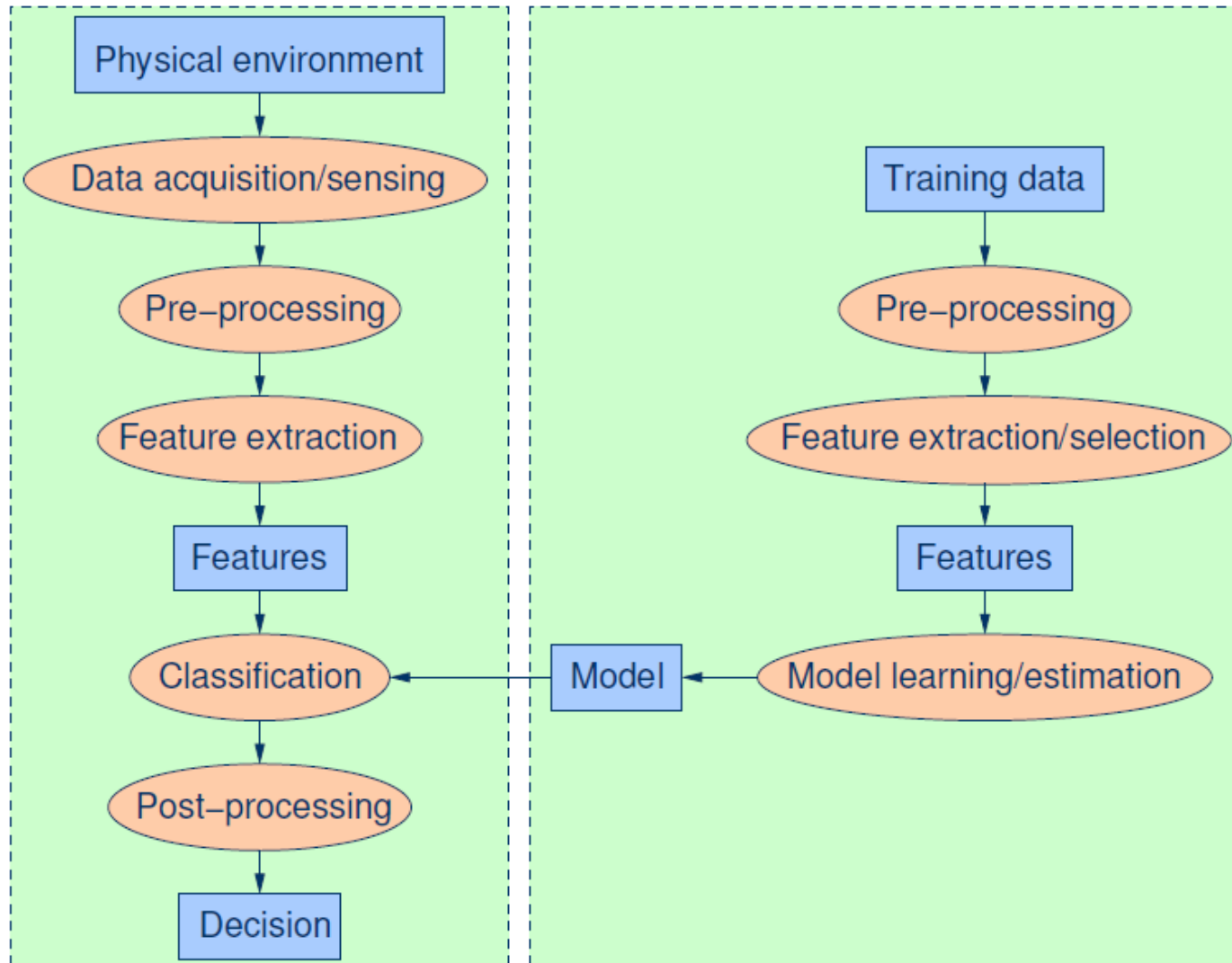
(b) 100 sample points

Today's Topics

- An example
- *Pattern recognition systems*
- The design cycle

Pattern Recognition Systems

- Object / process diagram of a pattern recognition system



Pattern Recognition Systems (cont.)

- Data acquisition and sensing
 - Measurements of physical variables
 - Important issues
 - Bandwidth
 - Resolution
 - Sensitivity
 - Distortion
 - SNR
 - Latency
 -

Pattern Recognition Systems (cont.)

- Pre-processing
 - Removal of noise in data
 - Isolation of patterns of interest from the background
- Feature extraction
 - Finding a new representation in terms of features

Pattern Recognition Systems (cont.)

- Modeling learning and estimation
 - Learning a mapping between features and pattern groups and categories
- Classification
 - Using features and learned models to assign a pattern to a category

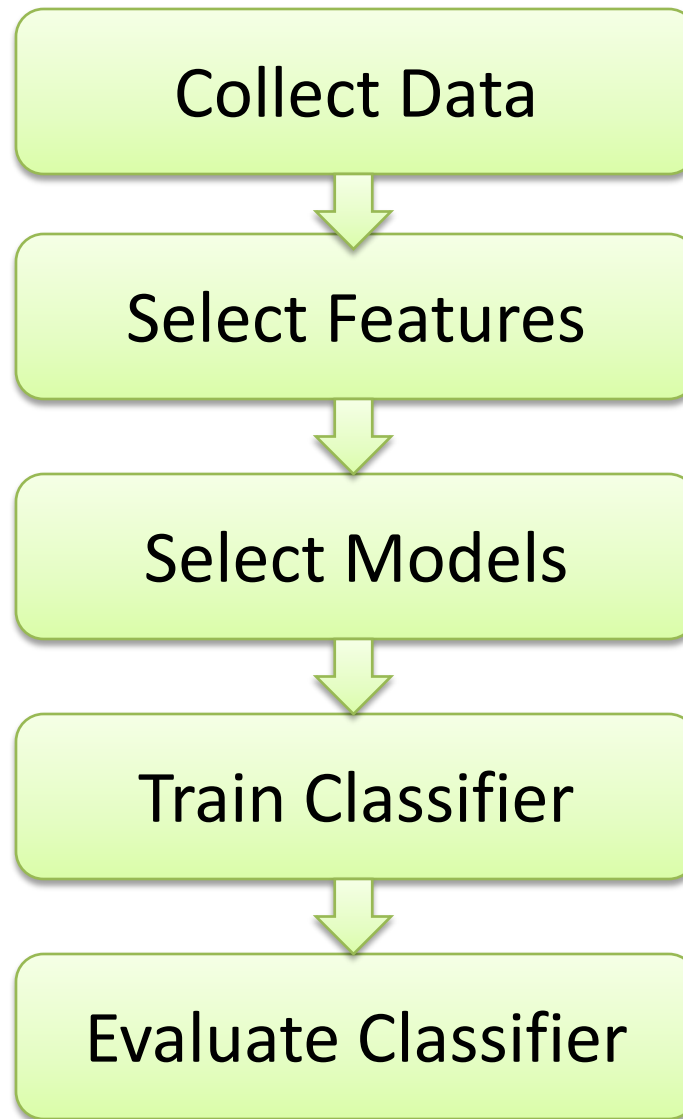
Pattern Recognition Systems (cont.)

- Post-processing
 - Evaluation of confidence in decisions
 - Exploitation of content to improve performance
 - Combination of experts

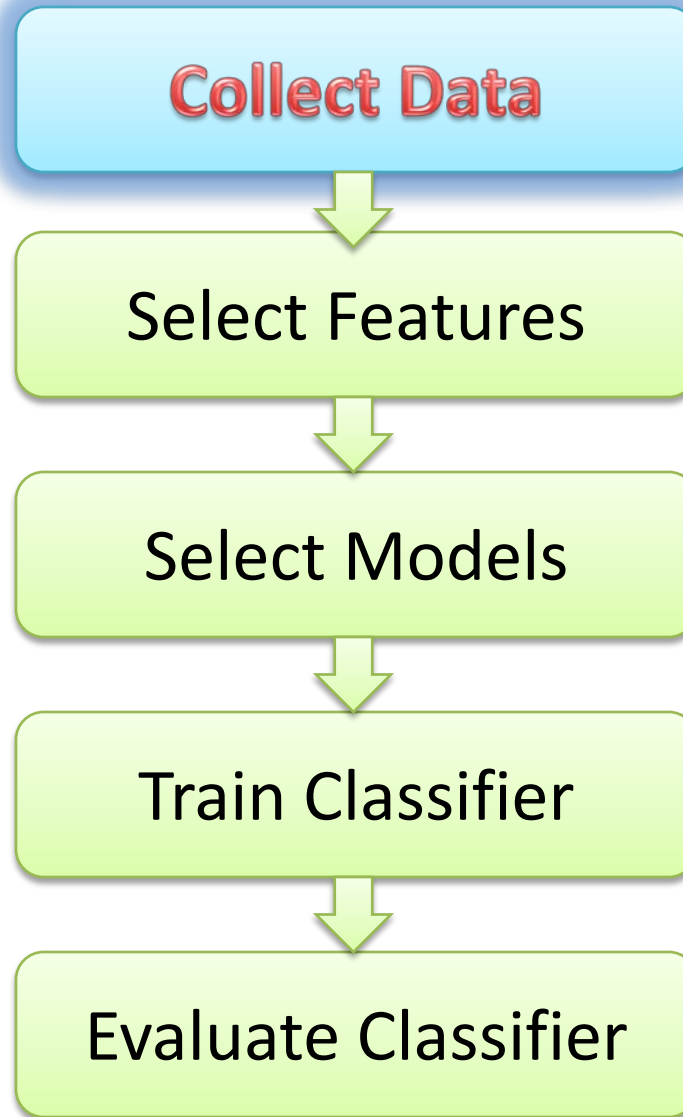
Today's Topics

- An example
- Pattern recognition systems
- *The design cycle*

The Design Cycle



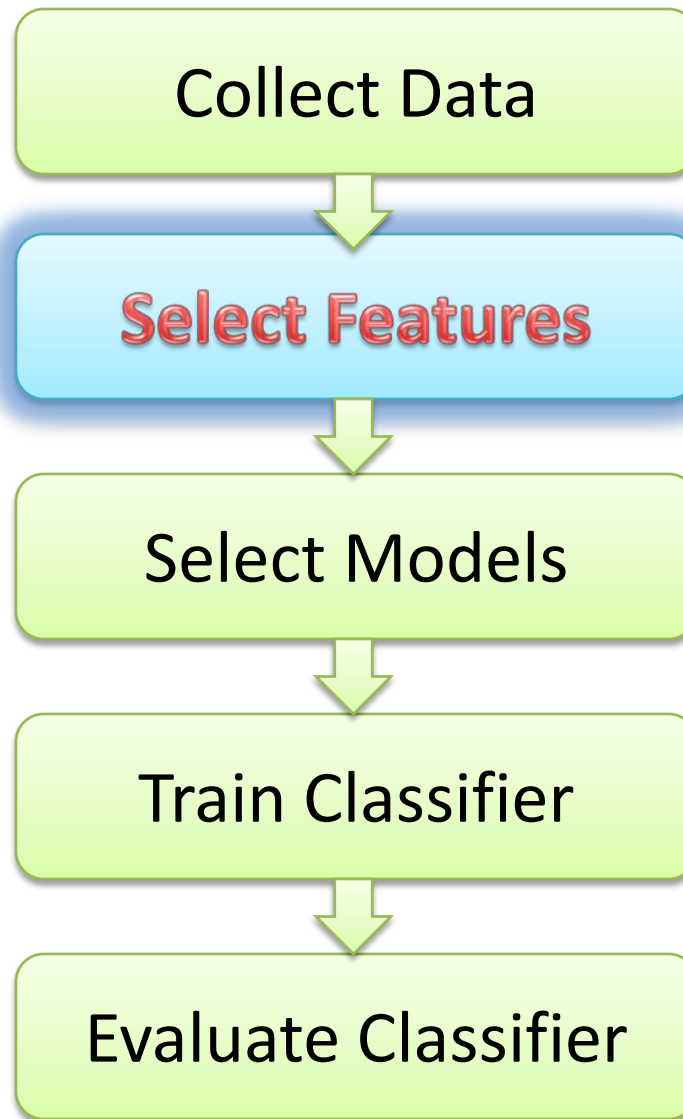
The Design Cycle



The Design Cycle (cont.)

- Data collection
 - Collecting training and testing data
 - How can we know when we have adequately large and representative set of samples?

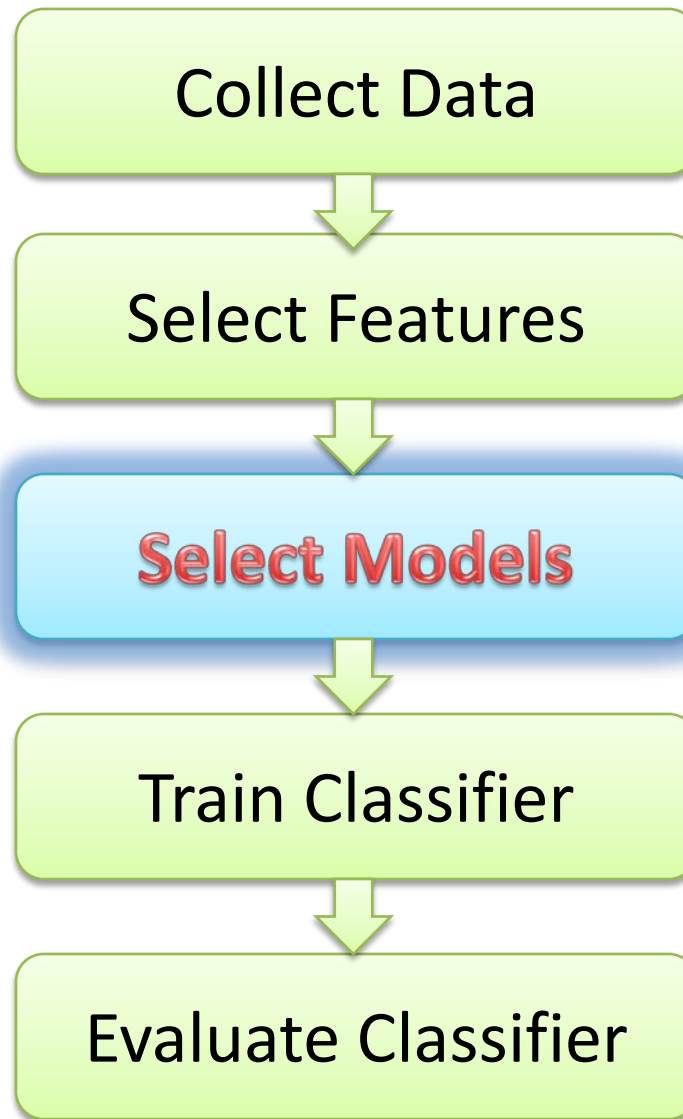
The Design Cycle



The Design Cycle (cont.)

- Feature selection
 - Domain dependence and prior information
 - Computational cost and feasibility
 - Discriminative features
 - Similar values for similar patterns
 - Different values for different patterns.
 - Invariant features with respect to translation, rotation and scale
 - Robust features with respect to occlusion, distortion, deformation, and variations in environment.

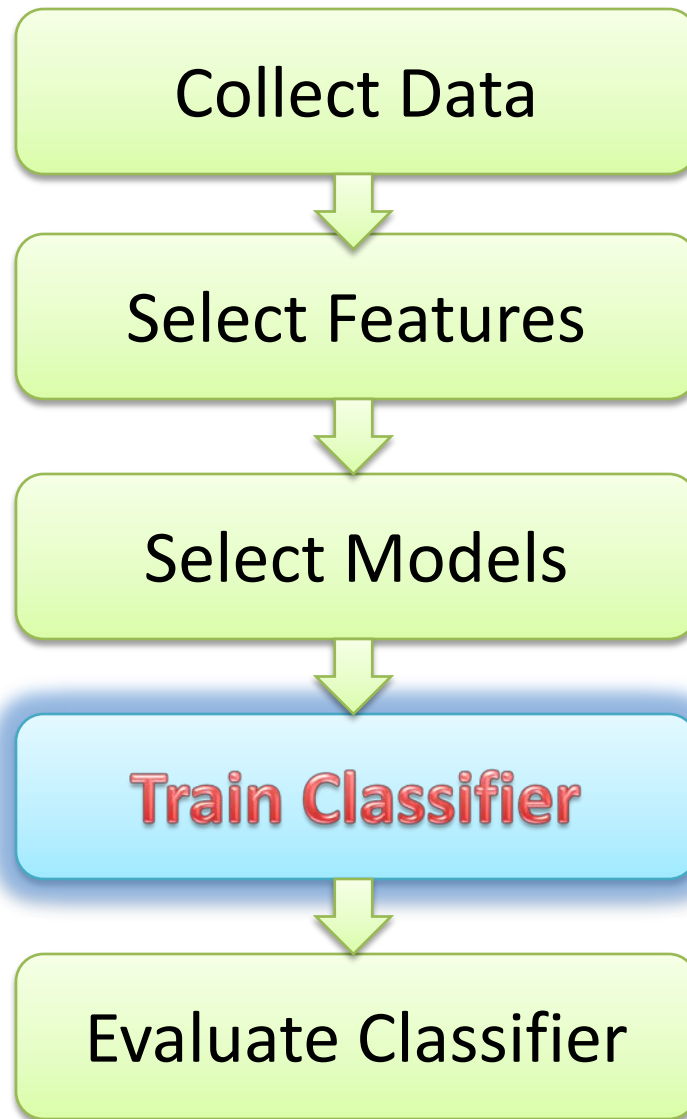
The Design Cycle



The Design Cycle (cont.)

- Model selection
 - Domain dependence and prior information
 - Definition of design criteria
 - Parametric vs. non-parametric models
 - Handling of missing features
 - Computational complexity
 - Types of models: templates, decision-theoretic or statistical, syntactic or structural, neural, and hybrid
 - How can we know how close we are to the true model underlying the patterns?

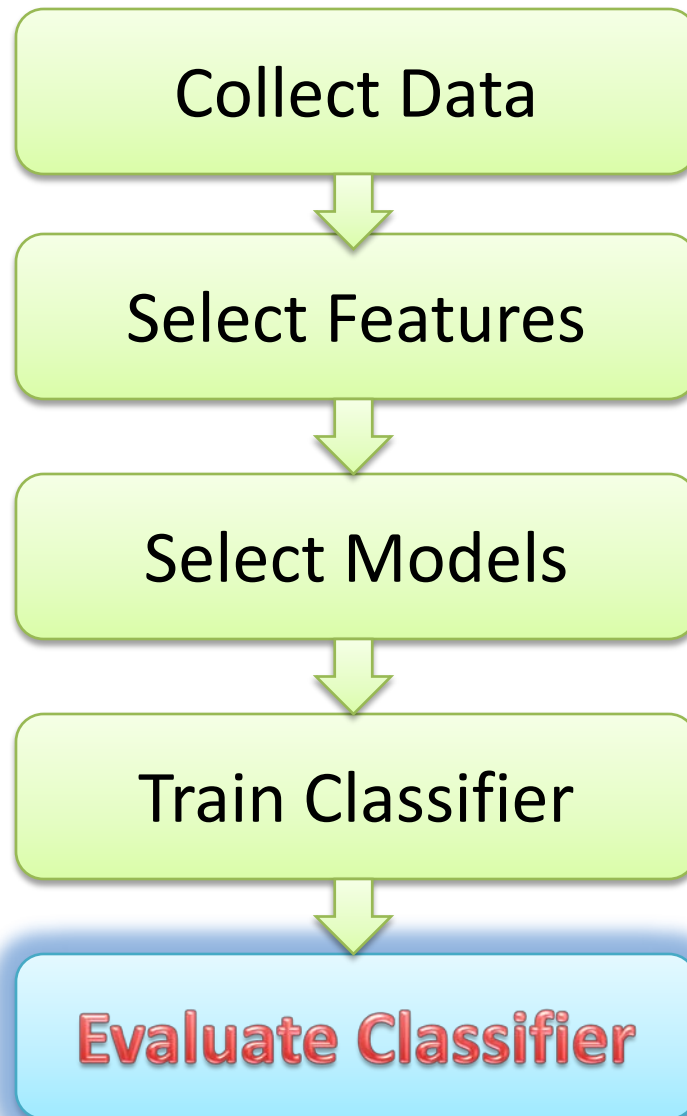
The Design Cycle



The Design Cycle (cont.)

- Training
 - How can we learn the rule from data?
 - 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
 - Reinforcement learning: no desired category is given but the teacher provides feedback to the system such as the decision is right or wrong.

The Design Cycle



The Design Cycle (cont.)

- Evaluation
 - How can we estimate the performance with training samples?
 - How can we predict the performance with future data?
 - Problems of overfitting and generalization.

Summary

- Pattern recognition techniques find applications in many areas: machine learning, statistics, mathematics, computer science, biology, etc.
- There are many sub-problems in the design process
- Many of these problems can indeed be solved
- More complex learning, searching and optimization algorithms are developed with advances in computer technology
- There remain many fascinating unsolved problems