

Introduction to Pattern Recognition

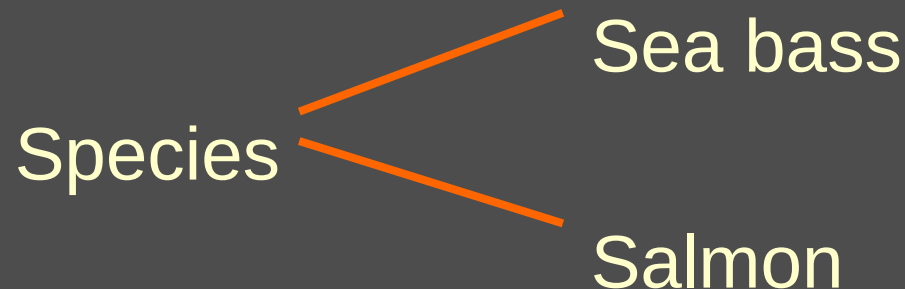
- Machine Perception
- An Example
- Pattern Recognition Systems
- The Design Cycle
- Learning and Adaptation
- Conclusion

Machine Perception

- Build a machine that can recognize patterns:
 - Speech recognition
 - Fingerprint identification
 - OCR (Optical Character Recognition)
 - DNA sequence identification

An Example

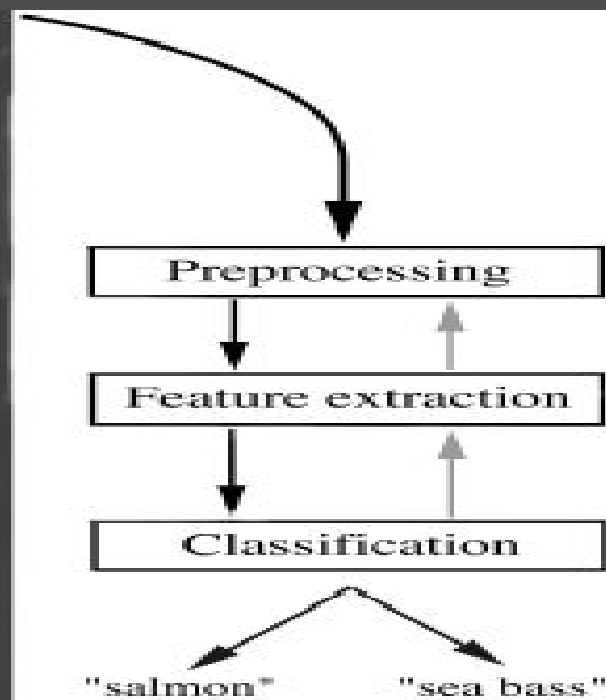
- “Sorting incoming Fish on a conveyor according to species using optical sensing”



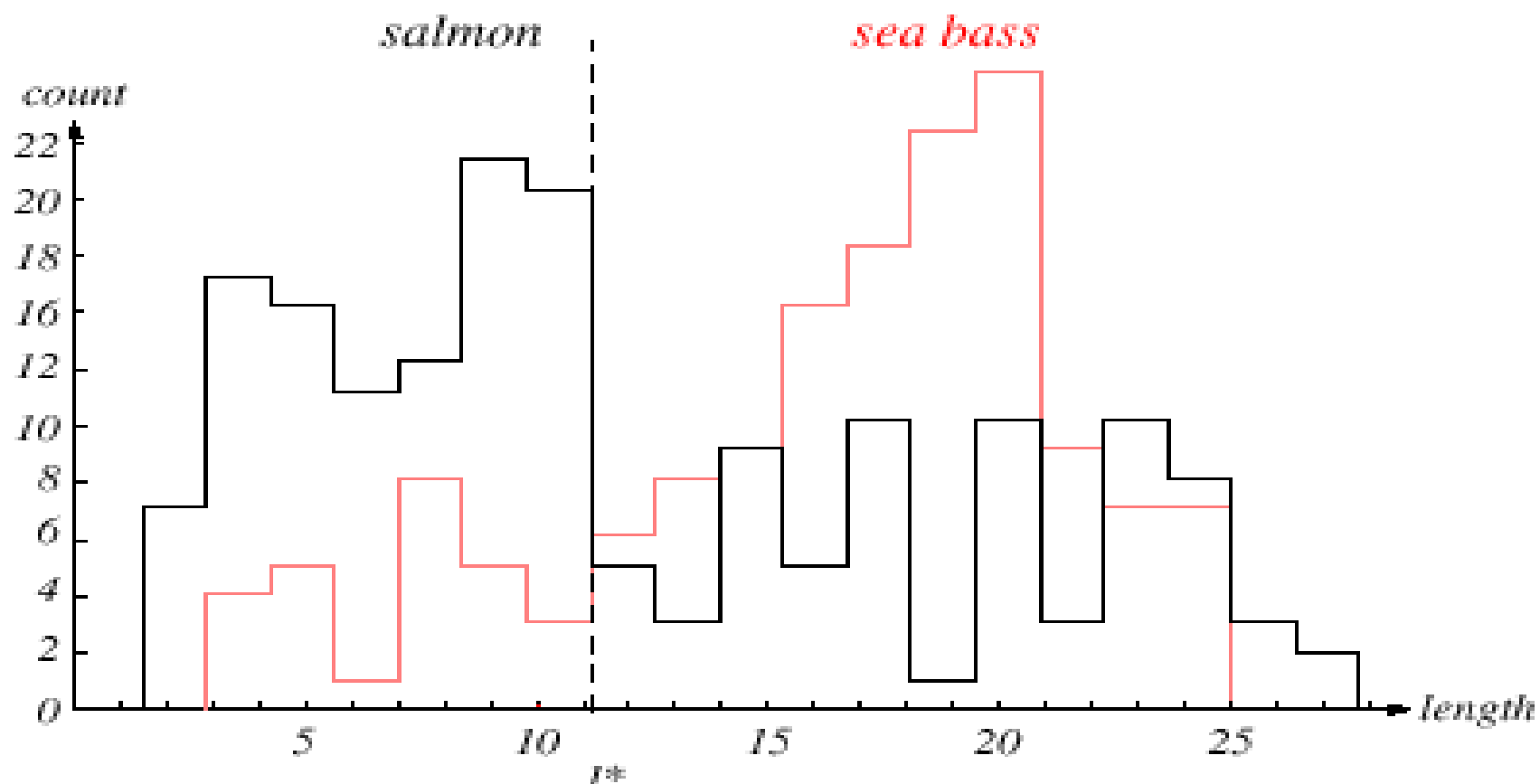
- Problem Analysis

- 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...
- This is the set of all suggested features to explore for use in our classifier!

- Preprocessing
 - Use a segmentation operation to isolate fishes from one another and from the background
- Information from a single fish is sent to a feature extractor whose purpose is to reduce the data by measuring certain features
- The features are passed to a classifier

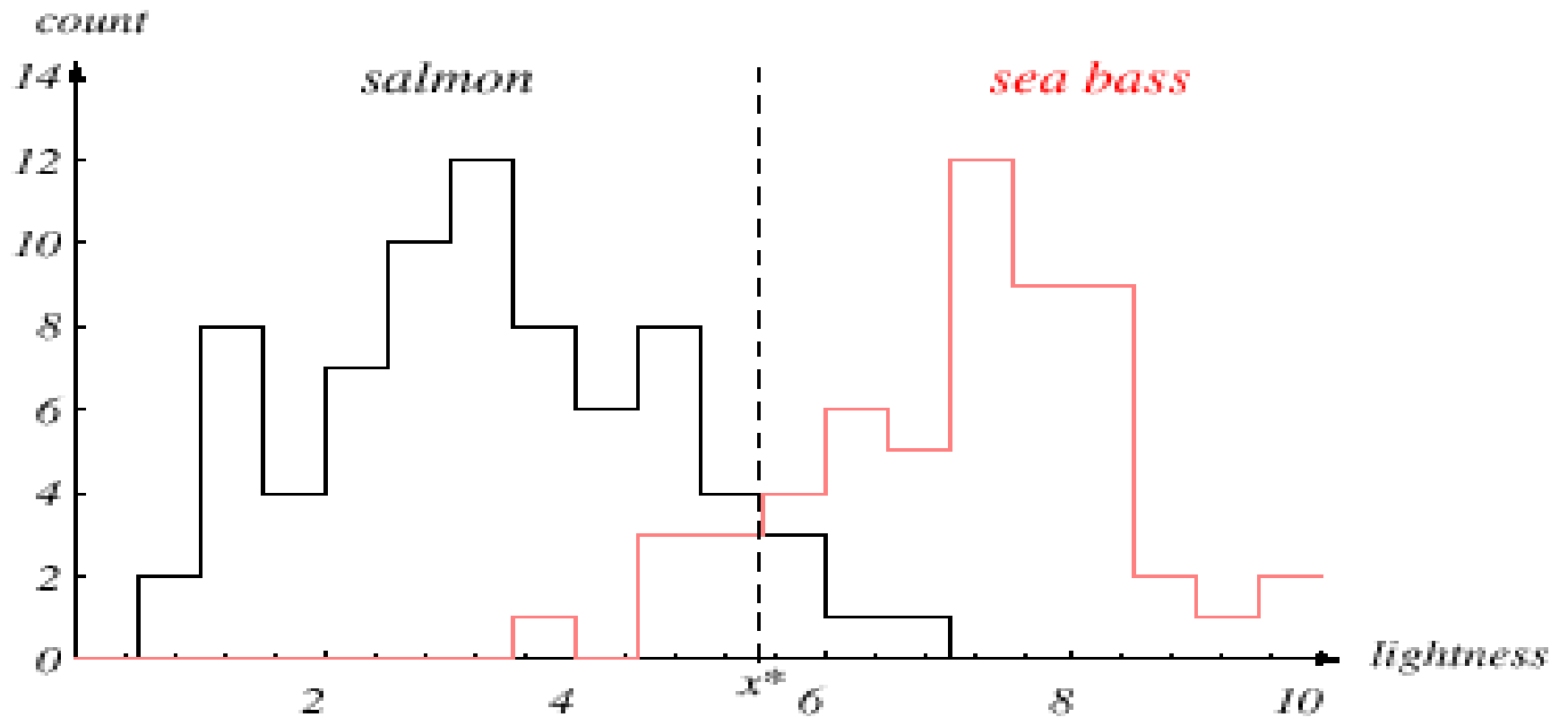


- Classification
 - Select the length of the fish as a possible feature for discrimination



The **length** is a poor feature alone!

Select the **lightness** as a possible feature.

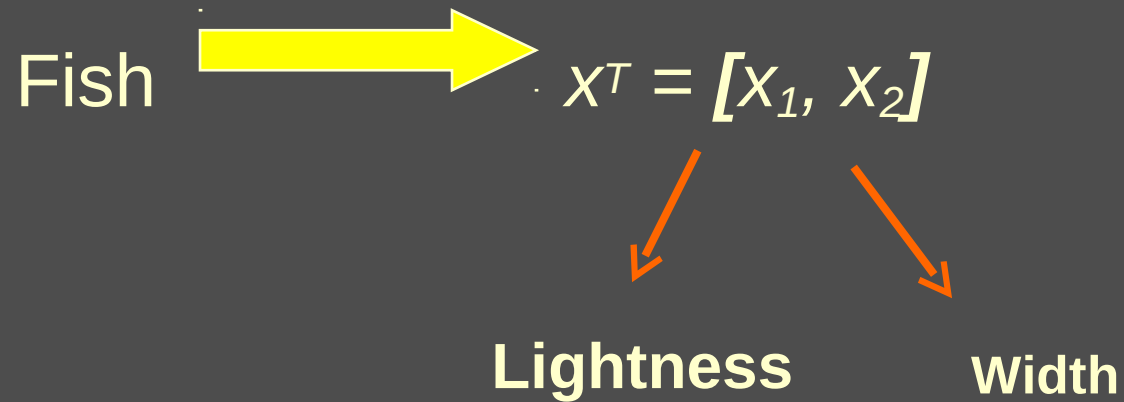


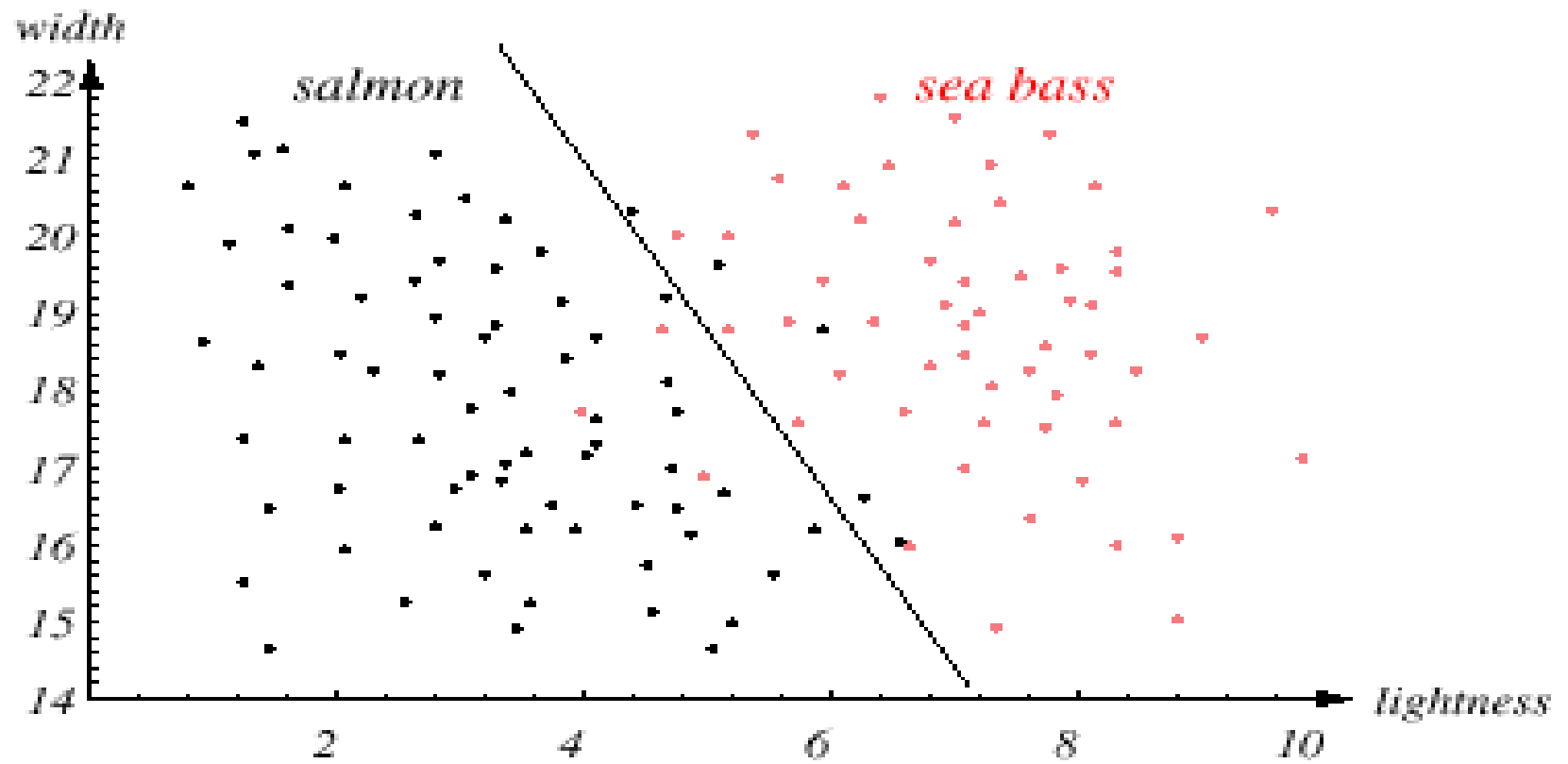
- Threshold decision boundary and cost relationship
- Move our decision boundary toward smaller values of lightness in order to minimize the cost (reduce the number of sea bass that are classified salmon!)



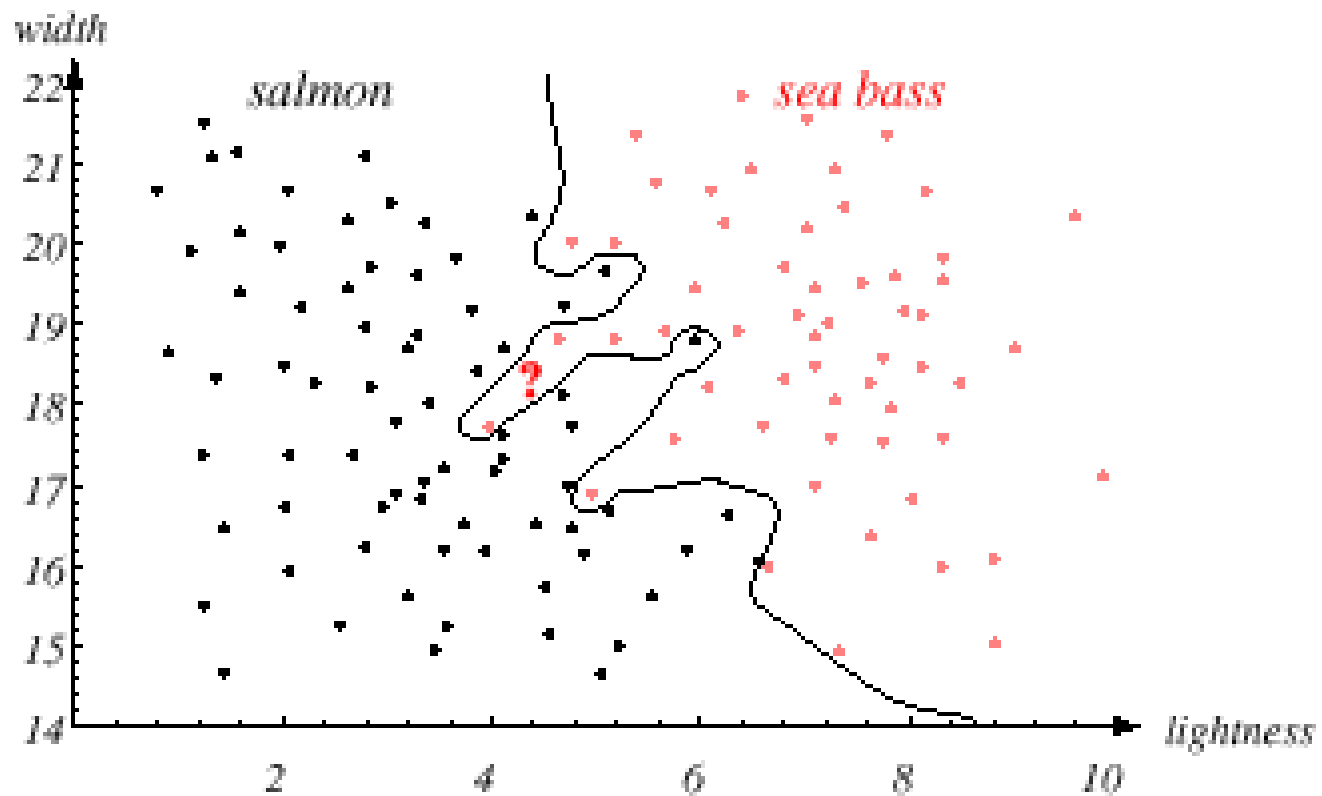
Task of decision theory

- Adopt the lightness and add the width of the fish





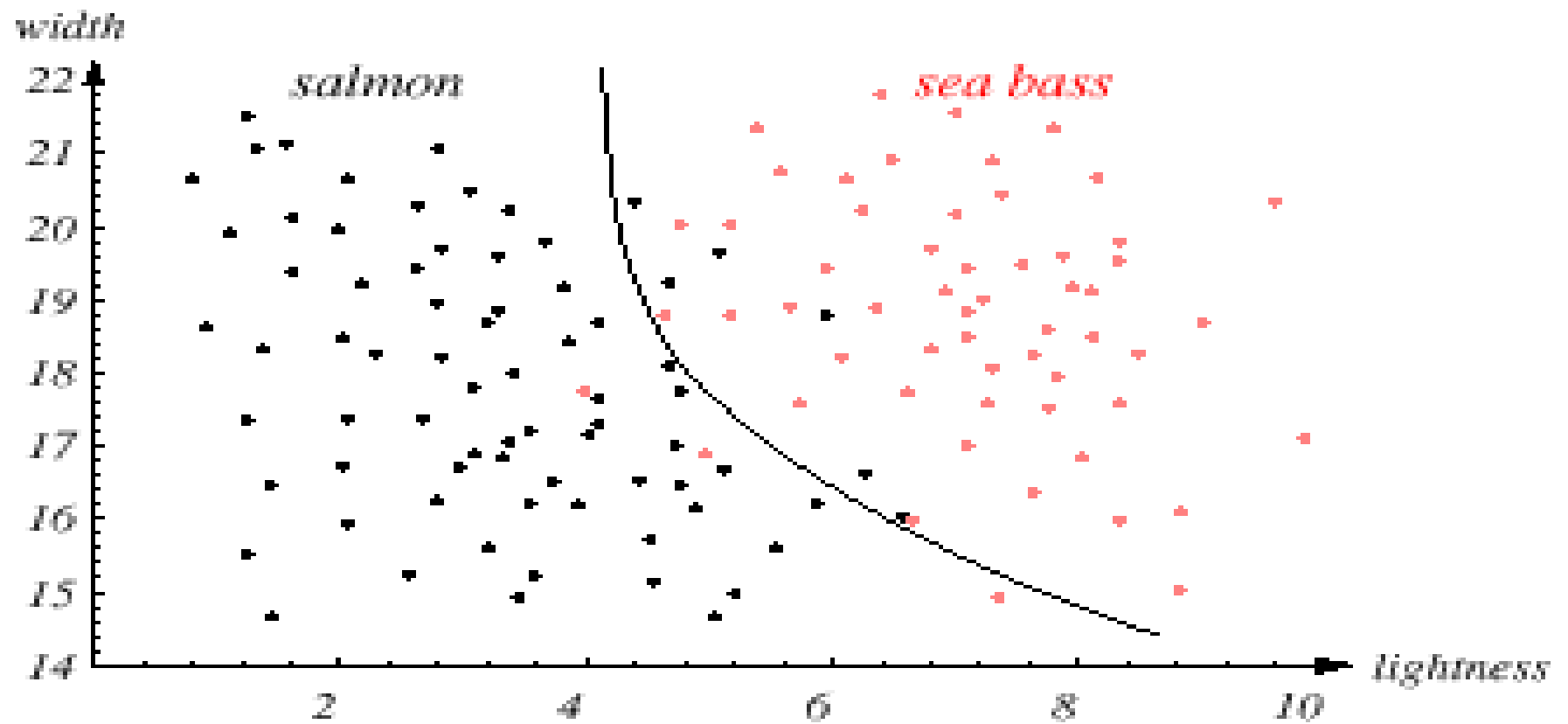
- We might add other features that are not correlated with the ones we already have. A precaution should be taken not to reduce the performance by adding such “noisy features”
- Ideally, the best decision boundary should be the one which provides an optimal performance such as in the following figure:



- However, our satisfaction is premature because the central aim of designing a classifier is to correctly classify novel input

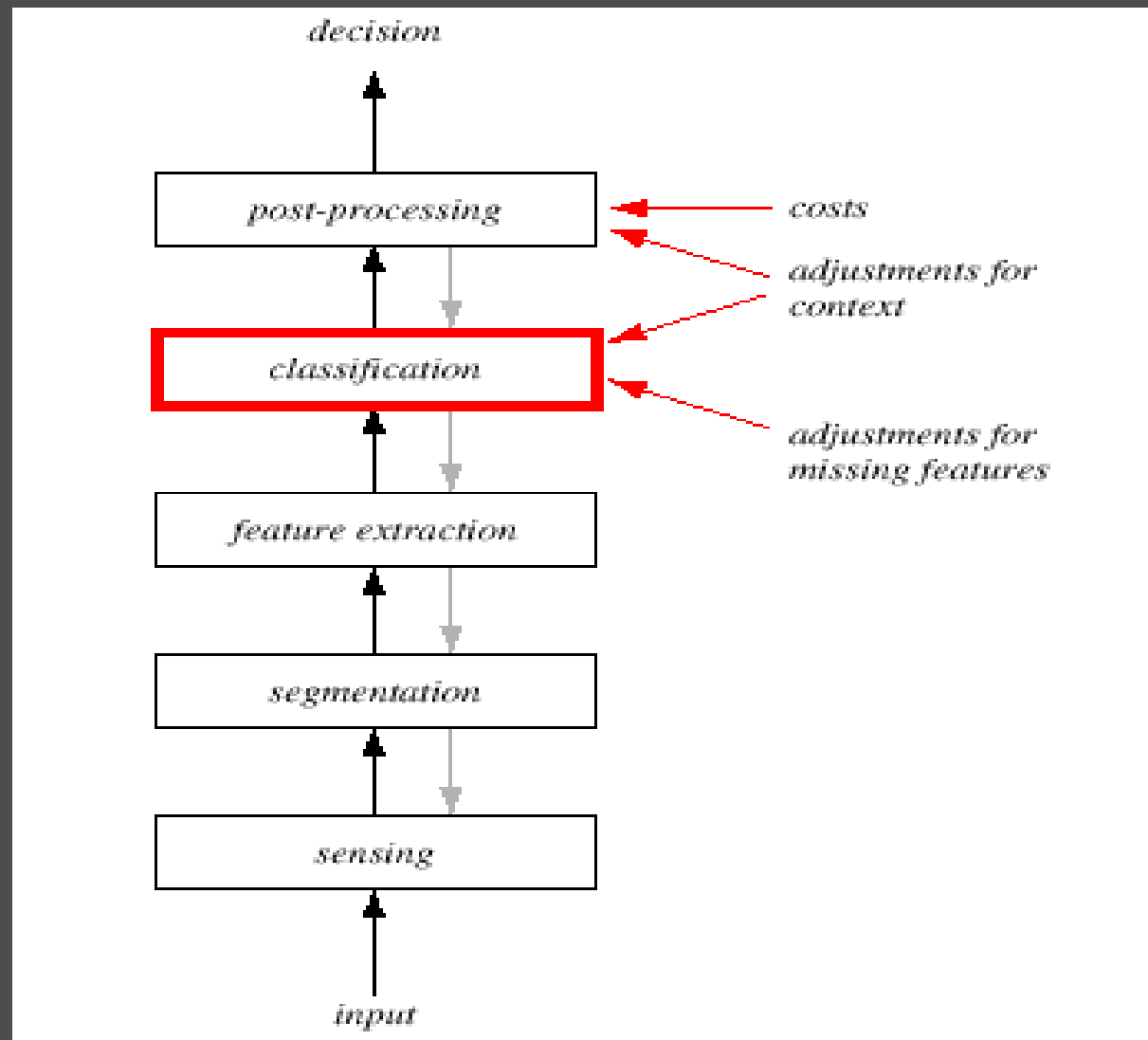


Issue of generalization!



Pattern Recognition Systems

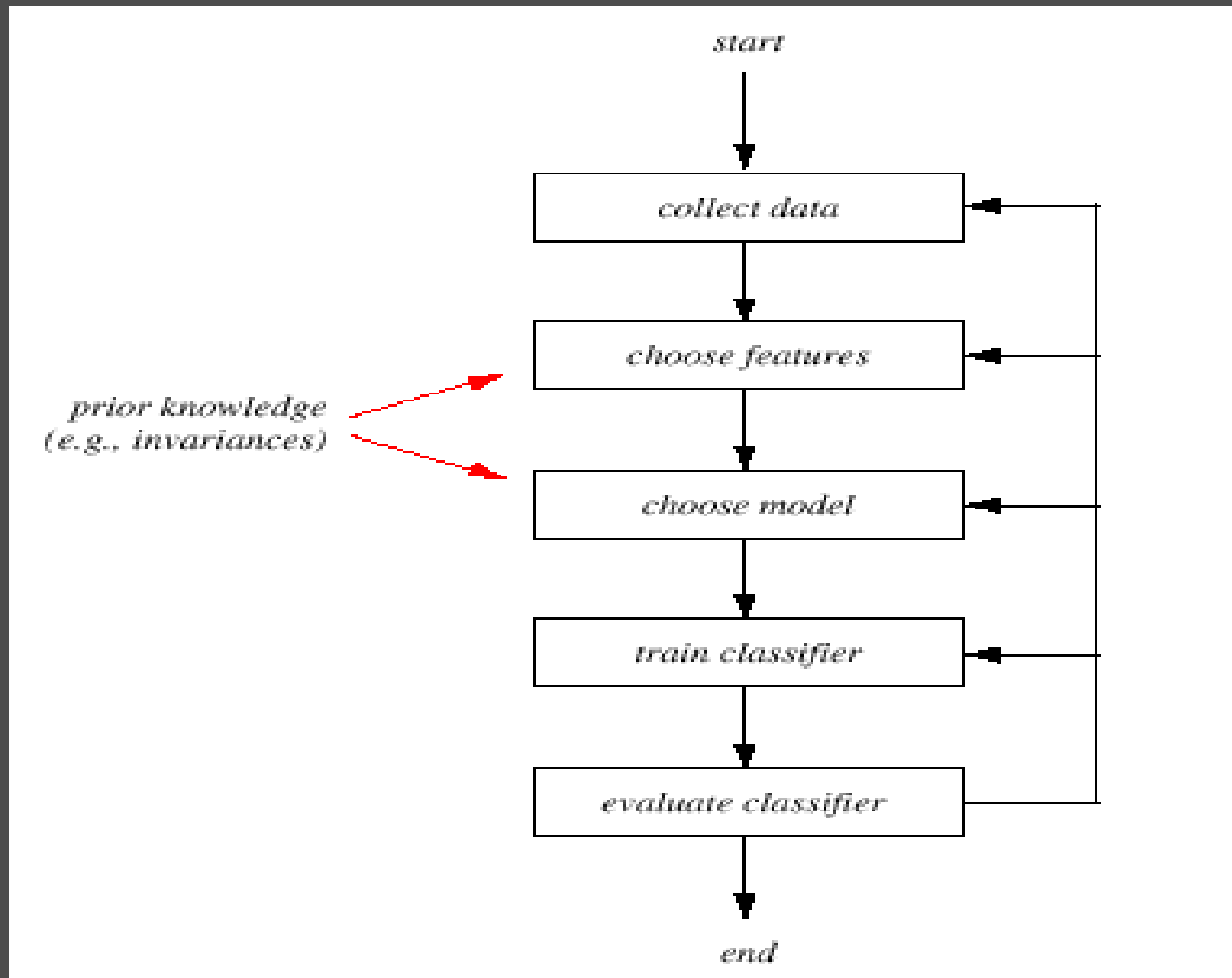
- Sensing
 - Use of a transducer (camera or microphone)
 - PR system depends of the bandwidth, the resolution sensitivity distortion of the transducer
- Segmentation and grouping
 - Patterns should be well separated and should not overlap



- 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
 - Exploit **context** input dependent information other than from the target pattern itself to improve performance

The Design Cycle

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



- 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. Simple to extract, invariant to irrelevant transformation insensitive to noise.

- Model Choice
 - Unsatisfied with the performance of our fish classifier and want to jump to another class of model

- 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 one

- Computational Complexity
 - What is the trade-off between computational ease and performance?
 - (How an algorithm scales as a function of the number of features, patterns or categories?)

Learning and Adaptation

- 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

Conclusion

- Reader seems to be overwhelmed by the number, complexity and magnitude of the sub-problems of Pattern Recognition
- Many of these sub-problems can indeed be solved
- Many fascinating unsolved problems still remain