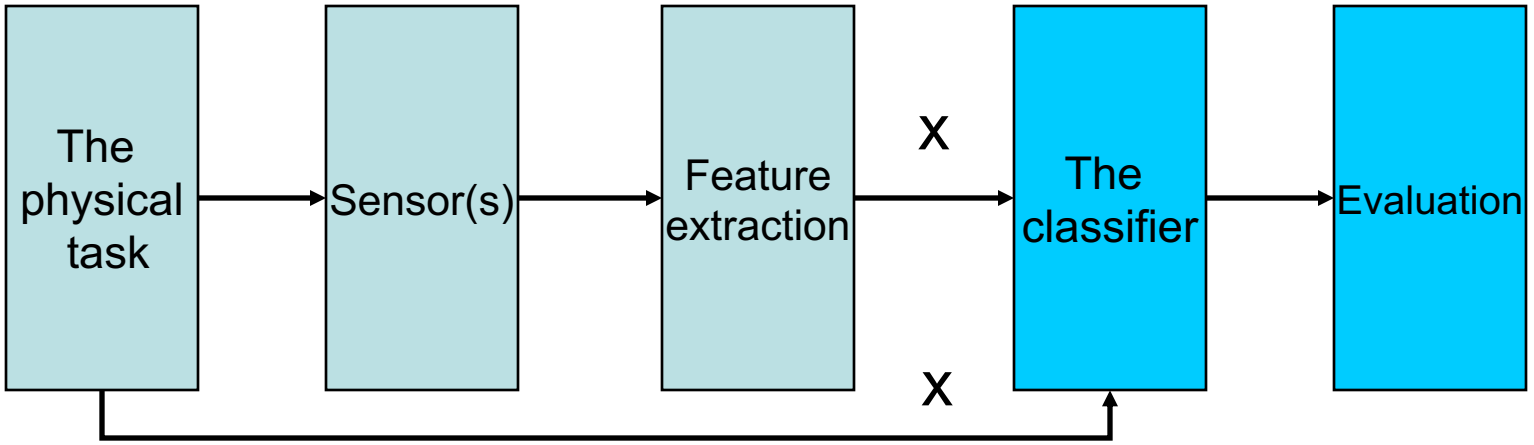
Introduction to Pattern Recognition

What is classification?

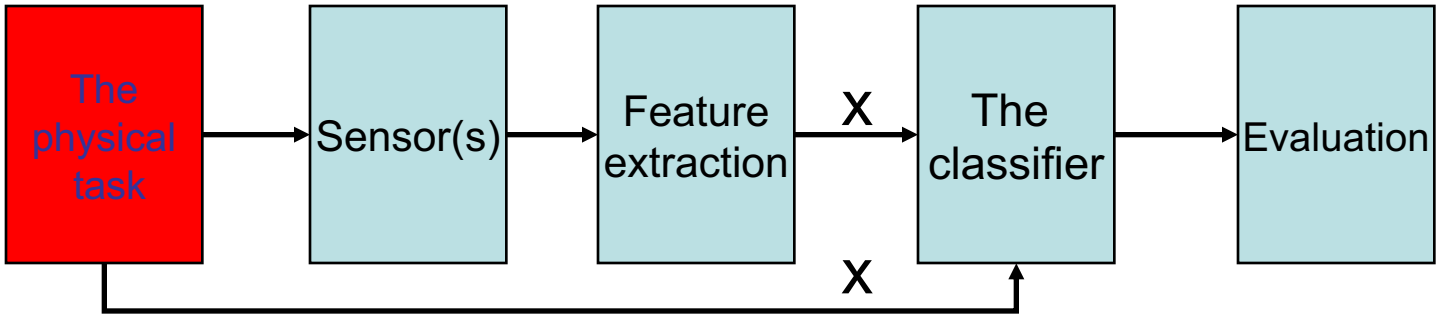
- Map a sensored signal into one of C classes.
- Human sensors and intelligence (HI) => Human classification
- Artificial (human-like) intelligence (AI) => Rule based machine classification
- Man-made sensors and math/statistics => Statistically based machine classification
-
- The last type is used in 90++% of the real world problems => Course focus is on this type

The big picture



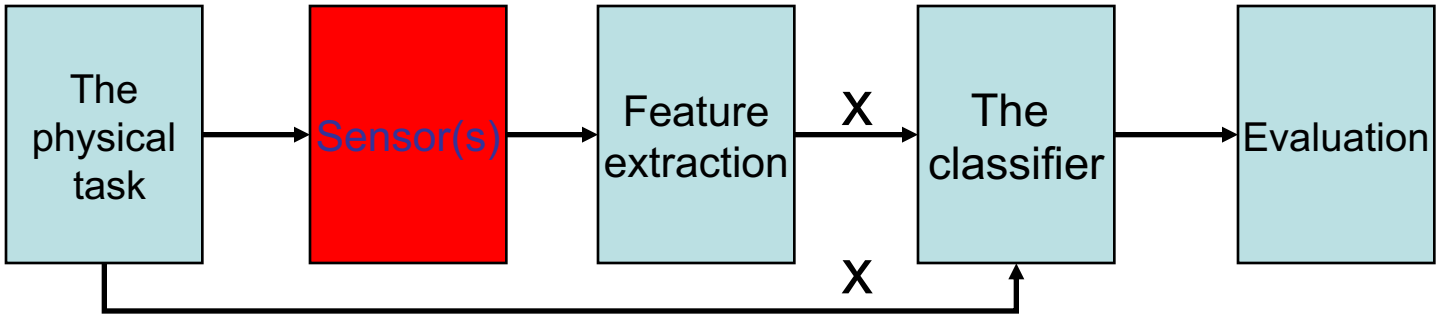
- Knowledge of all stages are mandatory!
- Weakest link determine performance.....
- Course focus : classification and evaluation

The physical task



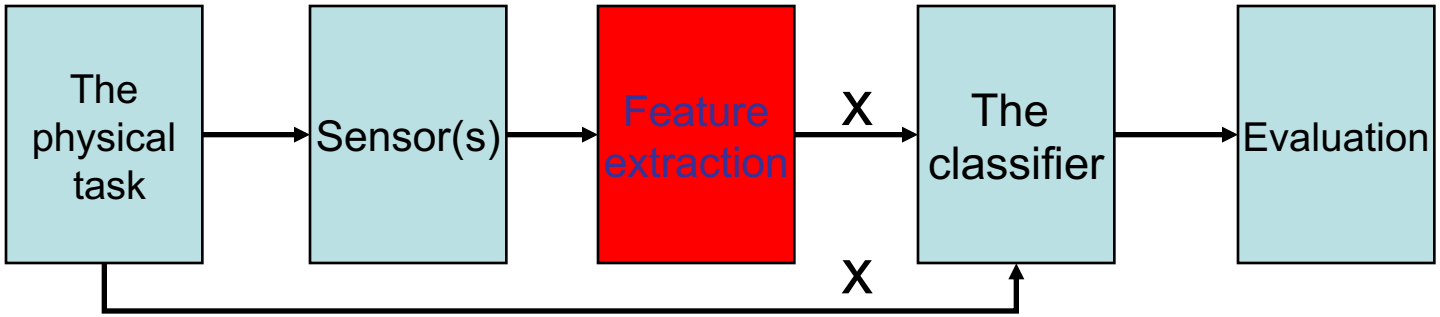
- Usually ”modeled” as a statistical source
- In some cases a ”deterministic” source
- Examples :
  - object recognition and tracking
  - medical diagnosis
  - speech recognition
  - instrumentation (temperature, pressure, wind, flow a.s.o.)
  - text recognition and parsing

Sensor(s)



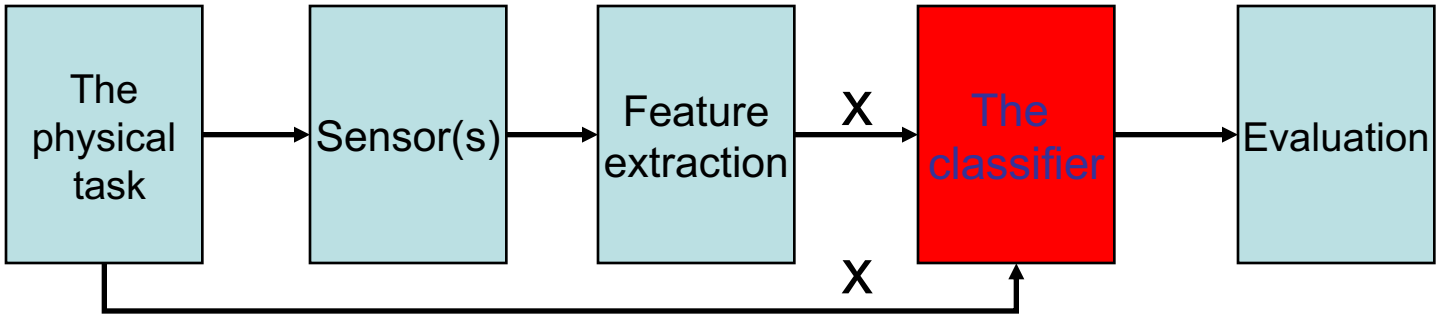
- Sensor selection and/or fusion
- Active vs. passive sensor (battery, energy, remoteness)
- Match to task (resolution, noise, aso.)
- Examples :
  - image/video
  - microphone
  - ultrasound sensors
  - antennas (electro-magnetic)
  - accelerator
  - flowmeter
  - strain gage

Feature extraction



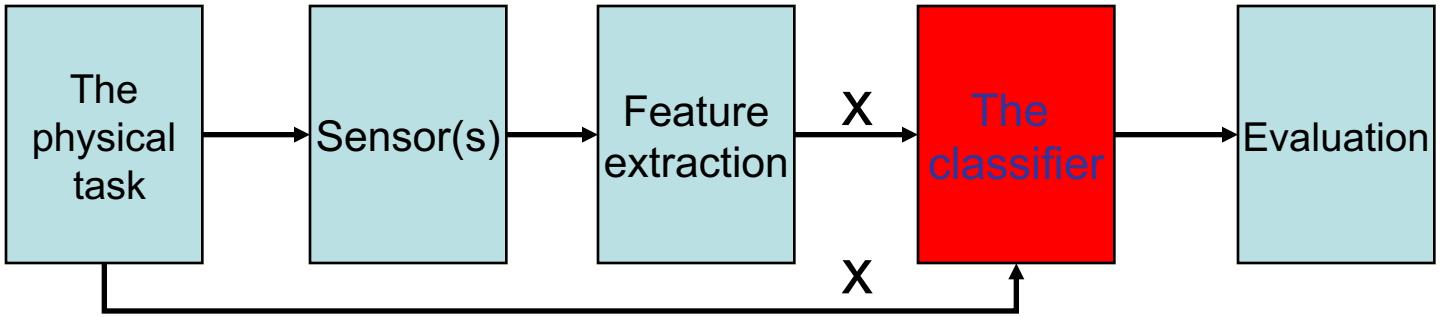
- Static or dynamic task -> feature form
- Extraction methods
- Feature quality criterias
  - class discriminability/separability
  - number of features
    - Large : separability(+), ”curse of dimensionality”(-)
    - Small : generalization(+)
- Feature selection and/or transform
- Feature correlation

The classifier - general part



- A **mathematical model** of the task given the chosen features!
  - Model mismatch is always a problem....
- A **theoretical optimal** classifier exists for any **statistical** task!
  - The Bayes Decision Rule (BDR) classifier...
  - Represents an (unknown?) upper bound on performance
  - The upper bound is **highly** dependent on :
    - Problem/task complexity
    - Choice of sensor and method for feature extraction (FE)
    - Always try out alternative FE-methods !!!

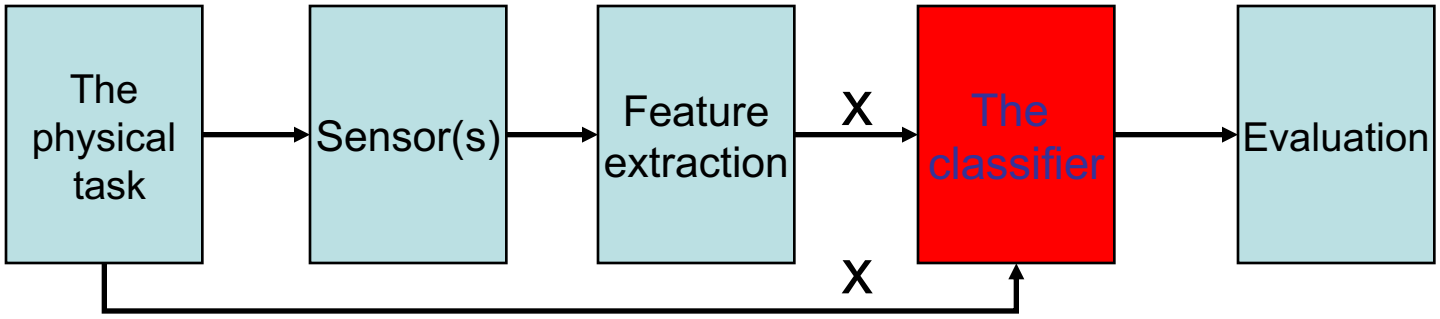
The classifier - type, structure etc.



- Two kinds of classification :
  - Recognition (x belongs to one of a given set of classes)
  - Clustering (also find the classes/clusters)

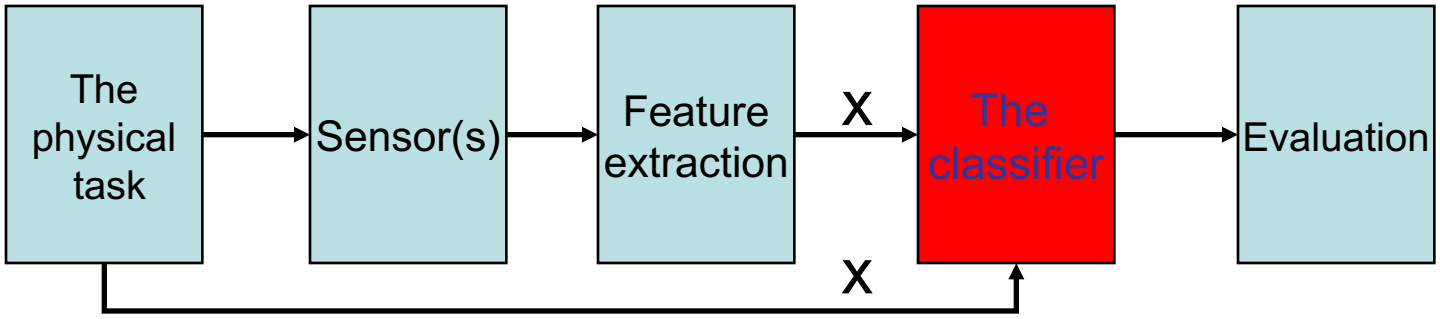
- Two kinds of classifier types :
  - Statistical or deterministic (my definition)
  - Choice not necessarily matched to kind of task!
  - Deterministic : a distortion/distance/similarity measure must be chosen
  - Statistical : the form of probabilities and densities must be chosen.

The classifier - type, structure etc.



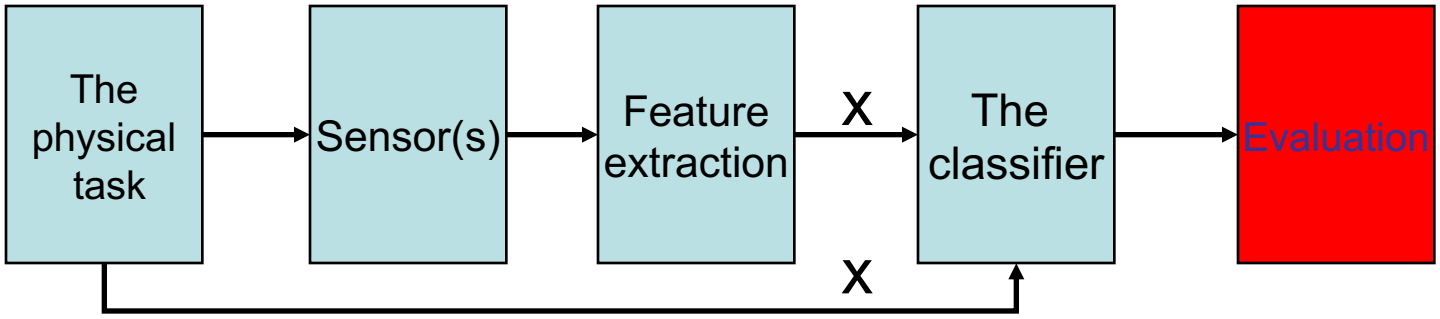
- Choices for a classifier
  - Deterministic : Template/Reference (KNN), Linear (discriminant)
  - Statistical : (Mixture) Gaussian
  - Classifier structure and complexity
    -

The classifier - training



- Classifiers must be trained!
  - Classifier parameter **estimation**
- Training set
  - Must be large enough(?) and ”representative”
  - Supervised : class labels are known
  - Unsupervised : class/cluster labels are not known
- Recognizers must be supervised trained and tested
- Clustering is in principle unsupervised training
- A good choice of training criteria is mandatory!

Evaluation



- Classifiers must be evaluated!
  - What is the real/"true" performance?
- Testset
  - Empirical error rate
  - 95% confidence error bound
- Verification of classifier hypotheses
- Small data set problem
  - "Leave-one-out technique"

Lecture content - 1

- Basic statistics and notation
- Class versus decision borders and regions
- The theoretical optimum BDR-classifier
- The practical suboptimal BDR-classifier
  - Plug in MAP-rule
  - Single Gaussian and Mixture Gaussian (GMM) cases
- The linear classifier
- The template based classifier
  - Distance metrics
  - Decision rules
- State of the art classifiers

Lecture content - 2

- Design/training of classifiers
- Plug in MAP and Maximum Likelihood (ML) training
  - Single Gaussian and GMM cases
- The linear classifier
  - The Minimum square error training and squashing
- The template based classifier
  - Template design using clustering

- Evaluation
  - Generalization
  - Error rate estimation and significance
  - The confusion matrix
  - Leave-one-out strategy for small data sets