Pattern Recognition

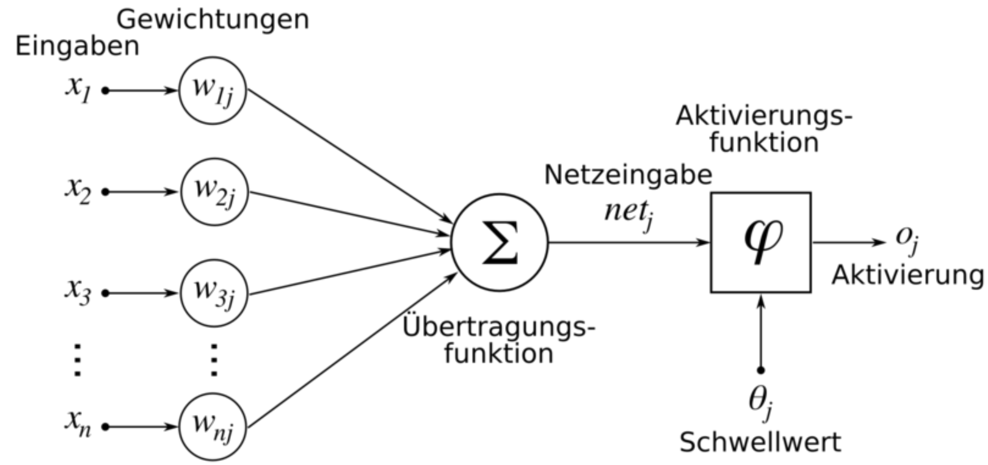
Summary

Topics

1. Neuron
2. Neural Networks
3. Clustering, SOM
4. Probability Theory
5. Bayesian Decision Theory
6. Maximum Likelihood Principle
7. Discriminative Learning
8. Support Vector Machines
9. Hinge Loss
10. Kernel-PCA
11. AdaBoost

Neuron

Neuron-model, geometric interpretation, linear classifier, weights, bias etc.



Simple Boolean functions

Learning:

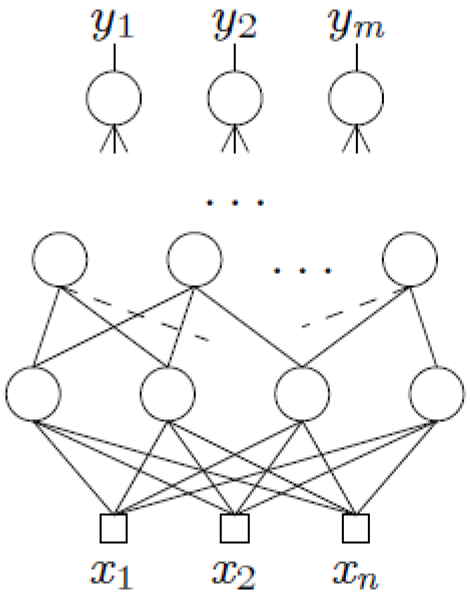
1. Perceptron algorithm: reduction to the system of linear inequalities, the algorithm, convergence proof
2. Kosinec algorithm

“Other” decision rules (i.e. quadratic, polynomial etc. – see seminars).

Neural Networks

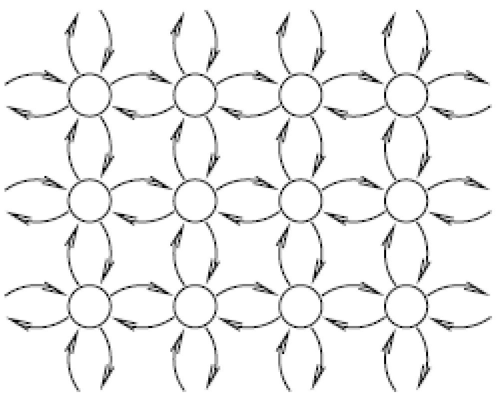
Multiclass case,
Fisher Classifier: reduction to the binary case

Feed-Forward Networks:



architecture, modeling capabilities,
learning – Error Back Propagation,
Applications: TDNN, convolutional networks

Hopfield-Networks:

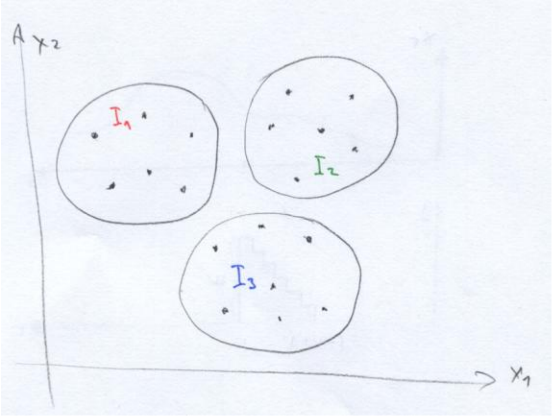


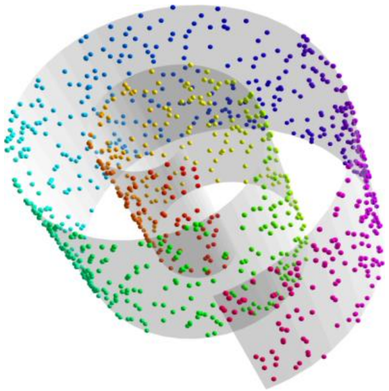
Structured output, energy, network dynamic,
external input, modeling principle

Clustering, Self-Organizing Maps

"Usual" clustering, K-Means Algrithm (batch and sequential)

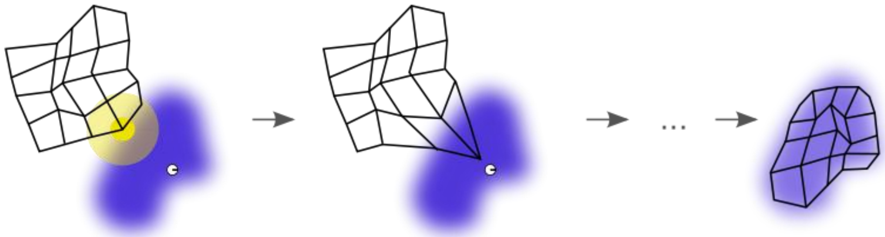
Variants: formulation without representatives



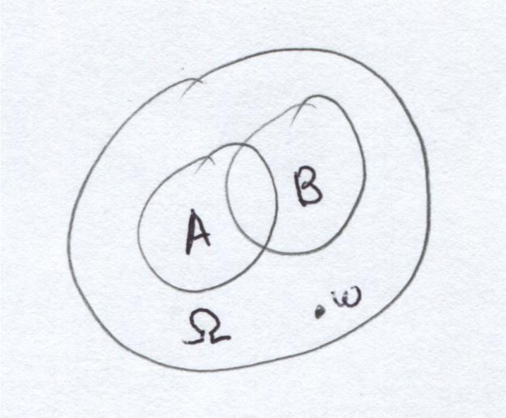


Self-Organizing Maps:
topology, neighborhood relation, network distances …

Sequential learning algorithm:



Probability Theory



Probability spaces:
(elementary) events, sigma-algebras,
practically relevant cases

Random variables:
cumulative distribution, probability distribution,
probability density, expectation

Random variables of higher dimension:
joint probability distribution (density), independence, conditional
probabilities, marginal distribution, Bayes' formula

Models for recognition:
classes and observations, recognition and learning

Bayesian Decision Theory

Decision, strategy, loss, Bayesian risk

$$R(e) = \sum_x \sum_k p(x,k) \cdot C\big(e(x), k\big) \rightarrow \min_e$$

Delta loss → Maximum A-posteriori decision

MAP for Gaussians → Linear classifier

Decision strategies with “rejection” option

“Metric” loss-functions – squared, absolute etc.

Additive loss-functions

Maximum Likelihood Principle

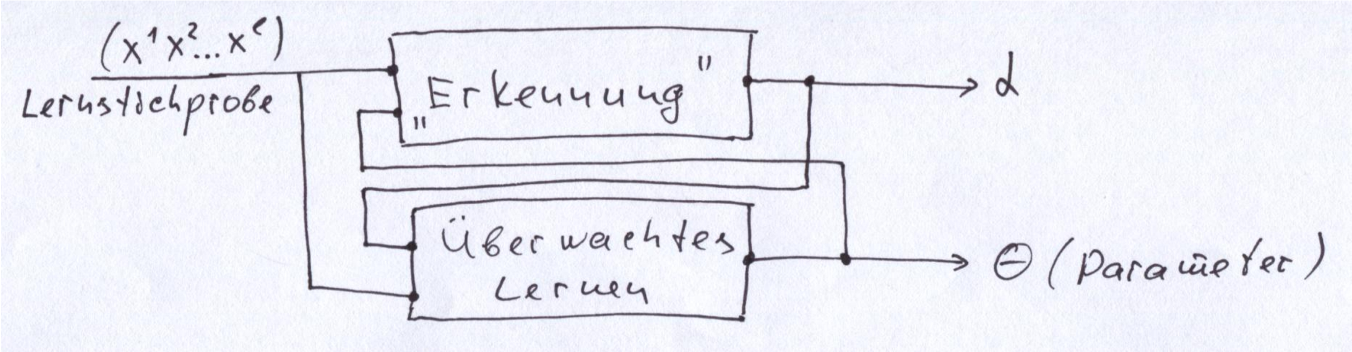
Supervised learning:

General discrete probability distributions, Shannon Lemma

Maximum Likelihood for Gaussians

Unsupervised learning:

Expectation Maximization Algorithm: derivation, the algorithm



Consistency, (un)biased estimation

Discriminative Learning

Discriminative models:

parameterized posterior probability distributions,
Maximum Conditional Likelihood

Gaussians → Logistic regression

Discriminant functions:

Gaussians → linear classifier, Empirical risk

Vapnik-Chervonenkis Dimension:

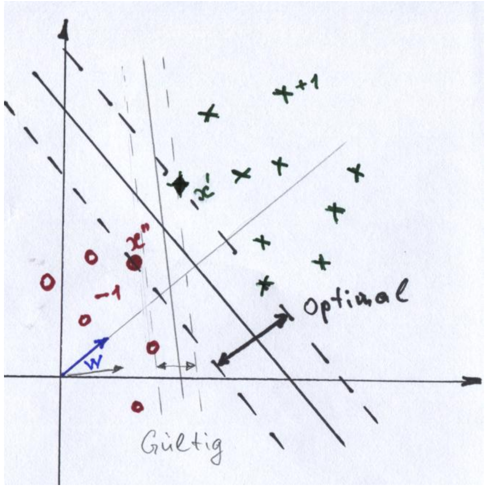
Recognition rate vs. generalization capabilities, overfitting

Support Vector Machines

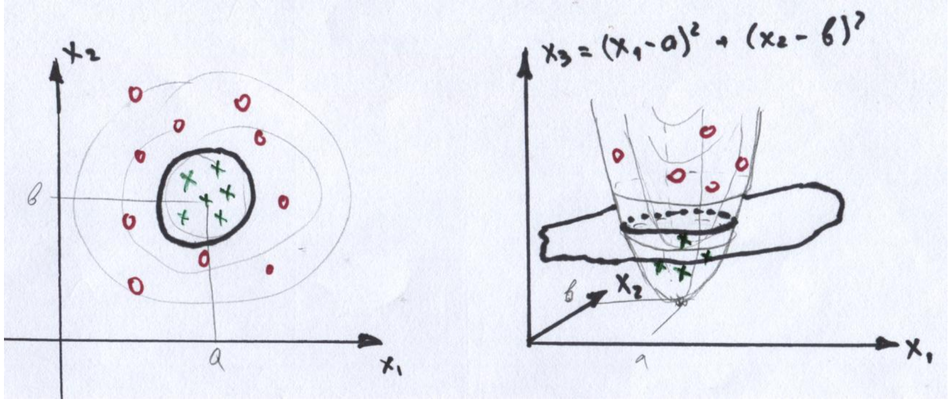
Linear SVM:

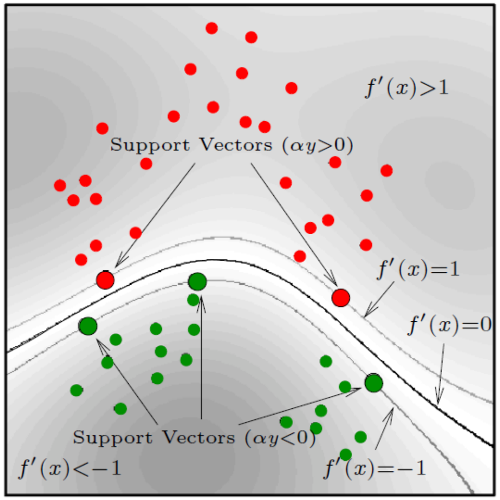
margin maximization, Lagrangian, dual variables

expressing all the stuff by scalar products



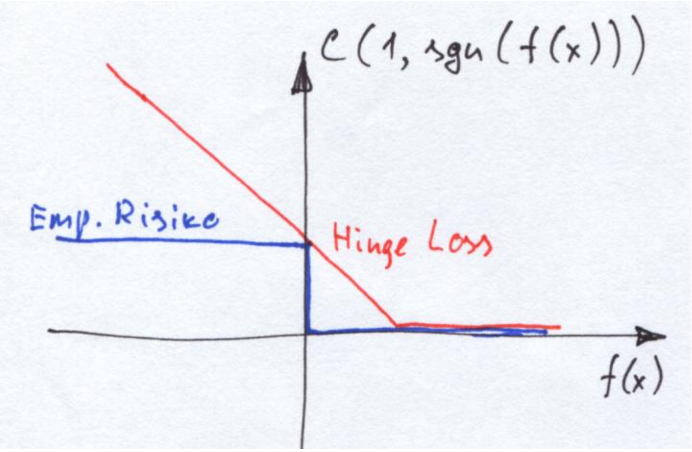
Feature spaces and kernels:





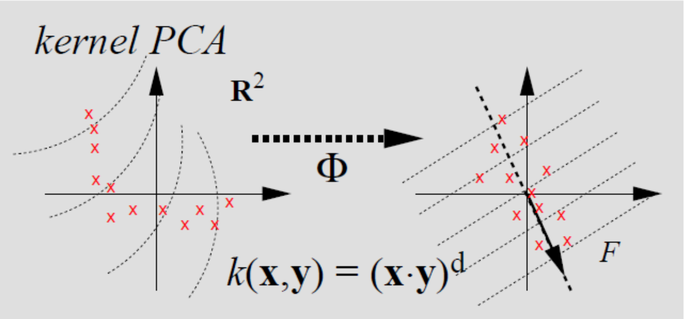
Hinge Loss, Kernel PCA

Hinge Loss:



Empirical risk minimization,
convex upper bound,
sub-gradient algorithm,
kernelization,
maximum margin vs. minimum loss,
a "unified" formulation

Kernel PCA:



PCA → expressing all by scalar products
PCA in feature space → kernels

AdaBoost

Combining weak classifiers in order to obtain a strong one

$$f(x) = \operatorname{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$$

Modeling capabilities

The algorithm:

Für $t = 1, \ldots, T$

1. Wähle (lerne) einen schwachen Klassifikator $h_t \in \mathcal{H}$ unter Berücksichtigung aktueller Gewichte $D^{(t)}$
2. Wähle α_t
3. Aktualisiere die Gewichte:

$$D^{(t+1)}(i) = \frac{D^{(t)}(i) \cdot \exp\big(-\alpha_t y_i h_t(x_i)\big)}{Z_t}$$

mit der Normierungskonstante Z_t so, dass $\sum_i D^{(t+1)}(i) = 1$.

WS 2013/2014

Computer Vision (Carsten Rother):

Combination BV(SS2013)+CV(WS2013/2014) is allowed for exams.

Machine Learning (Dmitrij Schlesinger):

Combination ME(SS2013)+ML(WS2013/2014) is not allowed.

Other courses (Holger Heidrich with others):

- Einführungspraktikum Computer Vision
- Komplexpraktikum Computer Vision
- Projektpraktikum Computer Vision
- Hauptseminar Bildanalyse

See “Image Processing: Summary” lecture and web-pages for details