



---

28.04.2017

# **INFO-4185 Grundlagen des Maschinellen Lernens**

---

Andreas Schilling



28.04.2017

Heute:

- Einführung
- Anwendungen
- Techniken



# Was ist Maschinelles Lernen?

---

28.04.2017

Lernen: Verhaltensanpassung durch

- Erwerb von Wissen und Fähigkeiten
- Das Wesentliche, die Essenz aus (möglichst wenigen) Beispieldaten extrahieren

Die Fähigkeit, aus Erfahrung zu Lernen, ist wesentlicher Teil jedes intelligenten Systems

Die Beschäftigung mit Maschinenlernen hilft, auch das menschliche Lernen besser zu verstehen



# Was ist Machine Learning?

---

28.04.2017

Voraussetzung:

- Verschiedene Situationen sind ähnlich

Frage:

- Was ist Ähnlichkeit



# Anwendungen:

---

28.04.2017

- Klassifizierung
- Regression
- Bestimmung von Parametern aus Daten
- Prädiktion
- Dimensionalitätsreduktion (Beseitigung von Redundanz, Extraktion von Wesentlichem)
- Clusteranalyse
- Modellierung der Datenentstehung
- Detektion von Ausreißern

auf der Grundlage von (Beispiel-)Daten



28.04.2017

## Automatische Unterscheidung zw. Lachs und Seebarsch

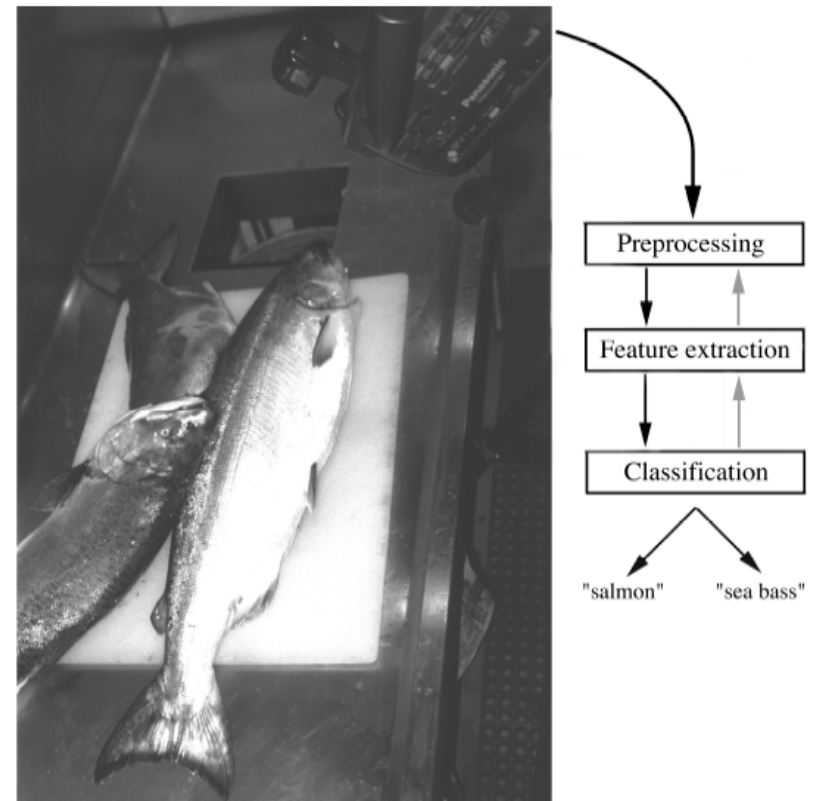


FIGURE 1.1. The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed. Next the features are extracted and finally the classification is emitted, here either "salmon" or "sea bass." Although the information flow is often chosen to be from the source to the classifier, some systems employ information flow in which earlier levels of processing can be altered based on the tentative or preliminary response in later levels (gray arrows). Yet others combine two or more stages into a unified step, such as simultaneous segmentation and feature extraction. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.



28.04.2017

## Gesichtsdetektion



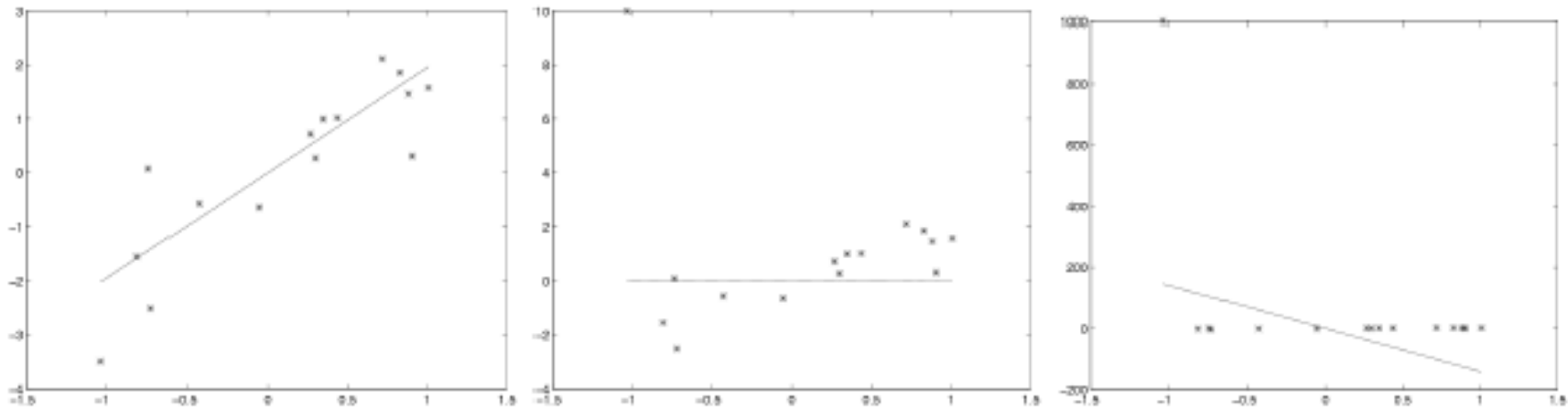
*Diplomarbeit W Kienzle 2003*

---

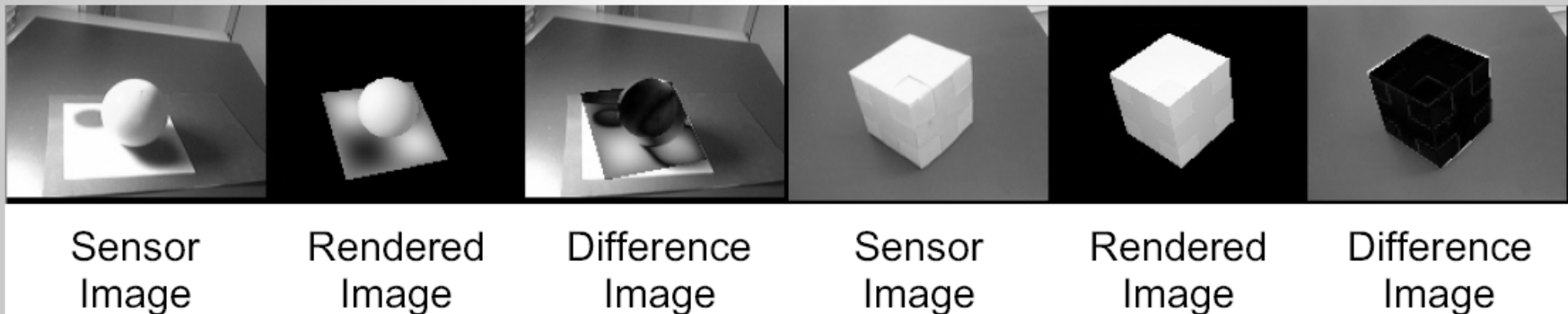
# Regression – Bestimmung von Parametern aus Daten



28.04.2017



## *Experiments: Pose & Illumination Estimation*

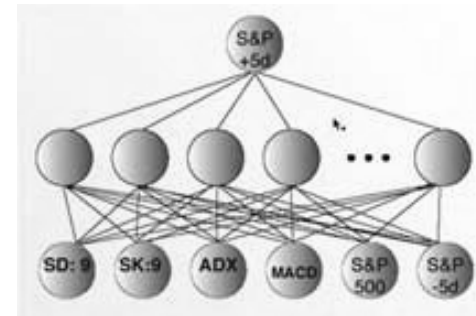






28.04.2017

## Börsenvorhersagen



### **Expert Prediction, Symbolic Learning, and Neural Networks: An Experiment on Greyhound Racing**

H. Chen<sup>1</sup>, P. Buntin, L. She, S. Sutjahjo, C. Sommer, D. Neely

### **Conclusion**

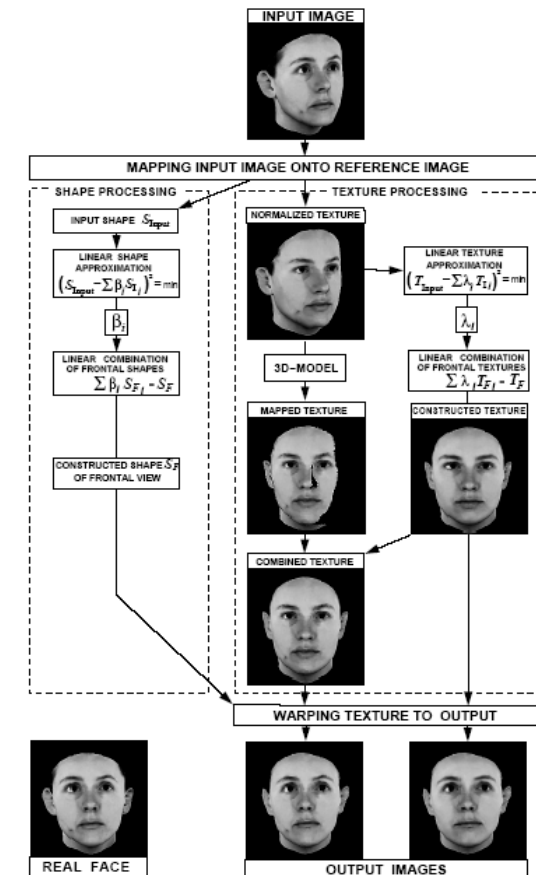
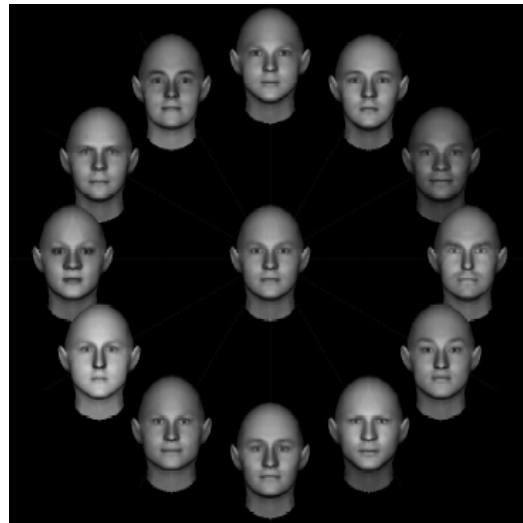
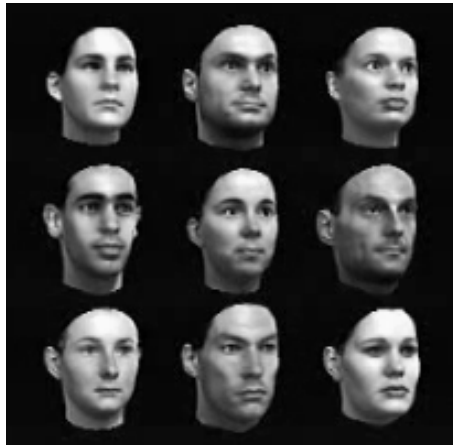
In terms of prediction accuracy and monetary payoff, both the Backpropagation network and ID3 performed better than human experts. Both algorithms

*IEEE EXPERT*, 1994



28.04.2017

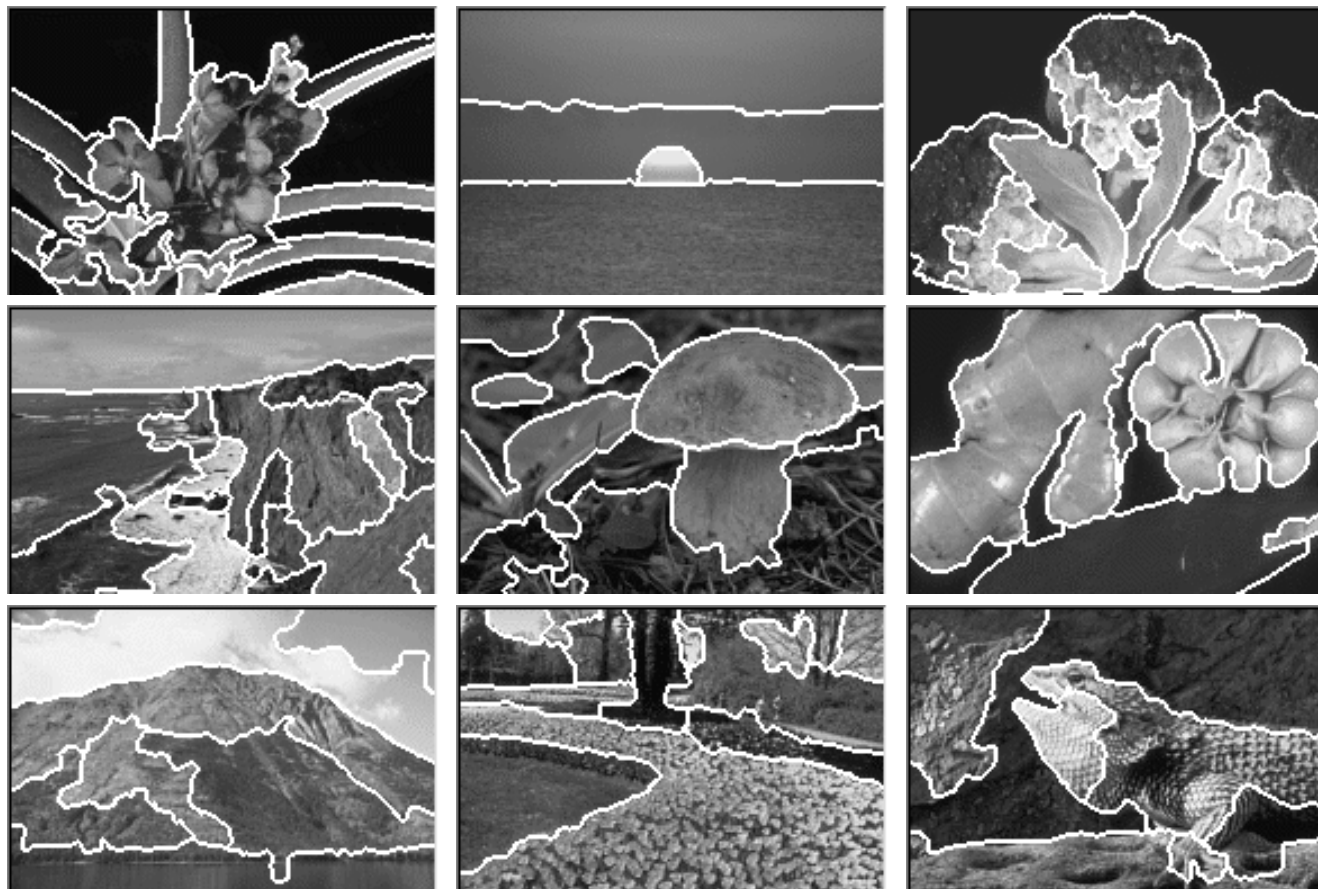
## Beispiel: Eigenfaces





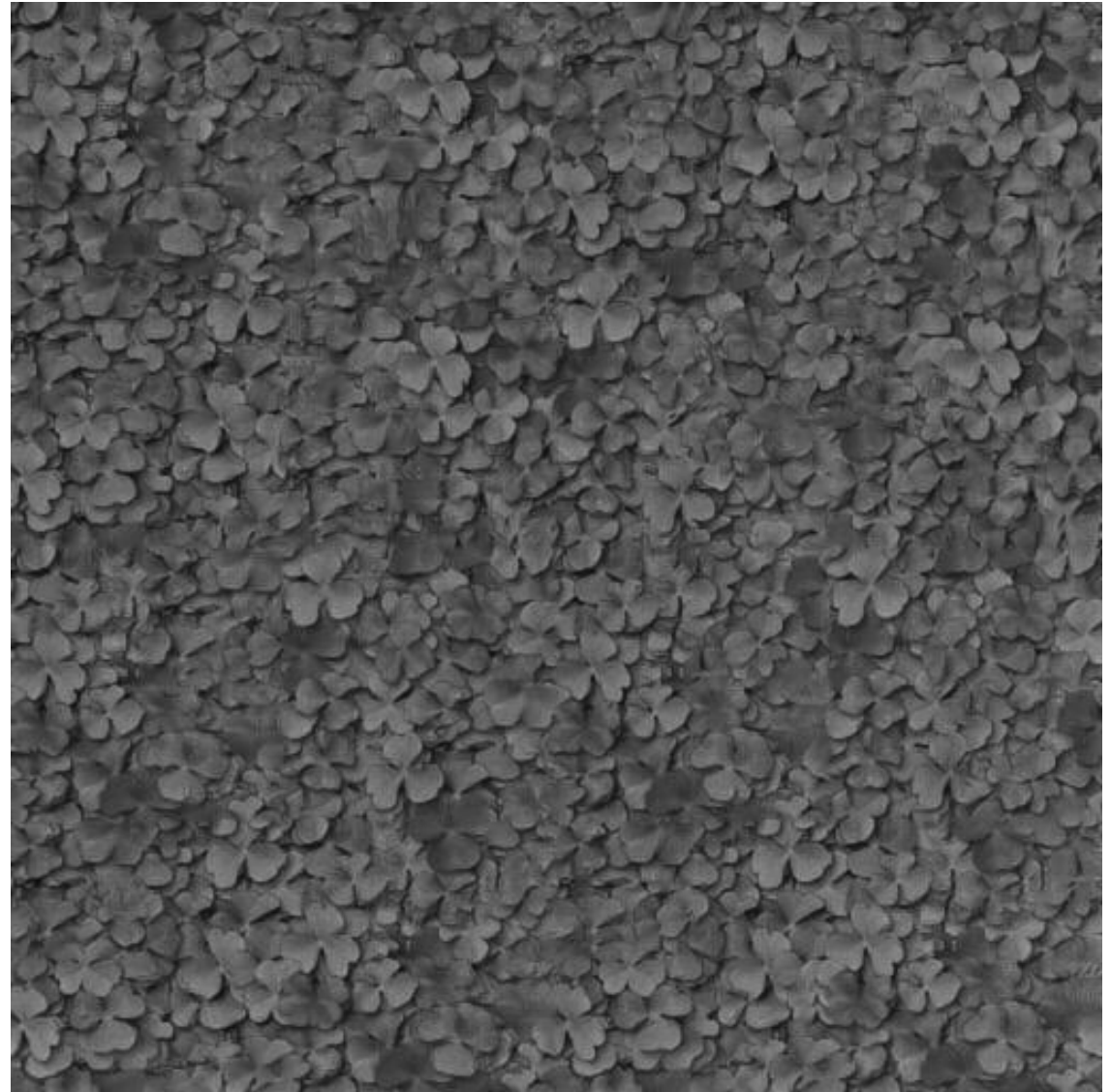
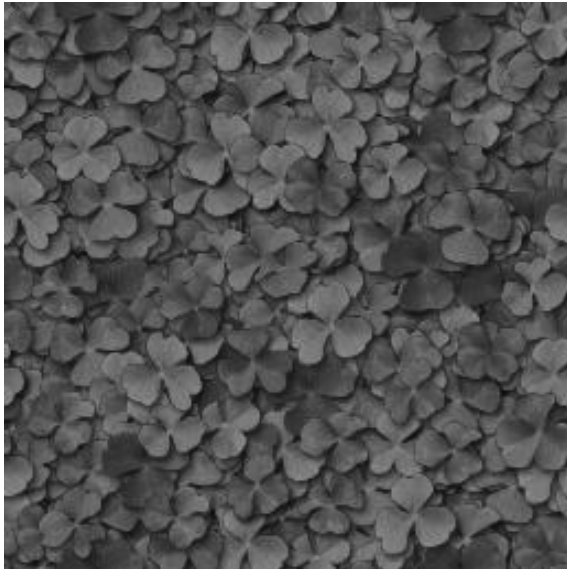
28.04.2017

## Beispiel: Segmentierung





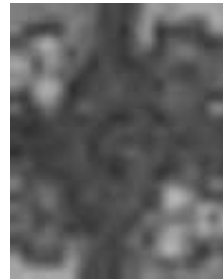
28.04.2017



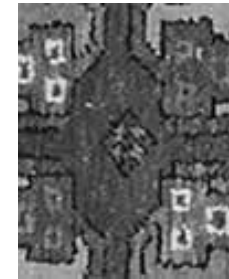


28.04.2017

Trainingsdaten:



A



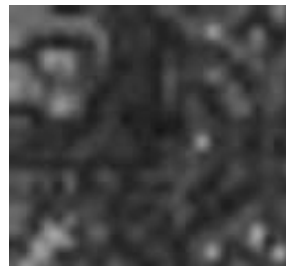
A'



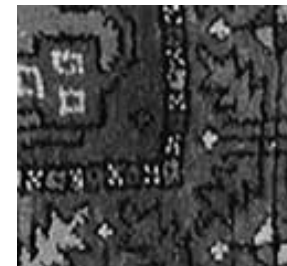
A



A'



A

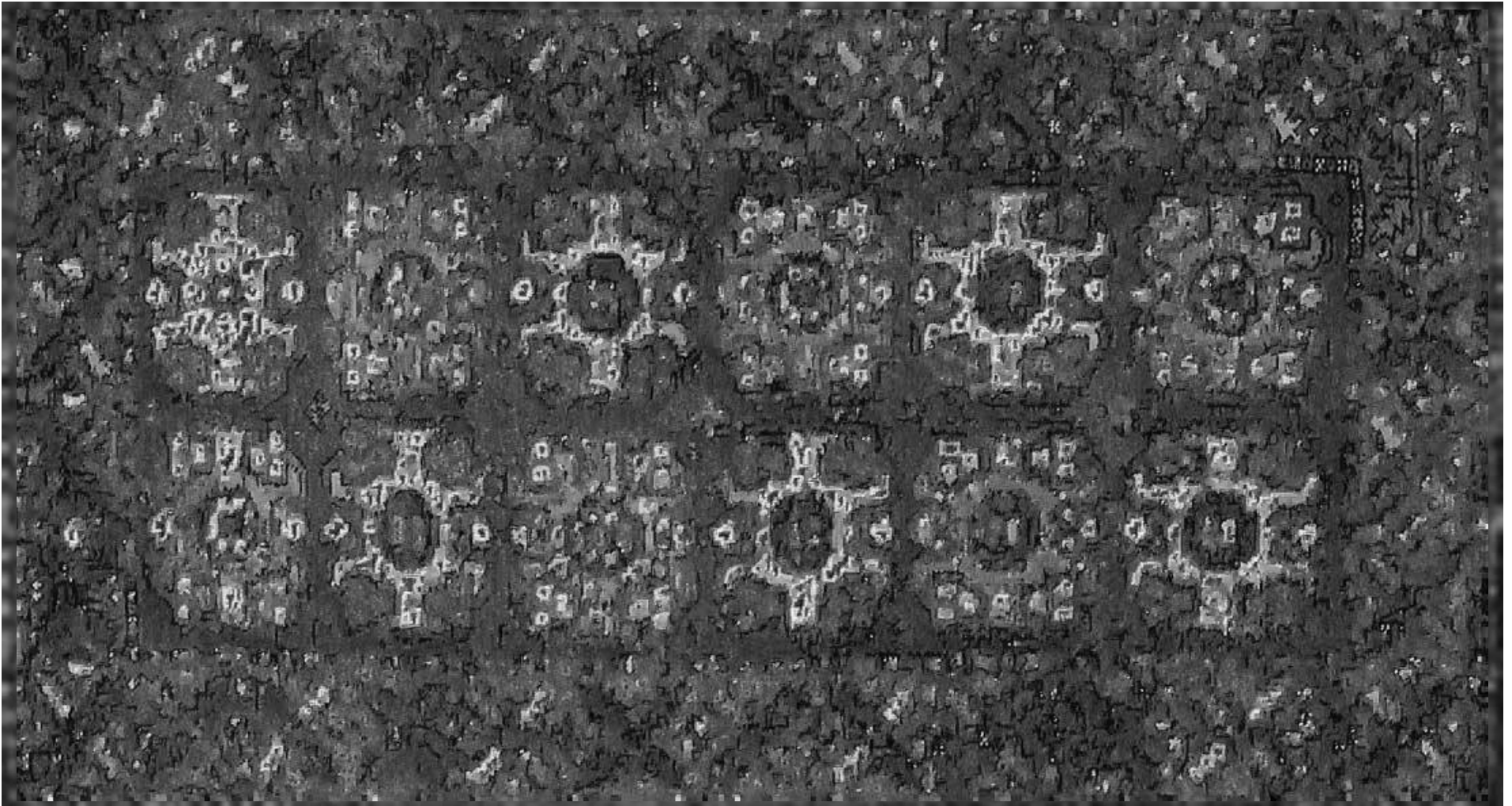


A'



28.04.2017

Testdaten:





28.04.2017

Trainingsdaten:







28.04.2017

Testdaten:





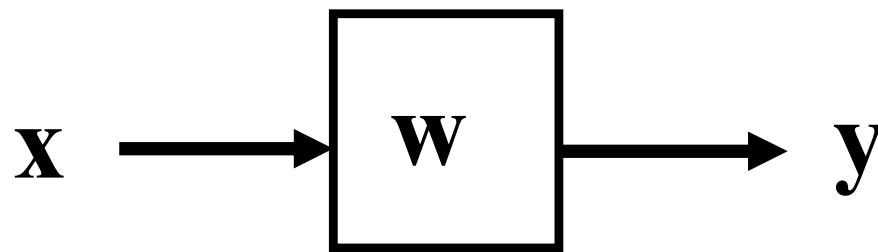


# Maschinelles Lernen: Was wird gelernt

28.04.2017

Die Aufgabe:

oft als Funktion darstellbar:  $y = f(x, w)$



$x, y$ : Eingabe, Ausgabe;  $w$ : Parameter

- Regression:  $y$  ist kontinuierlich
- Klassifikation:  $y$  ist diskret
- manchmal interessiert auch eine Verteilung  $p(y|x)$   
(Density Estimation)



28.04.2017

## Voraussage des Körperfettes durch den BMI

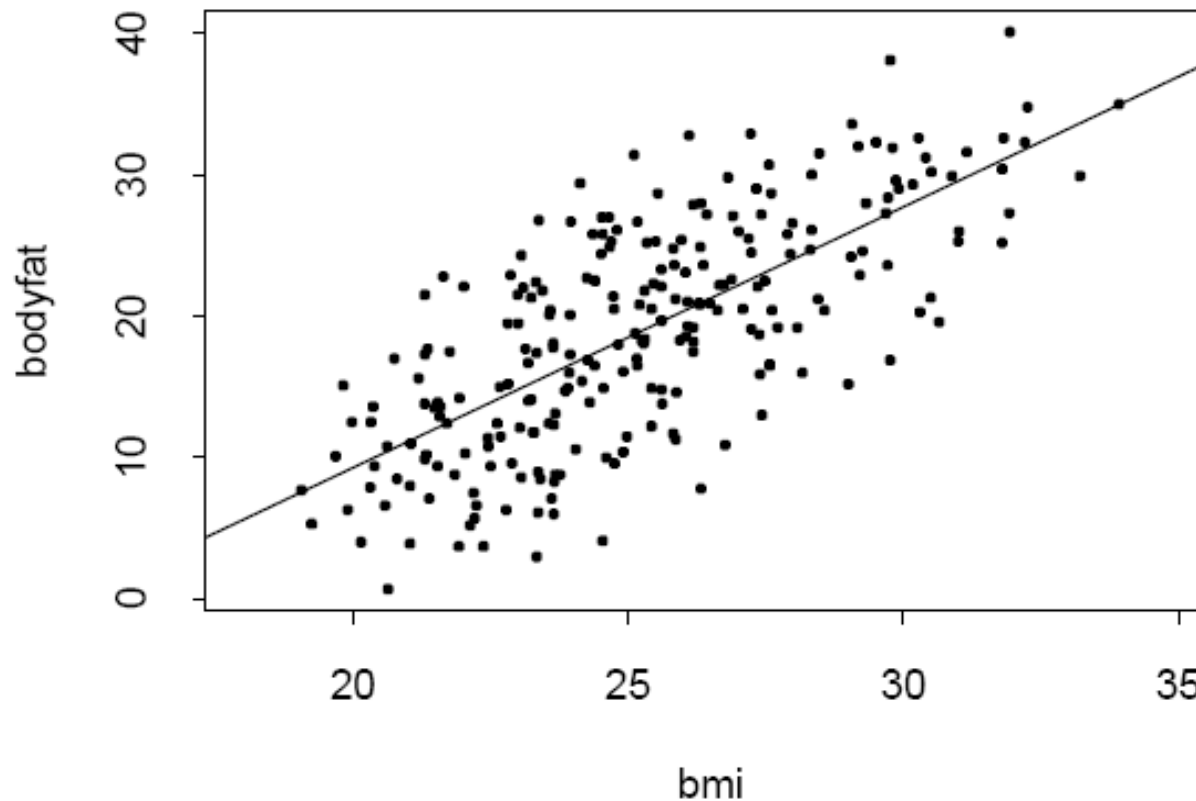
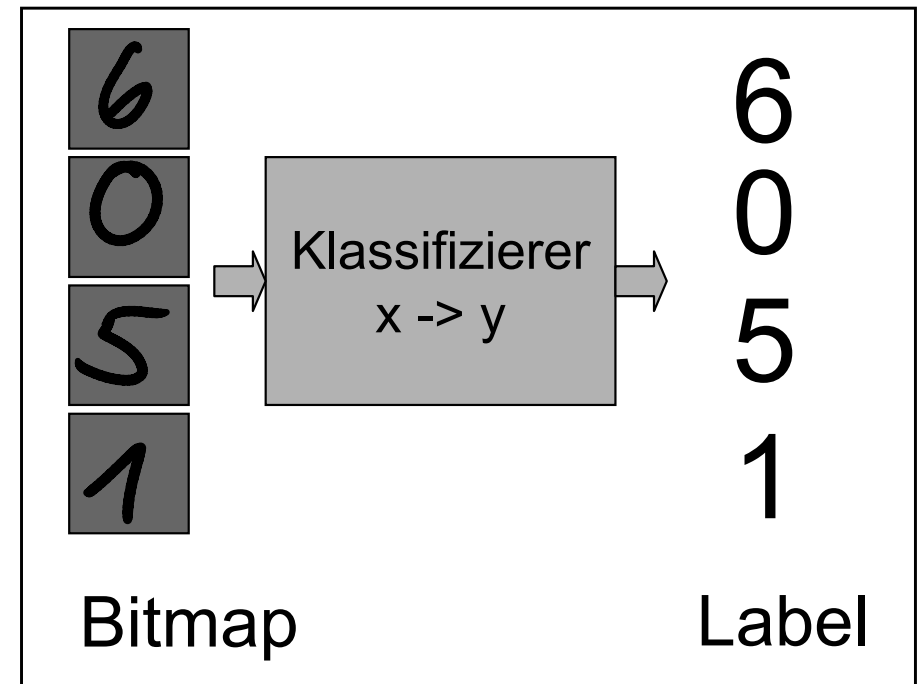


Abbildung aus: <http://www.unizh.ch/biostat/Lehre/korr-regr-1jk.pdf>

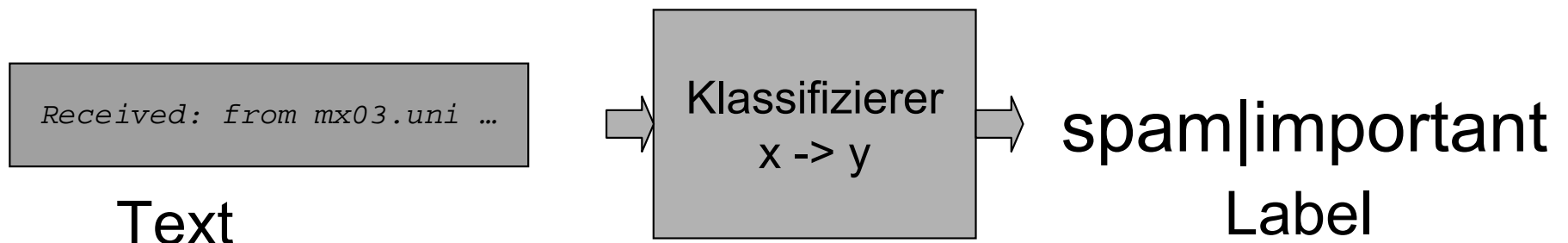
# Klassifikation: Beispiele

28.04.2017

- Zeichenerkennung



- spam filter





## Wie wird Erfolg gemessen?

- Maßzahl, z.B. Anteil korrekt klassifizierter Bilder
  - Bei Anwendung auf Trainingsdaten
  - Bei Anwendung auf neue Daten (Generalisierung)



# Maschinen Lernen: Womit wird gelernt

---

28.04.2017

Die Daten:

Supervised Learning:

- Daten mit Kennzeichnung (Ein- und Ausgabe)

Unsupervised Learning:

- Daten ohne Kennzeichnung

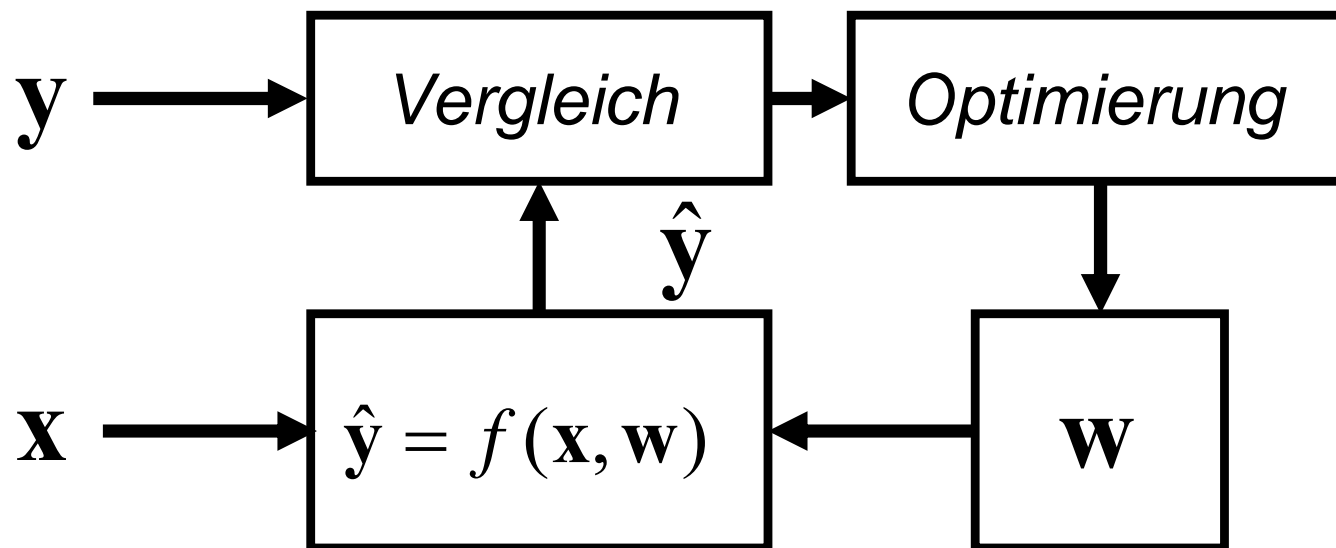


28.04.2017

Lernen ist Optimieren:

Maximierung der Performanz:

- Suche nach bestem Modell
- Suche nach besten Parametern





---

28.04.2017

# Grundlagen



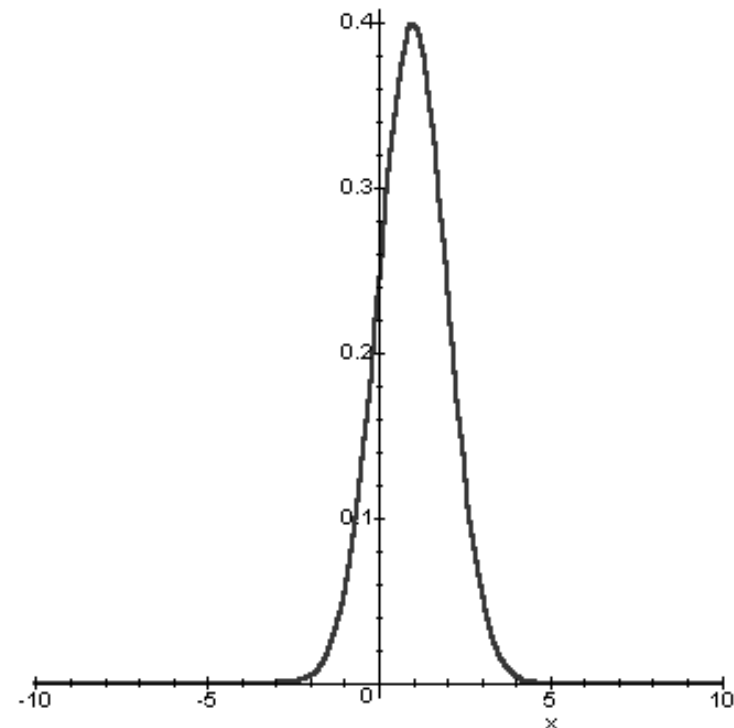
# Normalverteilung, Kovarianz

28.04.2017

## Normalverteilung:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

- innerhalb  $\pm\sigma$ :  $\sim 68\%$
- innerhalb  $\pm 2\sigma$ :  $\sim 95\%$
- innerhalb  $\pm 3\sigma$ :  $\sim 99.7\%$





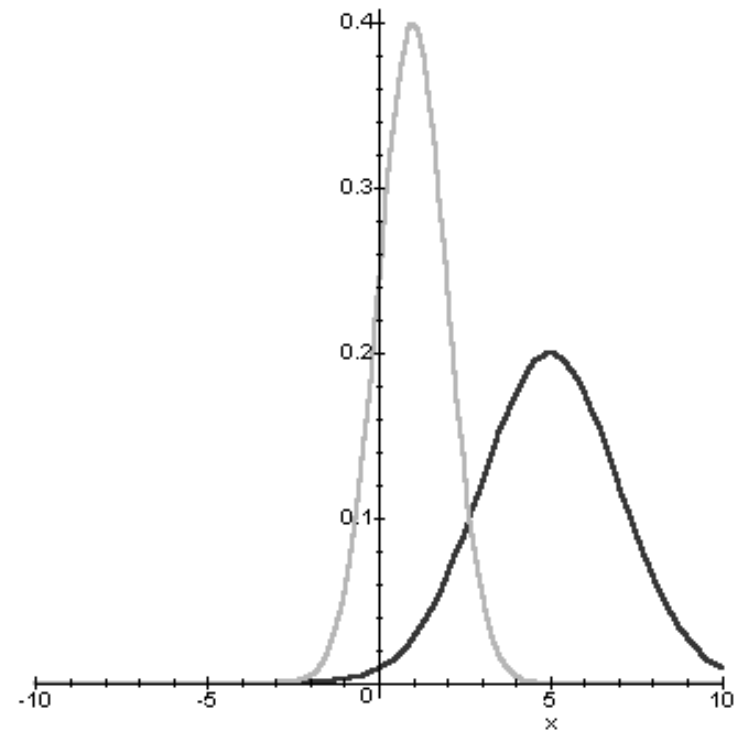


28.04.2017

Zwei Normalverteilungen:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-a)^2}{2\sigma_1^2}}$$

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-b)^2}{2\sigma_2^2}}$$



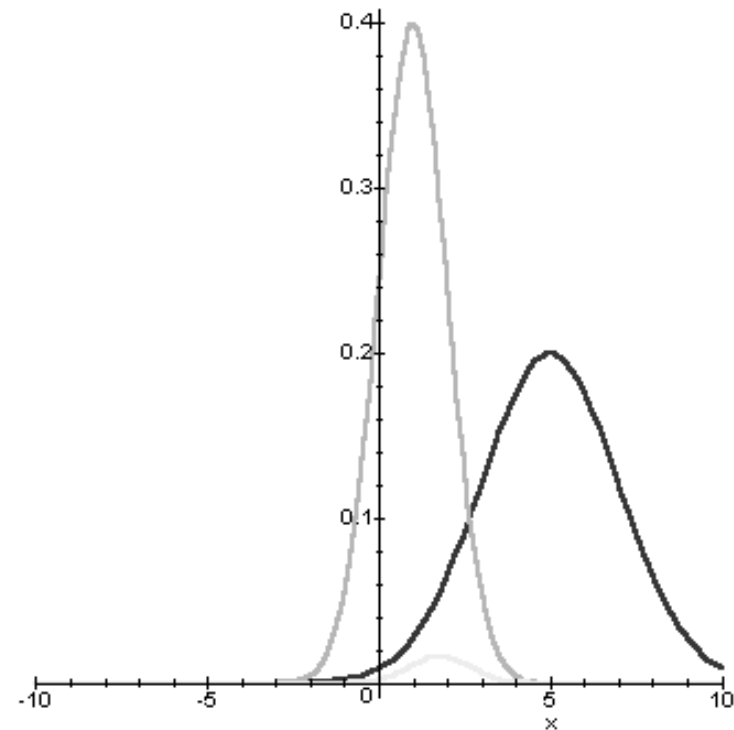


28.04.2017

## Kombination

durch Multiplikation

$$\begin{aligned} h(x) &= f(x) g(x) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(x-a)^2}{2\sigma_1^2} - \frac{(x-b)^2}{2\sigma_2^2}} \end{aligned}$$

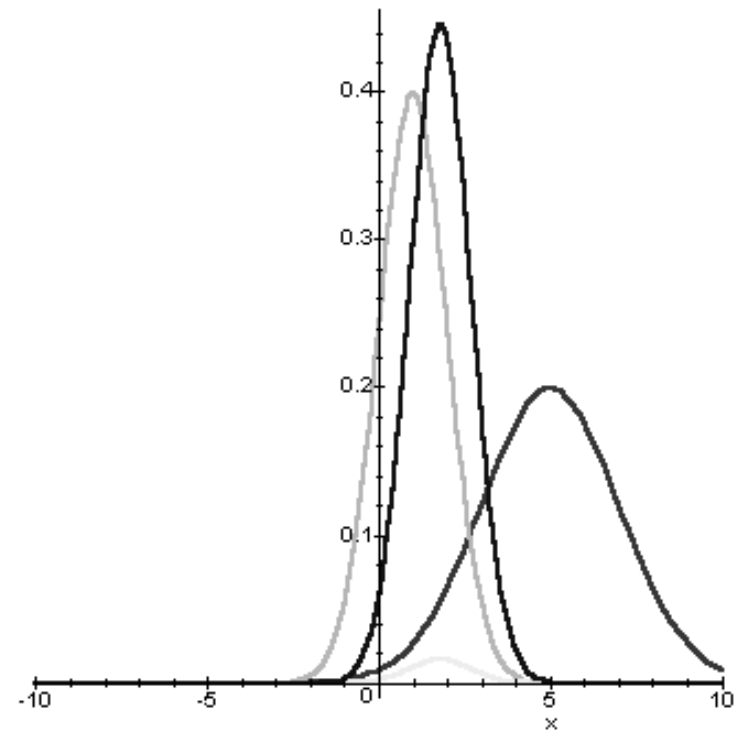




28.04.2017

Normierung von  $h(x)$ :

$$h_n(x) = \frac{1}{2\pi\sigma_3} e^{-\frac{(x-c)^2}{2\sigma_3^2}}$$
$$c = \frac{a\sigma_2^2 + b\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$
$$\frac{1}{\sigma_3^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

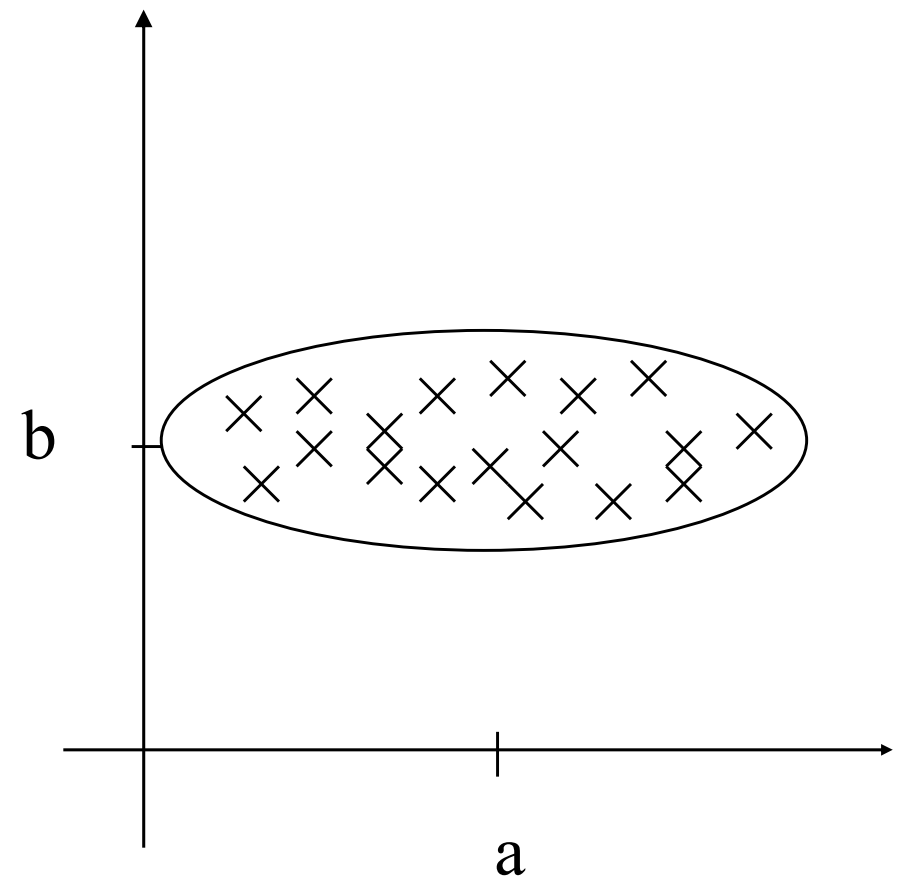




28.04.2017

2-Dimensional:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{(x-a)^2}{\sigma_x^2} + \frac{(y-b)^2}{\sigma_y^2}\right)}$$

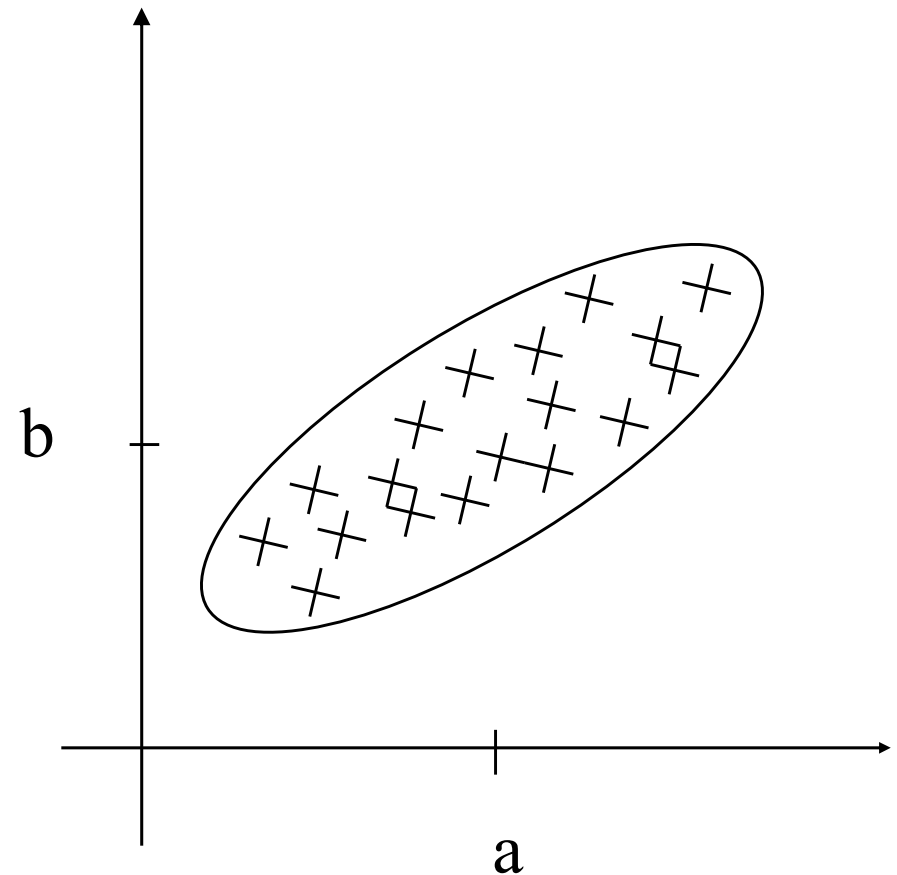




28.04.2017

2-Dimensional,  
nicht achsenparallel (mit  
Hilfe der Kovarianzmatrix  
 $[\sigma^2]=C$ ):

$$f(\mathbf{x}) = ce^{-\frac{1}{2}\mathbf{x}^T C^{-1} \mathbf{x}}$$





28.04.2017

Kovarianzmatrix kann berechnet werden als Erwartungswert der quadrierten Abweichungen vom wahrscheinlichsten Wert.

- Inneres Produkt (Skalarprodukt):  $\mathbf{y}^T \mathbf{x}$

$$\mathbf{y}^T \mathbf{x} = \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

- Äußeres Produkt:  $\mathbf{y} \mathbf{x}^T$

$$\mathbf{y} \mathbf{x}^T = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} = \begin{pmatrix} y_1 x_1 & y_1 x_2 & y_1 x_3 \\ y_2 x_1 & y_2 x_2 & y_2 x_3 \\ y_3 x_1 & y_3 x_2 & y_3 x_3 \end{pmatrix}$$



28.04.2017

## Kovarianzmatrix

$$\mathbf{d}_i = \begin{pmatrix} d_{x,i} \\ d_{y,i} \\ d_{z,i} \end{pmatrix} = \mathbf{x}_i - \bar{\mathbf{x}}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} - \begin{pmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{z}_i \end{pmatrix}$$
$$\mathbf{C} = \mathbb{E}[\mathbf{d}\mathbf{d}^T] = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} d_{x,i}^2 & d_{x,i}d_{y,i} & d_{x,i}d_{z,i} \\ d_{x,i}d_{y,i} & d_{y,i}^2 & d_{y,i}d_{z,i} \\ d_{x,i}d_{z,i} & d_{y,i}d_{z,i} & d_{z,i}^2 \end{bmatrix}$$



28.04.2017

Zusammenhängende Lehrveranstaltungen:

- Vorlesung + Übungen 2-stündig (6 LP)
- Praktikum im Wintersemester  
4-stündig (INFO-4162) für Master
- 6 LP Praktische Informatik

Benotung für BSc/MSc

- Vorlesung: Klausur am 28. Juli (Voraussetzung: Übungen)
- Übungen: Bonus für Klausur (max. 20%).

Kriterien für Übungen:

- alle Übungsblätter bearbeitet
- mindestens 50% der Punkte





28.04.2017



**Tobias Lang**

Raum C411

Tel. 78968

[t.lang@uni-tuebingen.de](mailto:t.lang@uni-tuebingen.de)



**Mathias Schickel**

Raum C411

Tel. 78968

[msch@fa.uni-tuebingen.de](mailto:msch@fa.uni-tuebingen.de)



Vorlesungsfolien, Papers usw.

- ILIAS-System
  - Informatik » Visual Computing
    - » INF4185 Grundlagen des Maschinellen Lernens
- Password:      evidence17