Gliederung

Gliederung

1. Verfahren / Übersicht

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

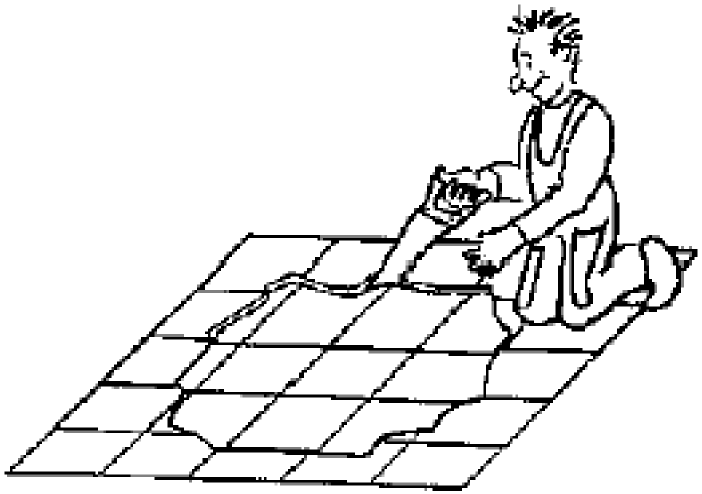
Klassifikationsschritte:

. Segmentierung,
. Merkmalsextraktion,
. Anordung im Merkmalsraum,
. eigentliche Klassifikation

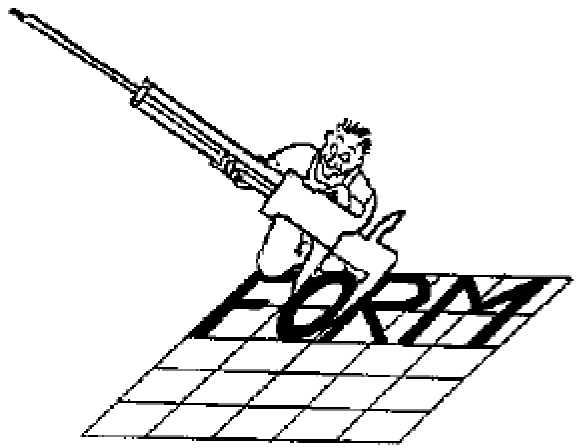
Klassifikationsverfahren (Auswahl):

. Ebenen-, Quader-, Hyperquader-Klassifikator,
. Hauptebenentransformation (Hilfsmittel für Ebenen- und Hyperquader-Klassifikatoren)
. Minimal-Distanz
. Maximum Likelihood

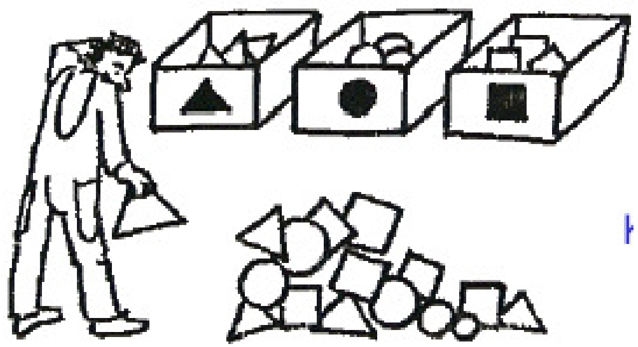
Klassifikation anschaulich



Segmentierung



Merkmalsextraktion



Klassifikation

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

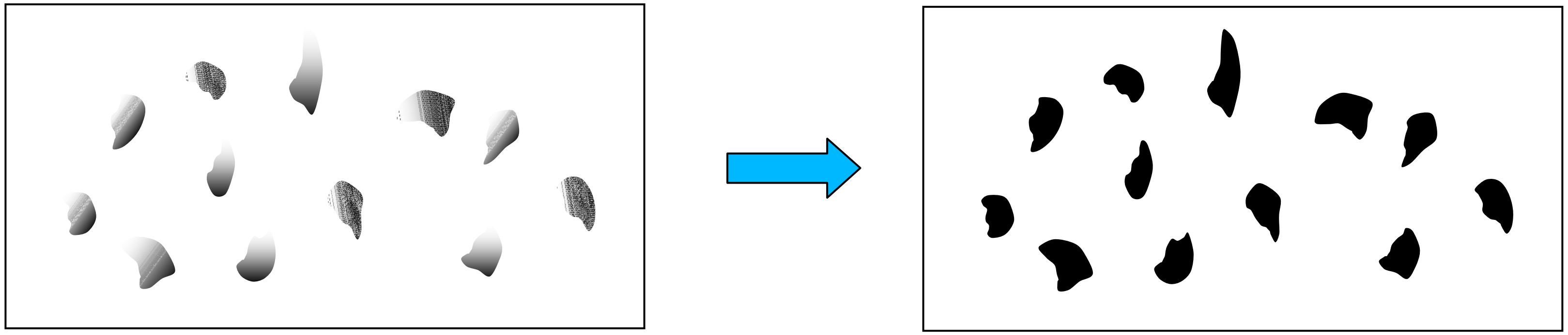
2. Segmentierung / ROI

Segmentierung:

Trennung von Objekt(en) und Hintergrund im gesamten Bild oder in ausgewählten Bereichen (ROI → region of interest)

Verfahren der Segmentierung (Auswahl), siehe auch bisherige Vorlesungskomplexe:

- Kantenextraktion / Konturverfolgung
- Punkt- bzw. lokale und globale Operatoren (Grauwertstatistik)
- Schwellwertverfahren (statisch bzw. adaptiv)
- Farben
- …

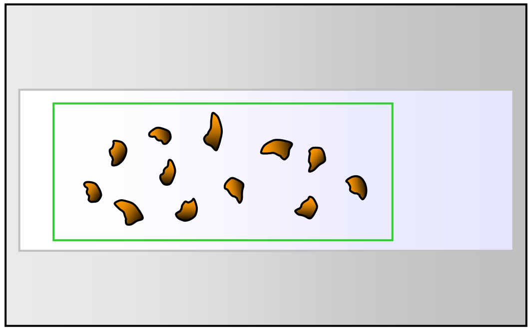


Verfahren zur Festlegung von ROI

manuell / interaktiv

An Labor- bzw. Entwicklerarbeitsplätzen,

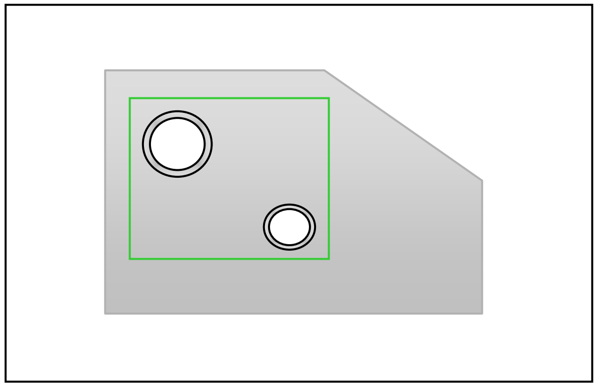
Bei Stichprobenprüfungen z.B. in Prüflabors zur Qualitätsicherung



Beispiel: Partikel auf Objektträger aus Glas

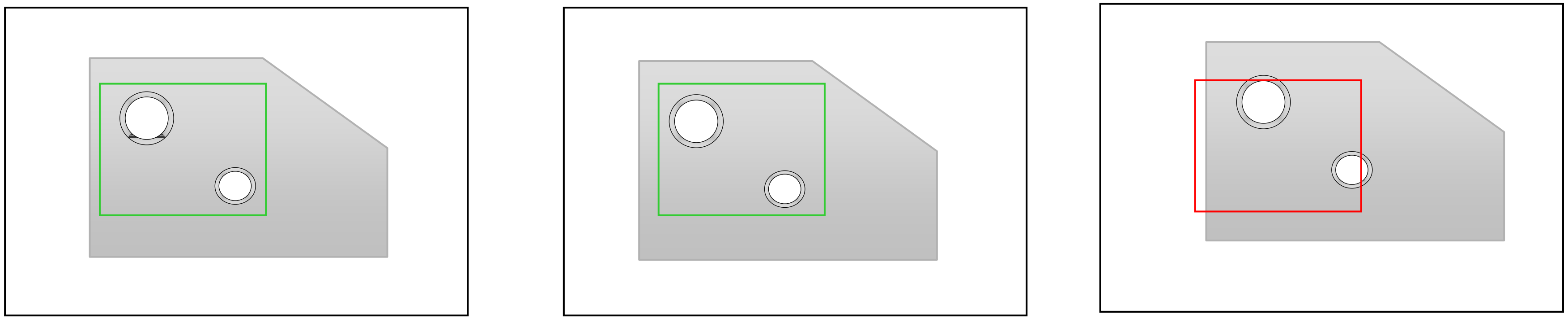
automatisch

bei der Prozesskontrolle in fortlaufenden Serienprüfungen, z.B. in Förderstrecken

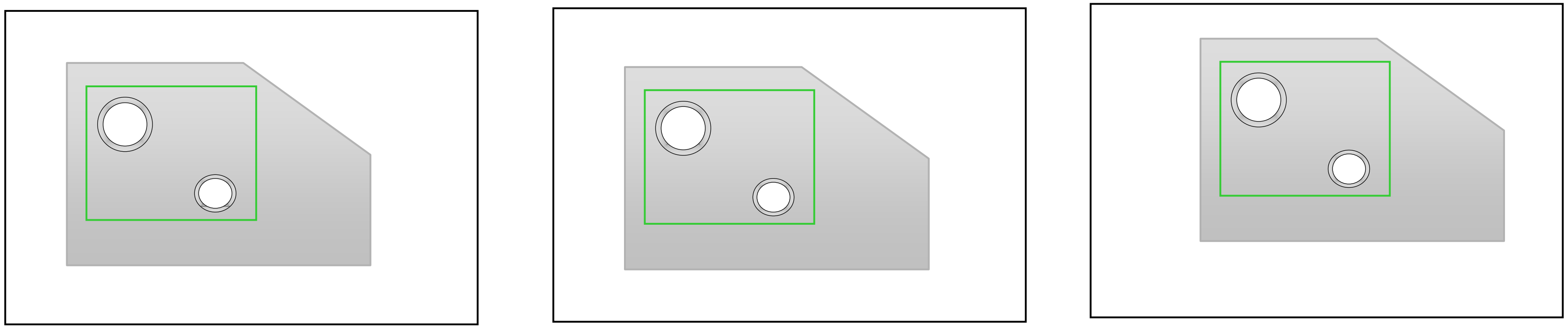


Beispiel: Werkstück mit Bohrungen

Automatische Festlegung von ROI



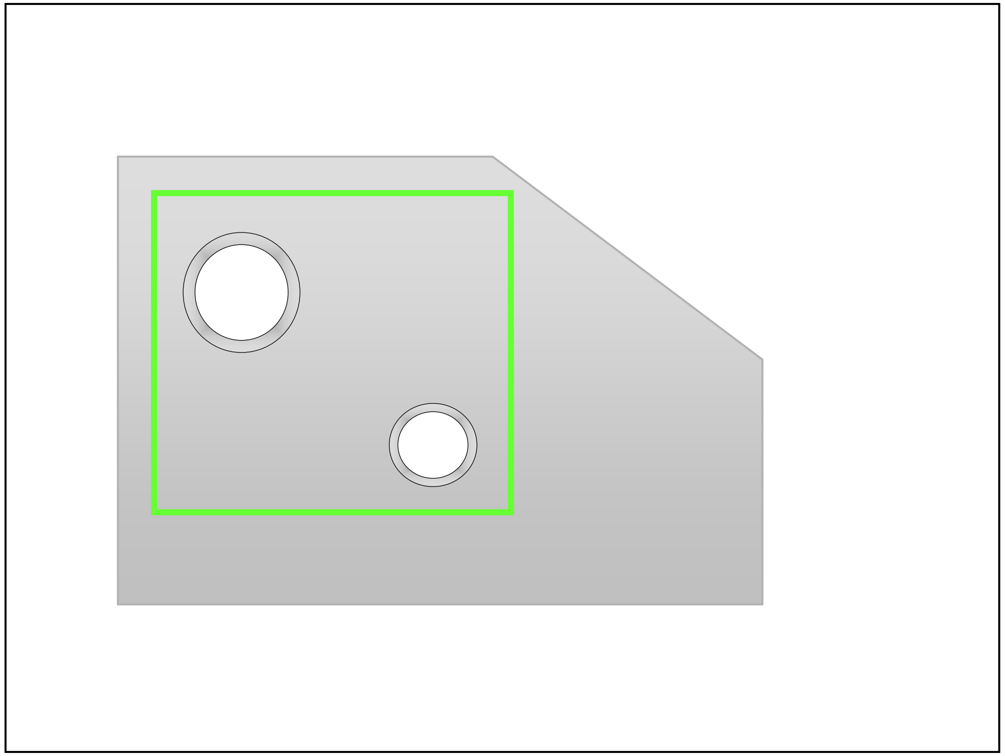
Statisch (feste Position innerhalb des Bildfeldes bei verschiedenen Bildern), mögliches Problem: Objekt außerhalb des ROI



dynamisch (Festlegung in Abhängigkeit von anderen Bildmerkmalen, z.B. Positionen von Justiermarken, Referenzkanten usw.), Nachteil: höherer Rechenaufwand

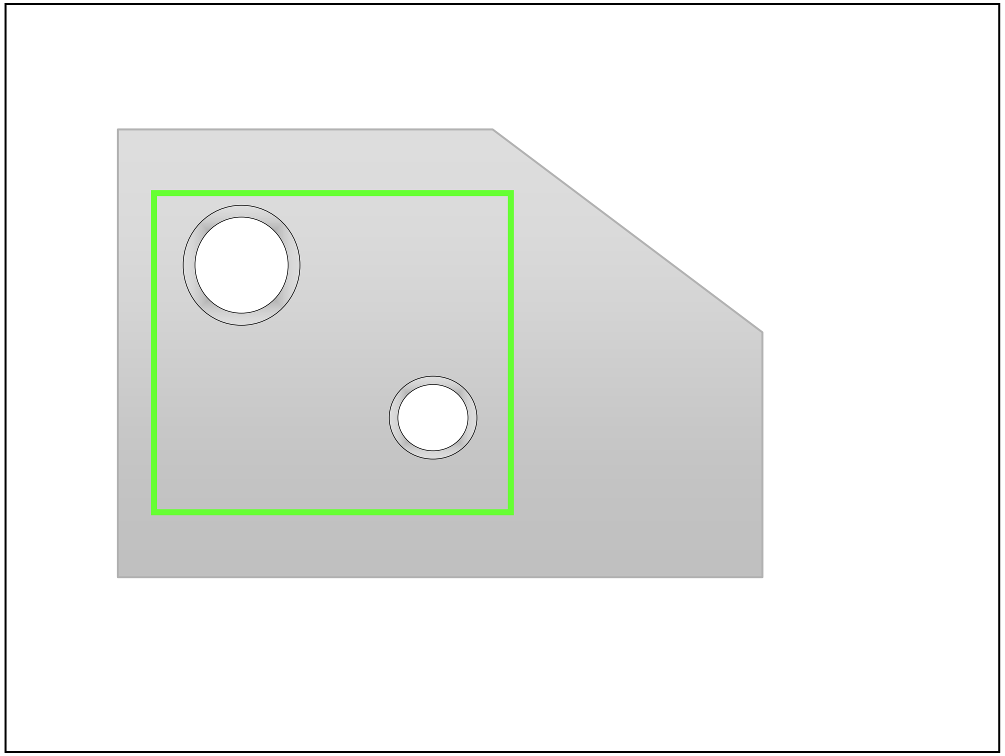
Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Automatische Festlegung von ROI



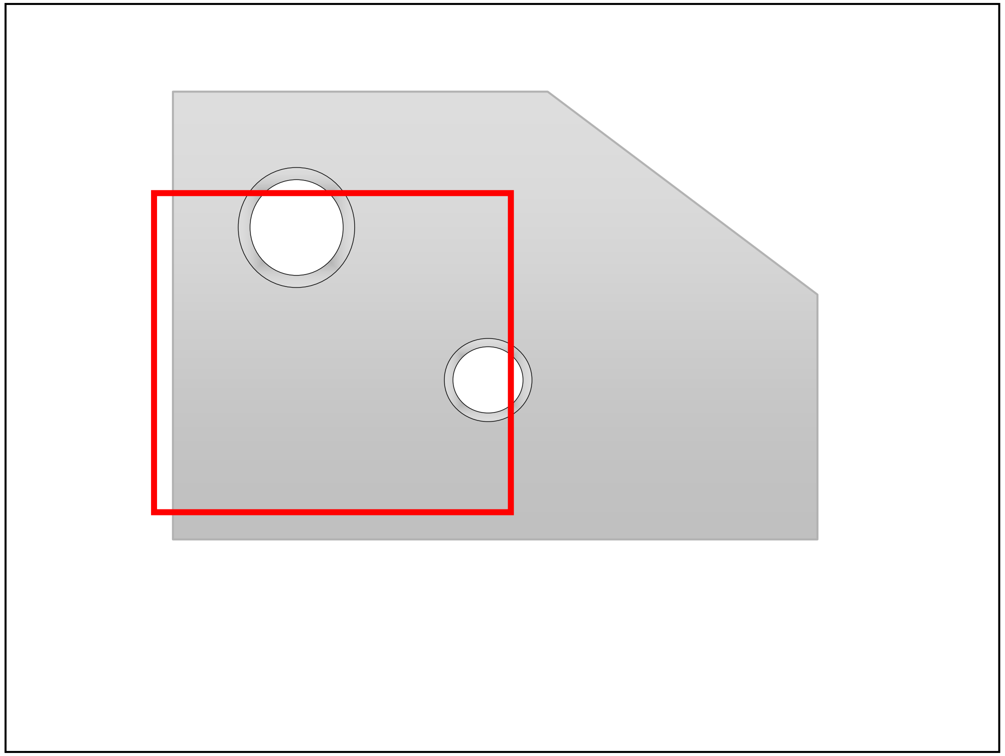
Objekte innerhalb des ROI

Automatische Festlegung von ROI



Objekte innerhalb des ROI

Automatische Festlegung von ROI



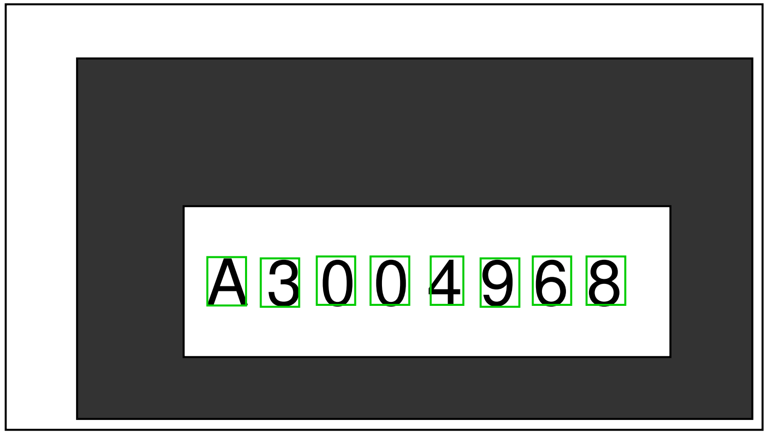
Objekte außerhalb des ROI !!!

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Segmentierung (Trennung von Objekt und Hintergrund) / Variante:

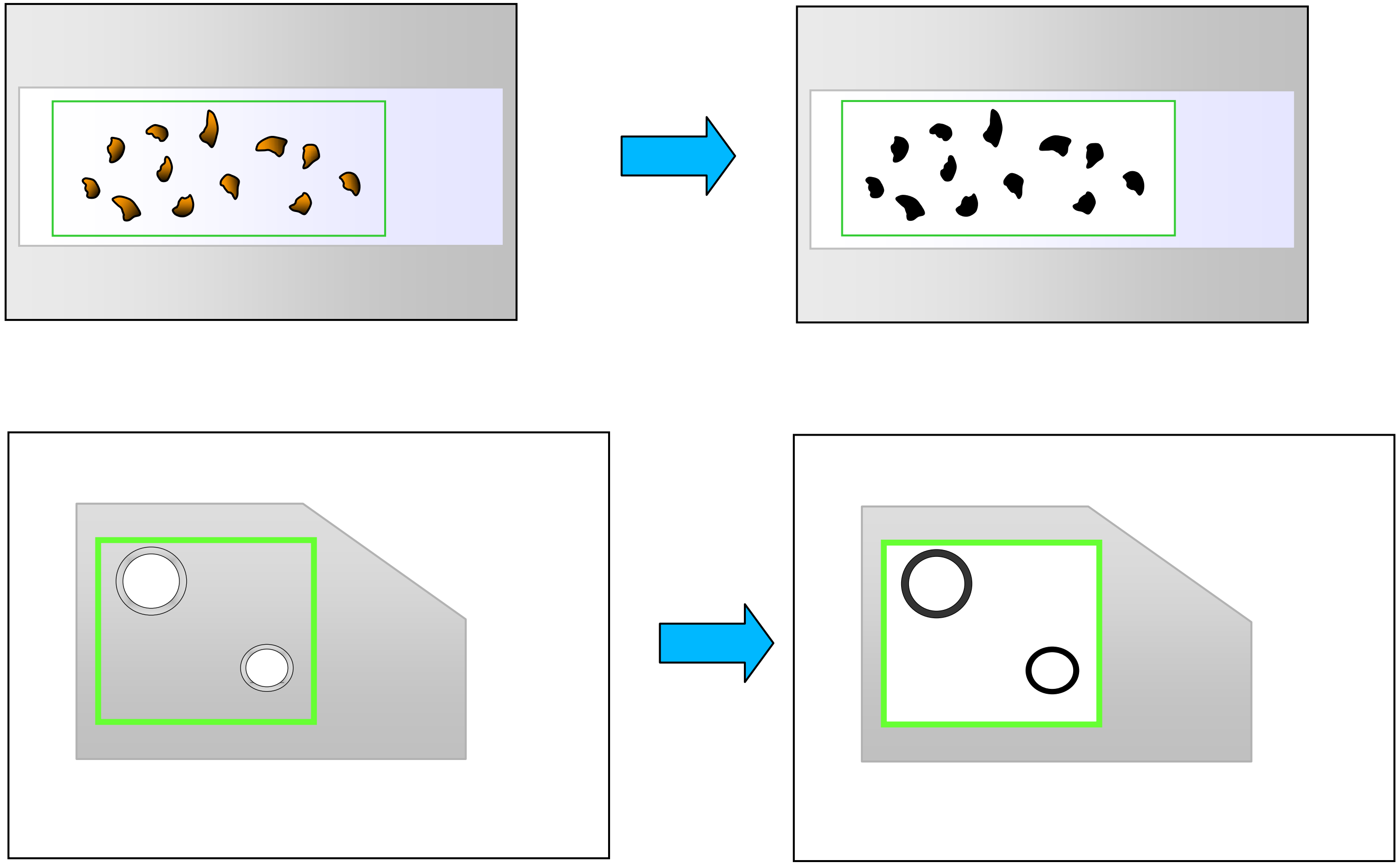
Definition von Bereichen innerhalb Objekte umschließender Flächenelemente (meist Rechtecke → Sub-ROI's); Ein ROI enthält dann nur ein Objekt

Anwendung z.B. bei der Klarschriftlesung



Beispiel der Segmentierung bei der Klarschriftlesung

Segmentierung (Trennung von Objekt und Hintergrund) im ROI



Beispiele: Trennung von Objekten und Hintergrund durch Anwendung der Grauwertstatistik (Binarisierung), jeweils in einer ROI

3. Klassifikation

3.1 Begriffe, grundlegende Sachverhalte

Klassifikation: Begriffe / grundlegende Sachverhalte

- Zuordnung von Objekten zu Klassen anhand der Merkmale (Merkmalsvektoren)
- Anforderungen an Objektmerkmale: signifikant, miteinander nicht korreliert

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Klassifikation: Begriffe / grundlegende Sachverhalte

1. Merkmalsextraktion: von einem Objekt wird ein Satz von n-Merkmalen C_ν ermittelt (Fläche, Form, Farbe, Textur...)

2. Merkmalsraum: Eintragen der Merkmale in n-dimensionalen Merkmalsraum; Beachte: n-Merkmale spannen n-dim. Merkmalsraum auf.

3. Merkmalsvektor: Von jedem Objekt gibt es einen Merkmalsvektor (Gesamtheit der Merkmale)

4. Merkmalscluster: Anhäufung von Objektmerkmalen im Merkmalsraum nennt man Cluster.

4. Eindeutige Klassenzuordnung: wenn sich die Cluster nicht überlappen, ist eine fehlerfreie Klassifikation möglich (andernfalls ist es das Ziel, die Klassifikationsfehler zu minimieren, z. B. Bayes-Klassifikator)

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.2 Phasen der Klassifikation

Klassifikation / Phasen

1. **Einlernphase (während der Anwendungsentwicklung oder Prozessanpassung):** Bildung von Clustern durch Merkmalsextraktion von bekannten Stichproben; Cluster beschreiben die Zugehörigkeit zu Klassen im Merkmalsraum

2. **Klassifikationsphase (im laufenden Prozess, in der Serienprüfung usw.)**
   Untersuchung unbekannter Objekte auf Zugehörigkeit zu Clustern bzw. Klassen

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Klassifikation / Phasen

Verschiedene Möglichkeiten des Einlernens

Überwachtes Lernen

Merkmale (Farbe, Form usw.) und Klassen sind bereits vorher weitestgehend bekannt. Durch genügend große Stichproben werden beim Anlernen die Clustergrenzen bestimmt

Unüberwachtes Lernen

Es sind keine Merkmale bekannt, welche die Objekte signifikant beschreiben. Ggf. ist auch die Anzahl möglicher Klassen nicht bekannt. Von Stichproben werden verschiedene Merkmale bestimmt und diese auf Clusterbildung untersucht. Merkmale mit signifikanter Clusterbildung werden dann in der Klassifikationsphase verwendet, die anderen Merkmale werden wieder verworfen.

Klassifikation: Einlernphase

Beispiel Klarschriftlesung

Teach In: z.B. Erfassung der Merkmale von Binärbildern, welche die Ziffern 0..9 präsentieren durch Untersuchung einer repräsentativen Anzahl von Stichproben

Datensatz																			
Muster	1	1	1	...	2	2	2	...	3	3	...	4	...	5	...	9	9	...	0	0	0
Bedeutung	"1"	"1"	"1"	...	"2"	"2"	"2"	...	"3"	"3"	...	"4"	...	"5"	...	"9"	"9"	...	"0"	"0"	"0"
Klasse (Nr)	1	1	1	...	2	2	2	...	3	3	...	4	...	5	...	9	9	...	10	10	10

Klassifikation / Einlernphase: mögliche Fehler

- **Gewählte Stichprobe zu klein → Merkmalsraum wird nicht vollständig abgedeckt (in der Klassifikationsphase können dann viele Objekte nicht klassifiziert werden),**

- **nahezu Wertgleichheit eines oder mehrerer Parameter für Objekte unterschiedlicher Klassen (Clustergrenzen zu dicht bzw. Clusterüberschneidungen) → Parameter ungeeignet, da zu große Überschneidungen mit anderen Parametern oder: anderen Klassifikator verwenden,**

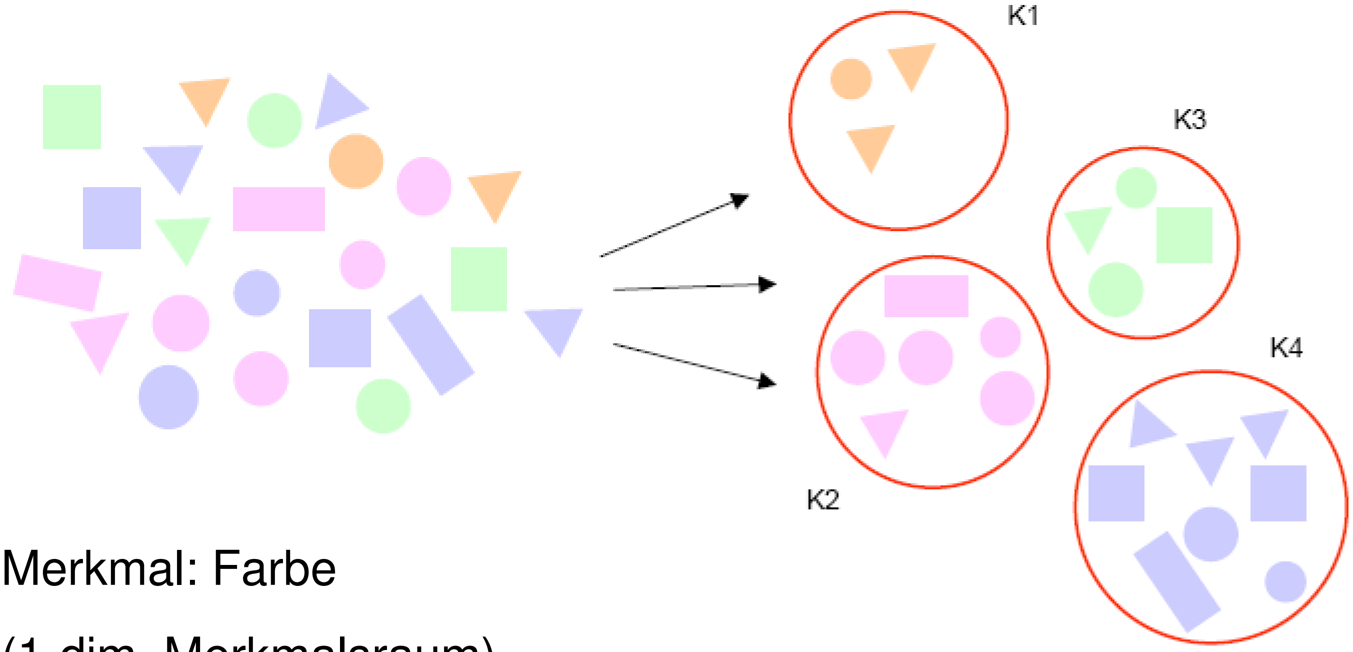
- **Abhängigkeit voneinander (z.B. sind bei Rechtecken mit konstanten Seitenverhältnis beide Seitenlängen als Merkmale voneinander abhängig, eine Seitenlänge kann also weggelassen werden),**

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.3 Klassifikation / einfache Beispiele

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

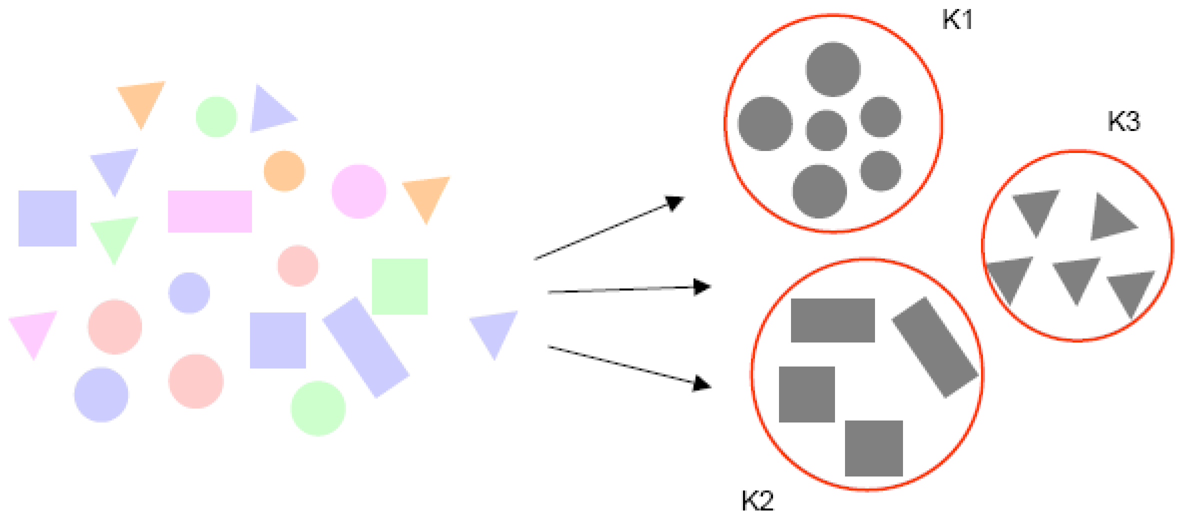
Klassifikation: Beispiele



Merkmal: Farbe

(1-dim. Merkmalsraum)

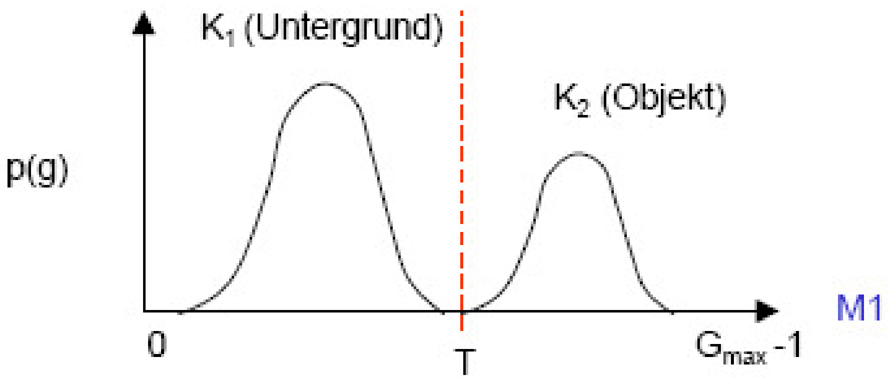
Klassifikation: Beispiele



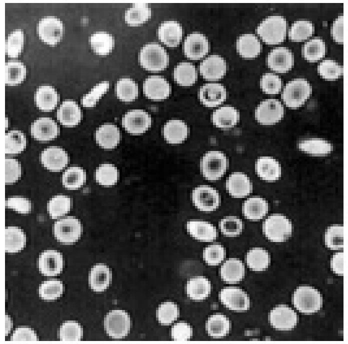
Merkmal: Form

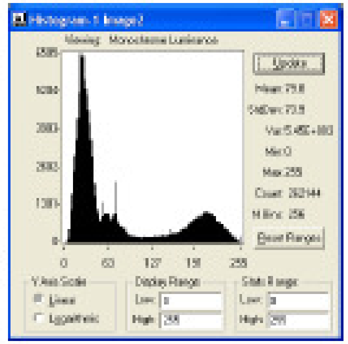
(1-dim. Merkmalsraum)

Klassifikation: Beispiele

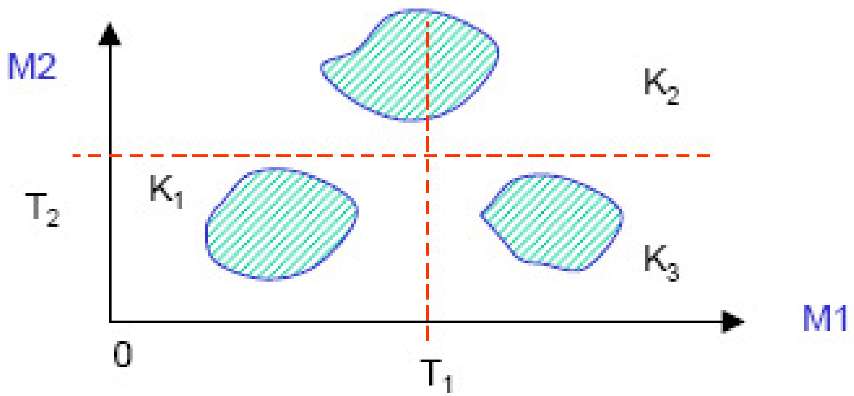


Trennung von Objekt (K2) und Hintergrund (K1)

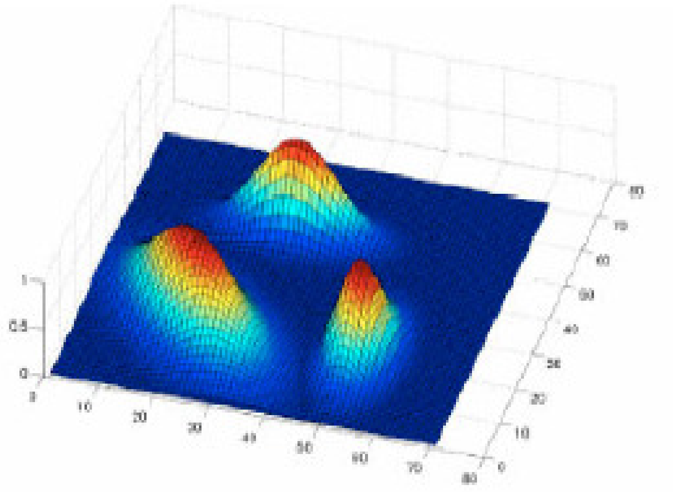




1 Merkmal → 1D-Merkmalsraum



Trennung von 3 verschiedenen Objekten (K1 .. K3)



2 Merkmale → 2D-Merkmalsraum!

Klassifikation: weiteres Beispiel

- siehe Vorlesungsschwerpunkt Merkmalsextraktion (Blutanalyse)

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.4 Klassentrennung / Merkmalsgüte

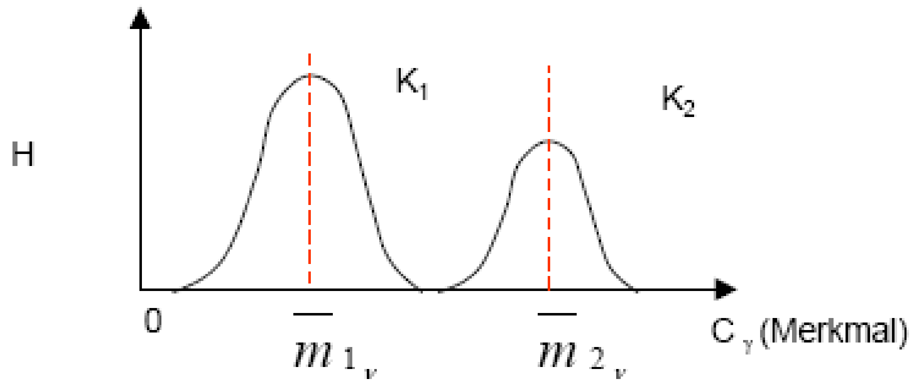
Merkmalsgüte

- Merkmalsgüte entscheidend für Klassifikationsgüte
- Gesucht ist eine Maßzahl für die Merkmalsgüte
- Beispiel:

  Gegeben sind zwei Klassen K_1 und K_2 die das Merkmal $C\nu$ (Merkmalsindex) besitzen.
  Die Klassen besitzen

  - den klassenbedingten Mittelwert $\overline{m}_{k\nu}$

  - die klassenbedingte Varianz $\sigma^2_{k\nu}$



Intuitiv einleuchtend:

Gute Klassentrennung, wenn:
- große Differenz der Mittelwerte
- kleine Streuungen

Gütemaß

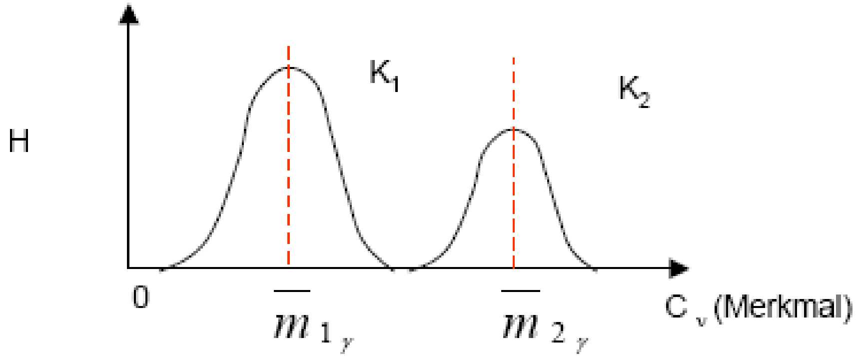
- Definition:

$$\boxed{G_\nu = \frac{\left(\overline{m}_{1\nu} - \overline{m}_{2\nu}\right)^2}{\sigma_{1\nu}^{\;2} + \sigma_{2\nu}^{\;2}}}$$

mit

$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} f_i$ Mittelwert

$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (f_i - \overline{m})^2$ Varianz

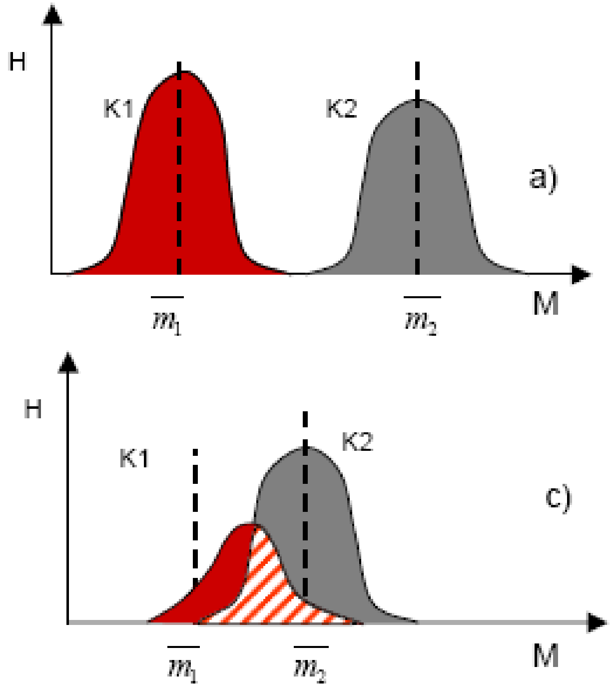


Ziel:

$\overline{m}_{1\nu} - \overline{m}_{2\nu}$ (Differenzen der Mittelwerte) $\Rightarrow$ Max

$\sigma_{1\nu}^{\;2} + \sigma_{2\nu}^{\;2}$ (Summe der Varianzen) $\Rightarrow$ Min

Beispiele:



Fälle für Klassentrennung:

a) eindeutige Trennung

b) c) bedingt trennbar (z. B. Bayes)

3.5 Klassifikatoren: Übersicht / Zielstellungen

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Klassifikatoren / Übersicht (Auswahl)

- Parallel-Epiped-Klassifikator (Rechteck-, Quader-, Hyperquader-Klassifikator,
  - Hauptachsentransformation (Hilfsmittel zur Vereinfachung bei Anwendung von Parallel-Epiped-Klassifikatoren)
- Nearest-Neighbourg-Klassifikator
- k-Nearest-Neighbourg-Klassifikator
- Minimum-Distanz-Klassifikator
- Statistische Klassifikatoren: Maximum-Likelihood-Klassifikator und Bayes-Klassifikator
- Klassifikation mittels LUT,
- Klassifikation mittels Neuronaler Netze oder Fuzzy-Logic

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Zielstellungen bei der Merkmalsauswahl

- Dimensionalität des Merkmalsraumes ↓
- Merkmalsanzahl ↓
- Relevanz des Merkmals ↑
- Bedingung: Unkorreliertheit der Merkmale

- hohe Merkmalsrelevanz → hohe Merkmalsgüte

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

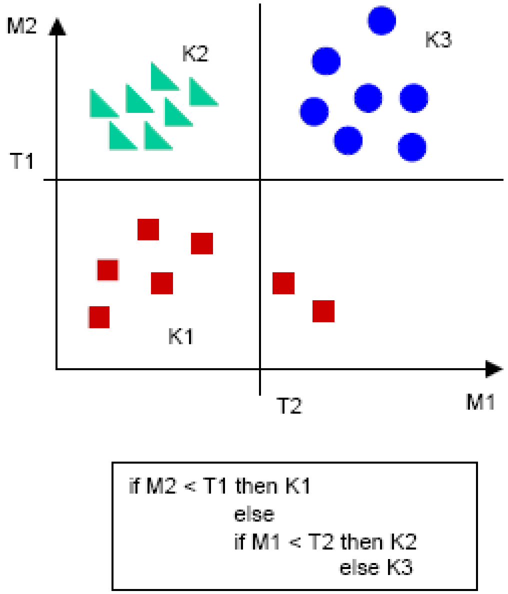
3.6 Ausgewählte Klassifikatoren (Beispiele)

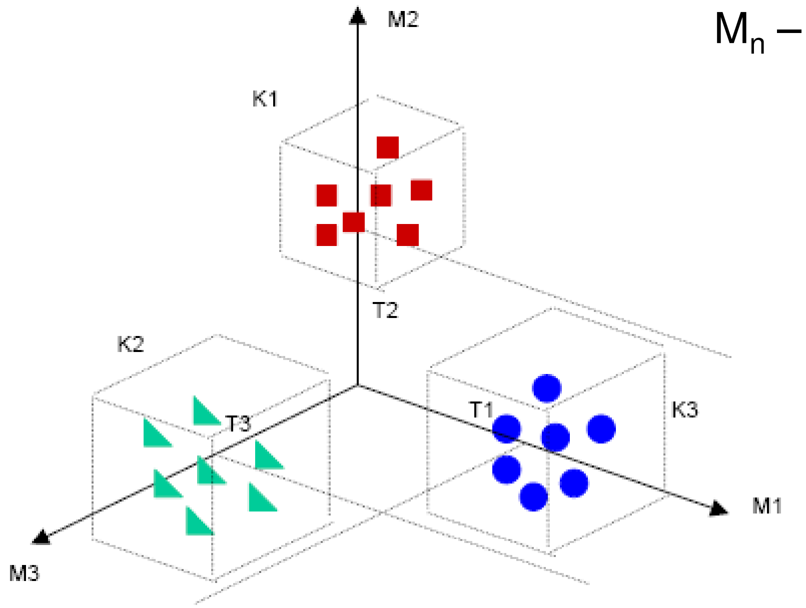
Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.6.1 Parallel-Epiped-Klassifikatoren

K_m – Cluster

M_n – Merkmale





Zweidimensional
Rechteck-Klassifikator

Dreidimensional
Quader-Klassifikator

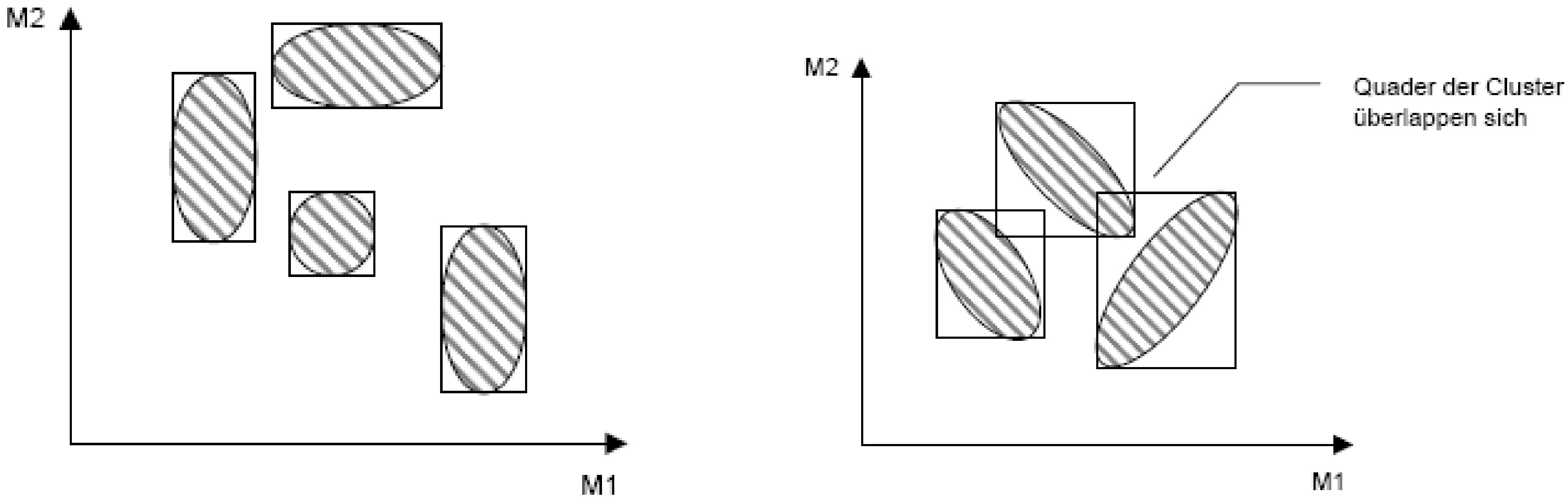
n-dimensional
Hyperquader-Klassifikator

Parallel-Epiped-Klassifikatoren: Vor- und Nachteile

- **geringer Rechenzeitaufwand**
- **Bei Clusterüberlappungen ist nicht immer eine eindeutige Klassifikation möglich**

Parallel-Epiped-Klassifikatoren

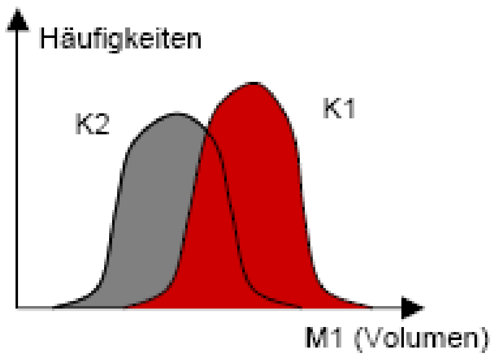
eignen sich vor Allem bei unkorrelierten (orthogonalen) Merkmalen

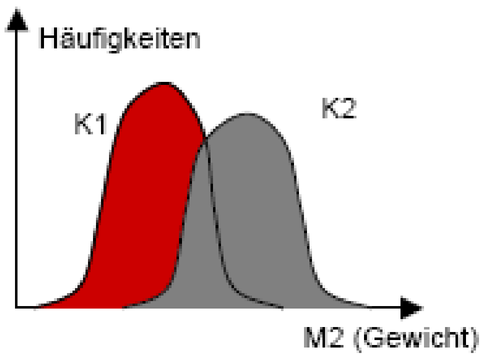


Hauptachsentransformation
(Bei der Anwendung von Parallelepiped-Klassifikatoren)

Beispiel

Aufgabe: Eine Erntemaschine soll Kartoffeln und Steine voneinander trennen.
Mit dem Merkmal Form ist kein Unterschied in Klassen möglich.
Betrachtet man die Merkmale: M1 (Volumen) und
M2 (Gewicht) so bekommt man zu folgenden Verteilungen:

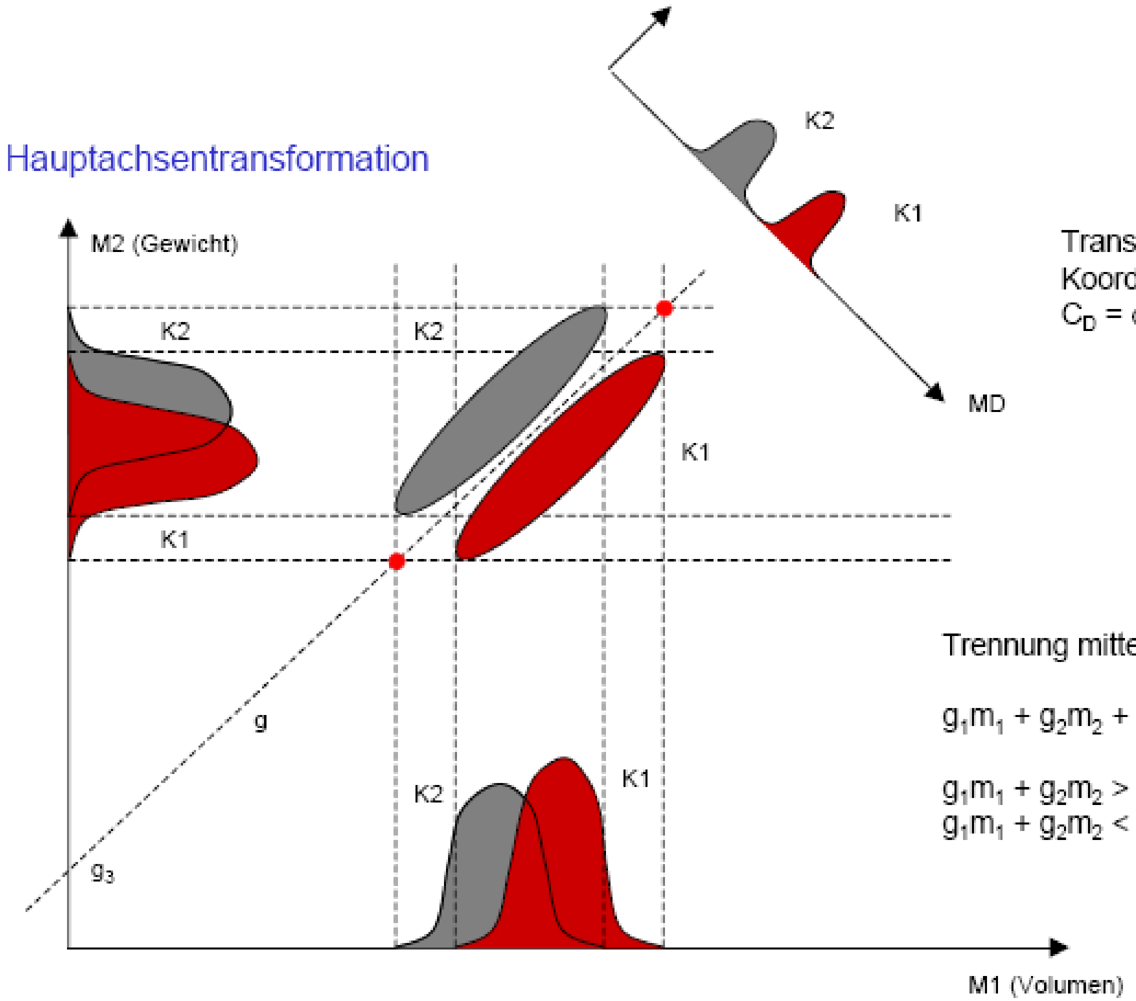




Man kann erkennen, dass sich diese Objekte auch in diesen Parametern nicht eindeutig unterscheiden.
Betrachtet man jedoch Volumen und Gewicht gemeinsam in einem 2D-Histogramm:
Volumen (M1) → X-Achse,
Gewicht (M2) → Y-Achse →erhält man folgende Cluster:

Hauptachsentransformation

(Bei der Anwendung von Parallelepiped-Klassifikatoren)



Transformation des Koordinatensystems
$C_D = \alpha\, c_1 + \beta\, c_2$

Trennung mittels Trennlinie g(M1, M2) möglich

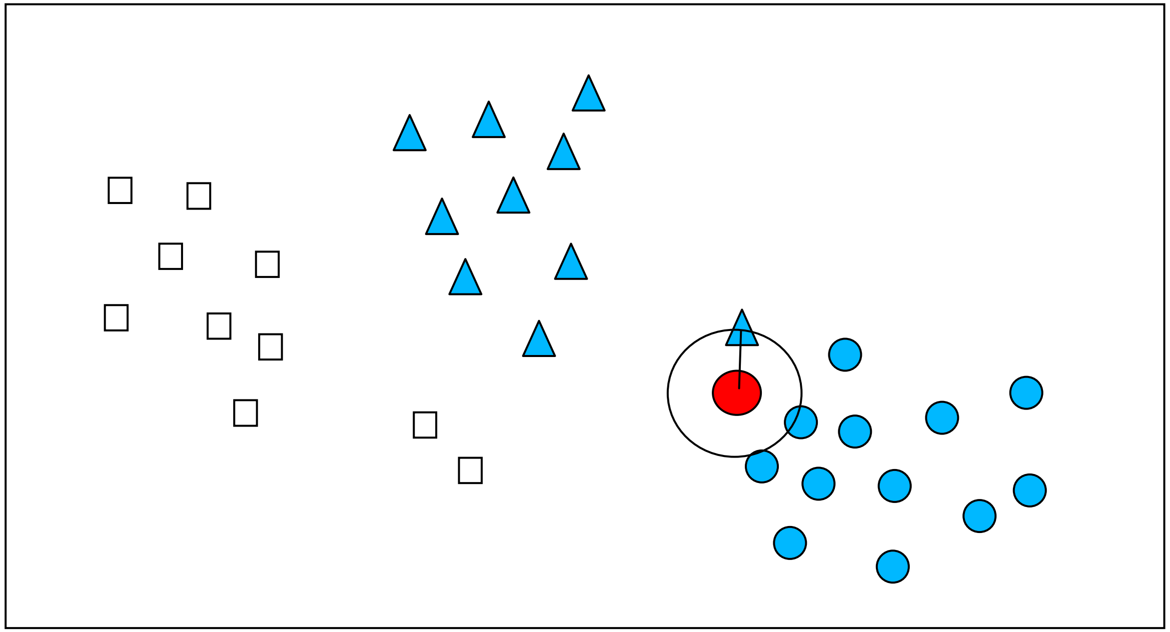
$g_1 m_1 + g_2 m_2 + g_3 = 0$

$g_1 m_1 + g_2 m_2 > 0 \rightarrow K2$
$g_1 m_1 + g_2 m_2 < 0 \rightarrow K1$

Hauptachsentransformation siehe u.a.:http://www.medialab.ch/archiv/pdf_studien_diplomarbeiten/da00/da2000-085_ChromatischeSteinklassifizierung/DokuHTML/klassifizierung.html#TOP

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

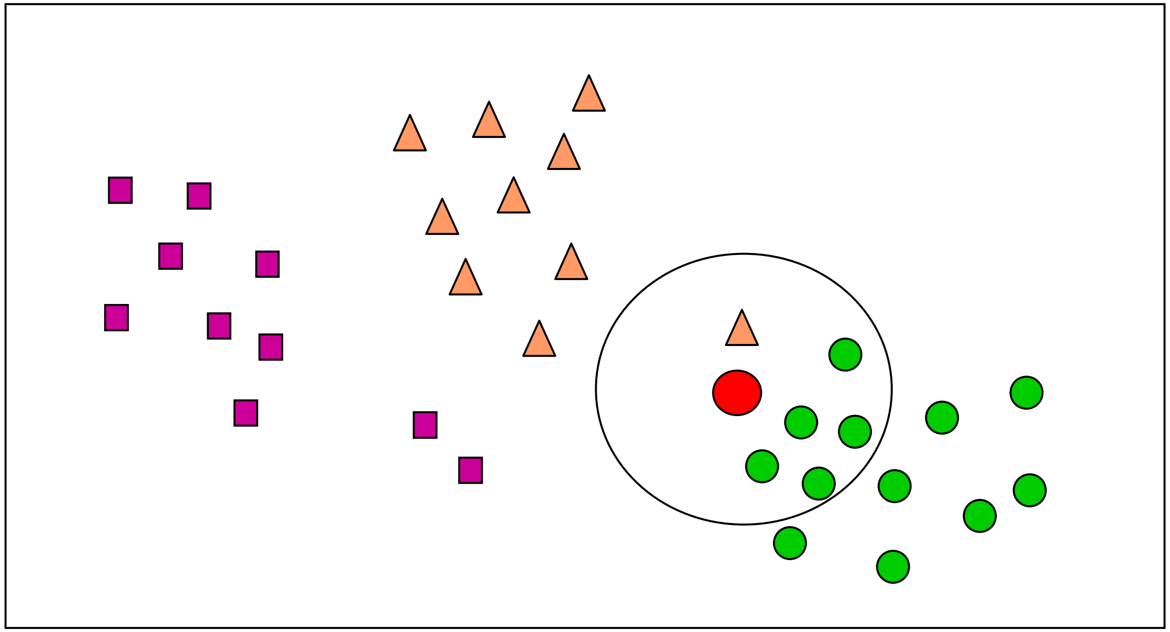
3.6.2 Nearest-Neighbourg-Klassifikator



Zuweisung zur Klasse des Objektes mit dem kleinsten euklidischen Abstand zum neuen Objekt

- **geeignet für überlappende Cluster**
- **Gefahr der Falschzuweisung**

3.6.3 K-Nearest-Neighbourg-Klassifikator



Beispiel: k=6

Untersuchung der k Objekte mit den kleinsten Abständen zum zu klassifizierenden Objekt; Zuordnung zur am häufigsten auftretenden Klasse dieser Objekte

- **geeignet für überlappende Cluster**
- **etwas stabiler gegenüber Falschzuweisungen**

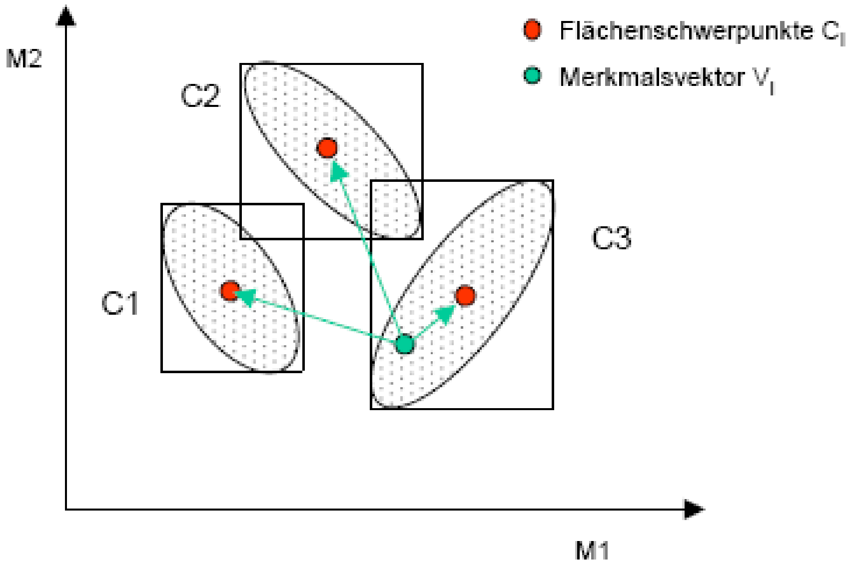
3.6.4 Minimum-Distanz-Klassifikator

(eignet sich auch bei korrelierten Merkmalen)

d) Minimum-Distanz-Klassifikator

Vorgehensweise:

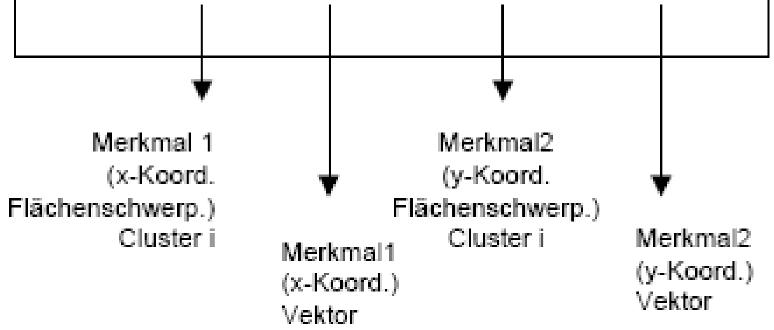
1. Bestimmung Flächenschwerpunkt C_i (M_1 , M_2) aller Cluster.
2. Bestimmung Abstand des akt. Merkmalsvektors V_j (M_1 , M_2) zu allen Flächenschwerpunkten.
3. Min. Abstand entspricht der Klassenzugehörigkeit.



Voraussetzung: Merkmale besitzen Normalverteilungen

Als Abstandmaß gilt der Euklidische Abstand

$$d_e = \sqrt{\left(M_{1C_i} - M_{1V_j}\right)^2 + \left(M_{2C_i} - M_{2V_j}\right)^2}$$



Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Minimum-Distanz-Klassifikator

Aufgabe: In einem 2D-Merkmalsraum (d.h. 2 Merkmale) werden 3 Cluster festgestellt mit folgenden Klassen-Flächenschwerpunkten:

$$C1: \quad M_{1_{C_1}} = 52 \quad M_{2_{C_1}} = 60$$

$$C2: \quad M_{1_{C2}} = 28 \quad M_{2_{C2}} = 38$$

$$C3: \quad M_{1_{C_3}} = 68 \quad M_{2_{C_3}} = 34$$

Neuer Merkmalsvektor: $M_{1_{V_1}} = 58 \quad M_{2_{V_1}} = 25$

Frage: Zu welchem Cluster (Klasse) gehört V_j ?

Lösung:

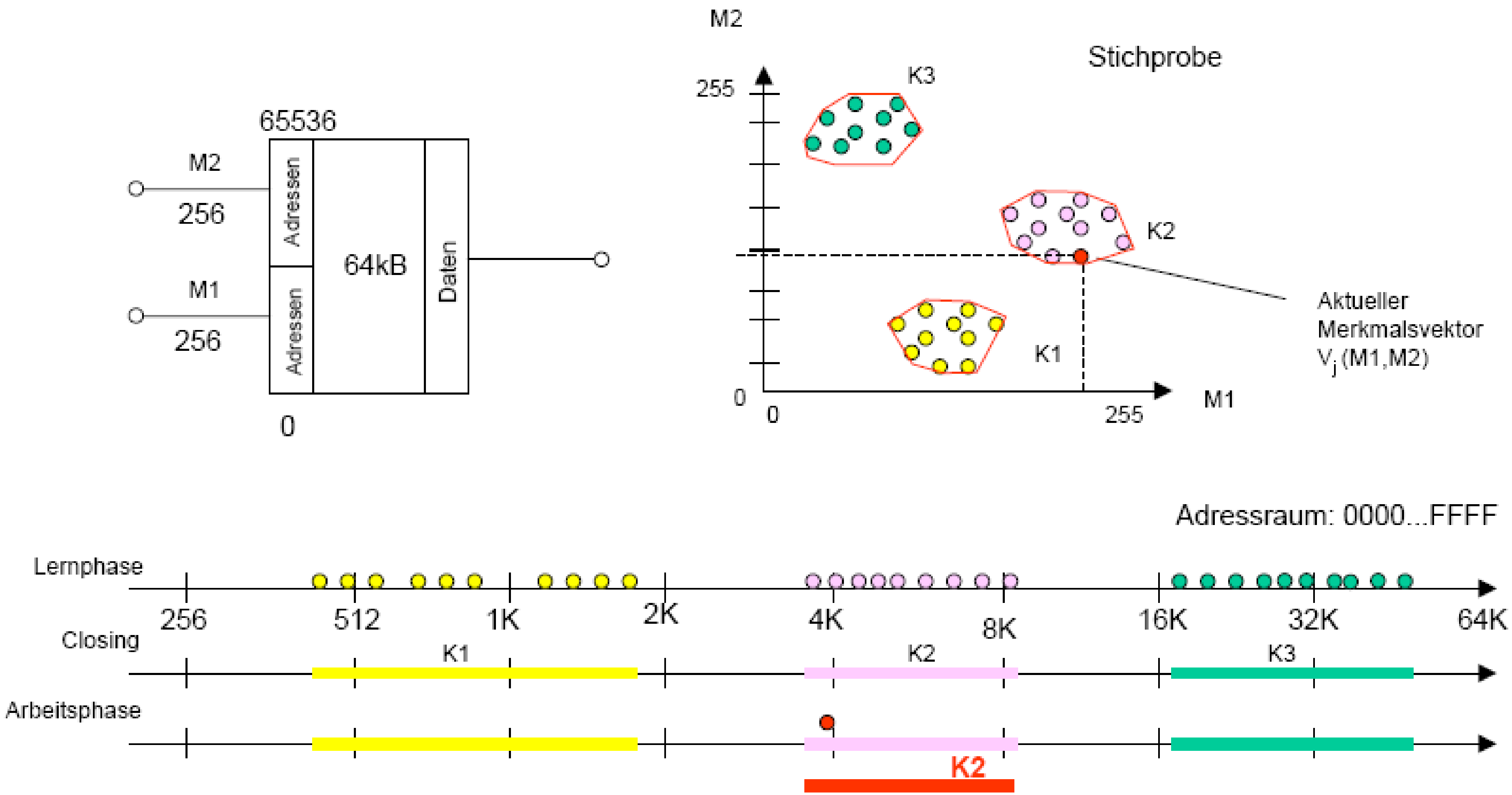
$$d_{e_1} = \sqrt{(52-58)^2 + (60-25)^2} = \sqrt{36+1225} = 35{,}5$$

$$d_{e_2} = \sqrt{(28-58)^2 + (38-25)^2} = \sqrt{900+169} = 32{,}7$$

$$d_{e_3} = \sqrt{(68-58)^2 + (34-25)^2} = \sqrt{100+81} = \underline{\underline{13{,}45}}$$

$$\mathrm{d}_{e_3} \rightarrow \text{Minimum} \Rightarrow \text{Vektor } V_j \text{ gehört zur Klasse3}$$

3.6.5 Klassifikation mittels LUT



Klassifikation mittels LUT / Veranschaulichung

Matrixspeicher:

Durch Anlegen der 2-dim. Adresse (Merkmalsvektor) erhält man mit dem Inhalt der Speicherzelle sofort die zugehörige Klasse

	0	1	2	3	4	…	250	251	252	253	254	255
0						:						
1		3	3	3		:						
2		3	3	3		:						
3		3	3	3		:						
4						:						
:	:	:	:	:	:	…	:	:	:	:	:	:
252						:			2	2	2	
253						:			2	2	2	
254						:			2	2	2	
255						:						

Beispiel:

der Merkmalsvektor [252;253] gehört zur Klasse 2

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Klassifikation mittels LUT / Veranschaulichung

Realisierung

- Entweder mittels Hardware (z.B. xy-adressierbarer Speicher)
- Oder per Software mittels ein- oder mehrdimensionaler Arrays

Klassifikation mittels LUT / Programmierbeispiel in C

Beispiel: siehe vorige Seite

```
void main()
:
:
int k, M[255][255];          Definition Klassenvariable k und Matrixspeicher M
:
k = M[252][253];
:
end.
```


Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.7 Statistische Klassifikatoren

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

3.7.1 Kurze Wiederholung zur Wahrscheinlichkeitsrechnung und mathematischen Statistik

Elementare Wahrscheinlichkeit (a-priori-Wahrscheinlichkeit)

$$p(E)=g/m$$

1. Die Wahrscheinlichkeit, bei einem Wurf eine 4 zu würfeln ist (bekannte Grundgesamtheit):

   p(4)=g/m= <u>1/6</u>

2. Die Wahrscheinlichkeit, bei einem Kartenspiel einen König zu ziehen ist (bekannte Grundgesamtheit):

   p(König)=4/32=1/8 =<u>0,125</u>

Additionssatz der Wahrscheinlichkeit „Entweder-Oder“ → OR

$$p=p(E_1)+p(E_2)\ldots+p(E_k)$$

Die Wahrscheinlichkeit, eine 3 oder eine 6 zu Würfeln, ist:

$p=1/6+1/6=\underline{1/3}$.

Die Wahrscheinlichkeit, einen König oder eine Dame zu ziehen, ist:

$p = 1/8 + 1/8 = \underline{1/4}$.

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Bedingte Wahrscheinlichkeit

Die Wahrscheinlichkeit des Ereignisses E_2 unter der Bedingung, dass E_1 schon eingetreten ist.

$$p(E_2/E_1)$$

Nachdem die 1. gezogene Karte ein König war und dieser nicht wieder zurückgegeben wurde, gibt die bedingte Wahrscheinlichkeit p(König2/König1) an, wieder einen König zu ziehen.

→ **p(Kö2/Kö1) = g/m = <u>3/31</u>**

Man sagt auch „W. $Kö_2$ geschnitten $Kö_1$".

Multiplikationssatz der Wahrscheinlichkeit für voneinander unabhängige Ereignisse

Die Wahrscheinlichkeit für das Auftreten zweier Ereignisse unter der Voraussetzung, dass diese voneinander unabhängig sind, ist

$$\mathbf{p(E_1 \& E_2) = p(E_1) \cdot p(E_{2/}).}$$

z.B. die Wahrscheinlichkeit für das zweimalige aufeinander folgende Würfeln einer „4“:

$$\mathbf{p(4 \,\&\, 4) = 1/6 \cdot 1/6 = \underline{1/36}}$$

z.B. die Wahrscheinlichkeit für das zweimalige aufeinander folgende Ziehen eines Königs unter der Vorauss., dass die erste Karte nach dem Ziehen wieder zurückgelegt wurde:

$$\mathbf{p(Kö \,\&\, Kö) = 1/8 \cdot 1/8 = \underline{1/64}}$$

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Multiplikationssatz der Wahrscheinlichkeit für voneinander nicht unabhängige Ereignisse

Die Wahrscheinlichkeit für das aufeinander folgende Eintreten zweier Ereignisse unter der Voraussetzung, dass diese voneinander nicht unabhängig sind (Verbund-wahrscheinlichkeit), ist

$$p(E_1 \& E_2) = p(E_1) \cdot p(E_2/E_1),$$

also das Produkt aus der elementaren Wahrscheinlichkeit des Eintretens von E_1 und der (bedingten)Wahrscheinlichkeit des Eintretens von E_2 unter der Voraussetzung, dass E_1 bereits eingetreten ist.

Beispiel: die Wahrscheinlichkeit für das zweimalige aufeinander folgende Ziehen eines Königs (ohne Zurücklegen der ersten Karte):

$$p(\text{Kö} \;\&\; \text{Kö}) = p(\text{Kö1}) \cdot p(\text{Kö2}/\text{Kö1}) = 1/8 \cdot 3/31 = \underline{0{,}012}$$

3.7.2 Maximum Likelihood-Klassifikator

(statistischer Klassifikator nach Thomas Bayes)

Allgemein gilt (die a priori Wahrscheinlichkeiten aller möglicher Ereignisse E_j müssen bekannt sein):

Die Wahrscheinlichkeit des Auftretens des Ereignisses E_i unter der Voraussetzung, dass das Ereignis E bereits eingetreten ist, beträgt (nach Bayes):

$$\boxed{p(E_i / E) = \frac{p(E_i) \cdot p(E / E_i)}{\sum_{j=1}^{k} p(E_j) \cdot p(E / E_j)}}$$

mit

$p(E_i)$	a-priori-W.
$p(E/E_i)$	bedingte W.
$p(E_i) \cdot p(E/Ei)$	Verbund-W.
$\sum_{j=1}^{k} p(E_j) \cdot p(E / E_j)$	Totale W.

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Maximum Likelihood-Klassifikator, angewendet auf das Ziehen einer bestimmten Karte aus einem von 2 möglichen Kartenspielen:

Gegeben sind zwei Kartenspiele K_1 und K_2; ***in letzterem sind die vier Buben durch 4 Könige ersetzt*** . Die a priori Wahrscheinlichkeiten des Zugriffs auf beide Kartenspiele seien gleichgroß (d.h. sie betragen jeweils 50%). Es werde ein König gezogen. Wie groß ist die Wahrscheinlichkeit, dass der König aus dem Kartenspiel K_1 stammt?

Es lässt sich formulieren:

$p(K_1) = p(K_2) = 0{,}5$ **(a priori Wahrscheinlichkeiten für den Zugriff auf eine der beiden Kartenspiele)**

$p(Kö/K_1) = 4/32 = 0{,}125$ **(Wahrscheinlichkeit, einen König zu ziehen unter der Voraussetzung des Zuges aus K_1)**

$p(Kö/K_2) = 8/32 = 0{,}25$ **Wahrscheinlichkeit, einen König zu ziehen unter der Voraussetzung des Zuges aus K_2)**

Wahrscheinlichkeit, aus einem der beiden Kartenspiele einen König zu ziehen, → ***totale Wahrscheinlichkeit:***

$p(Kö) = p(K_1) \cdot p(Kö/K_1) \quad + p(K_1) \cdot p(Kö/K_1) = 0{,}5 \cdot 0{,}125 + 0{,}5 \cdot 0{,}25 = 0{,}1875$

Nach Bayes beträgt dann die Wahrscheinlichkeit dafür, dass ein gezogener König aus dem Kartenspiel K1 stammt:

$$p(K_1/Kö) = \frac{p(K_1) \cdot p(Kö/K_1)}{p(K_1) \cdot p(Kö/K_1) + p(K_2) \cdot p(Kö/K_2)} = \frac{0{,}5 \cdot 0{,}125}{0{,}5 \cdot 0{,}125 + 0{,}50 \cdot 0{,}25} = \underline{\underline{0{,}33}}$$

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

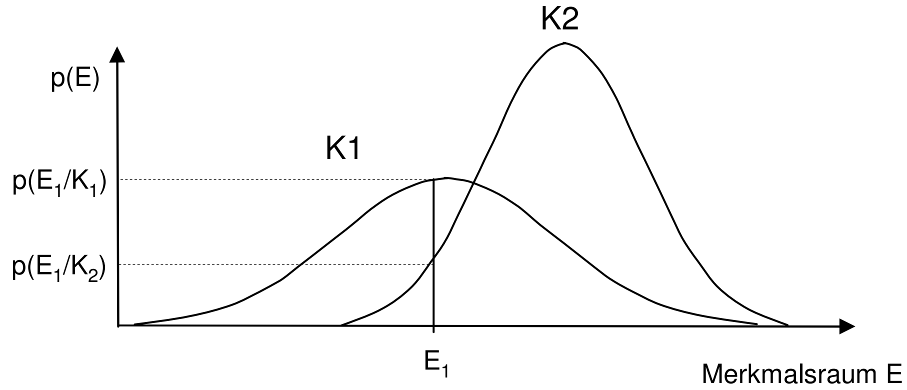
Maximum Likelihood-Klassifikator, angewendet auf einen eindimensionalen Merkmalsraum mit zwei Klassen und a priori-Wahrscheinlichkeiten für die Klassenzugehörigkeit

Die wahrscheinlichen Klassenzugehörigkeiten werden durch die Häufigkeitsverteilungen (Wahrscheinlichkeitsdichtefunktion) entlang der Merkmalsachse dargestellt.

Wahrscheinlichkeit der Zugehörigkeit zur Klasse K_i nach Bayes bei einem gemessenen Merkmal E (hier für E = E_1):

$$p(K_i / E_1) = \frac{p(K_i) \cdot p(E_1 / K_i)}{\sum_{j=1}^{k} p(K_j) \cdot p(E_1 / K_j)}$$

zwei Klassen: → j =1,2



---

p(K_i) a priori Wahrscheinlichkeit der Zugehörigkeit zur Klasse i

p(K_i/E) Wahrscheinlichkeit der Klassenzugehörigkeit zur Klasse i unter der Voraussetzung einer ermittelten Merkmalsgröße E

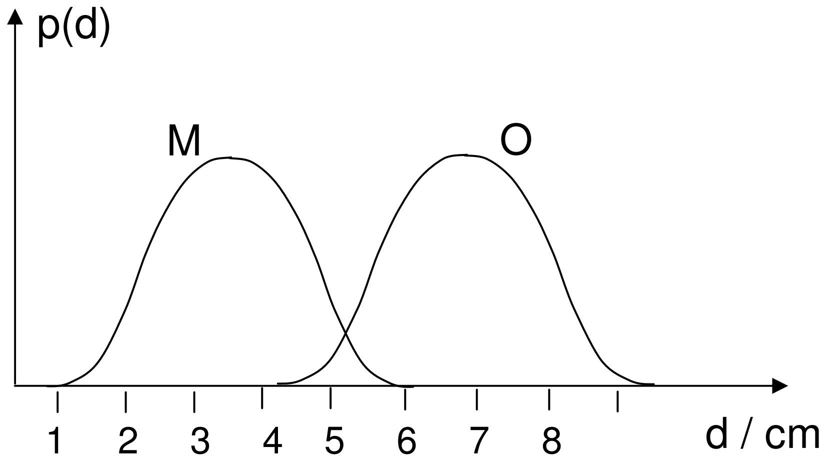
p(E/K_i) Wahrscheinlichkeit für das Auftreten einer Merkmalsgröße E unter der Voraussetzung der Zugehörigkeit zur Klasse i

Wahrscheinlichkeit der Zugehörigkeit zu einer der Klassen j .. k unter der Voraussetzung eines bestimmten Wertes von E (totale Wahrscheinlichkeit):

$$\sum_{j=1}^{k} p(K_j) \cdot p(E / K_j)$$

Maximum Likelihood-Klassifikator / Beispiel Obstsortieranlage

Früchte sind nach Orangen und Mandarinen zu sortieren. Als Parameter wird lediglich der Durchmesser herangezogen (eindimensionaler Merkmalsraum E). Durch Stichprobenmessungen wurden die folgenden Stichprobenverteilungen für die Klassen M (Mandarinen) und O (Orangen) ermittelt:



Weiterhin ist bekannt, dass durchschnittlich doppelt soviel Orangen wie Mandarinen anfallen. Demzufolge betragen die a priori-Wahrscheinlichkeiten für beide Klassen:

$$p(M) = 1/3 \text{ und } p(O) = 2/3.$$

Es ist die Wahrscheinlichkeit der Zugehörigkeit zu den beiden Klassen bei einer Frucht mit einem ermittelten Durchmesser d_1 zu bestimmen.

Maximum Likelihood-Klassifikator / Beispiel Obstsortieranlage / Forts.

Lösung

Nach der Bayesschen Formel gilt für die aposteriori Wahrscheinlichkeit, dass es sich um eine Mandarine (Klasse M) handelt, wenn der Durchmesser = d_1 ist:

$$p(M/d_1) = \frac{p(d_1/M) \cdot p(M)}{p(d_1/M) \cdot p(M) + p(d_1/O) \cdot p(O)}$$

Hierin sind $p(d_1/O)$ und $p(d_1/M)$ jeweis die bedingten Wahrscheinlichkeiten, dass die Früchte bei einem Durchmesser d_1 zur Klasse O bzw. M gehören (abzulesen aus den Stichproben-verteilungen) und p(O) bzw. p(M) die a priori-Wahrscheinlichkeiten der Klassenzugehörigkeiten. Der Nenner dient lediglich der Normierung auf eine maximale Wahrscheinlichkeit von 1.

Entsprechend gilt für die aposteriori Wahrscheinlichkeit der Zugehörigkeit zur Klasse O (Orangen) bei einem Durchmesser d_1:

$$p(O/d_1) = \frac{p(d_1/O) \cdot p(O)}{p(d_1/O) \cdot p(O) + p(d_1/M) \cdot p(M)}$$

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Maximum Likelihood-Klassifikator / Beispiel Obstsortieranlage / Forts.

Nochmal zur Erläuterung

Wahrscheinlichkeit des Eintretens des Ereignisses M unter der Voraussetzung, dass d1 bereits eingetreten ist (bedingte Wahrscheinlichkeit)

Produkt aus der Wahrscheinlichkeit des Eintretens des Ereignisses d1 unter der Voraussetzung, dass M bereits eingetreten ist (bedingte Wahrscheinlichkeit) und der apriori Wahrscheinlichkeit des Eintretens von M (Verbundwahrscheinlichkeit)

$$p(M/d_1) = \frac{p(d_1/M)\cdot p(M)}{p(d_1/M)\cdot p(M) + p(d_1/O)\cdot p(O)}$$

Hier:

Wahrscheinlichkeit der Zugehörigkeit eines Objektes zur Klasse Mandarinen unter der Voraussetzung, dass der Durchmesser d_1 ermittelt wurde

Hier:

Produkt aus der Wahrscheinlichkeit der Ermittlung eines Durchmessers d_1 unter der Voraussetzung, dass es sich beim zugehörigen Objekt um eine Mandarine handelt und der apriori Wahrscheinlichkeit der Zugehörigkeit zur Klasse Mandarinen

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!

Maximum Likelihood-Klassifikator / Obstsortieranlage mit Zahlenbeispiel

Gegeben sind die apriori Wahrscheinlichkeiten für Mandarinen mit 1/3 und für Orangen mit 2/3 sowie die relativen Häufigkeiten (siehe Diagramm). Es ist ein Objekt mit einem Durchmesser d von 5cm zu klassifizieren.

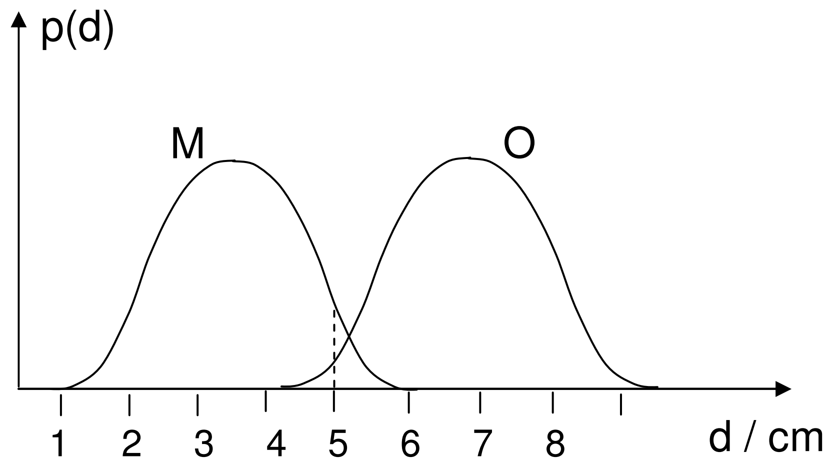


Abb: relative Häufigkeiten für das Auftreten des Durchmessers d bei Mandarinen (Klasse M) und Orangen (Klasse O)

Lösung:

Aus dem Diagramm wird abgelesen:

- p(5cm/M)=0,25 für Mandarinen;
- p(5cm/O) = 0,12 für Orangen.

Ermitteln der wahrscheinlichen Klassenzugehörigkeiten nach der Bayesschen Formel:

$$p(M/5cm) = \frac{0{,}25 \cdot 1/3}{0{,}25 \cdot 1/3 + 0{,}12 \cdot 2/3} \approx 0{,}51$$

$$p(O/5cm) = \frac{0{,}12 \cdot 2/3}{0{,}25 \cdot 1/3 + 0{,}12 \cdot 2/3} \approx 0{,}49$$

Demzufolge wird das Objekt der Klasse Mandarinen zugeordnet, wenn auch die ermittelte Wahrscheinlichkeit nur unwesentlich höher ist als diejenige der Zugehörigkeit zur Klasse Orangen.

Maximum Likelihood-Klassifikator / Verallgemeinerung auf m Merkmale (Merkmalsvektor $\vec{x}$) und n Klassen:

Die Wahrscheinlichkeit der Zugehörigkeit zu einer Klasse K_q ergibt sich dann zu:

$$p(K_q / \vec{x}) = \frac{\prod_{i=1}^{m} p(x_i / K_q) \cdot p(K_q)}{\sum_{k=1}^{n} \prod_{i=1}^{m} p(x_i / K_k) \cdot p(K_k)}$$

3.7.3 Bayes-Klassifikator

a) Wie Maximum-Likelihood-Klassifikator

b) Zusätzliche Definition eines „bedingten Risikos“ für eine Fehlklassifikation mit

$$R(K_i / \vec{x}) = \sum_{j=1}^{n} I_{ij} \cdot p(K_j / \vec{x})$$

-R(Ki/$\vec{x}$): „Verlustfaktor“ für die fälschliche Zuordnung eines Objektes der Klasse Ki zu einer anderen Klasse

- p(Kj/$\vec{x}$): aposteriori Wahrscheinlichkeit der Zugehörigkeit zur Klasse Kj (entsprechend Klassifikation nach dem Maximum Likelihood-Verfahren)

-I_{ij}: Kostenmatrix, Verlustmatrix: enthält in den Feldern i,j jeweils den quantifizierten Verlust, welcher entsteht, wenn man ein Objekt der Klasse K_i fälschlicherweise der Klasse K_j zuordnet.

c) Zuordnung jedes Objektes zur Klasse mit dem kleinsten bedingten Risiko („Bayes‘sche Entscheidungsregel“)

Anwendung des Bayes-Klassifikators auf das Beispiel Obstsortieranlage

Die mittleren Kosten für Falschzuweisungen (ermittelt z.B. durch Marketingspezialisten, BWL-Fachleute und/oder Statistiker) sollen betragen:

- für die Zuweisung eines Objektes der Klasse Mandarine der Klasse Orange: 5 ct.
- Für die Zuweisung eines Objektes der Klasse Orange der Klasse Mandarine: 10 ct.

Dann lässt sich folgende Kostenmatrix aufstellen (Klasse 1 Mandarinen, Klasse 2 Orangen):

$$\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{j \rightarrow} \\
 & & 1 & 2 \\ \hline
 & 1 & 0 & 5 \\ \hline
i \downarrow & 2 & 10 & 0 \\ \hline
\end{array}$$

l(i,j) in ct.

Es ergeben sich folgende Verlustfaktoren für die Zuweisung zu den beiden Klassen

$$R(M/d_1) = 0 \cdot p(M/d_1) + 5ct. \cdot p(O/d_1)$$
$$R(O/d_1) = 10ct. \cdot p(M/d_1) + 0 \cdot p(O/d_1)$$

und mit Zahlenwerten (siehe Beispiel unter 3.7.2):

$$R(M/d_1) = 0 \cdot 0{,}51 + 5ct. \cdot 0{,}49 = \underline{\underline{2{,}45ct.}}$$
$$R(O/d_1) = 10ct. \cdot 0{,}51 + 0 \cdot 0{,}49 = \underline{\underline{5{,}1ct.}}$$

Somit wird man auch in diesem Falle das Objekt, jetzt sehr eindeutig, der Klasse Mandarinen zuordnen, da der Verlustfaktor nur ca. halb so groß ist wie derjenige für die Zuweisung zur Klasse Orangen. Das Ergebnis würde allerdings im Falle einer anderen Kostenmatrix (diese wurde hier ja nur beispielhaft angenommen) auch ein völlig anderes sein.

Nur für Lehrzwecke zum internen Gebrauch an der FH Jena bestimmt!