

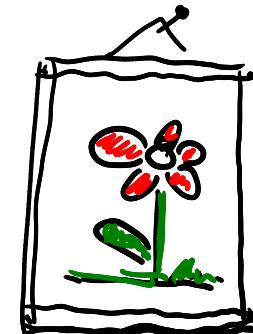
MACHINE LEARNING

I. Intro: Was ist Machine Learning?

 **FRAGE** Was ist das für ein Objekt??

Wie kommen wir zu unserer Entscheidung?

Wie haben wir gelernt so zu entscheiden?



Welche Definition würden wir formulieren?

- KRITERIEN

ENTScheidungs - Prozeß

- KLAR:
- wir studieren keine formalen Definitionen
 - wir haben viele Beispiele für Blumen, Bäume, Häuser, ... gesehen
 - wir übertragen unsere Erfahrungen auf neue Objekte



Def.:

Machine Learning befasst sich mit Verfahren, die zum Ziel haben
...
Inferenz zu automatisieren

1.1) Induktives & deduktives Schließen

VERSUCH

DEDUKTIV

- gehe vom Allgemeinen zum Speziellen
- ziehe logische Schlüsse aus allg. Erkenntnissen

INDUKTIV

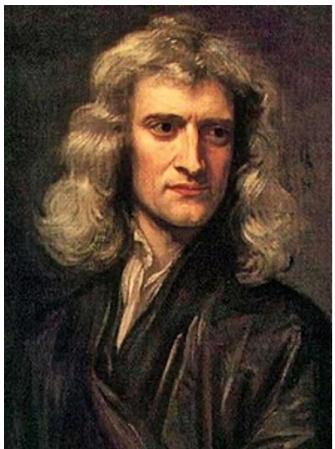
- gehe vom Speziellen zum Allgemeinen
- leite allgemeine Regeln aus Beobachtungen ab

MASCHINE

Newton (1643-1727)

bewies das Theorem mit

Hilfe von
mathemati-
schen und
physikalischen
AXIOMEN



Theorem

Planeten bewegen
sich auf Ellipsenbahnen
um die Sonne!



Kepler (1571-1630)

verfasste Beobachtungen
aus und
kam so
auf das
Theorem
(kein Beweis!).

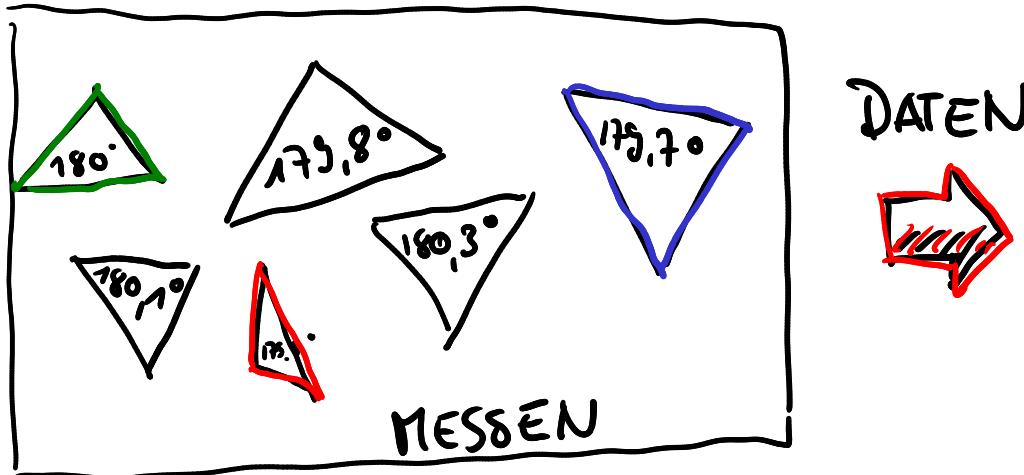


Lernen aus Daten vs. Einsicht

(3)

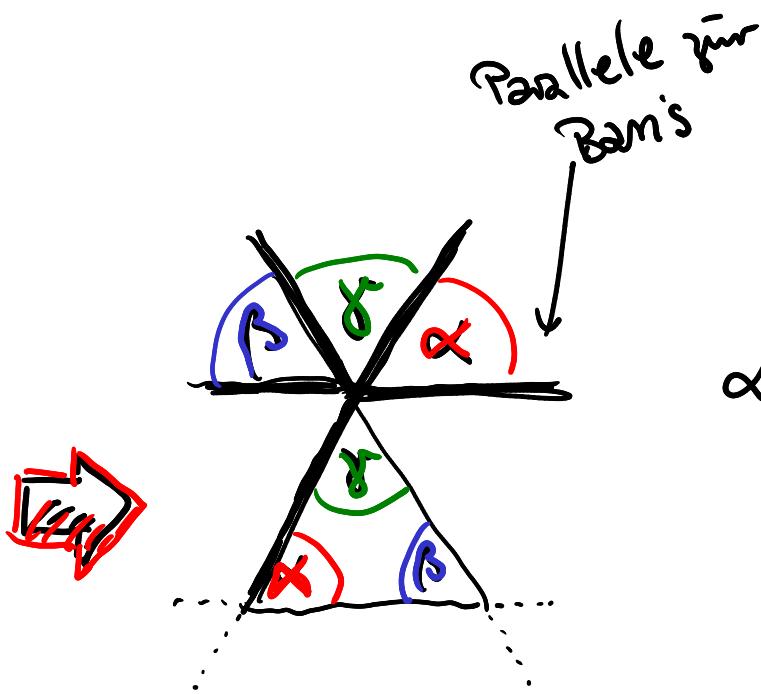
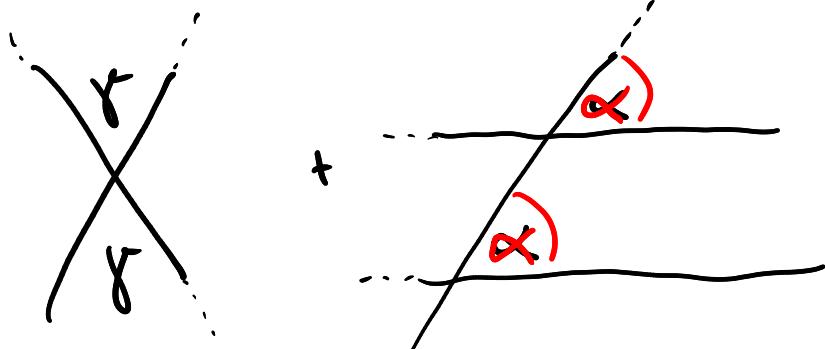
Bsp. 1) Die Winkelsumme im Dreieck ist 180°

$$\alpha + \beta + \gamma = 180^\circ$$



Zu Mittel ergibt sich eine Summe von $180^\circ \approx \alpha + \beta + \gamma$

AXIOME DER GEOMETRIE



$$\alpha + \beta + \gamma = 180^\circ$$

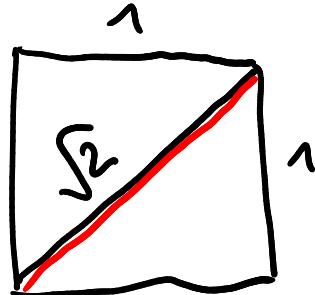
T



EXAKT

und für alle \triangle dieser Welt gültig !!

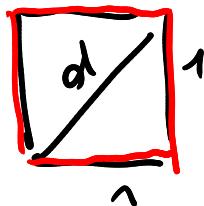
Bsp. 2)



Länge der Diagonale ist $\sqrt{2}$ und eine irrationale Zahl

④

→ Mengen mit unterschiedlicher
Nähergenauigkeit



⇒

$$\begin{aligned}d &= 1,4 \\d &= 1,41 \\d &= 1,414 \\&\vdots\end{aligned}$$



← sind alle rational

→ d ist rational

geht nicht!

SATZ von Pythagoras (570 - 510 v. Chr.)

$$d = \sqrt{1^2 + 1^2} = \sqrt{2}$$

→ Angenommen $\sqrt{2}$ ist rational $\Rightarrow \sqrt{2} = \frac{p}{q} \Rightarrow p^2 = 2q^2$
 $\Rightarrow p$ und p^2 sind gerade $\Rightarrow p = 2m \Rightarrow 4m^2 = p^2 = 2q^2 \Rightarrow q, q^2$ gerade
aber dann könnte man bei p/q mit 2 kürzen ↴

Beobachtung: Machine Learning kann (noch) nicht verwendet werden um
Einsicht zu erlangen. Das leistet nur die menschliche Intelligenz!
Schlussfolgerungen aus Daten sind nur so wert wie die Daten selbst (sh. Bsp. 2)

Def: Machine Learning - revisited

(5)

Machine Learning befasst sich mit Methoden, die induktive Inferenz automatisieren.

• Beobachtungen

- Trainingsbeispiele als Input für Lernalgorithmen
- Muster werden vom Algorithmus automatisch in den Daten gefunden
- Schlussfolgerungen in Form von allgemeinen Hypothesen in Bezug auf die Datengesamtheit werden abgeleitet
- Anwendung der Hypothesen auf unbekannten Daten ermöglicht Prognosen (Regression, Klassifikation, ...)

1.2) Szenarios des maschinellen Lernens

Unsupervised Learning

ZIEL: Lerne die Datenstruktur nur auf Basis von Inputs

KEIN Target



Supervised Learning

ZIEL: Lerne eine Transferfunktion $f(X) = Y$ basierend auf Input\Output Beispielen TARGET behaumt

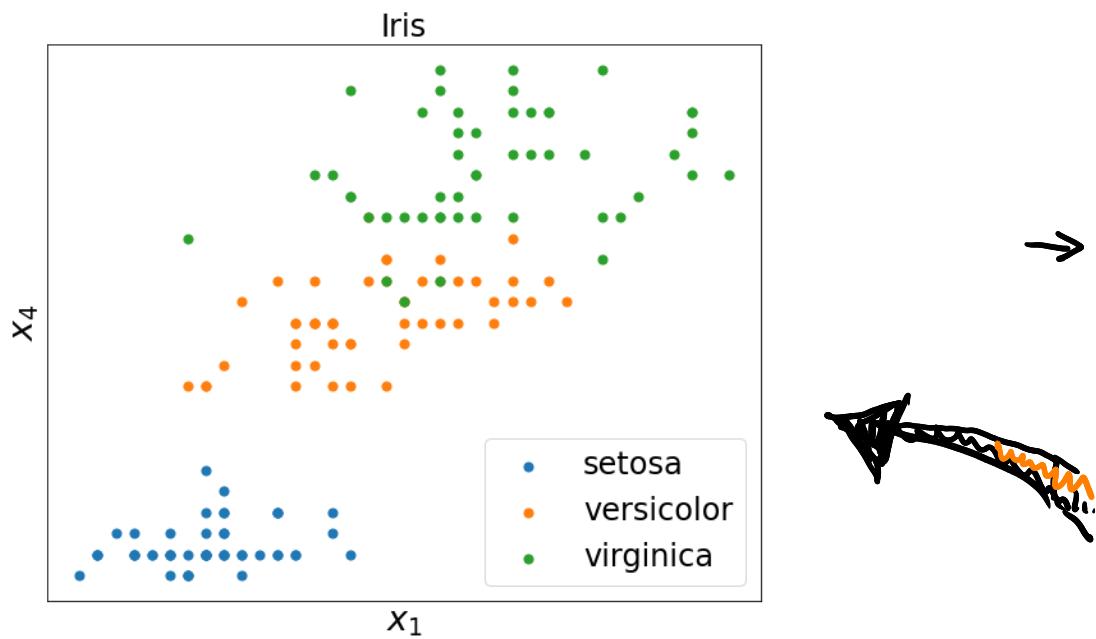
Reinforcement Learning

ZIEL: Entwickle einen "Agenten", der seine Performance basierend auf Interaktionen mit der Umgebung optimiert. Info über den Systemzustand \rightarrow REWARD-Funktion

1.2.1) Supervised Learning (Unser Haupt-Topic 😊)

7

Bsp.: Iris-Daten



→ Betachte nur die Features X_1 und X_4
→ damit wird eine graphische Darstellung möglich

SCATTER PLOT

- 4 Attribute X_1, \dots, X_4 (Features)
- 3 Klassen
- 50 Bsp. pro Klasse

FARBE repräsentiert die Klasse

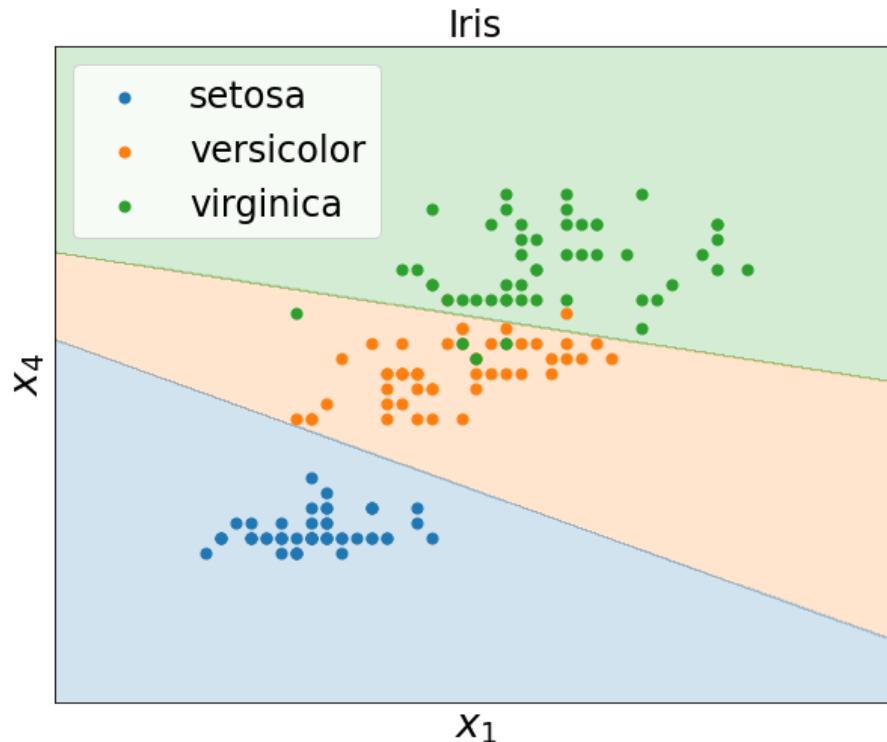
 ZIEL: Klassifikation!

gegeben: Input - Output Samples

$$S: (x_1, y_1), \dots, (x_n, y_n) \in \mathbb{R}^2 \times \{0, 1, 2\}$$

klassen

DATA



Finde eine Funktion

$h: \mathbb{R}^2 \rightarrow \{0, 1, 2\}$, die einen neuen Punkt (x_i, y_i) so gut wie möglich einer der 3 Klassen zuordnet

hier
eigentlich

Aufgabe: Zeite Regeln in Bezug auf x_1, x_4 ab (Entscheidungsgrenzen) mit denen man zuverlässig zuordnen kann!

Anwendungen:

1) Spam - Filter (Klassifikation)

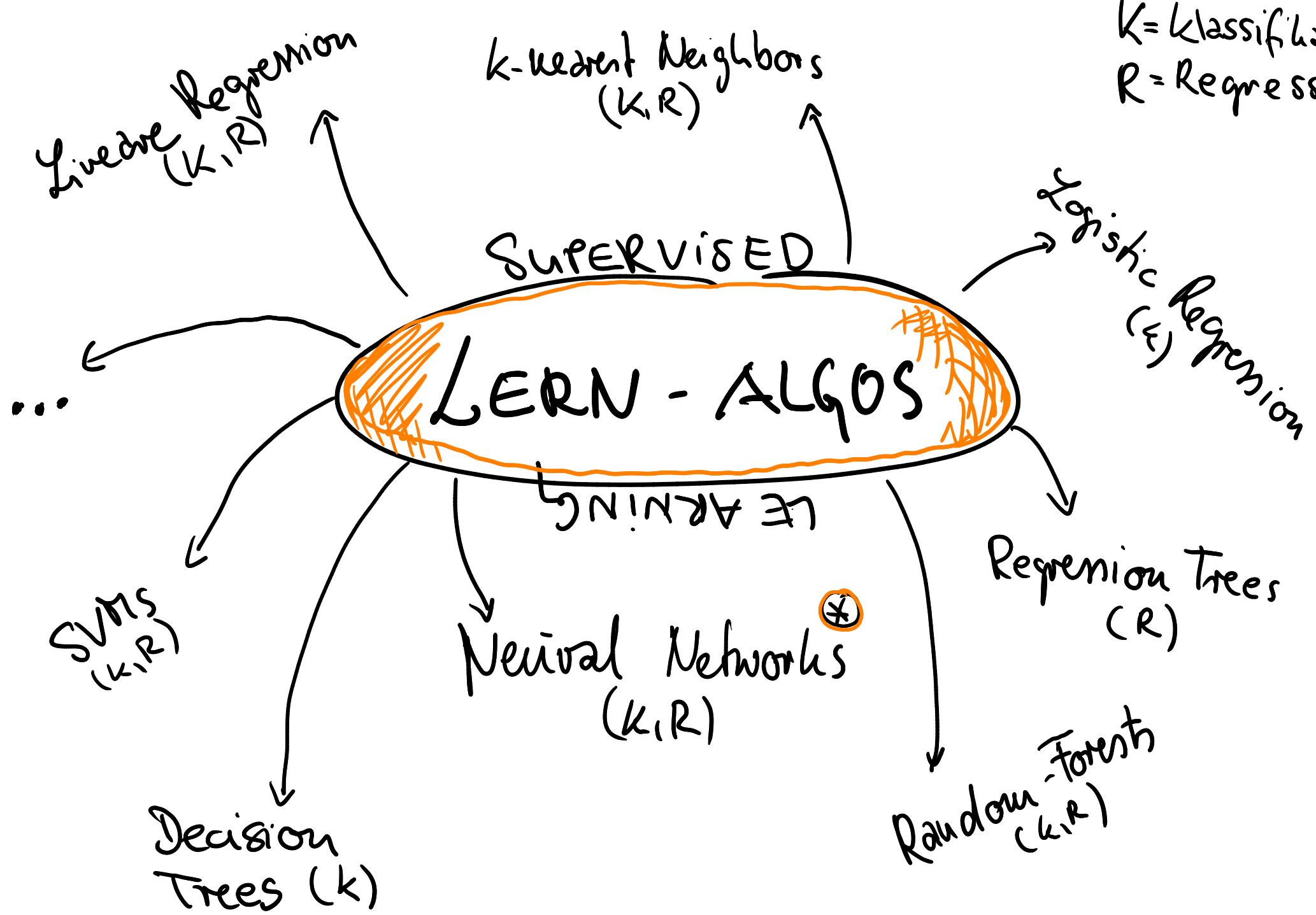
- Trainingsdaten: gelabelte Mails ($\text{spam} = 1$, no spam = 0)
- Modell: soll für unbekanntes neues Mail eine Aussage treffen

binär
↓

2) Preisprognose (Regression)

- Trainingsdaten: historische Preise & Zeitreihen bekannter Einflussgrößen
- Modell: lernt Zusammenhang zwischen erklärenden Größen und Zielgröße und leitet daraus Prognosen ab

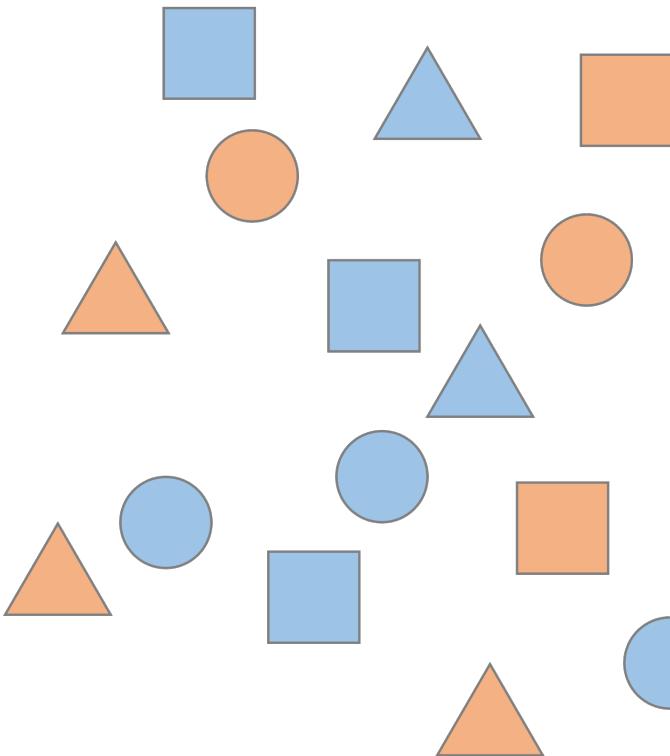
K = Klassifikation
R = Regression



* auch unsupervised möglich ...

1.2.2) Unsupervised Learning

Bsp.: farbige Formen



Gegeben

- Input: Form + Farbe
- Label: keins



- Was kann man aus den Daten lernen?
- Gibt es Strukturen?

CLUSTERING

- grappiere die Formen in Cluster

① Formen in gleichem Cluster \Rightarrow ähnlich

② Formen in verschiedenen Clustern \Rightarrow unähnlich

FRAGEN

Was heißt "ähnlich"?

Wie viele Cluster?

- Welcher Algo?

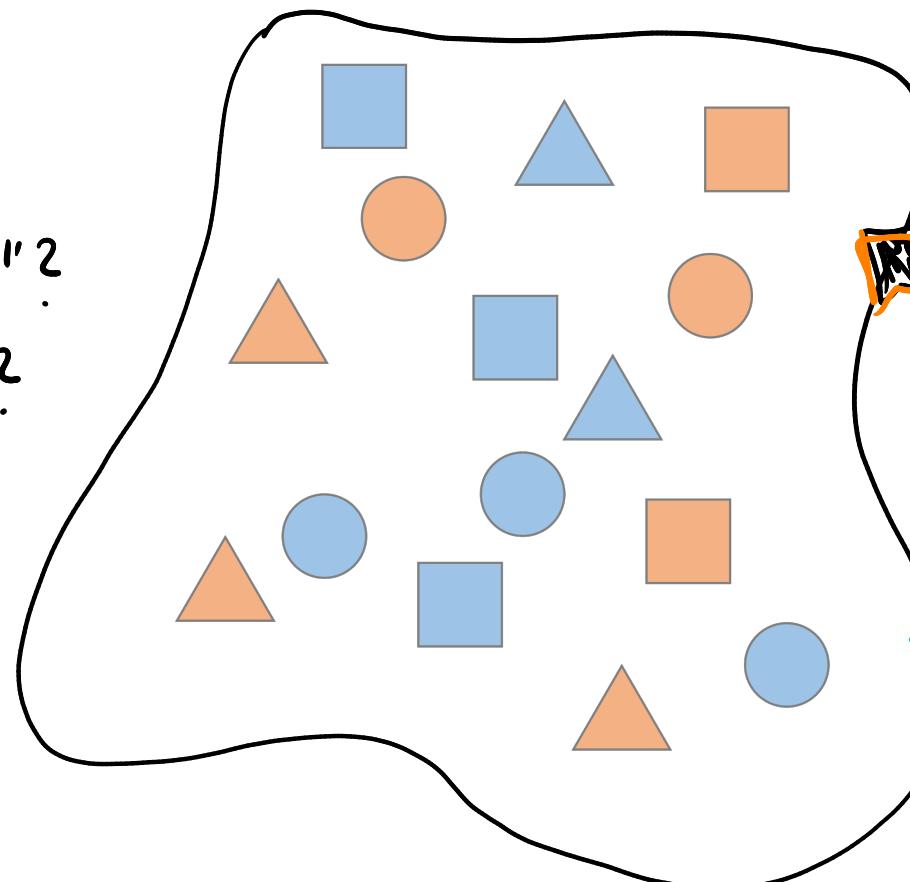
:

Lösung 1

C_1 = Quadrate

C_2 = Kreise

C_3 = Dreiecke



Lösung 2

C_1 = orange
 C_2 = blau

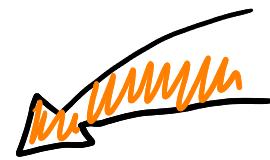
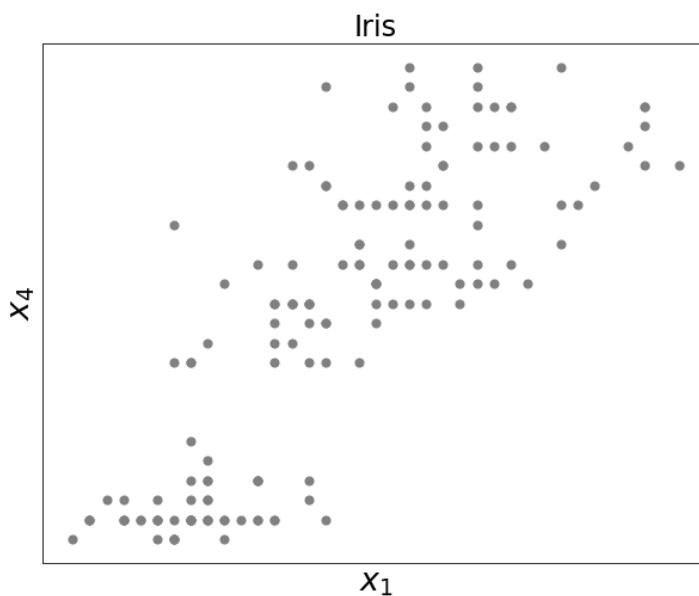
Lösung 3

C_1 = Quadrat blau
 C_2 = Quadrat orange
:
 C_6 = Dreieck orange

Ergebnis:

- ⇒ Lösung hängt von der Definition von "ähnlich" ab!
- ⇒ Lösung hängt vom Algorithmus ab

Bemerkung: Im ersten Beispiel ist es noch möglich das Ergebnis des Clusterings zu bewerten, aber ...



Iris - Daten, Features X_1, X_4
ohne Labels



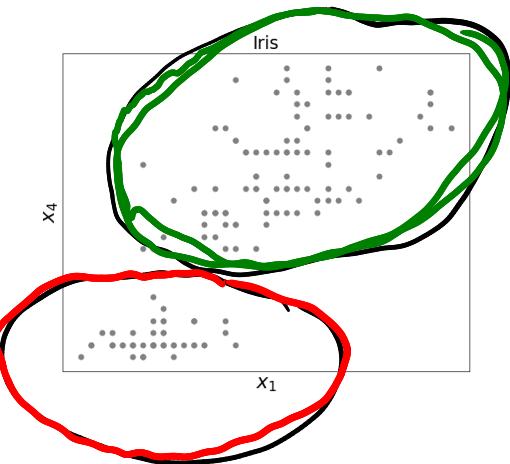
ZIEL
Daten in k Cluster
einteilen



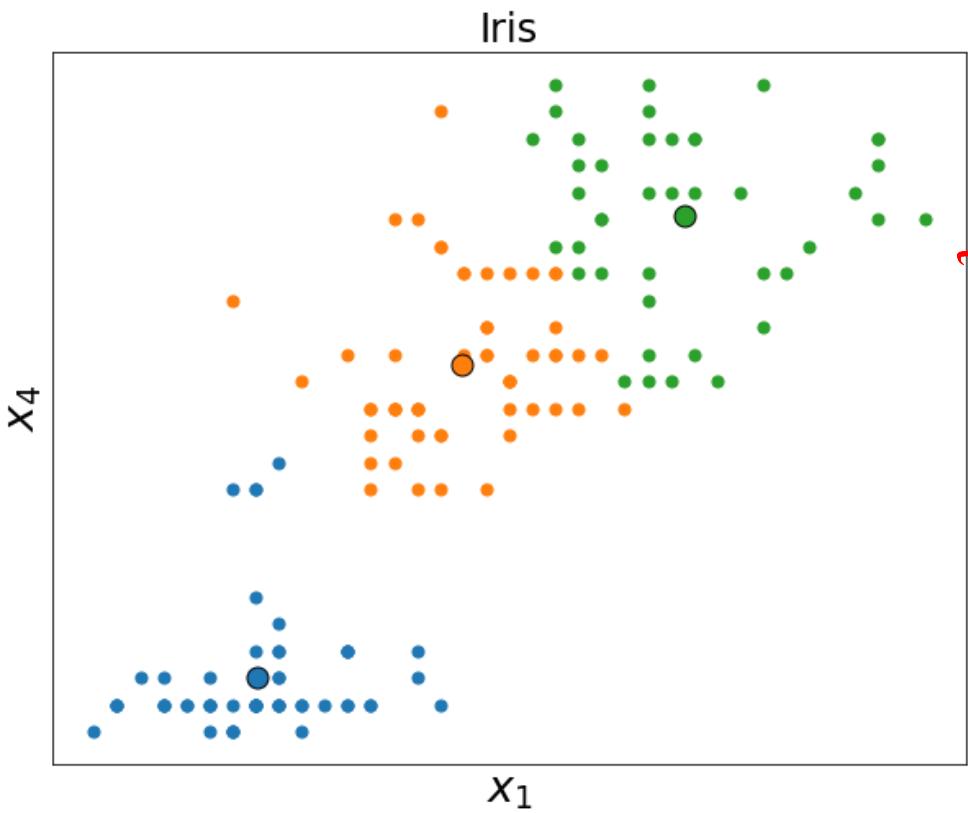
- wie viele Cluster?
- welcher Algorithmus?
- Bewertung? Man hat ja keine gelabelten Bsp.e ...

$\rightsquigarrow k=3$ Cluster

man sieht
wahrscheinlich
zunächst nur 2



14



k-means Clustering

x gehört zu Cluster c mit
Zentrum μ_c , wenn es näher zu
 μ_c liegt als zu den anderen
Zentren

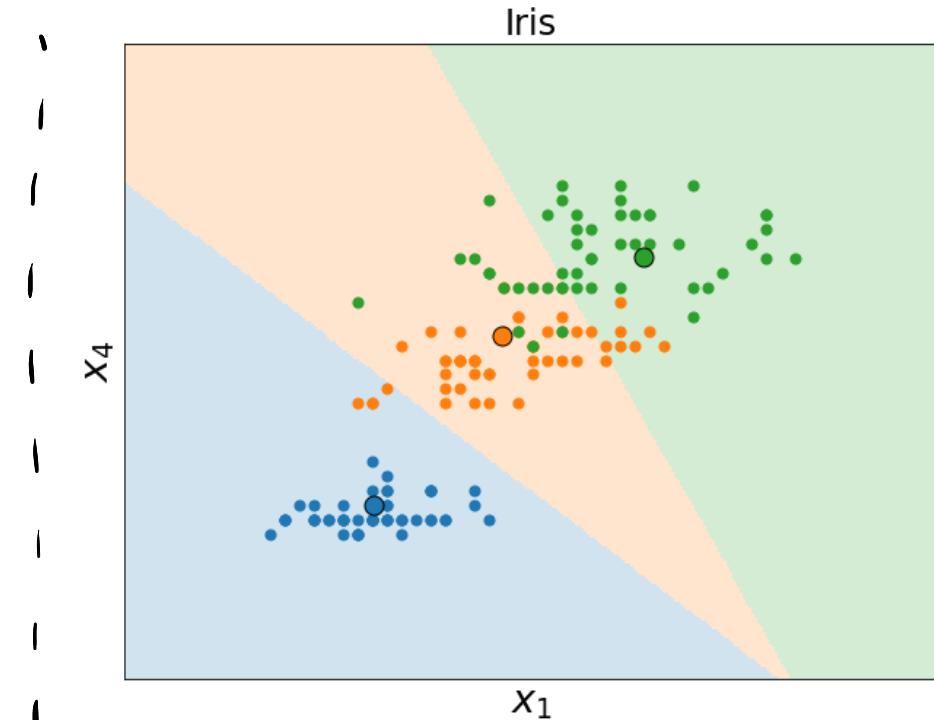
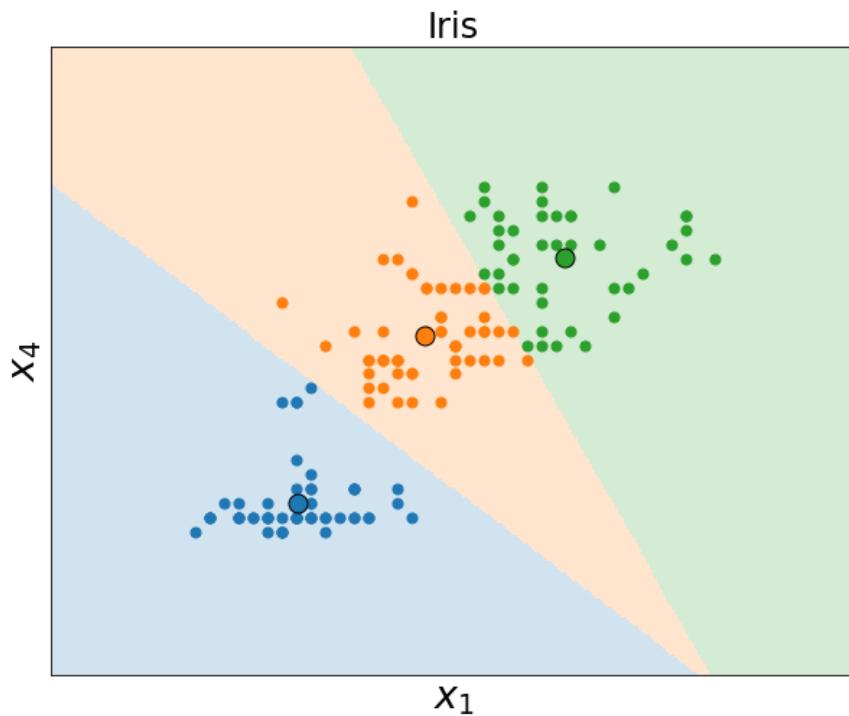
Anzahl $k=3$ der Cluster muss
vorgegeben werden ...



Bemerkung: Cluster sind keine Klassen! Wir hoffen nur, daß es passt

Ergebnis: Vergleich des Clusterings mit den echten Klassen

15



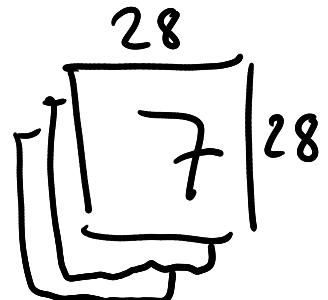
- Entscheidungsregionen des k-means sind farbig hinterlegt
- Cluster-Zentren = kringel

- wahre Klasse entspricht Farbe der Punkte
- blau ist gut getroffen 😊
- orange + grün schon weniger...

Bei niedrig-dimensionalen Daten kann man das Ergebnis immerhin noch graphisch auf Plausibilität hin überprüfen.

Schwierig wirds bei Daten im \mathbb{R}^n !

Bsp.: MNIST



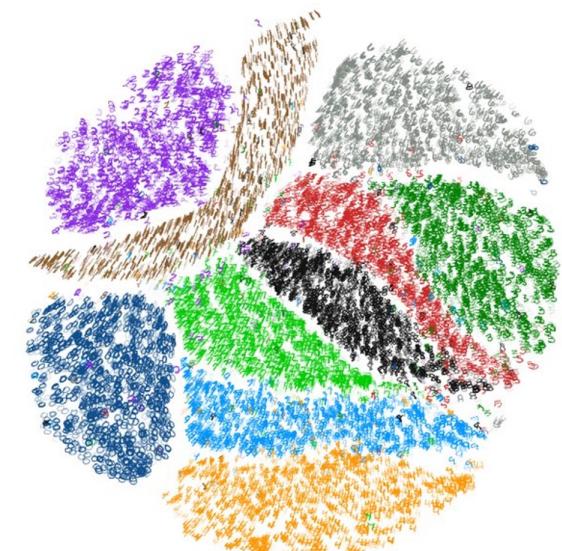
← handgeschriebene Ziffern 0, 1, ..., 9
← 28×28 Pixel

N.Pezzotti (2019)



Gibt es Strukturen, die typisch für die Klassen sind?

- Input: hochdim. Daten ohne Label im $\mathbb{R}^2 \Rightarrow$
- Für Visualisierung braucht man eine 2D \ 3D Repräsentation
- Struktur der Daten muss dabei möglichst erhalten bleiben.



00 11 22 33 44 55 66 77 88 99
00 11 22 33 44 55 66 77 88 99

Anwendungen:

- Image Segmentation:
Teile ein Bild in Regionen ein
- Analyse sozialer Netzwerke
Finde Communities
- Marktanalyse
- ...

Jordan (2010)

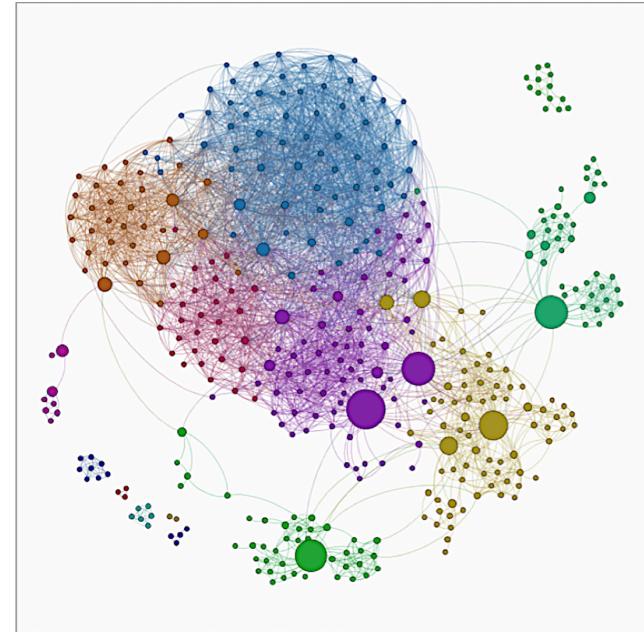


segment



(17)

R. Parkin (2014)



Visualisierung:

Man will eine niedrig-dimensionale Repräsentation der Daten finden, die "bedeutungsvoll" ist.

