



Algorithmen'n Blues

Prof. Dr. Katharina Zweig, TU Kaiserslautern

zweig@cs.uni-kl.de

Under CC-BY 3.0 SA

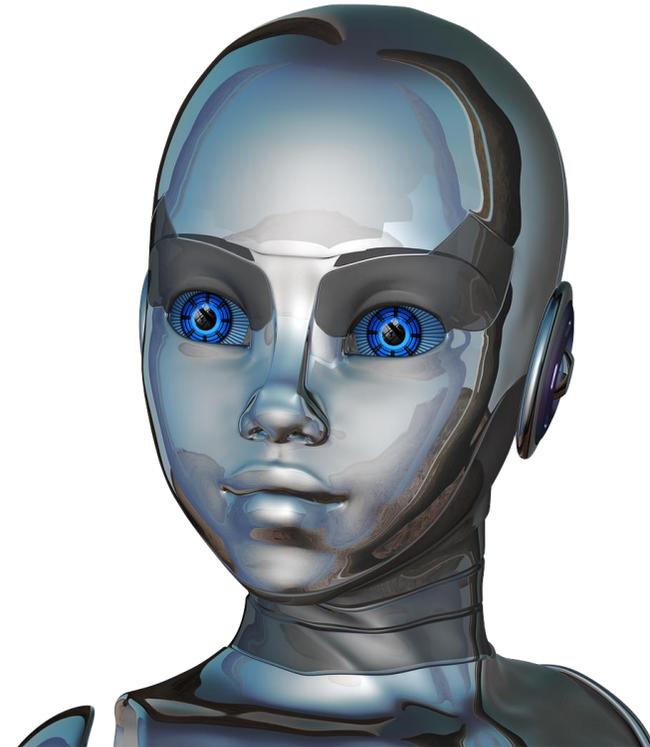
Die jeweiligen Bildrechte sind angegeben oder es handelt sich – nach bestem Wissen und Gewissen um CC 0 Bilder (meistens von Pixabay.com)

Können Computer
lernen?

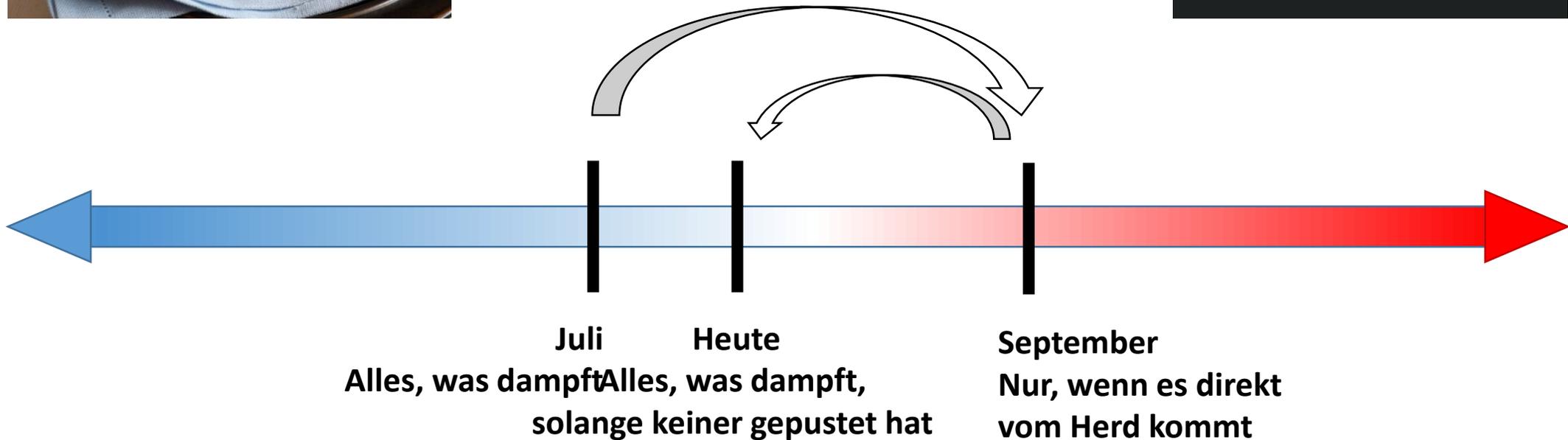


Was heißt Lernen?

- In derselben Situation ein vorher gezeigtes Verhalten wiederholen.
- In derselben Art von Situation das richtige Verhalten aus einer Reihe von Möglichkeiten auswählen.



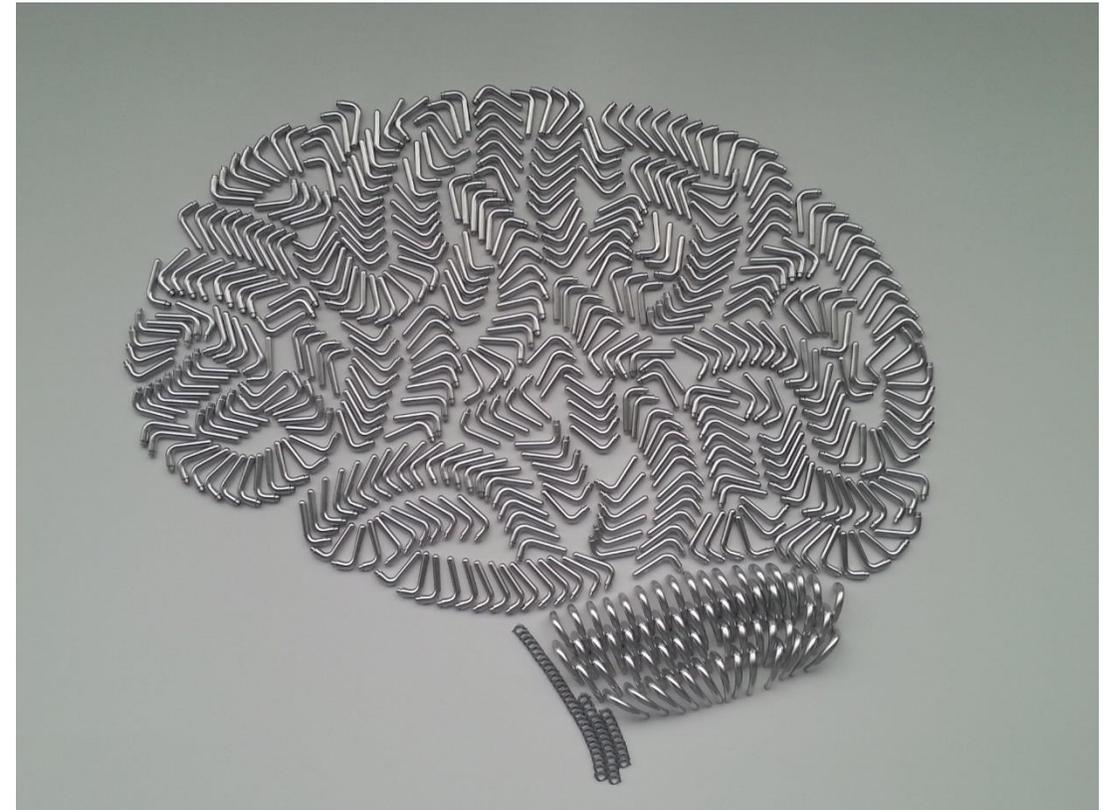
Sebastian lernt „heiss“ und „warm“



Sebastian lernt...



- Durch **Rückkopplung**: unerwartet heiß, unerwartet kalt
- Durch **Speicherung in einer Struktur**: in Neuronen und deren Verknüpfung.
- Durch **Generalisierung des Gelernten**.

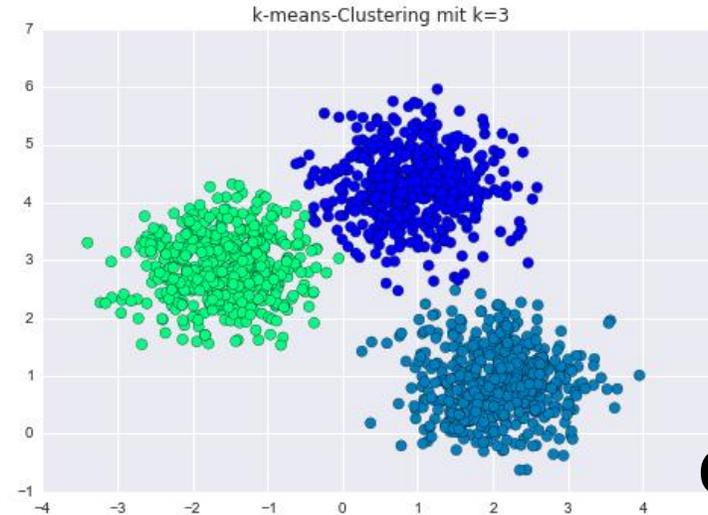
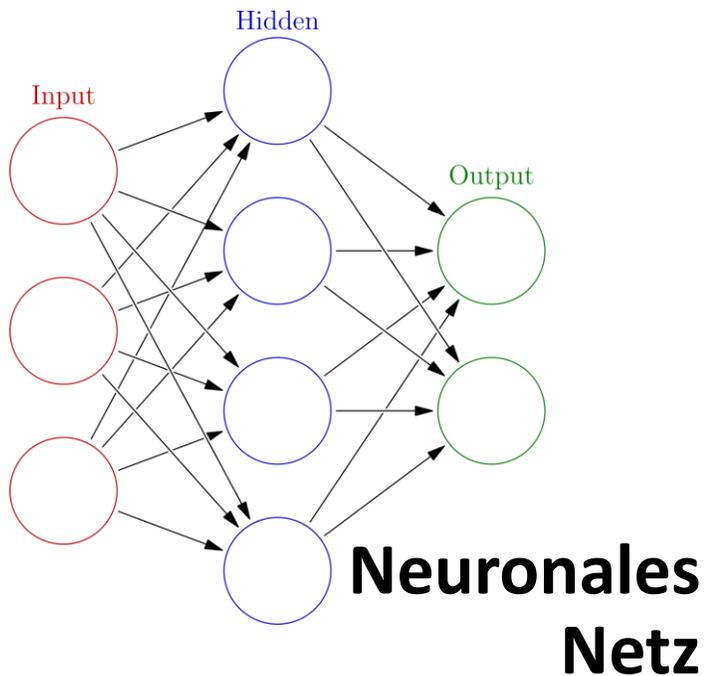


Computer lernen

Damit ein Computer lernen kann, benötigt er ebenfalls eine **Struktur**, um Gelerntes abzuspeichern.

Optimal auch **Rückkopplung**.

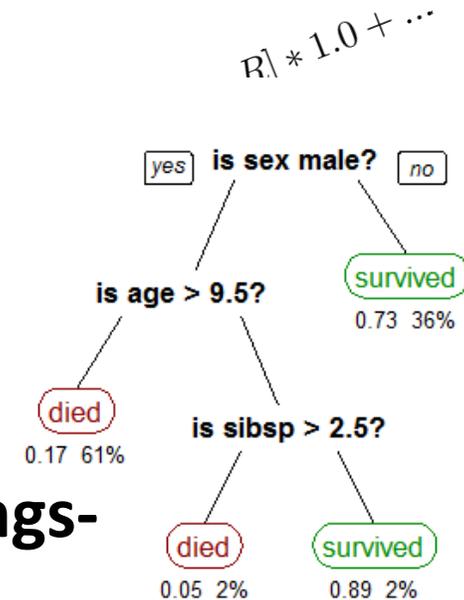
Er lernt **generelle Regeln**.



Formel

$$w_1 * \#V_h - w_2 * \#day_i V_h + w_3 * I[g = male]$$

Entscheidungs-bäume





Lernen durch Clustern

Was ist ein Rrrr, was ein Hiha und was ein Ts?



Hiha

Hiha



Hiha

Blaulicht



RRR

Auto



RRR



RRR



RRR



Ts



Ts



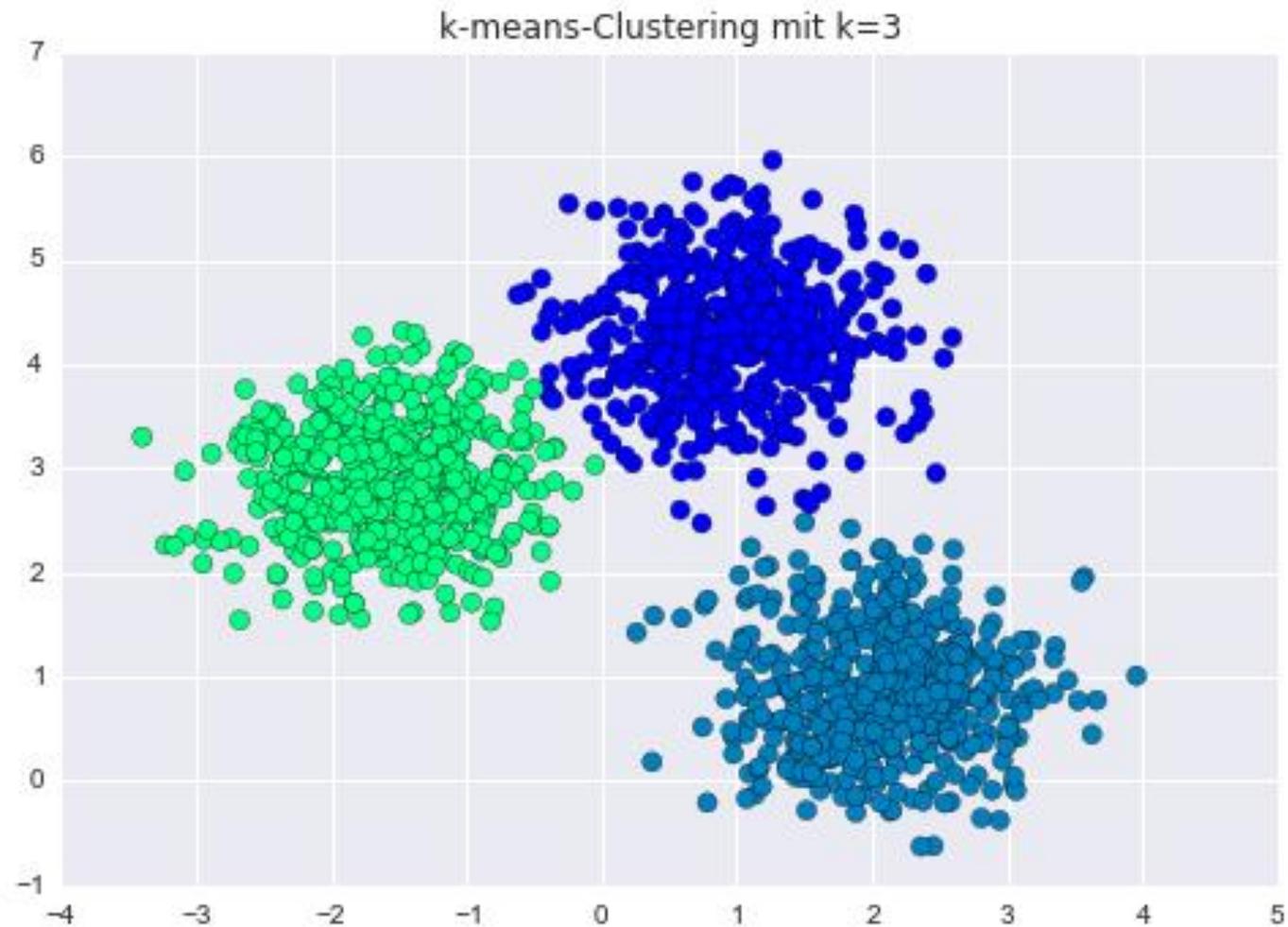
Ts

Zug



RRR

K-Means-Clustering



Vorhersagen + Fehler

- Mit gelernten Strukturen können Vorhersagen gemacht werden:
- Gegeben ein neuer Datenpunkt, zeigt die Struktur die Zugehörigkeit der Information



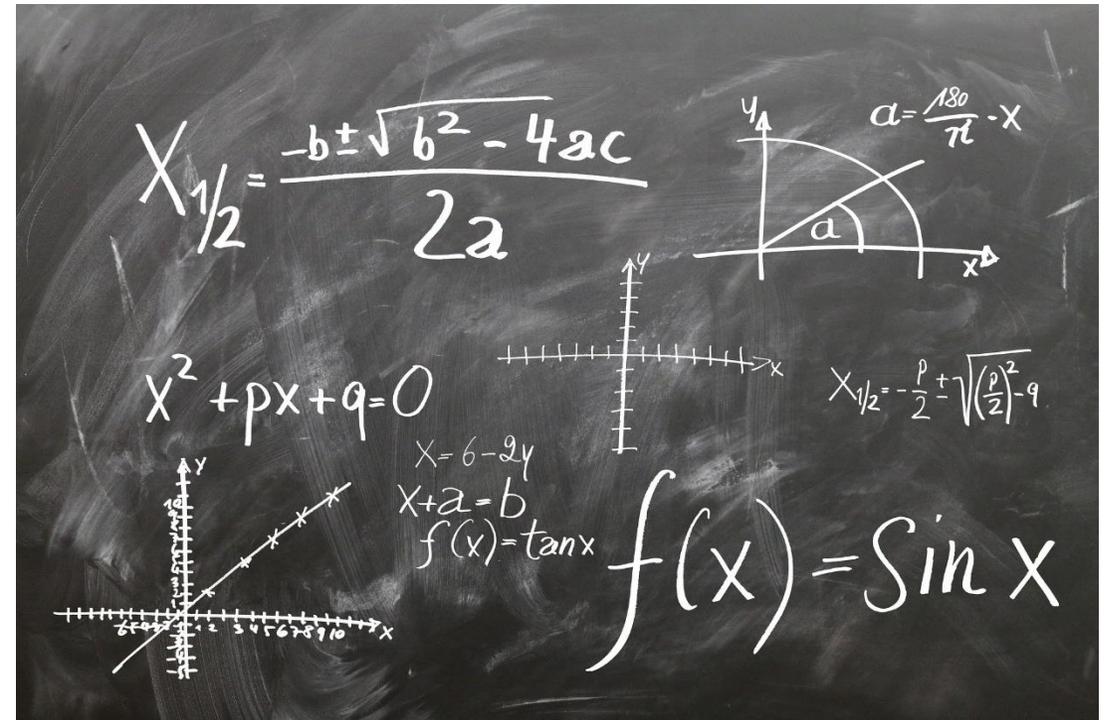


Warum hilft uns
maschinelles Lernen?

Lernen, lernen, lernen!



- Maschinelles Lernen sorgt für die beste Suchmaschine
- Findet die besten Produkte in riesigen Warenlagern, schlägt den besten Film, die neuesten Songs vor
- Findet neue Krebstherapien
- Sagt vorher, wer den Kredit zurückzahlen wird oder die Schule schmeißt



Menschliche Experten vs. Algorithmen



Menschliche Experten

1. Sehr lange Ausbildung, möglicherweise Akkreditierung, Fortbildungen.
2. Oftmals festgelegter/best practice Entscheidungsprozess.
3. Kann aber auch andere Informationen mitverarbeiten, sehr seltene Fälle entscheiden.
4. Protokoll der Entscheidungsfindung ist möglich, u.U. subjektiv
5. Oftmals hierarchisches Widerspruchssystem gegen Entscheidung
6. Kann vorurteilsbeladen sein
7. Entscheidung kann tagesabhängig sein

Entscheidungssysteme

1. Kann beliebig viele Datenpunkte verarbeiten
2. Immer festgelegter Entscheidungsprozess
3. Kann keine anderen Informationen mitverarbeiten, kann nur häufige Fälle entscheiden.
4. Protokoll der Entscheidungsfindung in eingeschränktem Sinne möglich
5. Gefahr des Monopols eines Anbieters
6. Kann vorurteilsbeladen sein
7. Immer dieselbe Entscheidung bei derselben Eingabe

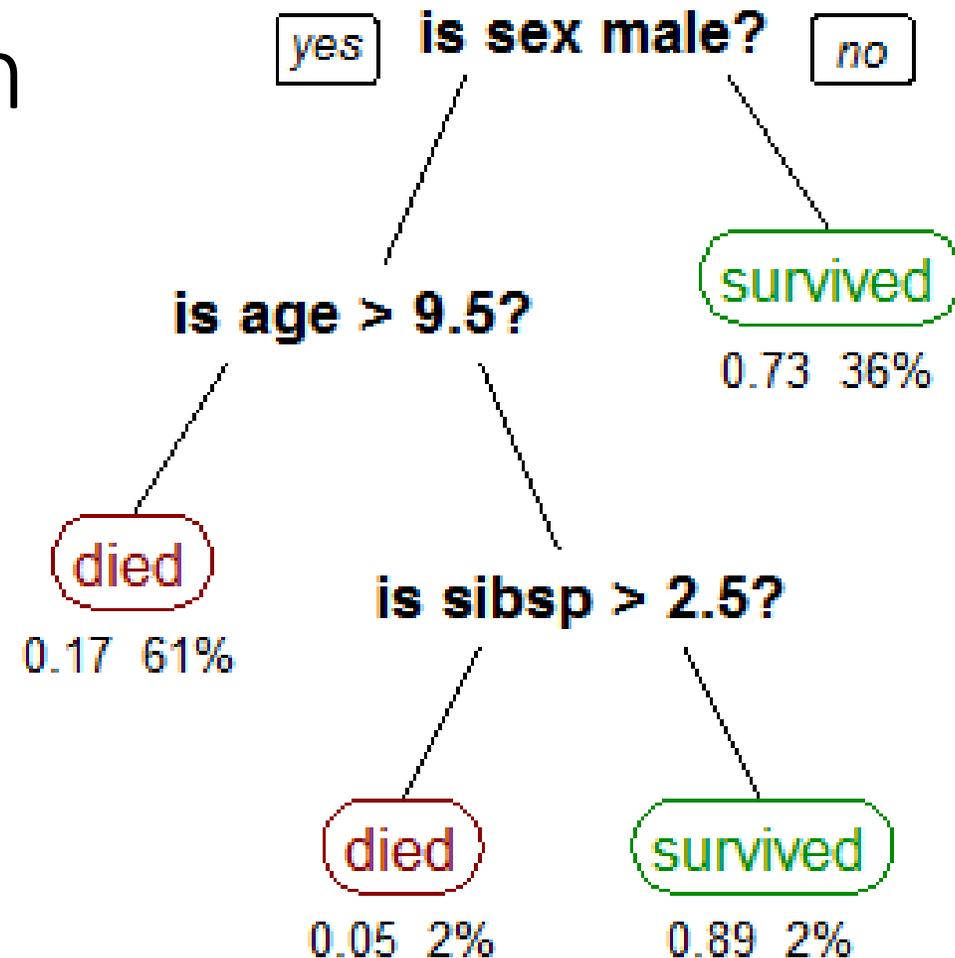
Lernen mit Entscheidungsbäumen



Lernen mit Entscheidungsbäumen

Der Baum zeigt die Überlebenswahrscheinlichkeit der Titanicpassagiere ("sibsp" bezeichnet die Anzahl der Geschwister an Bord (plus Ehegatten)). Die Zahlen unter den Blättern zeigen die Wahrscheinlichkeit und den prozentualen Anteil der entsprechenden Gruppe an der Gesamtheit.

(https://en.wikipedia.org/wiki/Decision_tree_learning, 28th of June, 2015)





Lernen mit Formeln

(und einem oder mehreren Schwellwerten)

Wie sagt man die Rückfallrate eines Verbrechens voraus?



https://de.wikipedia.org/wiki/Datei:Bundesarchiv_B_145_Bild-F083310-0001_Karlsruhe_Bundesverfassungsgericht.jpg

Von Bundesarchiv, B 145 Bild-F083310-0001 / Schaack, Lothar / CC-BY-SA 3.0, CC BY-SA 3.0 de, <https://commons.wikimedia.org/w/index.php?curid=5473096>

Predictive Policing



Wir haben schon
auf Sie gewartet!



Vorhersagen,
wann und wo
Straftaten
wahrscheinlich
sind.

Predictive Policing



Ein **Algorithmus**
hat mir geflüstert,
dass Du **fast** ein Krimineller bist.
Dann komm mal mit!



Aber auch: Vorhersagen,
ob ein Individuum
straffällig werden könnte!

Beispiel USA:

- 1) Oregon
- 2) Andere Bundesstaaten



Algorithmus

- Die Algorithmen designerinnen und -designer müssen nun entscheiden, welche der Daten vermutlich mit „Rückfallwahrscheinlichkeit“ korrelieren.
- Dies sollte am besten in einer einzigen Zahl münden, so dass man direkt sortieren kann.
- Je höher die Zahl, desto höher die Rückfallwahrscheinlichkeit.
- Beispiel Formel:

$$\begin{aligned} & 3 * \text{bisherige Verhaftungen} \\ & - 2 * \text{Anzahl Tage seit letzter Verhaftung} \\ & + 3 * (\text{Wenn Mann, dann 1, sonst 0}) \\ & + 2,5 * (\text{Wenn Raubüberfall, dann 1, sonst 0}) + \dots \end{aligned}$$

Allgemein



$$\begin{aligned} & w_1 * \text{bisherige Verhaftungen} \\ - & w_2 * \text{Anzahl Tage seit letzter Verhaftung} \\ + & w_3 * (\text{Wenn Mann, dann 1, sonst 0}) \\ + & w_4 * (\text{Wenn Raubüberfall, dann 1, sonst 0}) + \dots \end{aligned}$$

Der Computer bestimmt die Gewichte und bekommt ein Feedback (Rückkopplung), inwieweit die damit resultierende Bewertung tatsächlich mit dem (beobachteten) Verhalten übereinstimmt.

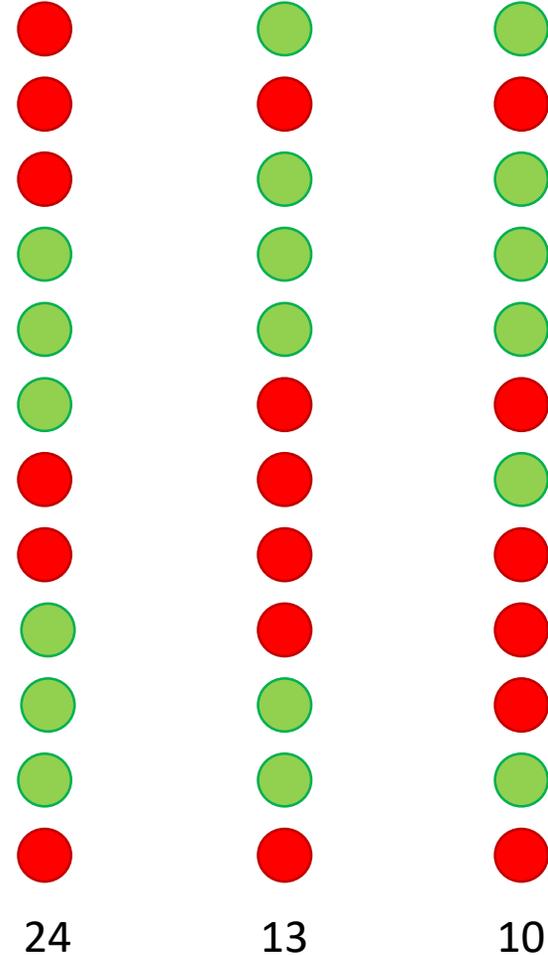


Qualität eines Algorithmus

„Lernen“ von Gewichten



- Algorithmus probiert Gewichte
- Bewertet jeweils, wie viele erwiesenermaßen Rückfällige möglichst weit oben stehen.
- Die Gewichtung, die das maximiert, wird für weitere Daten genommen.



Rote Kugeln symbolisieren rückfällige, grüne resozialisierte Personen.

Optimale Sortierung: Alle roten oben, alle grünen darunter.

Qualitätsmaß: Paare, bei denen die rote Kugel über der grünen einsortiert ist.

Oregon Recidivism Rate Algorithm

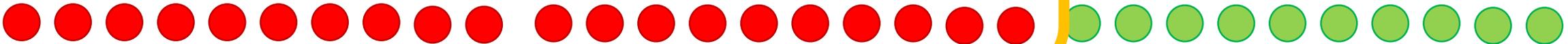
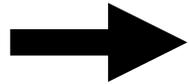


- Das oben genannte Qualitätsmaß dieses Algorithmus: 72 von 100 Paaren werden korrekt sortiert.
- Der in Oregon benutzte Algorithmus hat also, gegeben einen „Rückfall“ und einen „Nichtrückfall“, eine Chance von ca. 1:3 den Rückfall höher zu gewichten als den Nichtrückfall.
- Nur 25% aller so gemachten Prognosen sind falsch!
 - Das klingt doch ganz gut, oder?
- So werden aber keine Urteile gefällt!
- Problem: die Klassen sind ungleich verteilt!
 - 1000 Delinquenten
 - Ca. 2000 werden rückfällig

Optimale Sortierung



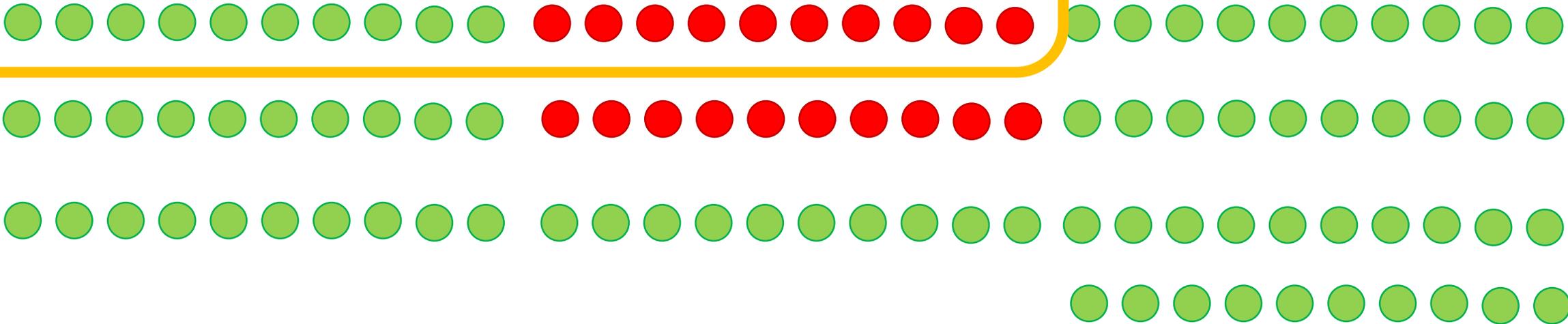
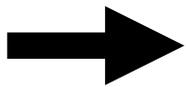
Erwartete 20% „Rückfällige“



Mögliche Sortierung eines Algorithmus mit dieser „Güte“ (75/100 Paaren)



Erwartete 20% „Rückfällige“



Das ist wie...



„Kaufen Sie diesen wunderbaren Wagen. TÜV? Brauchen Sie nicht! Und sehen Sie nur, die unglaublich gut erhaltenen Sommerreifen. Das ist noch Qualität!“



Rückfallvorhersagealgorithmus ist rassistisch (Propublica)



- In einer Studie von Propublica (anderer Algorithmus) war die Quote noch schlechter:
 - Nur 20% der (vorhergesagten) Gewalttäter begingen eine Straftat
 - Bei allen möglichen Straftaten war die Vorhersage etwas besser als ein Münzwurf.
 - Bei schwarzen Mitbürgern war die Vorhersage immer zu pessimistisch;
 - Bei weißen zu optimistisch.
- Northpoint Software ist eine Firma, der Algorithmus ist unbekannt.



Zweig'sche Regel

Algorithmen der künstlichen Intelligenz werden da eingesetzt, wo es **keine einfachen Regeln** gibt.

Sie suchen **Muster** in hoch-verrauschten Datensätzen.

Die Muster sind daher grundsätzlich **statistischer Natur**.

Versuchen fast immer, eine **kleine Gruppe** von Menschen zu identifizieren (Problem der **Unbalanciertheit**)

Wenn es **einfache Regeln zur Entscheidungsfindung** gäbe, **kennten wir sie schon**.

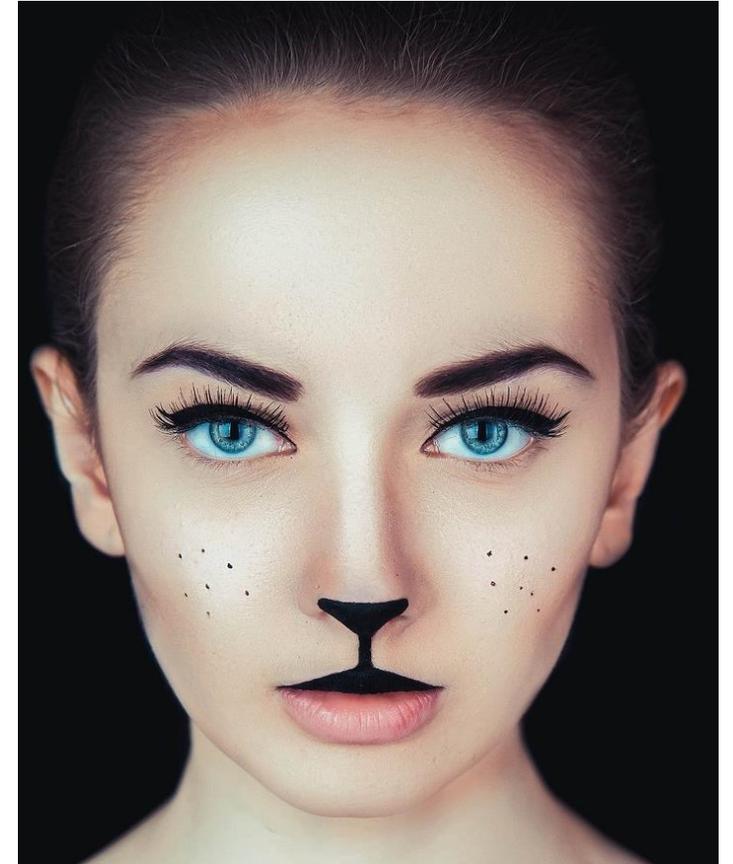


Statistische Vorhersagen über Menschen

Was bedeutet das eigentlich?

Zu 70% ein Krimineller....

- Wenn dieser Mensch eine Katze wäre und 7 Leben hätte, würde er in 5 davon wieder rückfällig werden...
- Nein!
- **Algorithmische Sippenhaftung**
 - Von 100 Personen, die „genau so sind wie dieser Mensch“, werden 70 wieder rückfällig;
 - Mitgefangen, mitgehungen;
 - In einer dem Delinquenten (der Delinquentin) völlig unbekanntem, algorithmisch bestimmten „Sippe“.



Probleme



- Aufmerksamkeitsökonomie der Richter und Richterinnen.
- „Best practice“ erfordert Nutzung der Software.
- Eine Nichtbeachtung der Empfehlung und gleichzeitige Fehleinschätzung wirkt viel schwerer als eine Beachtung der Empfehlung.
- Grundlegende Modellierung und Datenqualität kann schlecht sein.
- Der ins Gefängnis geschickte Delinquent **kann die Vorhersage prinzipiell nicht entkräften!**
 - Dies gilt auch für: Kreditvergaben, Bildungsangebote, Jobs, Personen, die von Drohnen erschossen werden oder als Terrorist eingesperrt werden, ...

Spielkampsche Regel

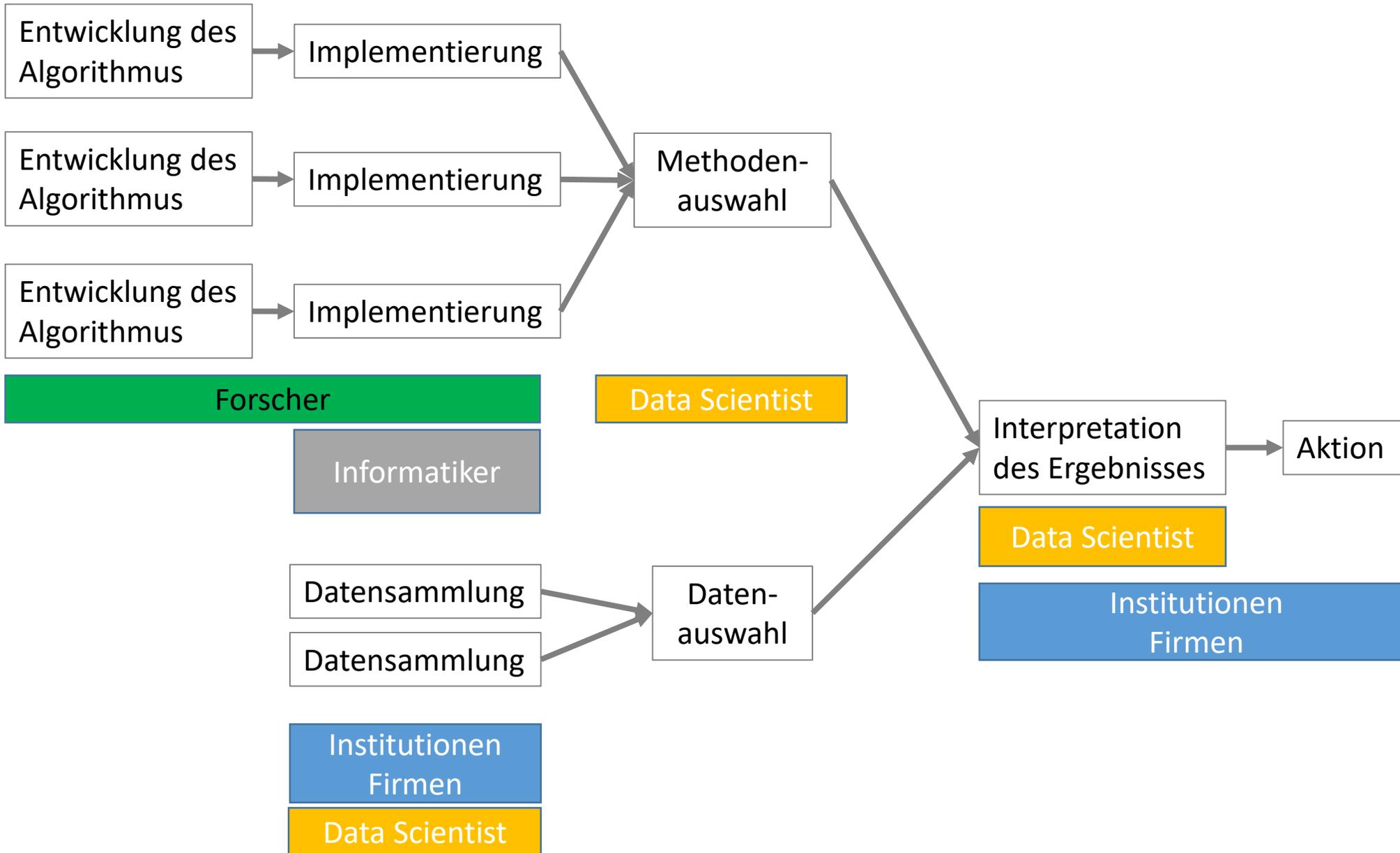


**Alle Algorithmen sind objektiv
Bis auf die von Menschen gemachten!**



Algorithmen in einer demokratischen Gesellschaft

Verkettete Verantwortlichkeiten



Wer überwacht die Auswirkungen auf die Gesellschaft?

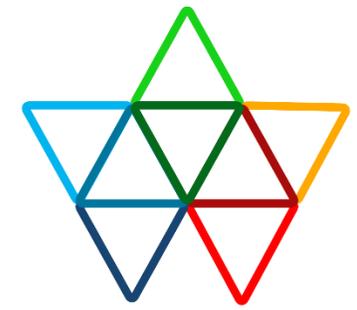
Medien?
Gesellschaft?
Politik?
Institutionen?
Firmen?
Recht?



Quis custodiet ipsos algorithmos

Der „Automated Decision Making“-TÜV vulgo: „Algorithmen TÜV“

Gründung von „Algorithm Watch“



ALGORITHM
WATCH



Lorena Jaume-Palasi, Mitarbeiterin im iRights.Lab



Lorenz Matzat, Datenjournalist der 1. Stunde, Gründer von lokaler.de, Grimme-Preis-Träger



Matthias Spielkamp, Gründer von iRights.info, ebenfalls Grimme-Preis-Träger, Vorstandsmitglied von Reporter ohne Grenzen.



Prof. Dr. K.A. Zweig, Junior Fellow der Gesellschaft für Informatik, Digitaler Kopf 2014, TU Kaiserslautern



Notwendige Eigenschaften

- Unabhängige Prüfstelle mit Siegelvergabe
- Möglichst auch mit Forschungsauftrag
- Identifikation der **kleinstmöglichen Menge** an zu überprüfenden Algorithmen
 - Die meisten Algorithmen sind harmlos;
 - Produkthaftung ermöglicht, dass andere, z.B. Versicherungen, Interesse an korrekten Algorithmen haben;
 - Wettbewerb ermöglicht, dass andere ‚neutralere‘ Algorithmen anbieten.
 - **Kein weiteres Innovationshemmnis!**
- **Non-Profit**

Beipackzettel für Algorithmen



Welches Problem „kuriert“ der Algorithmus?

Was ist das Einsatzgebiet des Algorithmus, was seine Modellannahmen?

Welche „Nebenwirkungen“ hat der Algorithmus?

Schlussformel



... zu Risiken und Nebenwirkungen der Digitalisierung befragen Sie bitte Ihren nächstgelegenen Data Scientist oder den deutschen Algorithmen TÜV.