

# Kinderleicht erklärt Wie funktioniert die KI?

EDU | days 2025



# Creative Commons Lizenz

- Die Materialien wurden teilweise vom Kurs [Die Welt der KI entdecken](#) und von [Science on Stage Deutschland](#) entnommen und adaptiert
- Der Kurs [Die Welt der KI entdecken](#) wurde von [Stefan Seegerer](#) und [Michaeli Tilman](#) entwickelt



# About

- AHS-Lehrer für M, INF, DGB
- Mitarbeiter eEducation Austria
- Referent PH Linz für den HLG DGB, Modul Programmieren
- Referent für die KinderUni, Science Holidays, DigiCamp, ...
- Ausbildung
  - *Software Engineering*
  - *Lehramt M / INF*
  - *Kulturwissenschaften mit Schwerpunkt Literaturwissenschaft*
  - *Bildungswissenschaften*
  - *Psychologie*
  - *Bildung und Medien*



# KI

- 2 Arten
  - Symbolische KI
    - Wissen wird symbolisch dargestellt und in Form von Wenn-Dann-Sonst-Regeln abgearbeitet
    - Die Logik kann von der KI selbstständig erweitert werden (logisches Schließen)
  - Maschinelles Lernen (Verstärkendes Lernen, Überwachtes Lernen, Unüberwachtes Lernen)
    - Die KI wird durch seine Umwelt / Trainingsdaten „trainiert“ und lernt
    - Deep Learning mittels neuronaler Netze ist eine besonders komplexe Art von maschinellem Lernen

# KI – Beispiele Symbolisch / ML

- Symbolische KI
  - Deep Blue (1996 / 1997)
    - Vorberechnung der folgenden Züge
    - Schach: 8 x 8 Feld →  $10^{44}$  mögliche Spielstellungen
    - Programmierer konnten erklären, wie Mensch geschlagen wurde
      - Riesige Eröffnungsdatenbank
      - Rechenpower zum Vorausberechnen der Züge
- Maschinelles Lernen
  - AlphaGo (2016 / 2017)
    - Training anhand von bestehenden Spieldaten bzw. Spielen gegen sich selbst
    - Go: 19 x 19 Feld →  $10^{170}$  mögliche Spielstellungen
    - Programmierer konnten nicht erklären, wie Mensch geschlagen wurde
      - Neuronales Netzwerk



# Symbolische KI

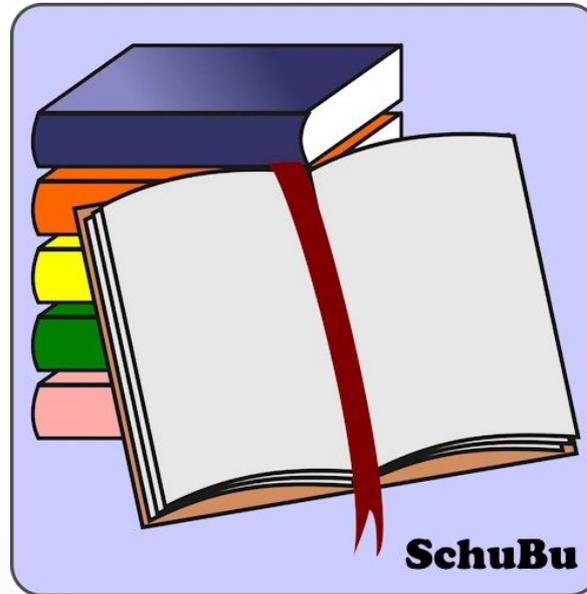
## Wissensrepräsentation und Regeln

- Wissen
  - Das Spielfeld ist 3 x 3 Felder groß
  - Freie Felder können mit O oder X besetzt werden
  - Es gibt 2 Spieler
- Regeln
  - Spieler sind immer abwechselnd an der Reihe
  - Ein Spieler setzt immer O, ein Spieler setzt immer X
  - Wenn ein Spieler 3 seiner Symbole in einer Reihe auf dem Spielfeld platziert hat, hat er gewonnen und das Spiel ist beendet
  - Ist das Spielfeld vollständig besetzt, ohne dass es 3 gleiche Symbole in einer Reihe gibt, endet das Spiel unentschieden

# Symbolische KI

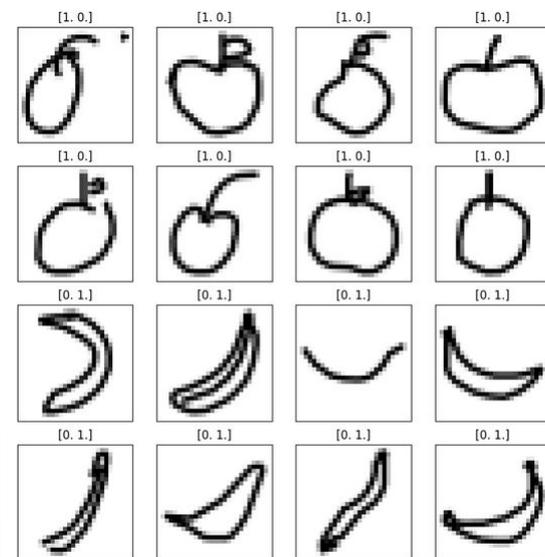
## Wissensrepräsentation und Regeln

- Beispiel Baumerkennung auf Basis von Wissen und Regeln



# Maschinelles Lernen: Lernen aus Nutzerverhalten

- Vorschläge von Streaming-Plattformen
  - Netflix
  - Amazon Prime
  - Spotify
  - Etc.
- Selbst ausprobieren:  
<https://quickdraw.withgoogle.com>

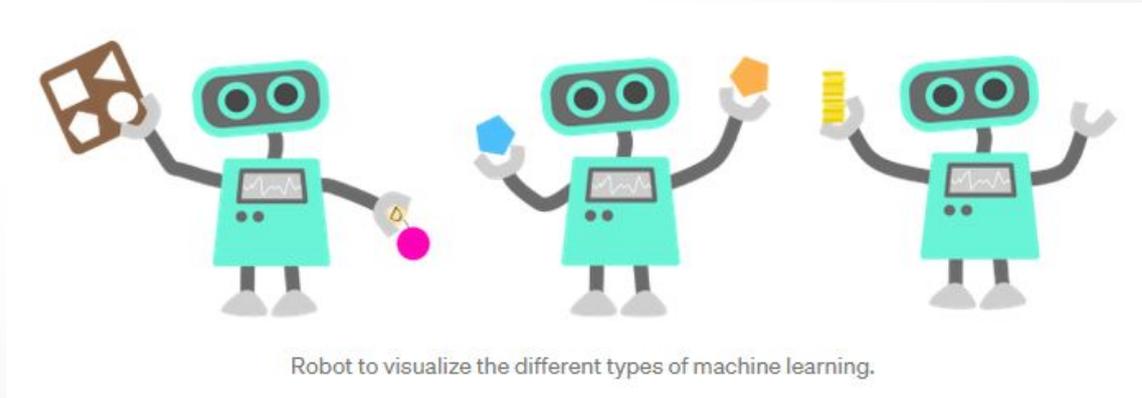


Auszug aus den Trainingsdaten

# Maschinelles Lernen: Kategorien

- Verstärkendes Lernen
- Überwachtes Lernen
- Unüberwachtes Lernen
- (Teilüberwachtes Lernen)

[Einstiegsvideo](#)



# Maschinelles Lernen: Verstärkendes Lernen

## Agent

- Ein Computerprogramm, das
  - seine Umwelt wahrnimmt,
  - in Abhängigkeit dieser Aktionen ausführt und
  - in der Lage ist, bis zu einem gewissen Grad unabhängiges und autonomes Verhalten zu zeigen

The concept of rational agents is central to AI: “An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators” (Russel & Norvig, 2010, p. 34). The vacuum-cleaner robot is a very simple form of an intelligent agent, but things become very complex and open-ended when we think about an automated taxi.

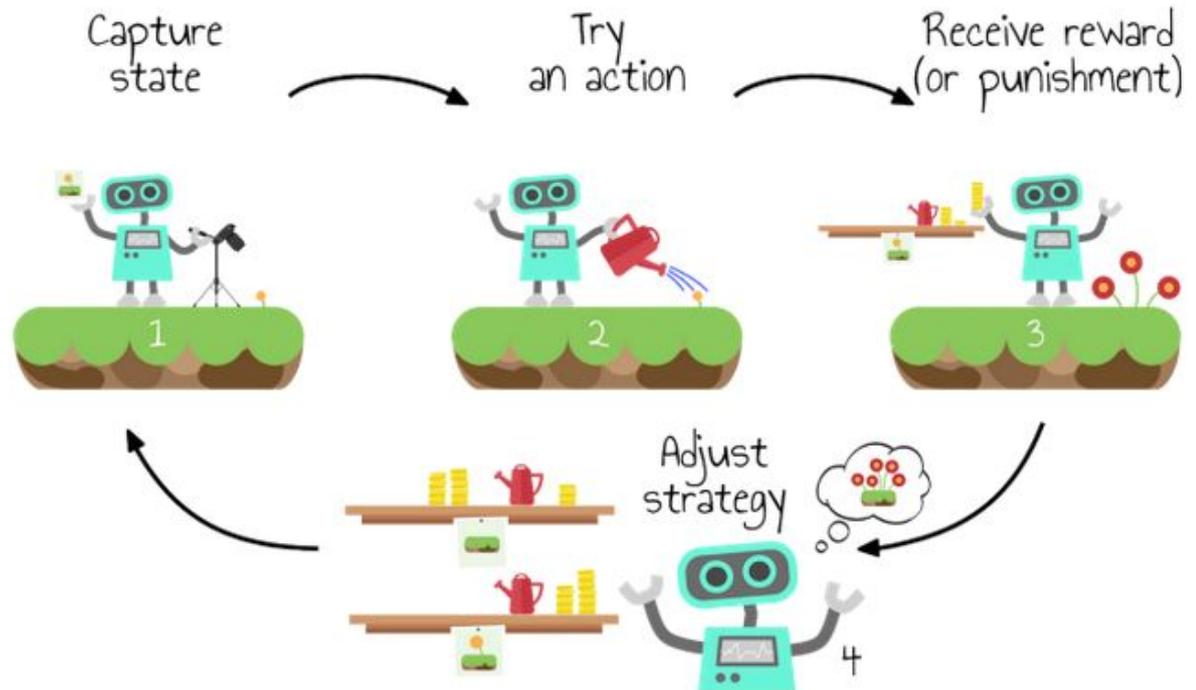


# Maschinelles Lernen: Verstärkendes Lernen

- Verstärkendes Lernen

- Zustand erfassen
- Aktion wählen und durchführen
- Belohnung oder Bestrafung erhalten
- Strategie optimieren

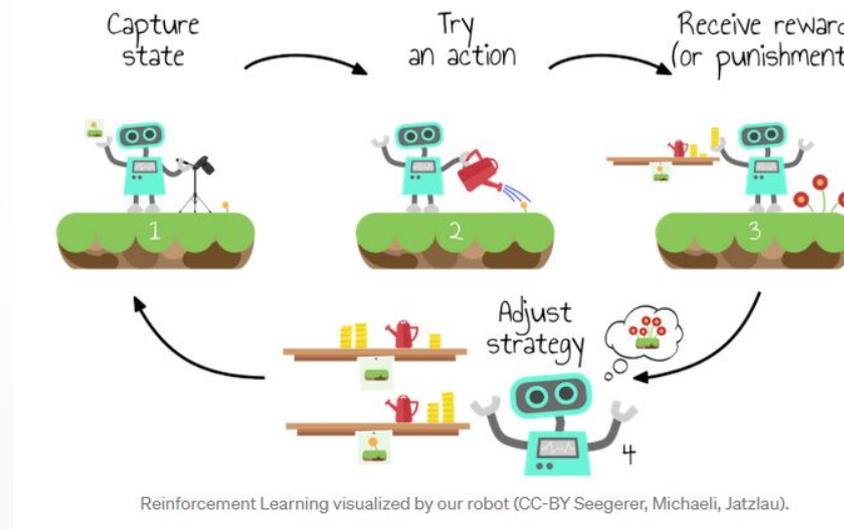
Vergleiche: Kinder!



Reinforcement Learning visualized by our robot (CC-BY Seegerer, Michaeli, Jatzlau).

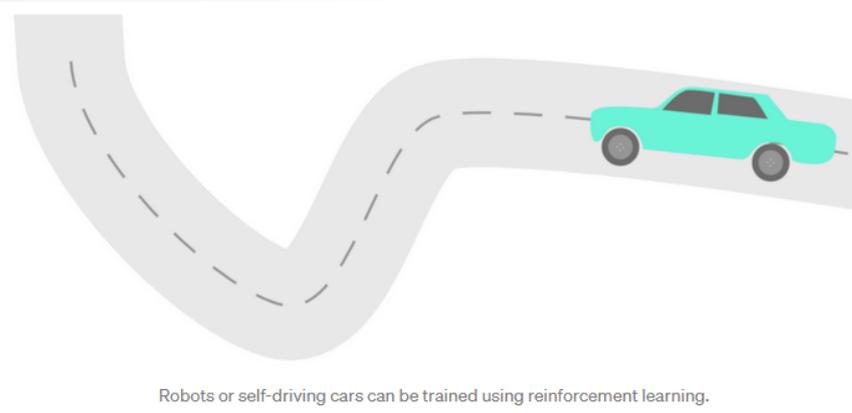
# Maschinelles Lernen: Verstärkendes Lernen

- Beispiel: Flappy Bird
- Beispiel: Schlag das Krokodil – Unplugged und **online**



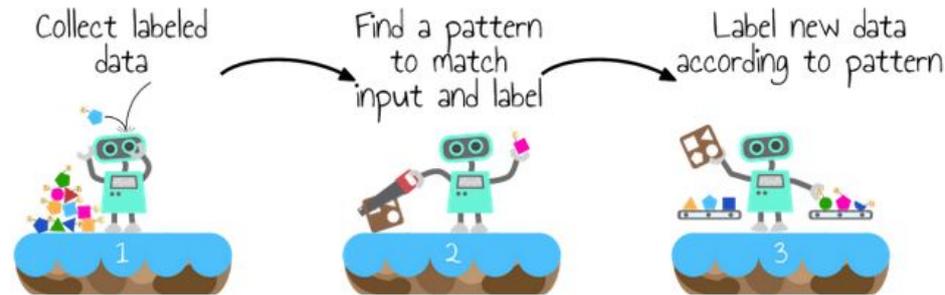
# Verstärkendes Lernen

- Anwendungsgebiete
  - Prozessoptimierung
  - Selbstfahrende Autos
  - ...



# Maschinelles Lernen: Überwachtes Lernen

- Gegeben:
  - Beschriftete Trainingsdaten
  - Achtung: Nicht zu viele Trainingsdaten, sonst „over-fitting“, d.h. die Regeln werden zu spezifisch, geringfügige Abweichungen werden nicht mehr erkannt
- Gesucht:
  - Ein Satz von Regeln auf Basis der Trainingsdaten
- Danach:
  - Test des Modells an neuen Daten, um die Güte des Modells zu bestimmen

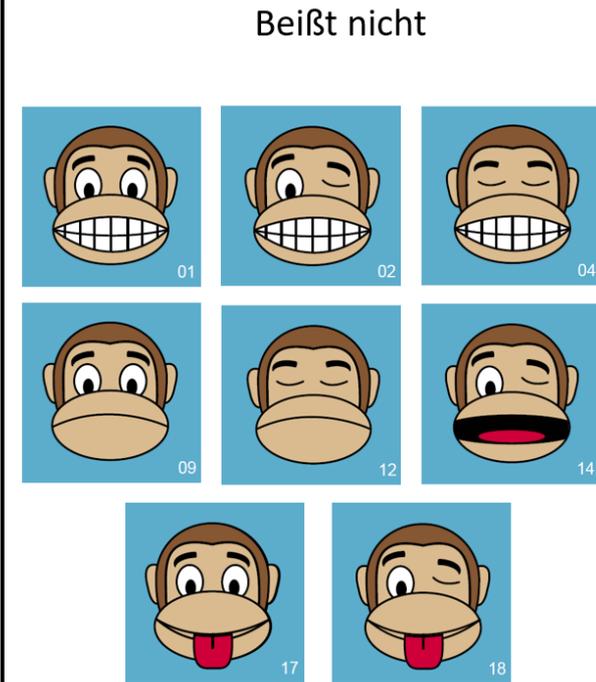
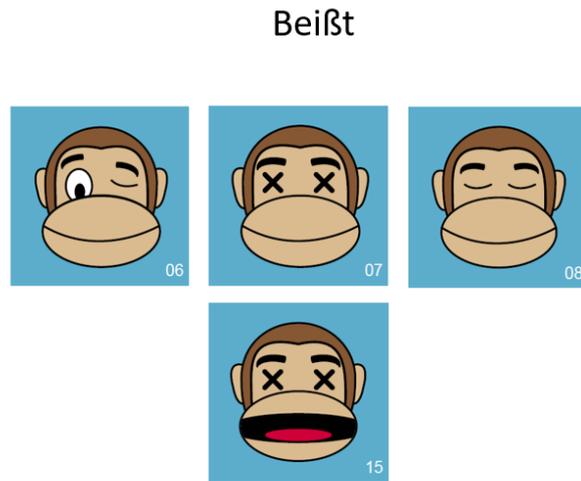


Supervised Learning visualized by a robot (CC-BY Seegerer, Michaeli, Jatzlau).

# Maschinelles Lernen: Überwachtes Lernen

## Einfache Variante

- Beschriftete (kategorisierte) Trainingsdaten
- Regeln müssen gefunden werden
  - Wann beißt der Affe?
- Online-Test



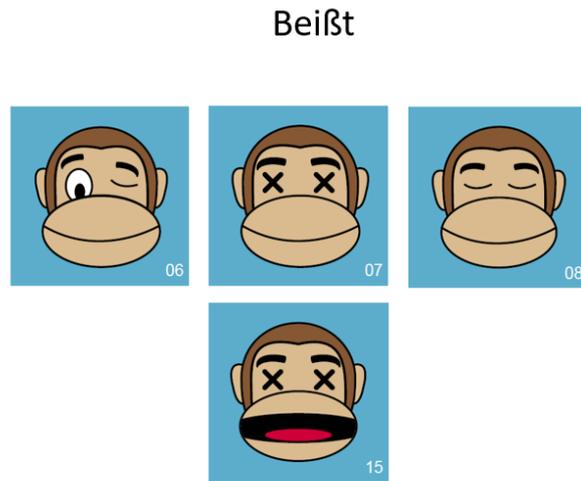
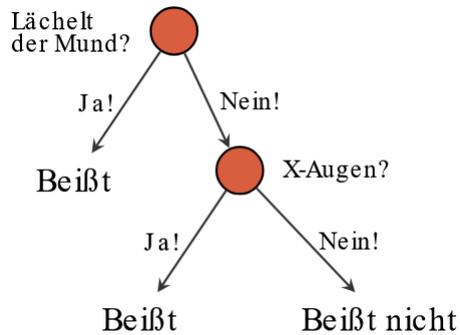
Annabel Lindner  
Stefan Seegerer



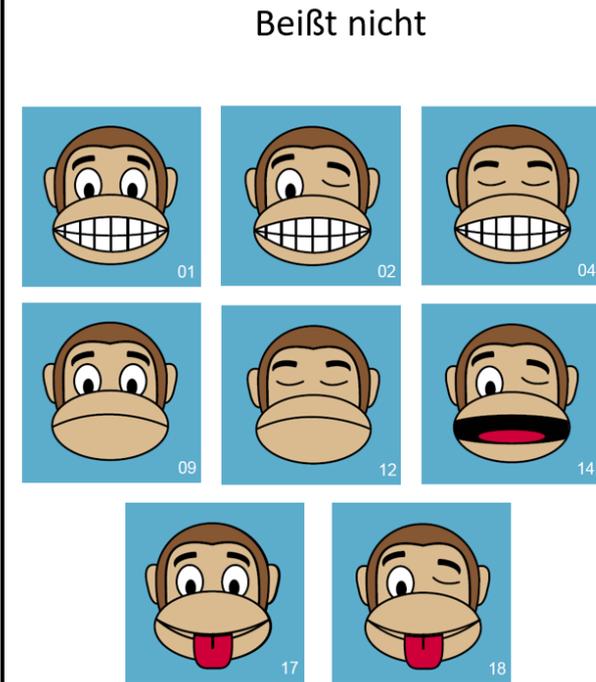
# Maschinelles Lernen: Überwachtes Lernen

## Einfache Variante

- Affe beißt, wenn...
  - er mit geschlossenem Mund lächeltODER
  - er X-Augen besitzt



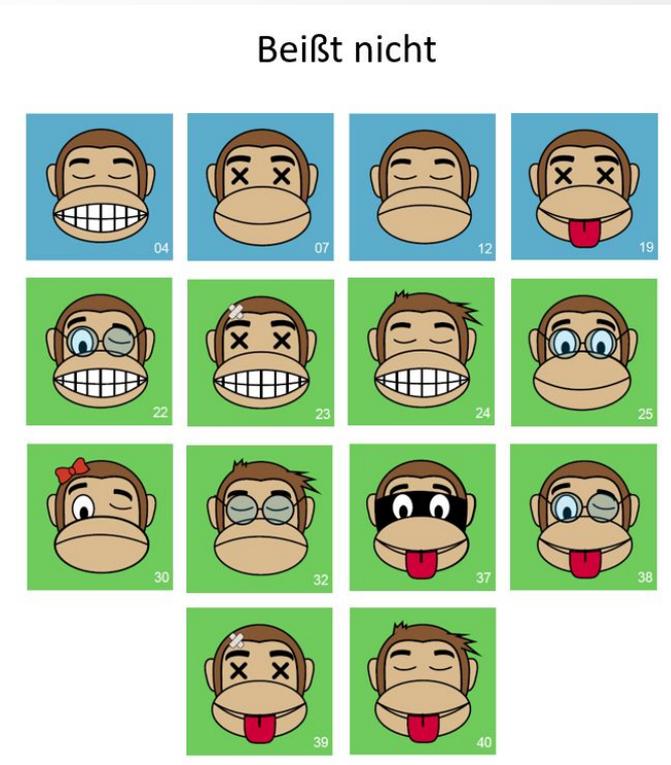
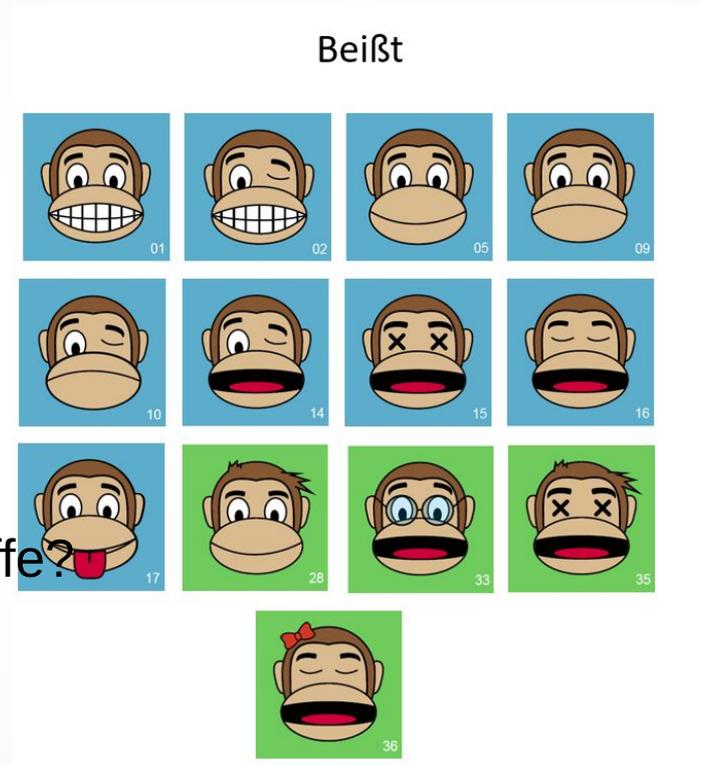
Annabel Lindner  
Stefan Seegerer



# Maschinelles Lernen: Überwachtes Lernen

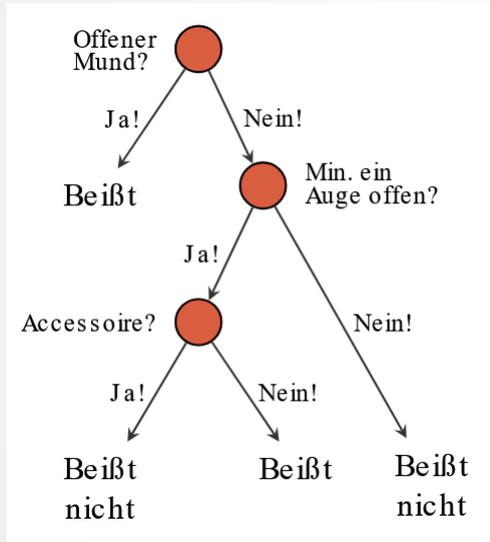
## Schwierige Variante

- Fortgeschritten
- Beschriftete (kategorisierte) Trainingsdaten
  - Wann beißt der Affe?
- Regeln müssen gefunden werden
- Online-Test

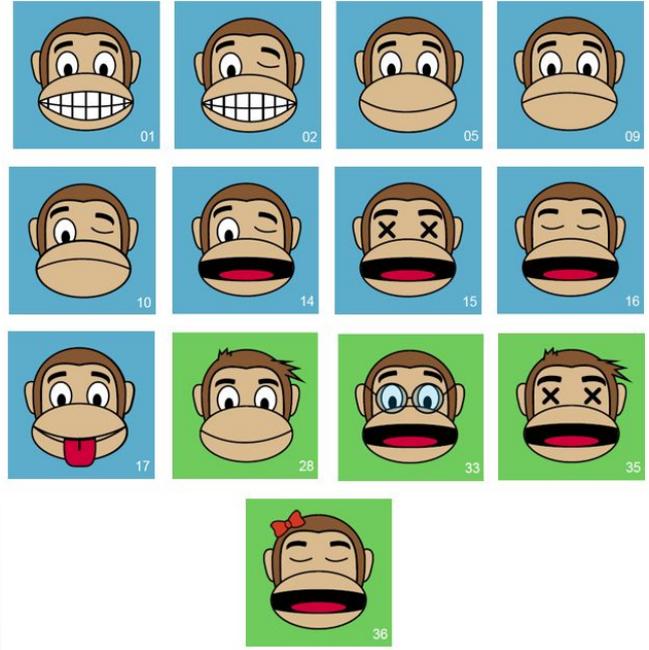


# Maschinelles Lernen: Überwachtes Lernen

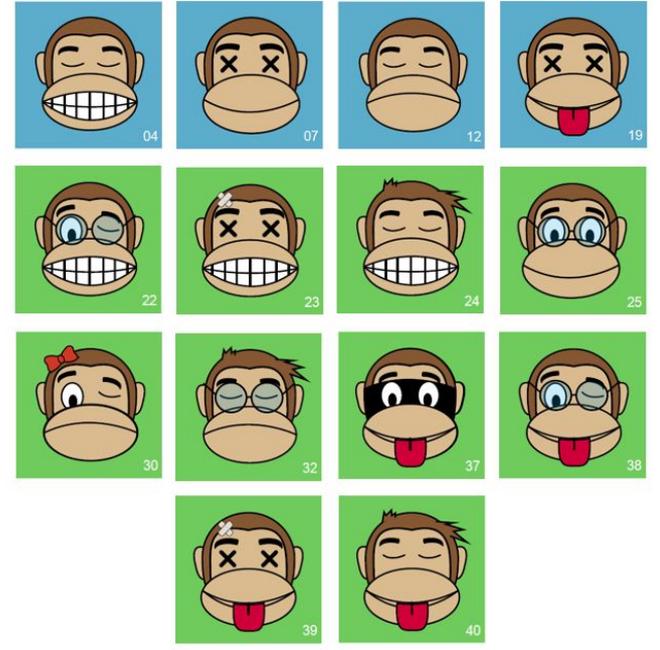
## Schwierige Variante



Beißt

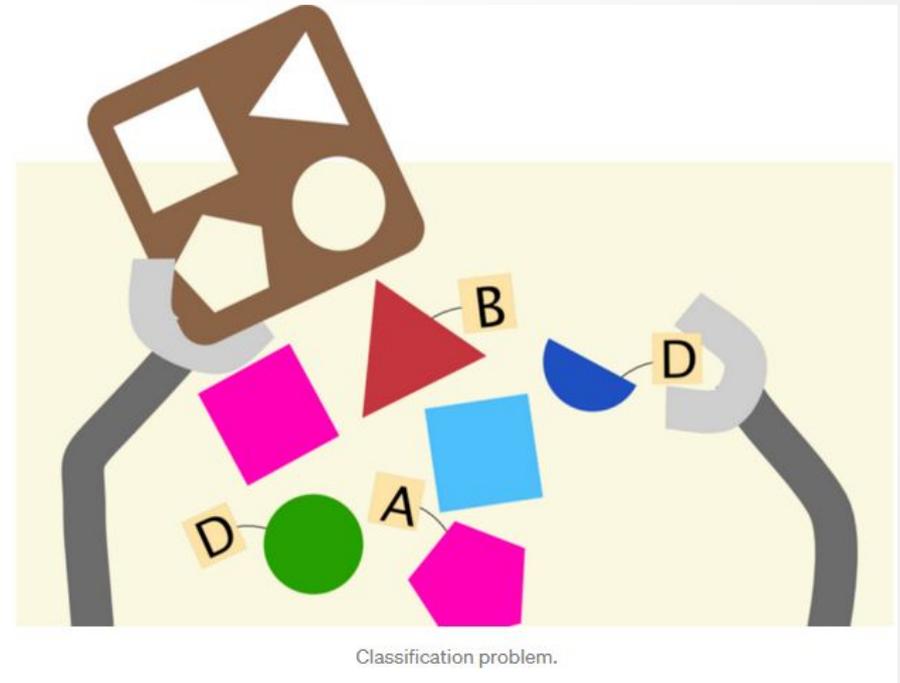


Beißt nicht



# Maschinelles Lernen: Überwachtes Lernen

- Webseite
  - Lernt durch Trainingsdaten ...
    - Bilder
    - Töne
    - Posen
  - ... zu klassifizieren



# Maschinelles Lernen: Überwachtes Lernen

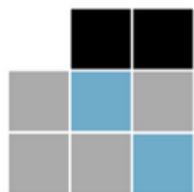
- Qualitätskontrolle anhand der vorgenommenen Klassifikationen

- *Richtig positiv / True positive (TP): Fälle, in denen das Modell **ja** vorhergesagt hat und die tatsächliche Klasse ebenfalls **ja** war.*
- *Richtig negativ / True negative (TN): Fälle, in denen das Modell **nein** vorhergesagt hat und die tatsächliche Klasse ebenfalls **nein** war.*
- *Falsch positiv / False positive (FP): Fälle, in denen das Modell **ja** vorhergesagt hat und die tatsächliche Klasse aber **nein** war.*
- *Falsch negativ / False negative (FN): Fälle, in denen das Modell **nein** vorhergesagt hat und die tatsächliche Klasse aber **ja** war.*

# Maschinelles Lernen: Überwachtes Lernen

## Genauigkeit (accuracy)

Die Genauigkeit gibt an, wie viele aller Fälle korrekt klassifiziert wurden, also:



$$= \frac{\text{\#korrekt vorhergesagt}}{\text{\#insgesamt vorhergesagt}}$$

$$\frac{41 + 49}{100} = \frac{90}{100} = 90\%$$

$$\frac{2 + 990}{1000} = \frac{992}{1000} \approx 99\%$$

$$\frac{8 + 21}{36} = \frac{29}{36} \approx 81\%$$

		Vorhergesagt		$\Sigma$
		Katze	Hund	
Tatsächlich	Katze	41	7	48
	Hund	3	49	52
$\Sigma$		44	56	100

		Vorhergesagt		$\Sigma$
		krank	gesund	
Tatsächlich	krank	2	8	10
	gesund	0	990	990
$\Sigma$		2	998	1000

		Vorhergesagt		$\Sigma$
		Spam	kein Spam	
Tatsächlich	Spam	8	2	10
	kein Spam	5	21	26
$\Sigma$		13	23	36

# Maschinelles Lernen: Überwachtes Lernen

## Trefferquote (recall)

Die Trefferquote (recall) misst, wie viele der positiven Fälle richtig vorhergesagt wurden:



$$= \frac{\text{\#Klasse 1 richtig vorhergesagt}}{\text{\#gehören tatsächlich zu Klasse 1}}$$

$$\frac{41}{48} \approx 85\%$$

$$\frac{2}{10} = 20\%$$

$$\frac{8}{10} = 80\%$$

		Vorhergesagt		$\Sigma$
		Katze	Hund	
Tatsächlich	Katze	41	7	48
	Hund	3	49	52
$\Sigma$		44	56	100

		Vorhergesagt		$\Sigma$
		krank	gesund	
Tatsächlich	krank	2	8	10
	gesund	0	990	990
$\Sigma$		2	998	1000

		Vorhergesagt		$\Sigma$
		Spam	kein Spam	
Tatsächlich	Spam	8	2	10
	kein Spam	5	21	26
$\Sigma$		13	23	36

# Maschinelles Lernen: Überwachtes Lernen

## Präzision (precision)

Die Präzision gibt an, wie gut (oder wie präzise) das Modell positive Fälle erkennt, also beispielsweise wie viele der als Spam eingestuften Mails wirklich Spam sind.



$$\frac{\text{#Klasse 1 richtig vorhergesagt}}{\text{#Klasse 1 insgesamt vorhergesagt}}$$

$$\frac{41}{44} \approx 93\%$$

$$\frac{2}{2} = 100\%$$

$$\frac{8}{13} \approx 62\%$$

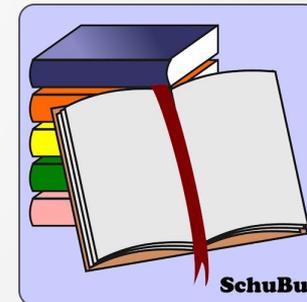
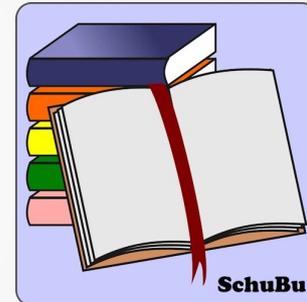
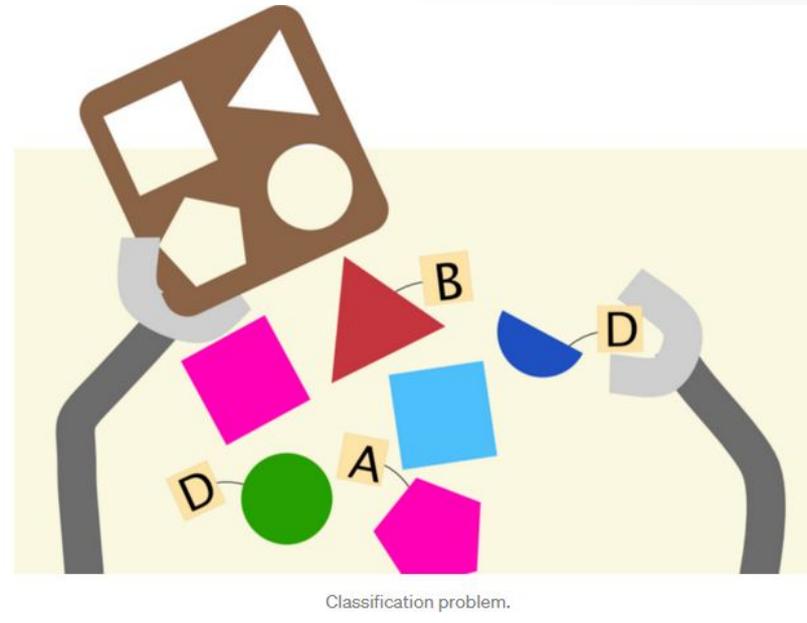
		Vorhergesagt		$\Sigma$
		Katze	Hund	
Tatsächlich	Katze	41	7	48
	Hund	3	49	52
$\Sigma$		44	56	100

		Vorhergesagt		$\Sigma$
		krank	gesund	
Tatsächlich	krank	2	8	10
	gesund	0	990	990
$\Sigma$		2	998	1000

		Vorhergesagt		$\Sigma$
		Spam	kein Spam	
Tatsächlich	Spam	8	2	10
	kein Spam	5	21	26
$\Sigma$		13	23	36

# Maschinelles Lernen: Überwachtes Lernen

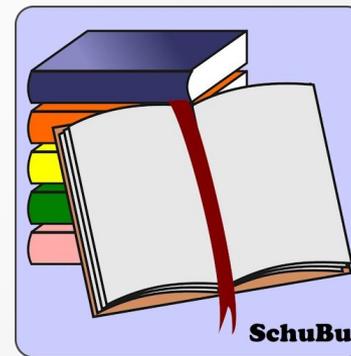
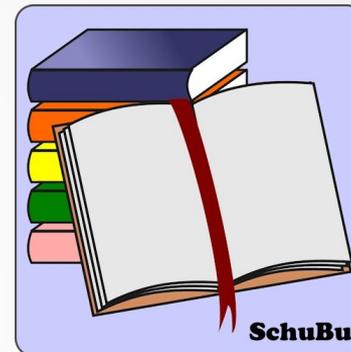
- Anwendungsgebiete
  - Klassifizierungen
    - z.B. Spam / kein Spam



# Maschinelles Lernen: Überwachtes Lernen

## Problem: „Schwache“ Trainingsdaten

- KI sieht „nur“ Pixel
- KI immer nur so gut wie die Qualität der Trainingsdaten
  - Für die Qualität der Trainingsdaten verantwortlich: Data Scientist
- Bias
  - Verzerrung von Trainingsdaten ([Video](#))
- Beispiel
  - Gesichtserkennung erwünscht
  - T-Shirt-Farbe kann Bias darstellen
    - Person X in den Trainingsdaten immer rotes Shirt
    - Person Y in den Trainingsdaten immer grünes Shirt
    - Was, wenn nun Person X im Realfall ein grünes Shirt trägt?



# Maschinelles Lernen: Überwachtes Lernen

## Problem: Wer beschriftet (annotiert)?

- Datenarbeiter:innen
  - Zumeist aus Entwicklungsländern
  - Beispiele
    - Syrien – Satellitenbilder – Klassifizierung Baum, Auto, Straße, etc.
      - Selbstfahrende Autos, Drohnen, autonom agierende Waffen → Psyche?
    - Kenia – gewaltvolle Texte klassifizieren
      - Um ChatGPT mitteilen zu können, was es *nicht* sagen soll → Psyche?
  - Gegenbewegung:
    - Verein „responsible annotation“
      - z.B. Autisten als Annotierer
    - Faire Arbeitsbedingung
    - <https://trainingstation.at>



LÖSUNG ANZEIGEN

LÖSUNG VERBERGEN

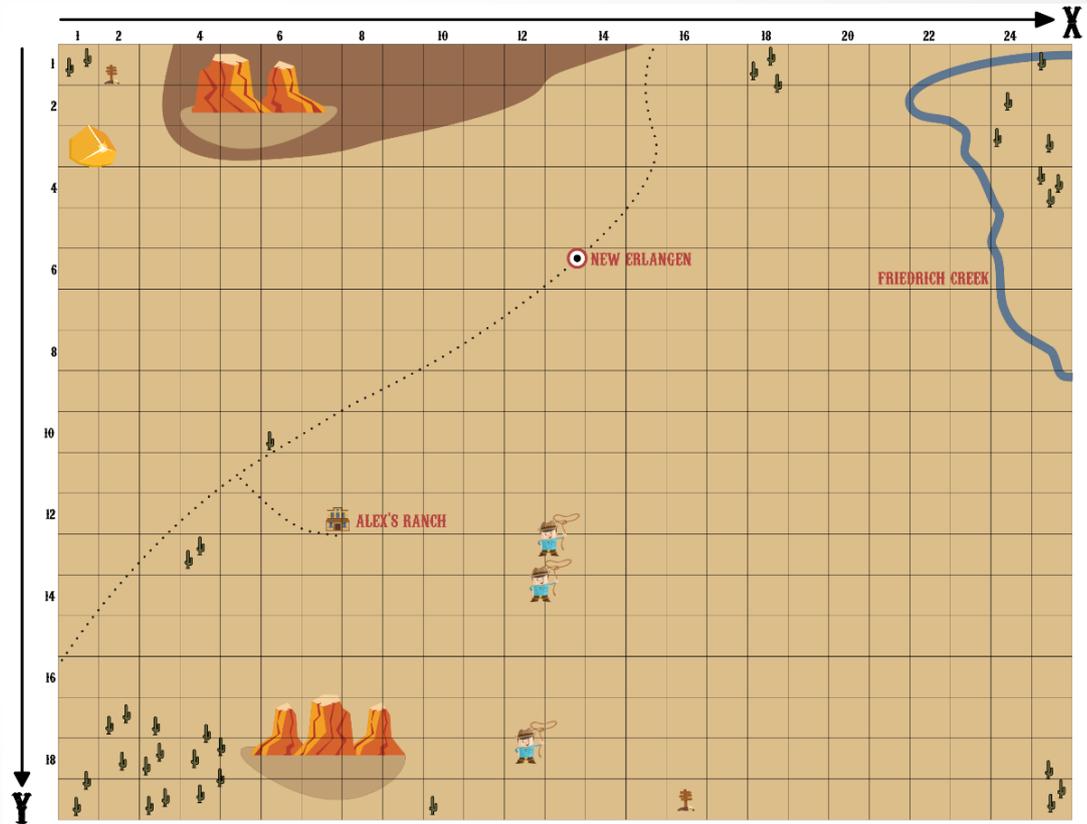
GENAUIGKEIT BERECHNEN

91 %

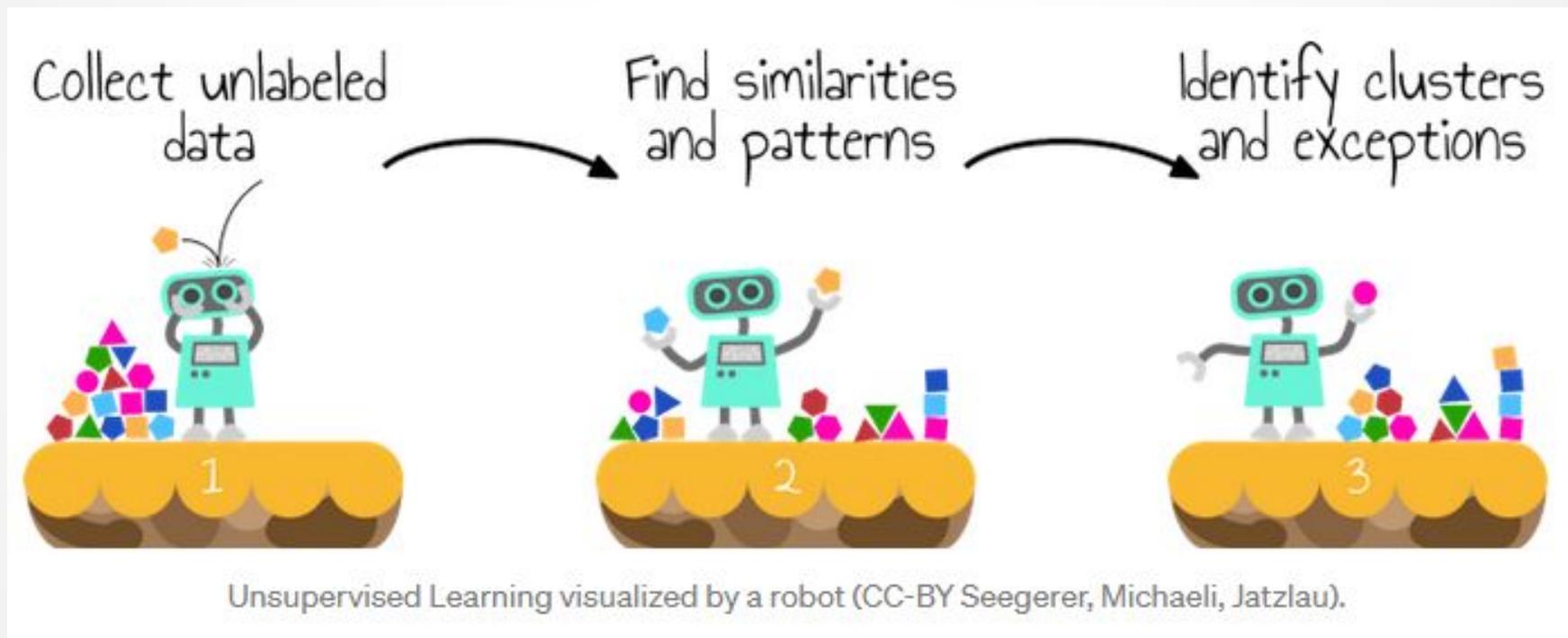


# Maschinelles Lernen: Unüberwachtes Lernen

- Ersten 3 Runden
  - Setze ein Männchen auf das gefundene Gold
- Weitere Runden
  - Wähle jenes Männchen, das am nächsten zum neu gefundenen Gold steht
  - Ziehe es die Hälfte der Entfernung zum Gold Richtung Gold
- Am Ende
  - Haben die Männchen die 3 Gold-Felder gefunden?

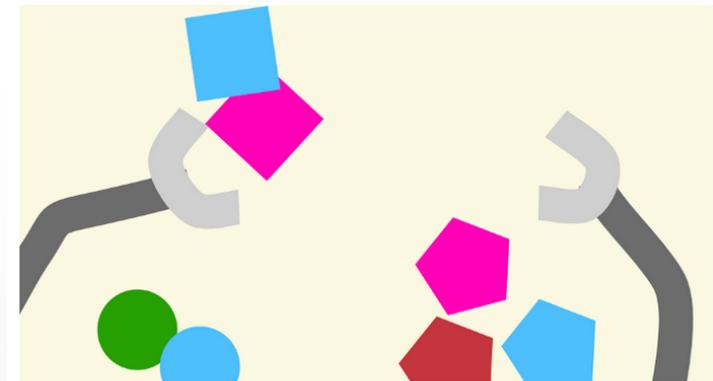
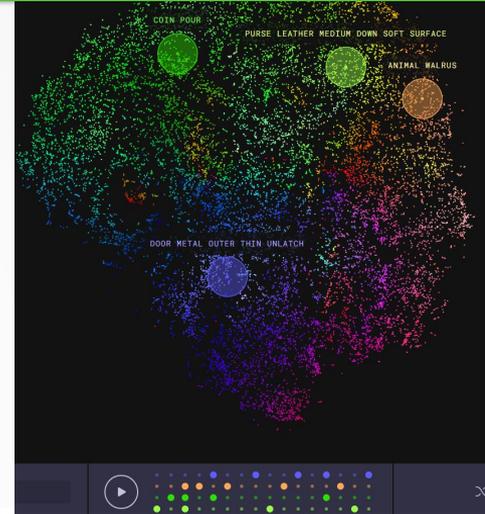


# Maschinelles Lernen: Unüberwachtes Lernen



# Maschinelles Lernen: Unüberwachtes Lernen

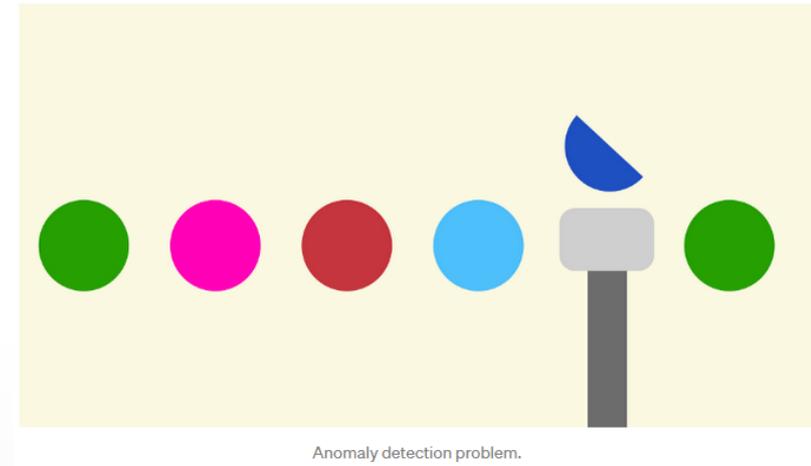
- Anwendungsgebiete
  - Cluster-Analyse (meistens mehr als 2 Dimensionen)
    - Um Kunden in Gruppen einzuteilen
      - Dimensionen
        - Alter
        - Anzahl Einkäufe
        - Durchschnittliche Summe pro Einkauf
        - Geschlecht
        - ...
    - Um Texte Themengebieten zuzuordnen
    - ...



Cluster analysis problem.

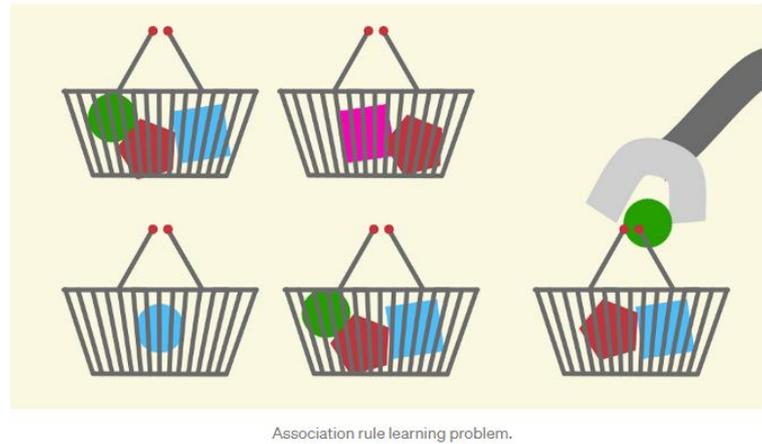
# Maschinelles Lernen: Unüberwachtes Lernen

- Anwendungsgebiete
  - Anomalie-Entdeckung
    - Etwa was den Netzwerk-Traffic betrifft
    - Oder Zahlungen mit einer Kreditkarte



# Maschinelles Lernen: Unüberwachtes Lernen

- Anwendungsgebiete
  - Beziehungs-Detektion
    - Zum Aufspüren von Wenn-Dann-Sonst-Regeln
    - z.B. Wer Produkt A kauft, kauft sehr wahrscheinlich auch Produkt F



# Maschinelles Lernen: Unüberwachtes Lernen

## Problem: Woher kommen die Daten?

- Abgreifen von Online-Daten: Webscraping
  - Zeitungsartikel
  - Social-Media-Profile
  - Urheberrechtsproblematiken? Persönlichkeitsrechte?
- Bezahlung von Menschen für das Hochladen von ...
  - Gesichtern
  - Sprachproben
  - Etc.

# Maschinelles Lernen: Unüberwachtes Lernen

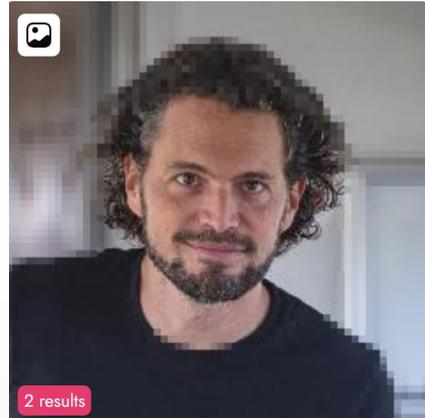
## Anwendung: Bildersuche

<https://pimeyes.com/>

← Face search [edit photo\(s\)](#) ×

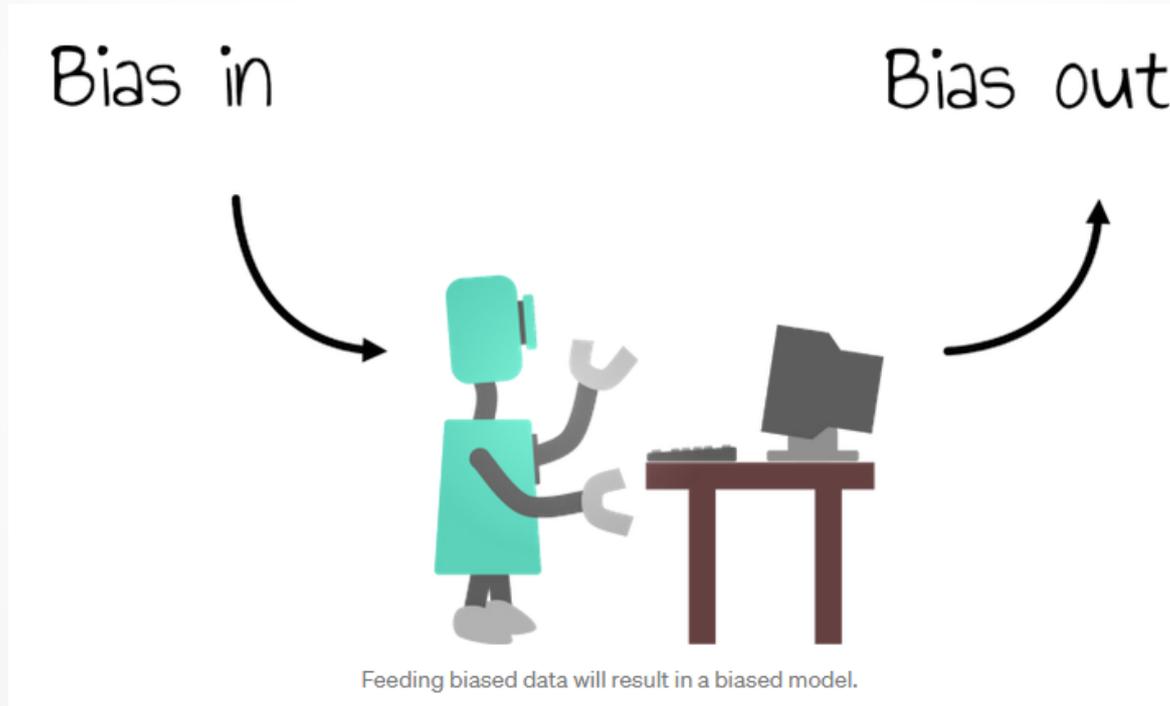


+  
Add more photos for better results



# Problematiken von ML

- Qualität ist stark von den verfügbaren Daten abhängig



# Problematiken von ML

- Teilweise sind die Ergebnisse schwer nachvollziehbar



It can be difficult to reconstruct how a result of a machine learning model was obtained.

# Deep Learning: Neuronale Netze

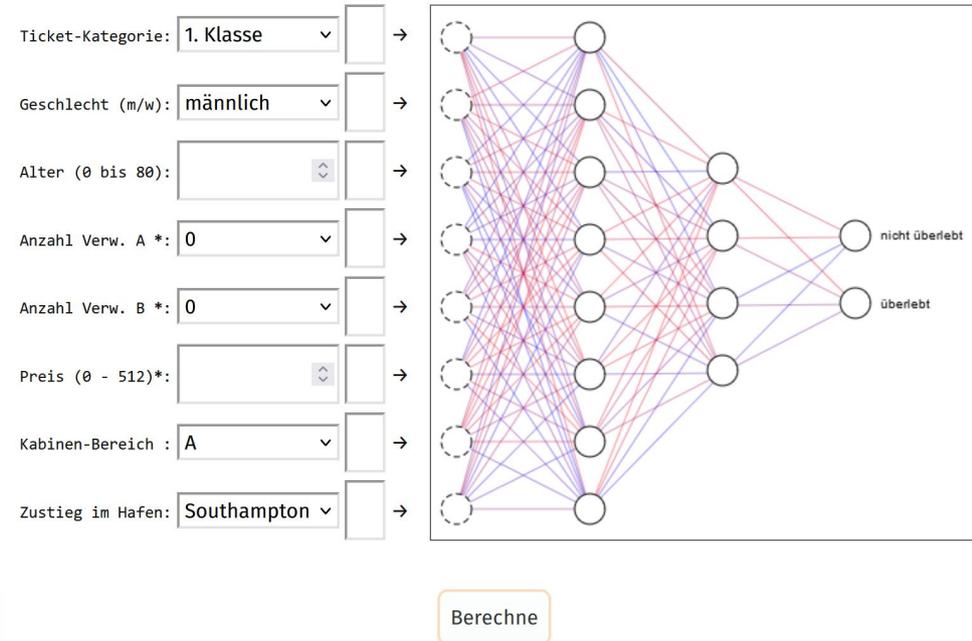
- Vom menschlichen Gehirn inspiriert
  - Unser Gehirn besteht aus Nervenzellen, sogenannten Neuronen
    - Aktuelle Schätzung: ca. 86 Milliarden Neuronen
  - Ein Neuron ist eine einfache Zelle, die eingehende Signale aufnimmt und – unter bestimmten Bedingungen – ein Signal an andere Neuronen ausgibt
  - Jedes Neuron reagiert anders auf eingehende Signale und kann diese Reaktion im Laufe der Zeit auch anpassen → das ist der Schlüssel zum neuronalen Lernen

[Einstiegsvideo](#)



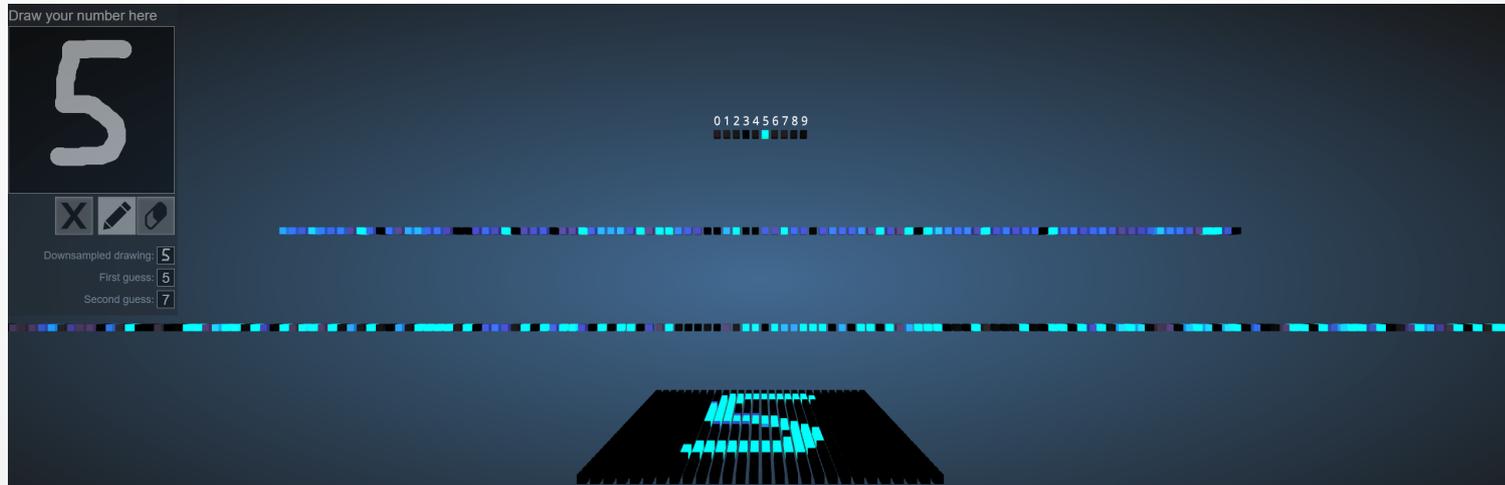
# Deep Learning: Neuronale Netze

- In Computerprogrammen kann man dies nun nachbilden
  - Viele Neuronen ergeben dann ein neuronales Netz
    - Zwischen Eingabeschicht und Ausgabeschicht liegen dann beliebige viele Zwischenschichten
    - Viele Zwischenschichten → Deep Learning



# Neuronale Netze – Beispiel Ziffernerkennung

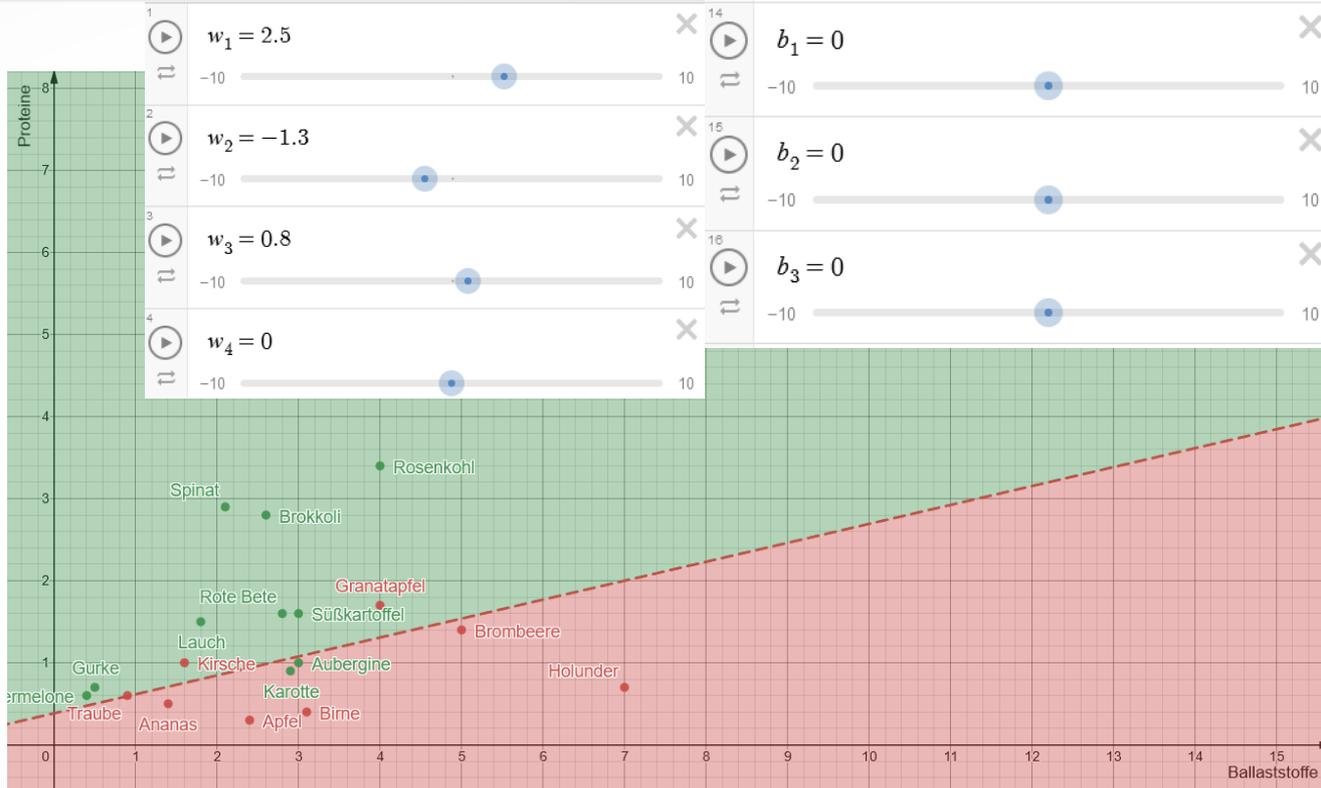
- 784 Eingabe-Neuronen (Pixel)
- 300 Neuronen in der ersten Zwischenschicht
- 100 Neuronen in der zweiten Zwischenschicht
- 10 Ausgabe-Neuronen (Zuweisung zu einer Ziffer)



# Neuronale Netze – Beispiel Obst / Gemüse

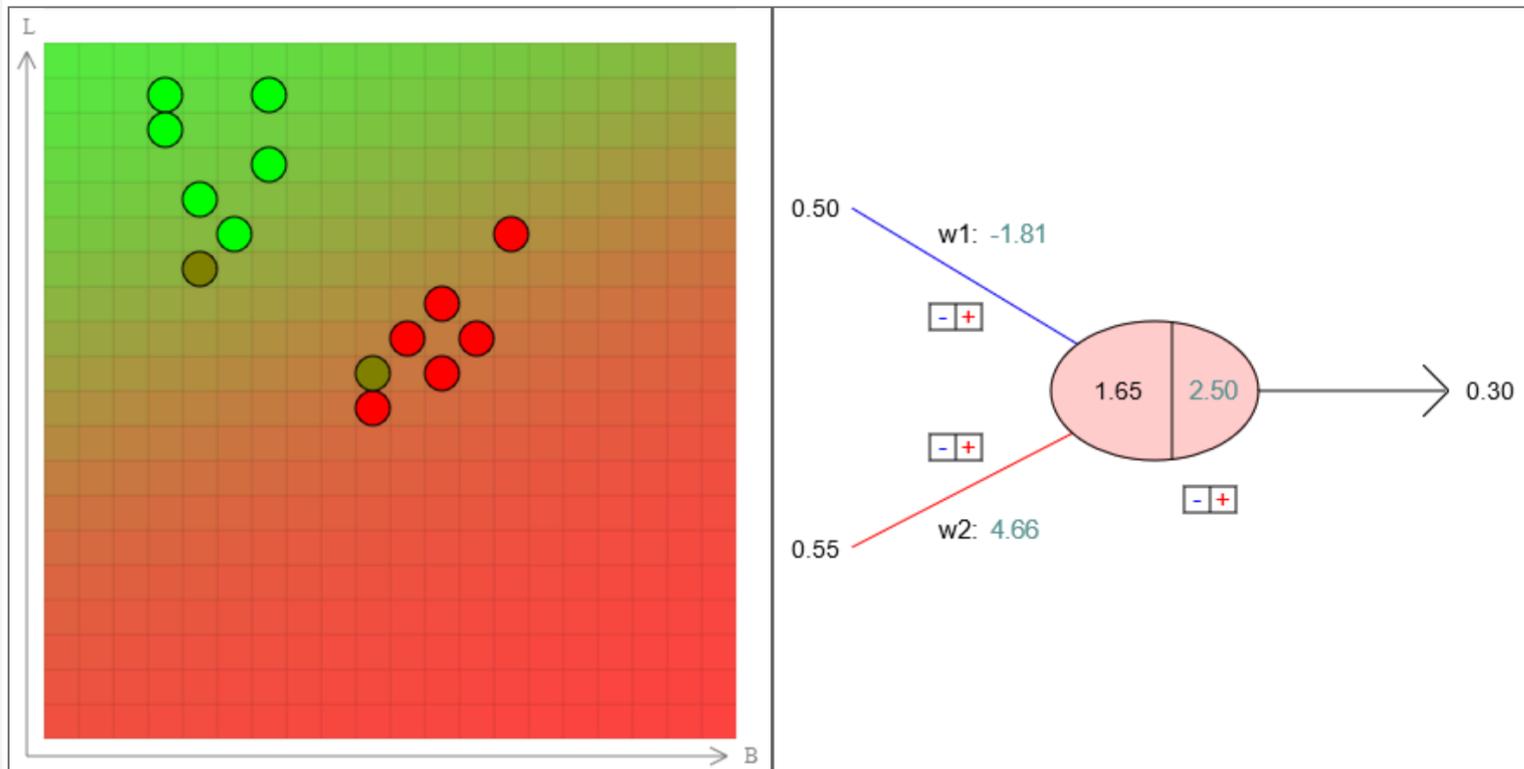
- Gewichte
  - Änderung  $k$
- Bias
  - Änderung  $d$

Probiere selbst!



# Neuronale Netze

## Beispiel Raupe / Marienkäfer - händisch

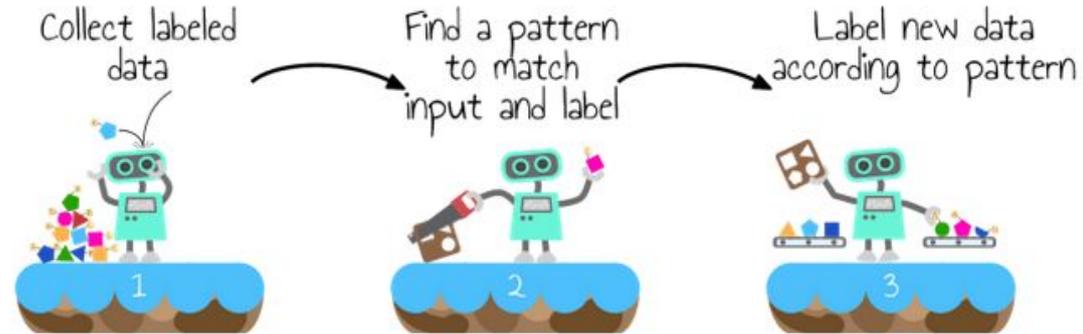


Vorhersage (69%): Marienkäfer (Breite: 5 / Länge: 5.5)

$0.50 \cdot -1.81 + 0.55 \cdot 4.66 = 1.65$  (kleiner als der Schwellwert, also Marienkäfer)

# Neuronale Netze

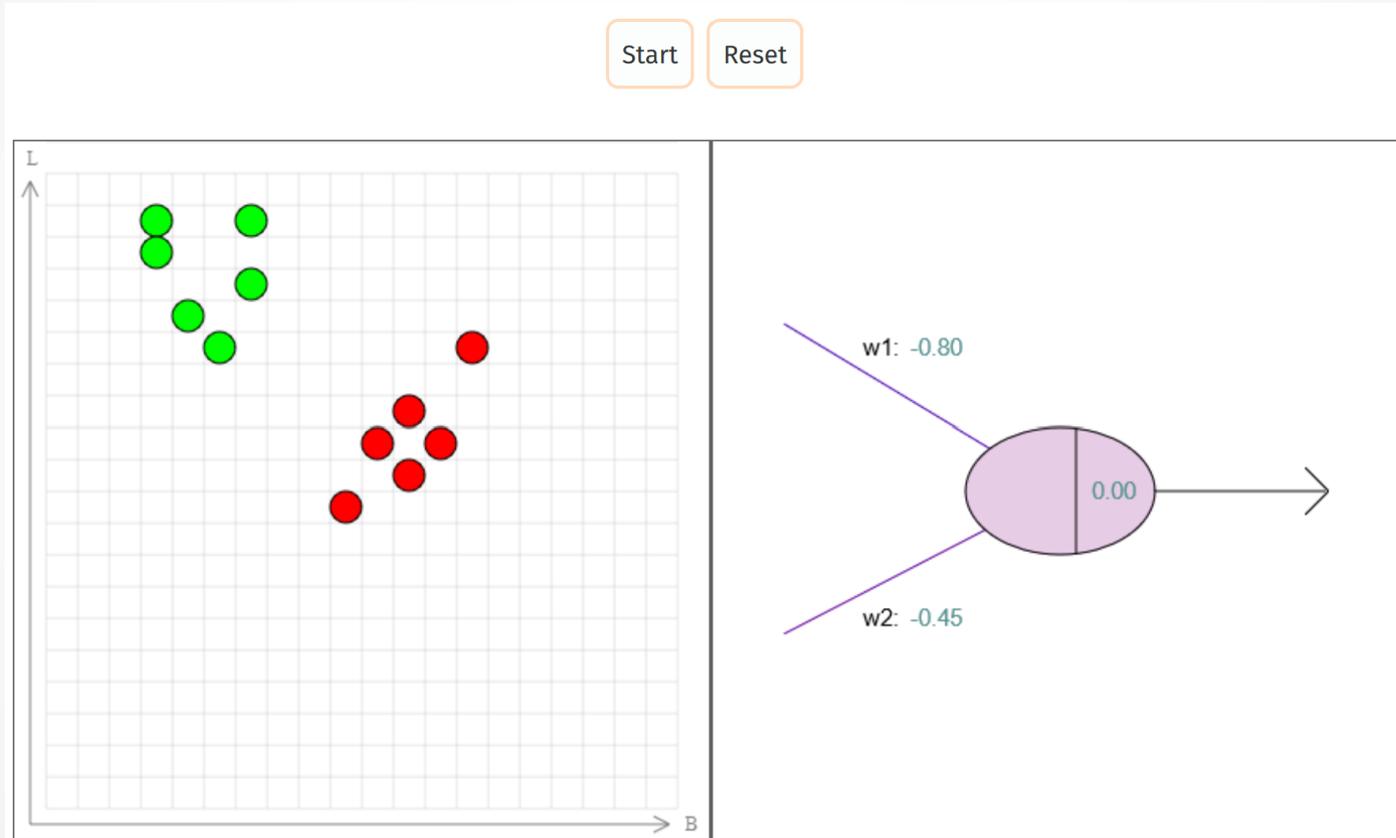
- Problem
  - Experimentelles Auffinden der richtigen Gewichte und Bias ist mühsam
- Lösung
  - Überwachtes Lernen, um Gewichte und Bias zu finden
    - Beschriftete (klassifizierte) Trainingsdaten
    - Regeln (Gewichte und Bias) werden gesucht
    - Überprüfung der Regeln an Testdaten



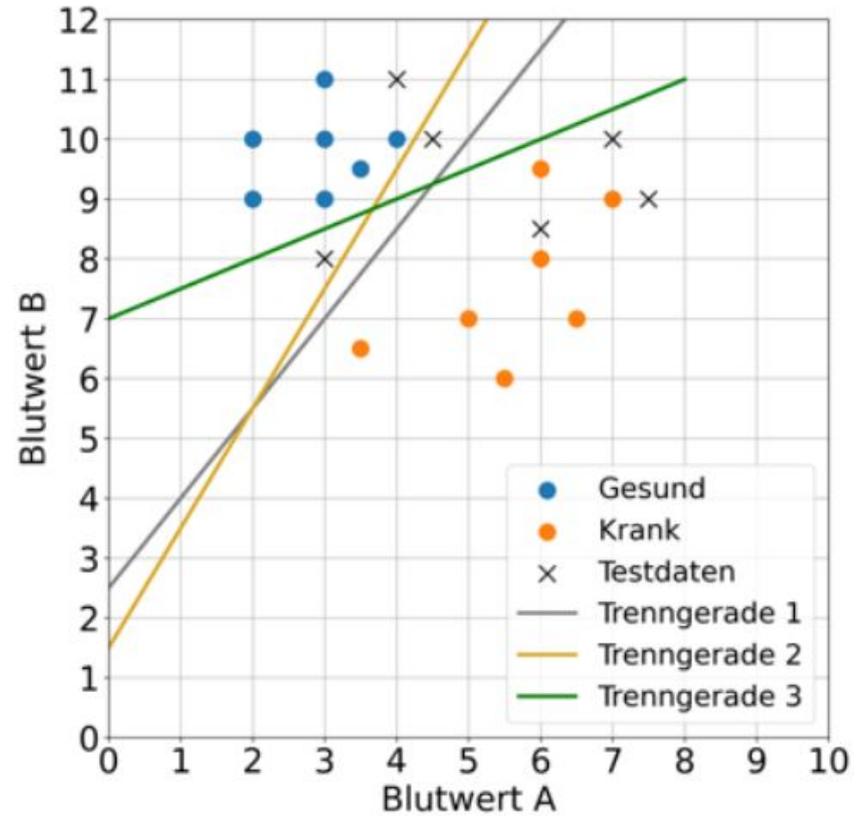
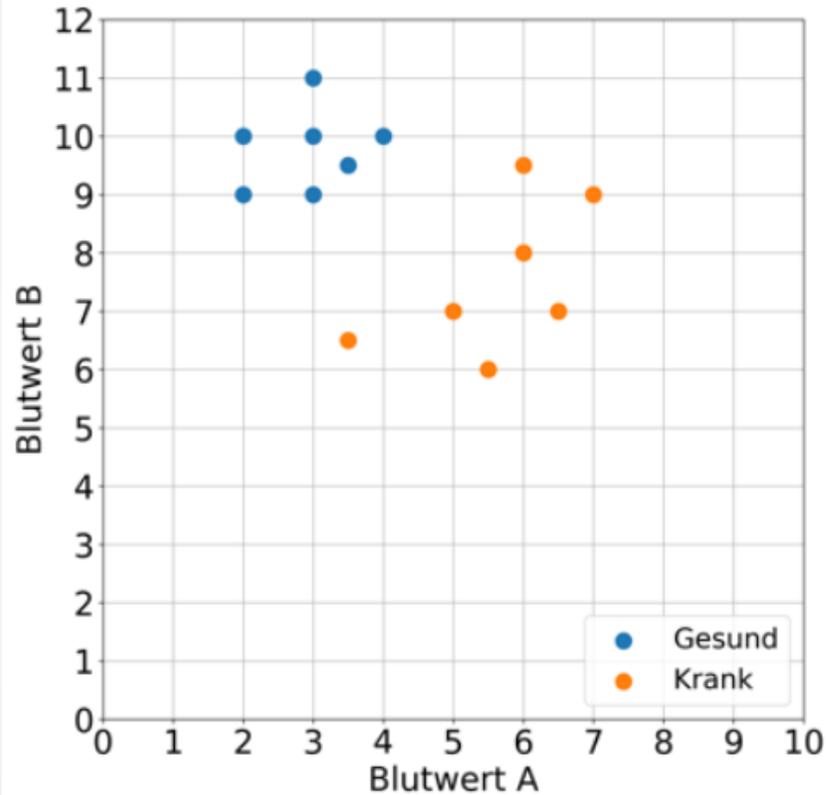
Supervised Learning visualized by a robot (CC-BY Seegerer, Michaeli, Jatzlau).

# Neuronale Netze

## Beispiel Raupe / Marienkäfer - automatisiert



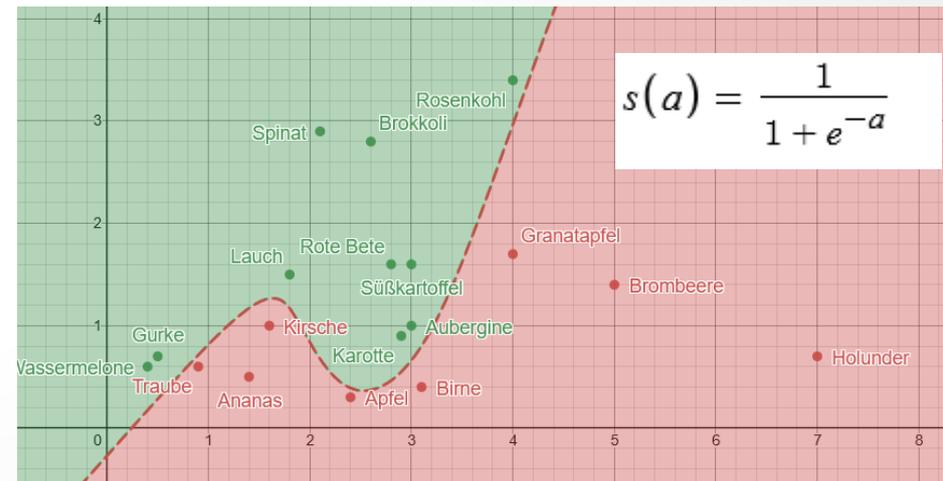
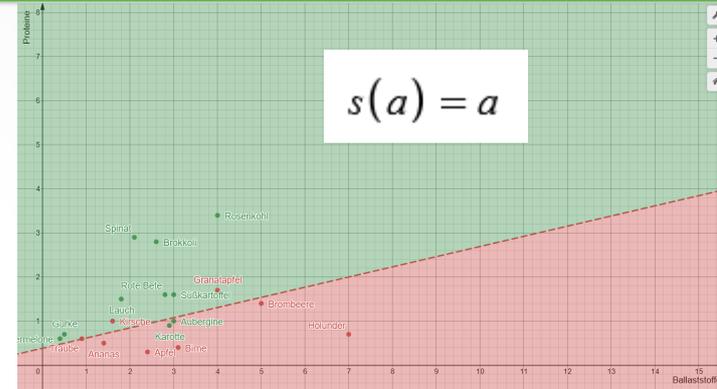
# Arbeitsblatt Klassifizierungsproblem



# Neuronale Netze – Beispiel Obst / Gemüse

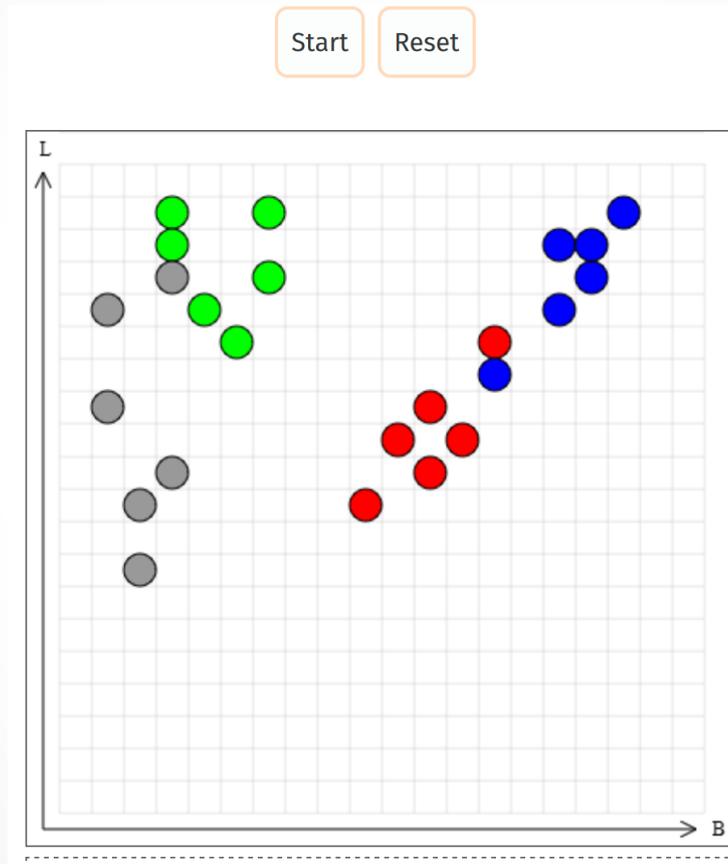
- Gerade reicht nicht, um die Daten richtig abzugrenzen
- Wechsel auf eine nichtlineare Aktivierungsfunktion
- Adaptierung der Gewichte

Probiere selbst!



# Neuronale Netze

## Beispiel mit mehreren Tieren automatisiert



# Unterschied Algorithmus / Neuronales Netz

- Algorithmus
  - Die Güte des Ergebnisses ist von der Qualität des programmierten Algorithmus abhängig
  - Ergebnis bei gleichem Input immer dasselbe
    - z.B. Größter Gemeinsamer Teiler
- Neuronales Netz
  - Die Güte des Ergebnisses ist von der Qualität der Daten abhängig, mit der das neuronale Netz trainiert wurde
  - Ergebnis bei gleichem Input nicht immer dasselbe
    - z.B. ChatGPT



$$f(x) = y$$

Güte des Algorithmus

Algorithmengetrieben



$$f(x) = y$$

Güte der Daten

Datengetrieben

# Zusammenfassendes Kurz-Video

Kurz-Video



# Workshop-Unterlagen



[www.openlearning.at](http://www.openlearning.at)