



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



CHAI
Humanities-Centered AI



Foto: UHH/Lutsch

Magnus Bender

Was haben Agenten mit Künstlicher Intelligenz (KI) zu tun?

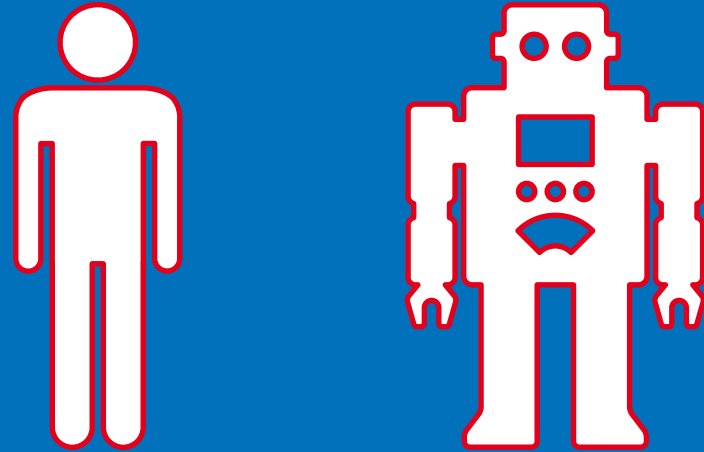
„Die Verarbeitung und Generierung natürliche Sprache ist mit ChatBots wie ChatGPT oder Gemini in aller Munde. Aber was genau ist KI eigentlich? Wie unterscheidet sich KI von maschinellem Lernen und DeepLearning? Und was haben Agenten damit zu tun?“

Überblick

- **Aufgabe:** Verarbeitung natürlicher Sprache, z.B., zur Informationsrecherche
- **Applikationen:** ChatBots, z.B., ChatGPT oder Gemini
- **Methoden:** Maschinelles Lernen und/ oder DeepLearning
- ⇒ KI ?
- **Was?** Unterstützung eines Menschen
- **Wie (extern)?** Interaktion über natürliche Sprache
 - Eingaben
 - Ausgaben
- **Wie (intern)?** Maschinelles Lernen, DeepLearning, Transformer, ...
- ⇒ Agent ?

Über mich

- Wissenschaftlicher Mitarbeiter
 - Universität Hamburg
 - Fakultät für Geisteswissenschaften
 - Fachbereich Philosophie
 - Humanities-Centered Artificial Intelligence (CHAI)-Institut
 - uhh.de/chai
- Abitur 2016
- Studium der Informatik an der UzL
 - Bachelor 2019
 - Master 2021
- Promotion @ UzL, *laufend*
 - Natürliche Sprachverarbeitung
- Wiss. Mitarbeiter @ UzL 2021 - 2024



I. Ein *intelligentes* System?

Ein Rasenmähroboter soll ...

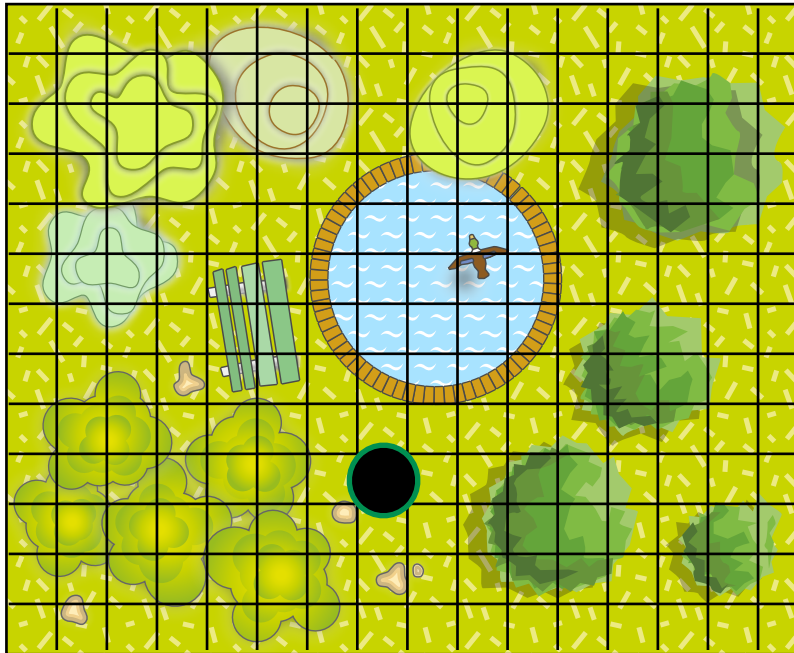


- Rasen mähen
 - Nur auf der Wiese, dort aber überall
 - Nicht im Blumenbeet oder Teich
 - Keine Menschen oder Tiere stören
- Alleine zurück finden und aufladen
- Passende Zeitpunkte bestimmen
 - Kein Regen, Schnee, Ruhezeit
 - Ausreichend Rasen gewachsen

Rasenmähroboter von Holger Casselmann aus https://commons.wikimedia.org/wiki/File:Automower_HC1.jpg veröffentlicht unter CC BY-SA 3.0 (<https://creativecommons.org/licenses/by-sa/3.0/>)

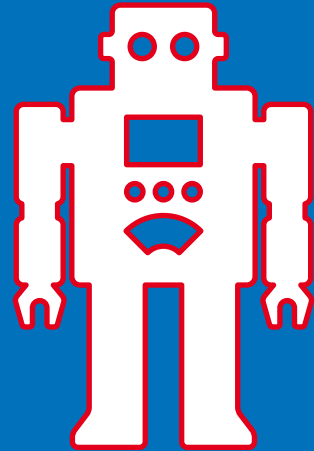
Was?

Formalisiert



- **Ziel:** Rasen mähen, *dabei*
 - Menschen unterstützen und nicht stören
- **Fähigkeiten:**
 - Schneidemesser (de-)aktivieren
 - Nach rechts, links, oben, unten fahren
- **Techniken:**
 - Sensoren zur Erkennung der Umgebung
 - Planung des Mähvorgangs
 - Durchführen des Plans, dabei Änderungen/ Einflüsse berücksichtigen

Wie?

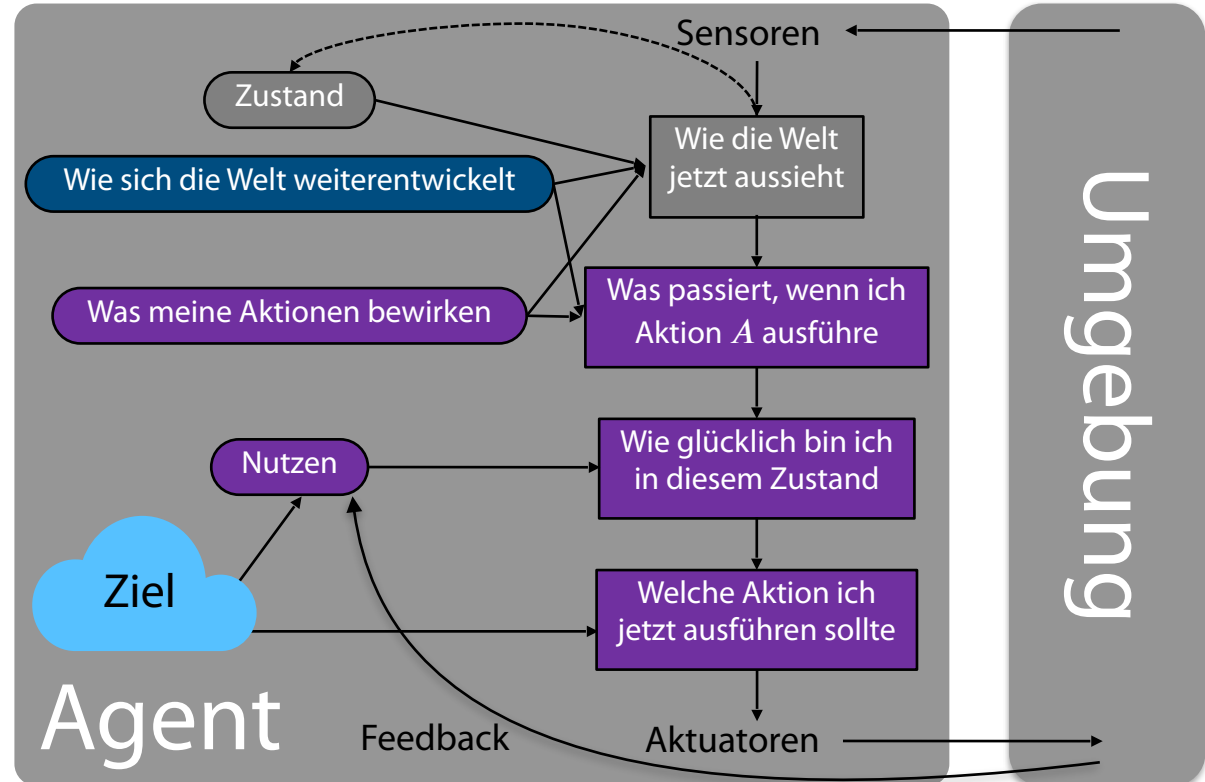


II.

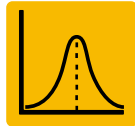
Intelligente Agenten

Agenten

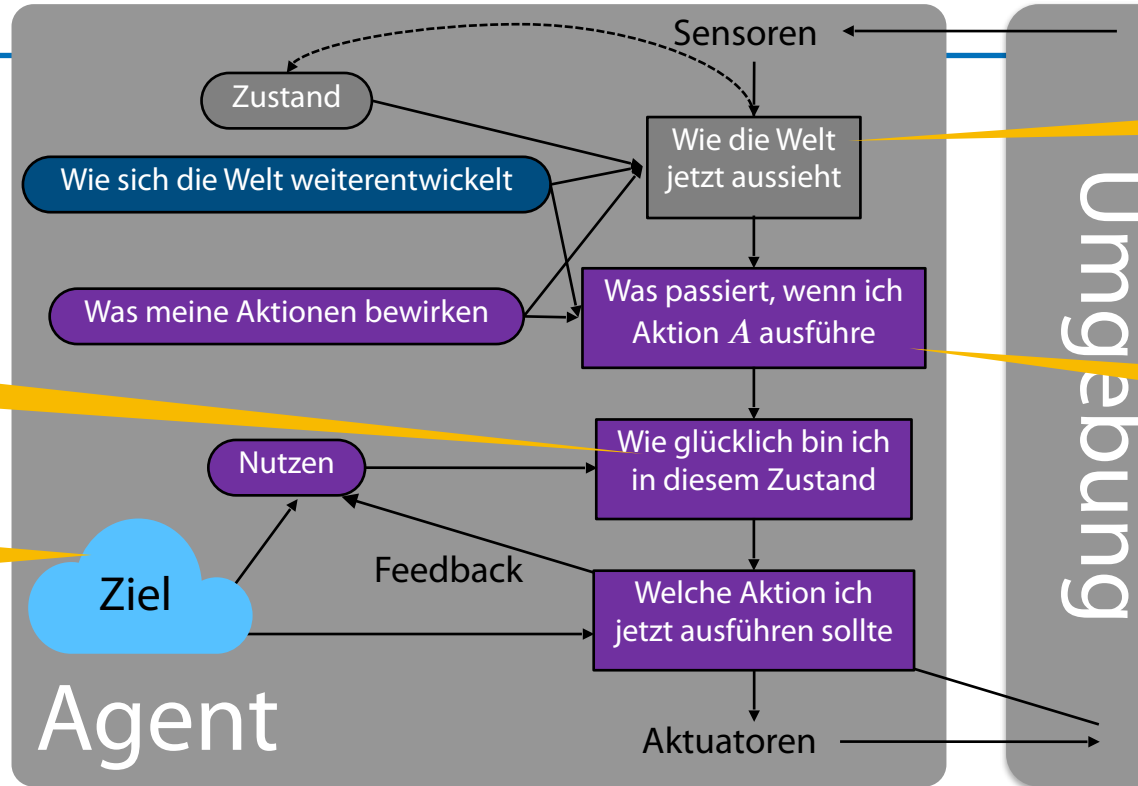
- Wissenschaft der intelligenten Systeme
- Agenten
 - Haben/bilden Ziele
 - Sensoren/Aktoren
 - Handlungsplanung
 - Lernen zur Laufzeit
- Agenten interagieren mit Menschen (and anderen Agenten)
 - Ziele der Agenten beeinflussbar



Regression



Reinforcement Learning
(Ziel und Nutzen aus Feedback „lernen“)

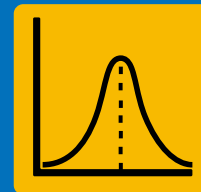
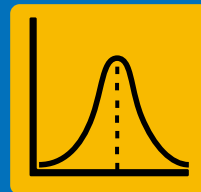


Bilderkennung,
Klassifikation



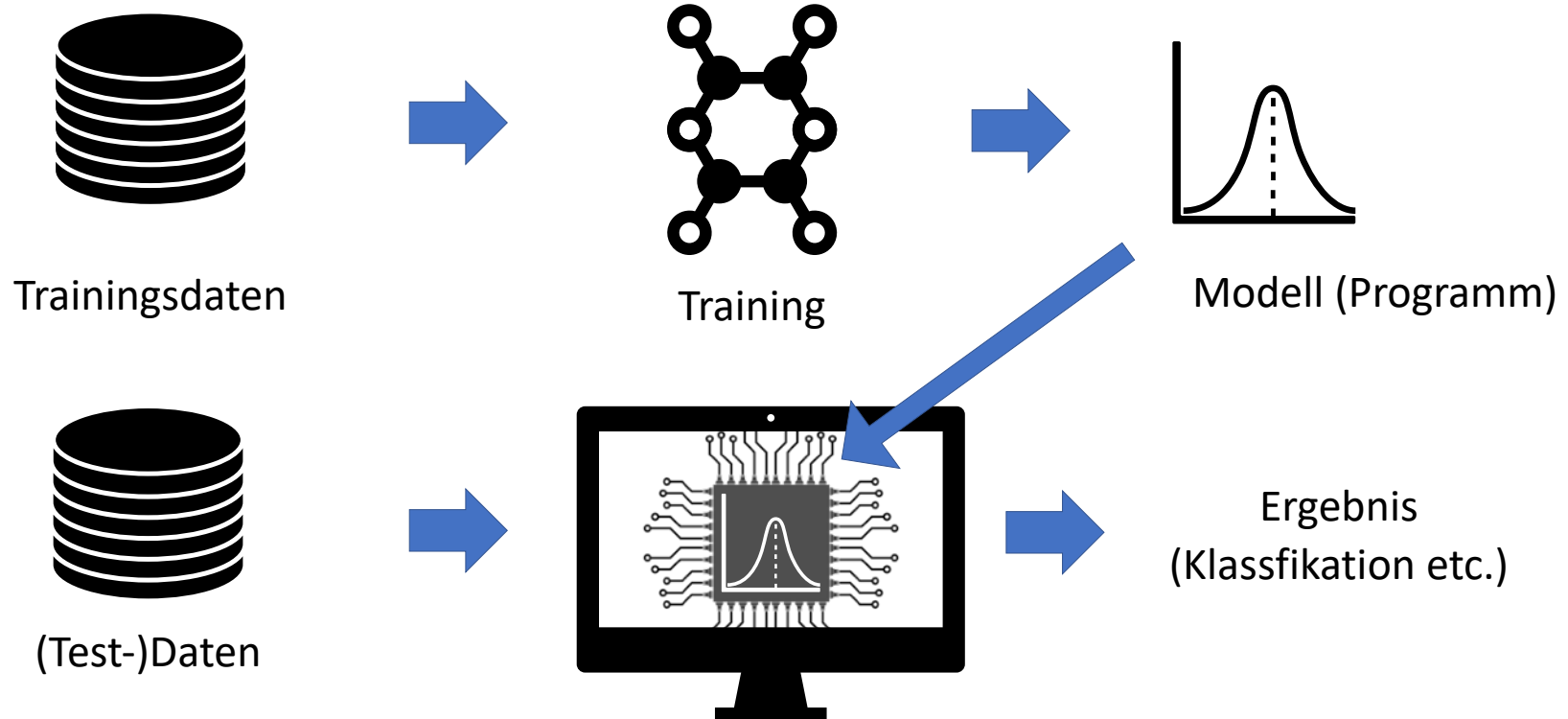
Causal Models
(Intervention)

Causal Models (Counterfactuals)
"Was wäre, wenn ich x getan hätte?"



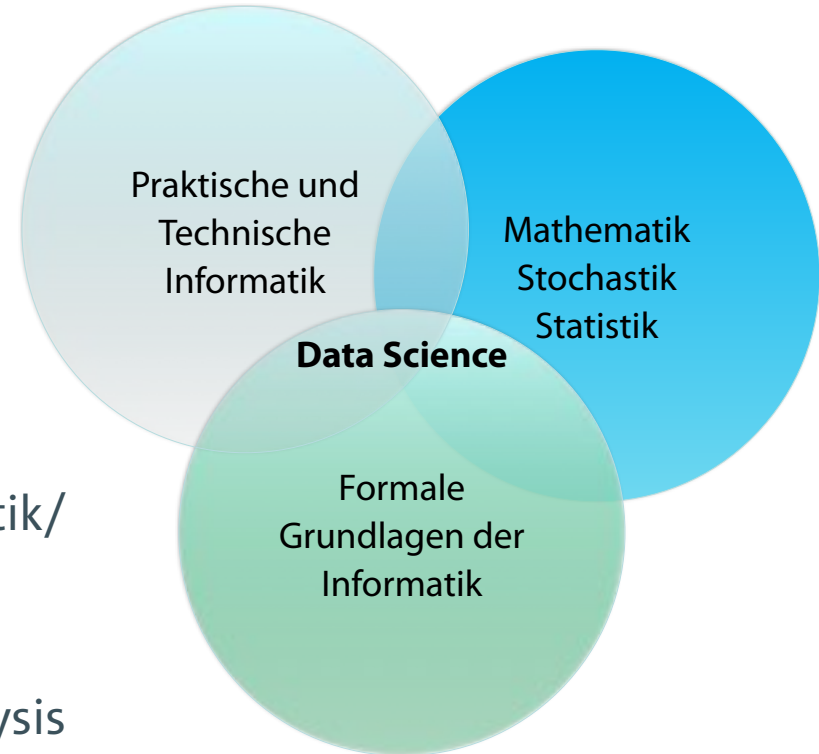
III.

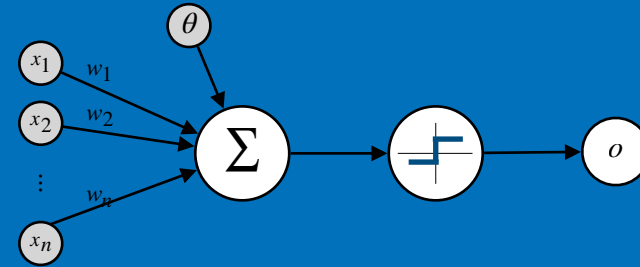
Maschinelles Lernen



Data Science

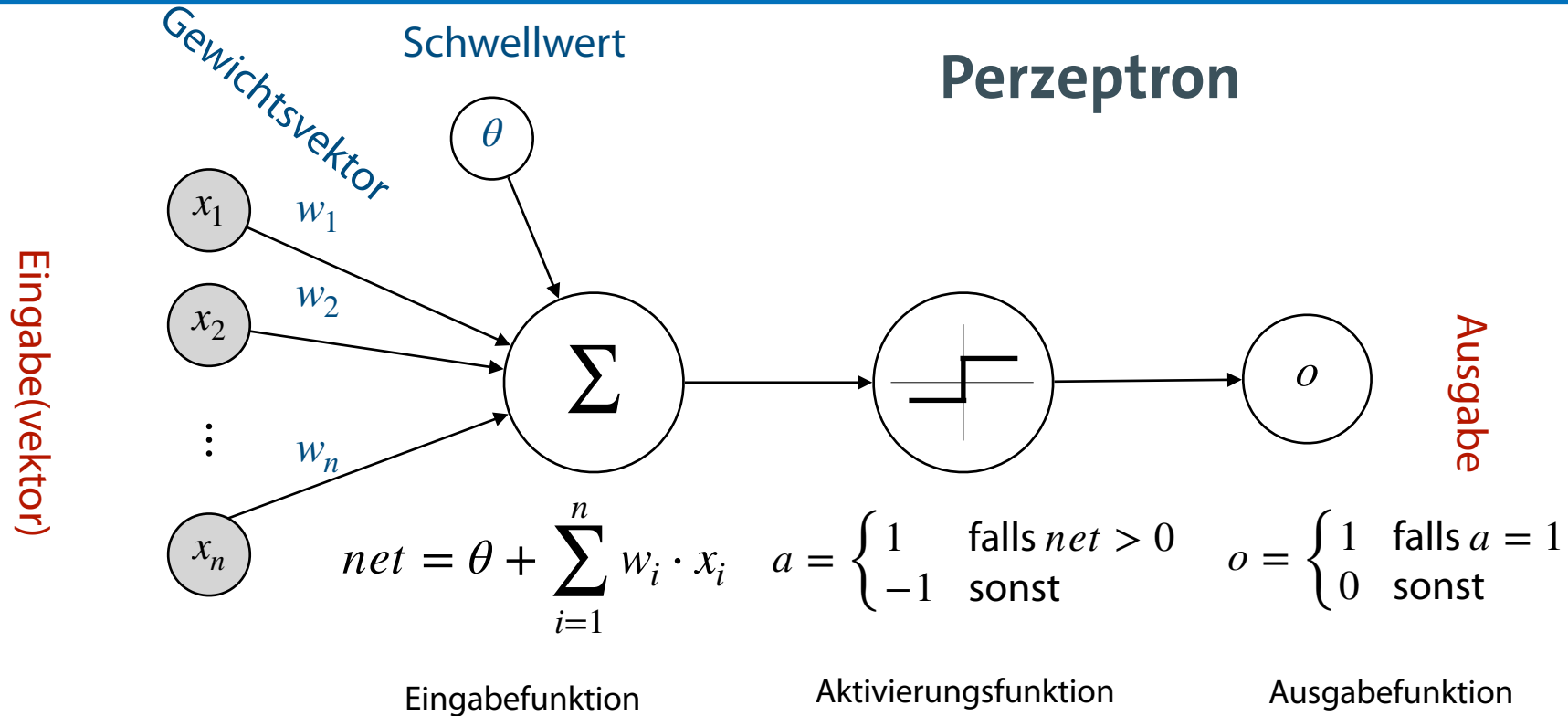
- Extraktion von Wissen aus Daten (u.a. Graphdaten)
- Begriff schon vor 60 Jahren für Informatik vorgeschlagen
- Entwicklung innovativer Konzepte in den Bereichen Logik, Datenbanken und Stochastik/Statistik (Datenanalyse und Wissensentdeckung)
- Verwendung von linearer Algebra und Analysis

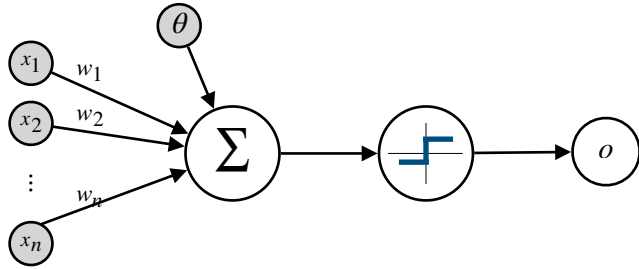




IV. a.

Perzeptron

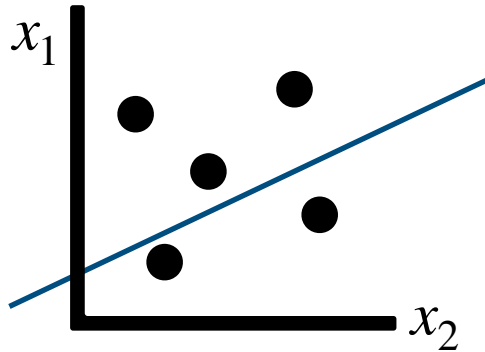




Zusammengefasst

$$o(\vec{x}) = \begin{cases} 1 & \text{falls } \vec{x} \cdot \vec{w} > \theta \\ 0 & \text{sonst} \end{cases}$$

Lineare Funktion



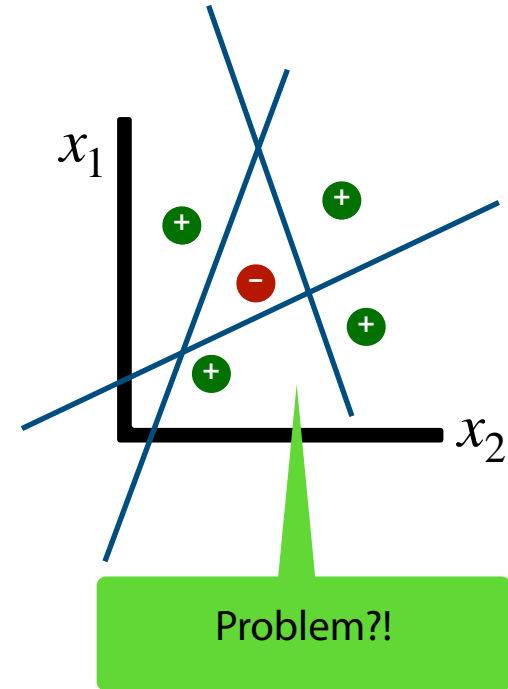
Perzeptron

Lineare Funktion zur Klassifikation

- Eingabe eines mehrdimensionalen Vektors $\vec{x}$
- Ausgabe $o \in \{0,1\}$ (häufig auch $o \in \{-1,1\}$)

Parameter

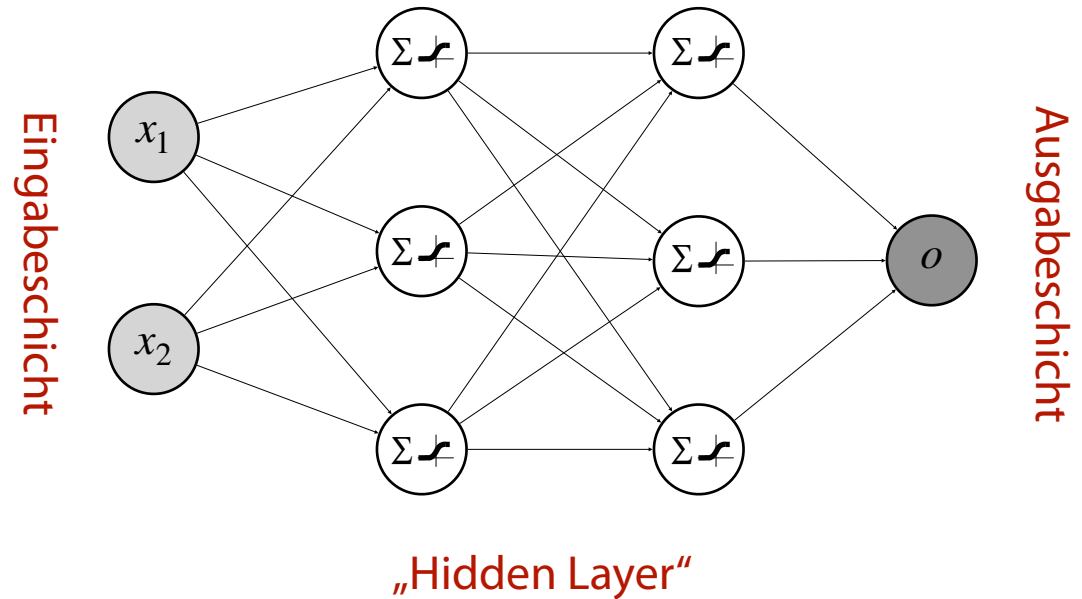
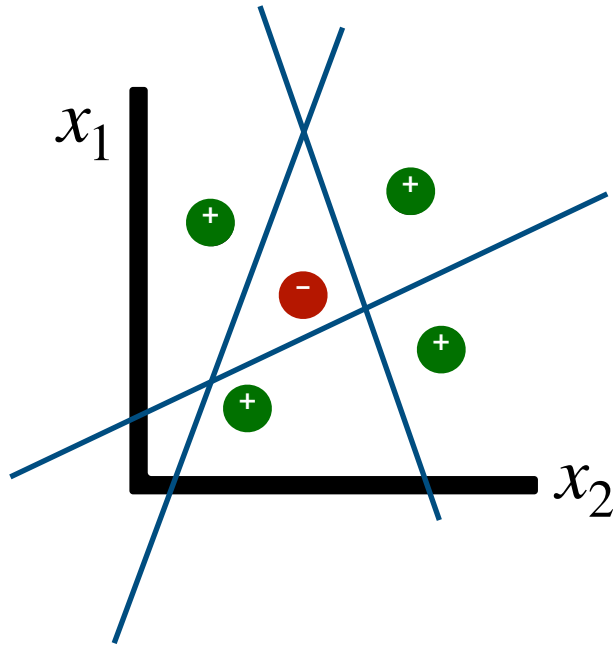
- Gewichte $\vec{w}$ und Schwellwert θ
- *Müssen bestimmt/ gelernt werden*



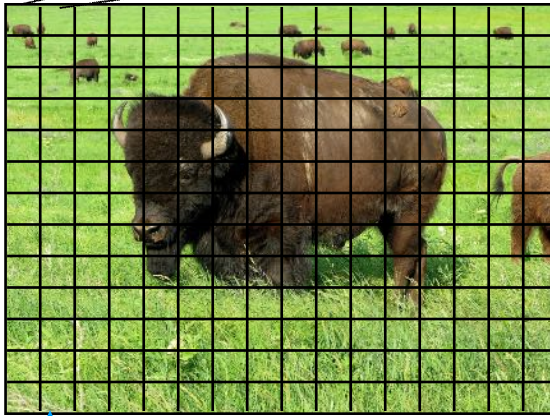
IV. b.

Perzeptron Netzwerke – Deep Learning

Kombination von Perzeptrons

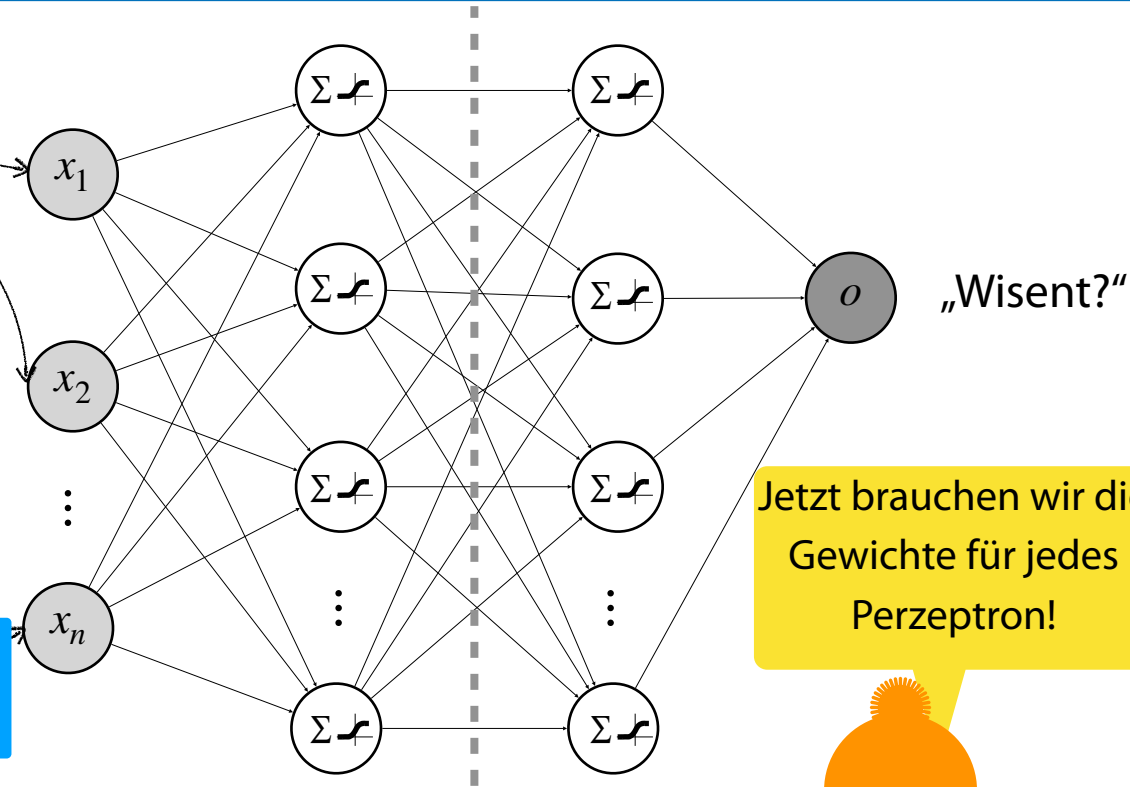


Beispiel

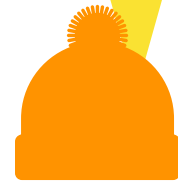


- Gegeben viele Bilder verschiedener Tiere
- Ergebnis zu „Wisent?“ für jedes Bild

https://commons.wikimedia.org/wiki/File:American_Bison_9.jpg

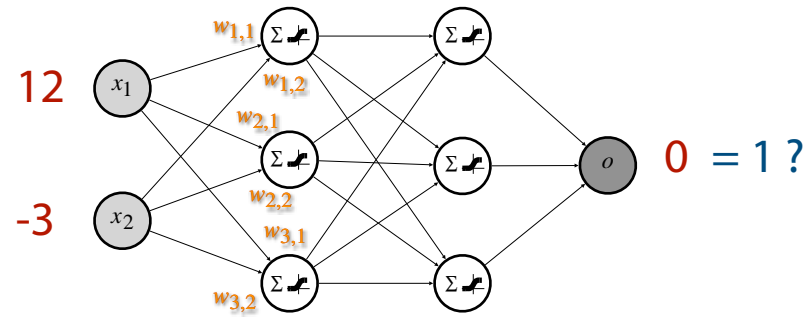


Jetzt brauchen wir die
Gewichte für jedes
Perzeptron!



Lernen der Gewichte (Idee)

1. Gewichte zufällig initialisieren
2. Optimieren der Gewichte
 - i. Ergebnis (Ausgabe) für einen Trainingsdatensatz bestimmen
 - ii. Vergleich mit Zielausgabe
 - iii. Gewichte gemäß Fehler anpassen
 - iv. Wiederholen, bis Fehler ausreichend klein
3. Netzwerk mit neuen Daten testen



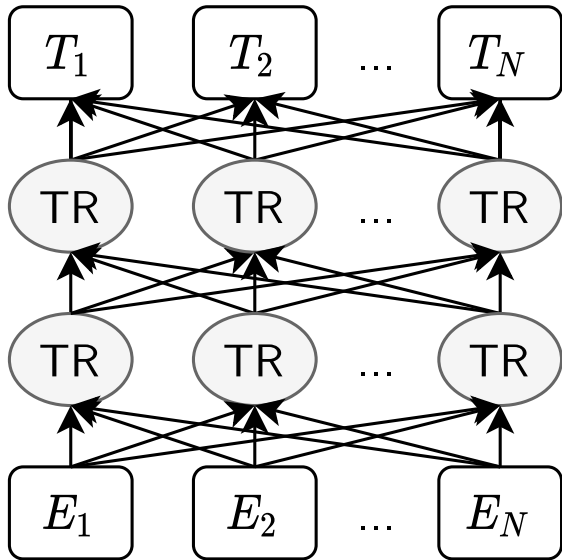
Zufällig oder geht das besser?



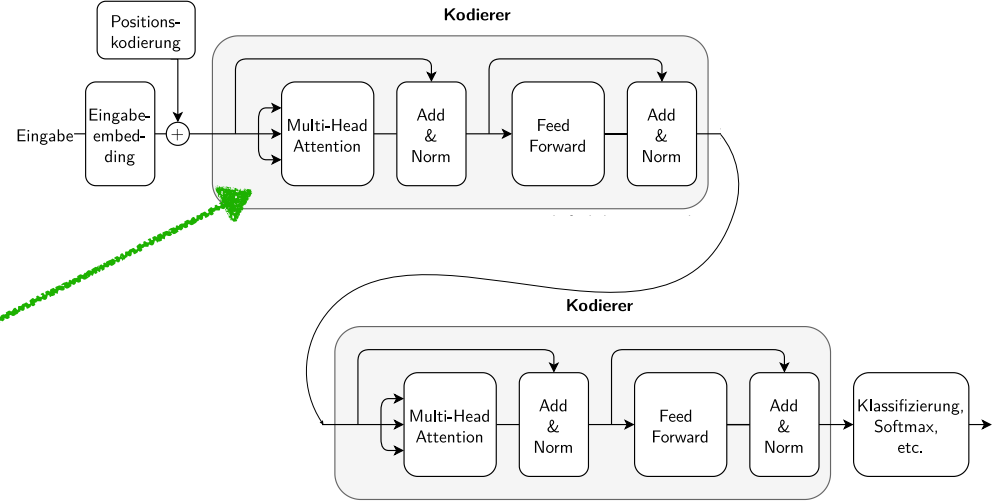
IV. c. Transformer Sprachmodelle



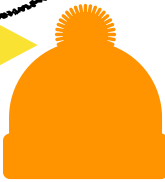
Bidirectional Encoder Representations



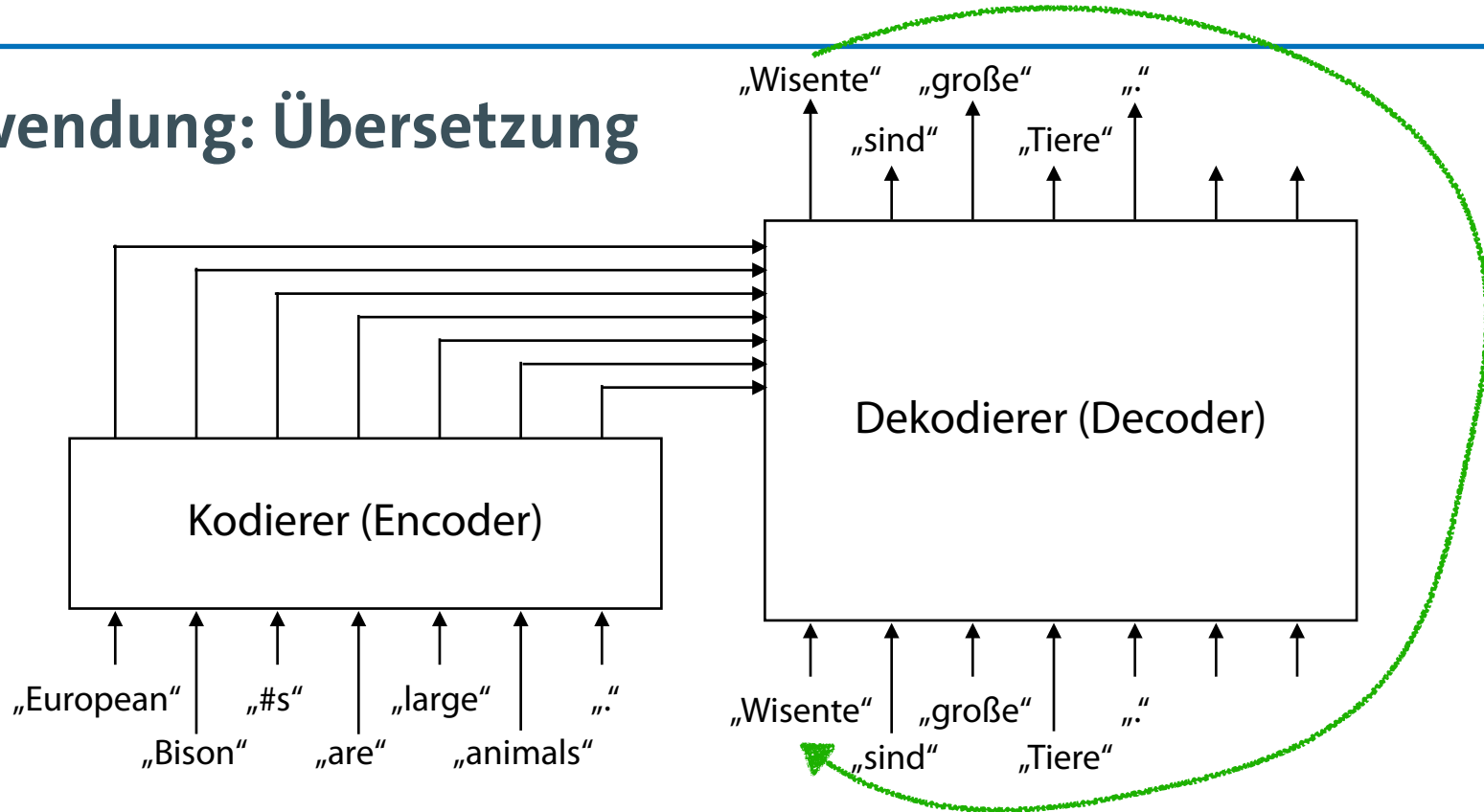
Devlin et al., „Pre-training of Deep Bidirectional Transformers for Language Understanding“, NAACL-HLT 2019



Weiterhin Grundidee des Netzwerks,
 nur viel aufwendigere Module als
 Perzeptrone → Module in Modulen ...



Anwendung: Übersetzung



Chat GPT

Fine-Tuning von GPT 3.5 (kostenfrei) oder GPT 4 (Abo)

- GPT 4 ist „multimodal“, d.h.,
nicht nur Text

Zusätzlich

- Überwachtes Lernen
- Reinforcement Learning
(Feedback)

<https://chat.openai.com/chat>

OP

What are the differences to GPT?



ChatGPT is a variant of GPT, but there are a few key differences between the two models.

1. Training data: ChatGPT is specifically trained on conversational data, such as dialogues and chat logs, whereas GPT is trained on a broader range of internet text. This makes ChatGPT better suited for generating responses in a conversational context.
2. Fine-tuning: ChatGPT is fine-tuned on conversational tasks, such as question answering and dialogue generation, whereas GPT is fine-tuned on a wide range of language tasks such as language translation, summarization and text generation.
3. Model size: ChatGPT is a smaller model than GPT, it has fewer parameters and therefore requires less computational power to run.
4. Output: ChatGPT is trained to generate human-like text specifically for chatbot applications, it can be fine-tuned to understand and generate text that is specific to a certain conversational topic (e.g. customer service, technical support, e-commerce) while GPT is more general and can be fine-tuned to generate text on a wide range of topics.

In summary, ChatGPT is a smaller, more specialized version of GPT, which is specifically designed for conversational tasks and fine-tuned on conversational data.





Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

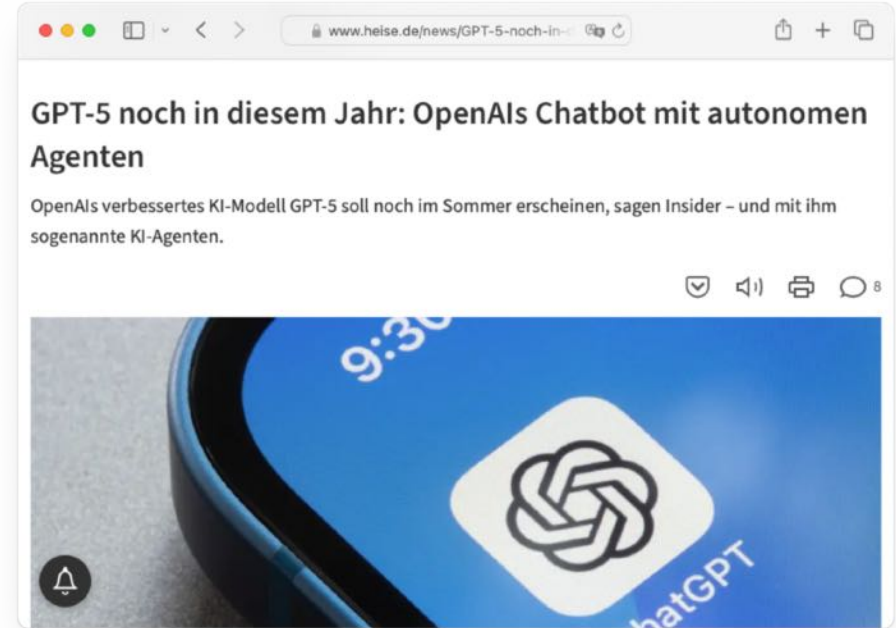


CHAI
Humanities-Centered AI

V. Fazit

Fazit

- Agent
- Künstliche Intelligenz
- Maschinelles Lernen
- DeepLearning
- Transformer (BERT, GPT)
- ChatBot (ChatGPT, Gemini)
- Data Science



<https://www.heise.de/news/GPT-5-noch-in-diesem-Jahr-OpenAIs-Chatbot-mit-autonomen-Agenten-9662010.html>