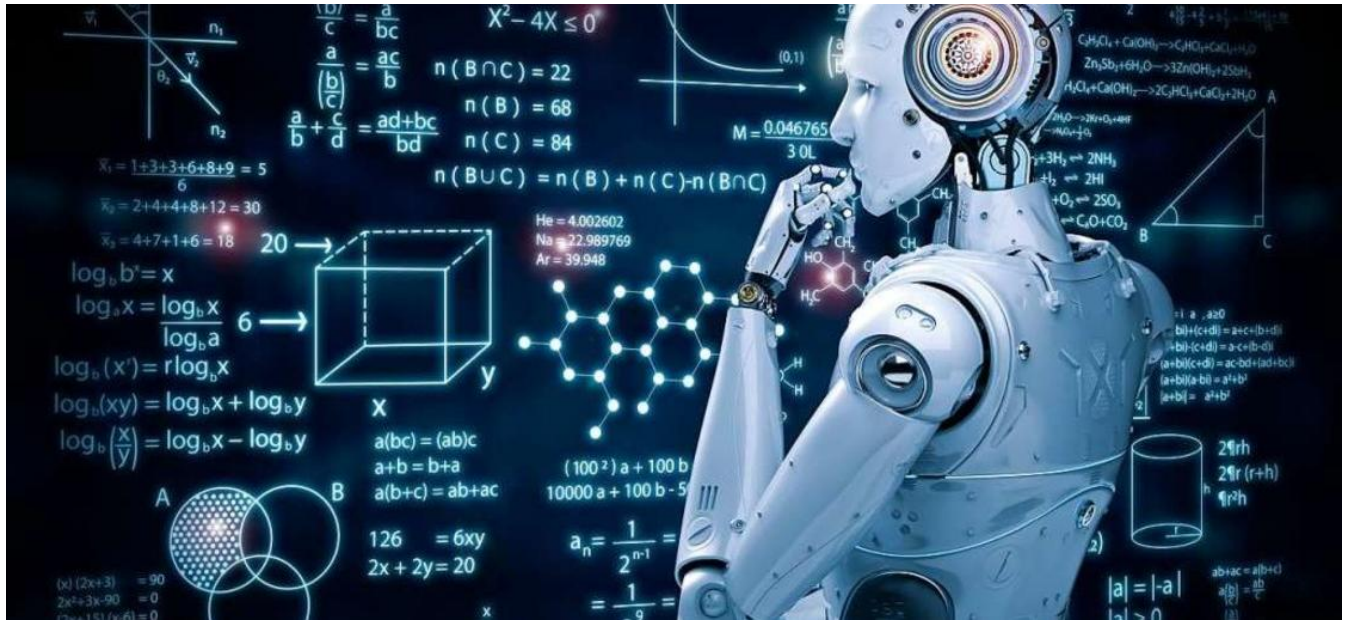
HOCHSCHULE HEILBRONN



› AI FOR BUSINESS APPLICATIONS

VORLESUNGSBEGLEITENDE LITERATUR

HOCHSCHULE HEILBRONN

HILBERT, S., KRAUS, E., & LINDL, A. (2025). *Machine Learning: Eine Einführung für Psychologie, Geistes- und Sozialwissenschaften* (Quantitative Sozialforschung). Wiesbaden: Springer Fachmedien Wiesbaden.

KOCH, D., KOHNE, A., & BRECHBÜHLER, N. (2025). *Prompt Engineering im Unternehmen – eine Einführung*. Wiesbaden: Springer Gabler

Grafiken wurden teilweise mit LLMs erstellt.

KLAUSUR

HOCHSCHULE HEILBRONN

- Keine Open-Book-Klausur
- Aber: Selbstgeschriebenes Cheat Sheet (1 Blatt, Vorder- + Rückseite)

INHALT

HOCHSCHULE HEILBRONN

› 1. EINLEITUNG

KÜNSTLICHE INTELLIGENZ DEFINITIONEN

HOCHSCHULE HEILBRONN

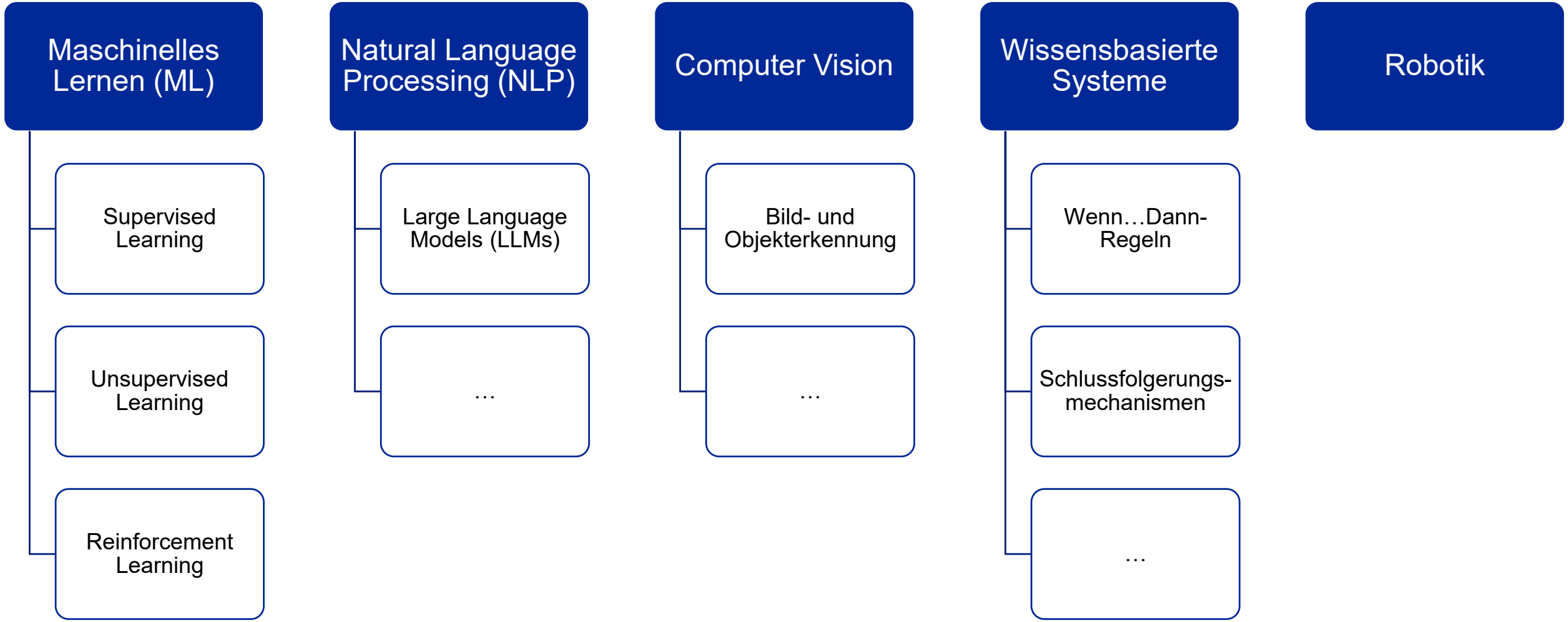
- „Künstliche Intelligenz ist die Wissenschaft und Technik, intelligente Maschinen zu bauen, insbesondere intelligente Computerprogramme.“
  (McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*)

- „Künstliche Intelligenz ist die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren.“
  (https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt)

- „KI ist die Entwicklung von Agenten, die ihre Umgebung wahrnehmen und Handlungen ausführen, um ihre Ziele zu erreichen.“
  (Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach.*)

KÜNSTLICHE INTELLIGENZ TEILBEREICHE

HOCHSCHULE HEILBRONN



MASCHINELLES LERNEN: PYTHON

Installieren Sie Anaconda für die Nutzung von Python:

https://www.anaconda.com/download

› 2. PROMPT ENGINEERING

PROMPT ENGINEERING
LLMS: ABGRENZUNG VON GEN-AI

HOCHSCHULE HEILBRONN

Generative Künstliche Intelligenz (GenAI):

> lernt auf Basis von Trainingsdaten statistische Muster und

> erzeugt daraus neue Inhalte

Beispiele:

> DALL-E (text-to-image)
  Link: https://openai.com/de-DE/index/dall-e-3/

> ChatGPT (text-to-text)
  Link: https://chatgpt.com/

> MusicLM (text-to-music)
  Link: https://musiclm.com/

PROMPT ENGINEERING
LLMS: ABGRENZUNG VON GEN-AI

HOCHSCHULE HEILBRONN

Large Language Models (LLMs):

> spezifische Unterkategorie der generativen KI

> große neurale Netze mit Milliarden von Parametern (trainierbare Werte)

> reagieren auf *Prompts* (= Eingabeaufforderungen)

> Text fortsetzen, Fragen beantworten, Übersetzen, Zusammenfassen, Programmieren,…

PROMPT ENGINEERING
LLMS: ABGRENZUNG VON GEN-AI

HOCHSCHULE HEILBRONN

Beispiele *Large Language Models (LLMs):*

- ChatGPT (OpenAI)
- Claude (Anthropic)
  Link: https://claude.ai/new
- Gemini (Google)
  Link: https://gemini.google.com/app?hl=de
- Mistral (Open Source)
  Link: https://mistral.ai/
- LlaMA (Meta)
  Link: https://www.llama.com/

PROMPT ENGINEERING
LLMS: FUNKTIONSWEISE

HOCHSCHULE HEILBRONN

Hauptkomponenten LLM:

> Tokenisierung

> Transformer-Architektur

Siehe:

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. Advances in Neural Information Processing Systems.

Link: https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

PROMPT ENGINEERING
LLMS: TOKENISIERUNG

HOCHSCHULE HEILBRONN

Bevor ein Transformer Text verarbeiten kann, muss er den Text in kleine Einheiten zerlegen.

> Diese Einheiten heißen *Tokens.*

> Nicht unbedingt Wörter, sondern Wortteile, Silben oder sogar einzelne Zeichen

> Beispiel (vereinfachtes Byte-Pair-Encoding-Verfahren):
  - Satz: *„Chatbots sind hilfreich.“*
  - Tokens: ["Chat", "bot", "s", "sind", "hilf", "reich", "."]

Tokenisierung = Übersetzung von Sprache in maschinenlesbare Bausteine.

PROMPT ENGINEERING
LLMS: TRANSFORMER-ARCHITEKTUR

HOCHSCHULE HEILBRONN

1. Eingabe (Tokens → Vektoren)

> Jedes Token wird in einen Vektor eingebettet (sogenanntes Embedding).

> Zusätzlich: Positionskodierung → Modell weiß, in welcher Reihenfolge Wörter stehen.

2. Self-Attention (Aufmerksamkeitsmechanismus)

> Bei jedem Wort Berücksichtigung alle anderen im Satz, um Relevanz zu bestimmen

> Beispiel:
In *„Der Hund jagt die Katze, weil sie schnell ist“* muss das Modell erkennen, dass „sie“ sich auf „Katze“ bezieht – dank Attention klappt das.

PROMPT ENGINEERING
LLMS: TRANSFORMER-ARCHITEKTUR

HOCHSCHULE HEILBRONN

3. Feedforward-Netzwerk

> Nach der Attention-Layer:
Vektoren durchlaufen kleine neuronale Netze, die Muster verstärken oder abschwächen.

4. Mehrere Schichten (Stacking)

> Transformer bestehen aus zahlreichen solcher Attention- + Feedforward-Blöcken.

> Je tiefer (mehr), desto „intelligenter“ die Repräsentation von Sprache.

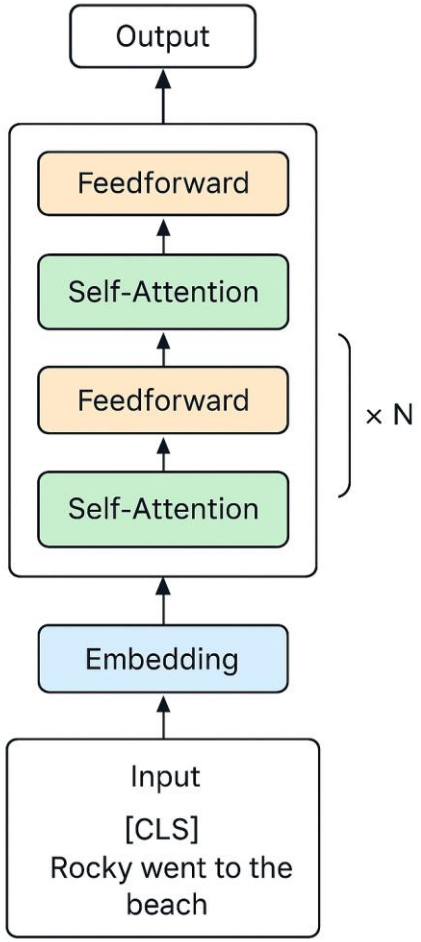
5. Ausgabe (Decoding)

> Modell berechnet, welches Token am wahrscheinlichsten als Nächstes kommt.

> Beispiel: Prompt „Die Sonne geht im ___ auf“ → Wahrscheinlichstes Token: „Osten“.

PROMPT ENGINEERING
LLMS: TRANSFORMER-ARCHITEKTUR

HOCHSCHULE HEILBRONN



PROMPT ENGINEERING
LLMS: STÄRKEN & SCHWÄCHEN

HOCHSCHULE HEILBRONN

✅ Stärken von LLMs

> Breites Wissen und Vielseitigkeit

> Sprachverständnis & Kontextverarbeitung

> Kreativität & Ideenfindung

> Anpassbarkeit

> Produktivitätssteigerung

❌ Schwächen von LLMs

> Halluzinationen (Faktenfehler)

> Begrenztes Kontextfenster

> Bias & Ethikprobleme

> Keine echte Logik oder Weltmodell

> Kosten & Ressourcen

> Abhängigkeit von Datenqualität

> Pfadabhängigkeit

PROMPT ENGINEERING
PROMPT-DESIGN-PRINZIPIEN

HOCHSCHULE HEILBRONN

1. Klarheit & Präzision
2. Kontextbereitstellung
3. Rollen- und Stilanweisung [optional]
4. Formatvorgabe
5. Iteratives Prompting
6. Begrenzung
7. Explizite Aufgabenstellung

PROMPT ENGINEERING
PROMPT-TECHNIKEN

HOCHSCHULE HEILBRONN

Grundlegende Prompt-Techniken

1. Zero-Shot Prompting:
   Frage oder Aufgabe, ohne Beispiel oder zusätzliche Anweisung.

2. One-Shot Prompting:
   Beispiel als Orientierung, bevor die eigentliche Aufgabe gestellt wird

3. Few-Shot Prompting:
   Mehrere Beispiele, um das gewünschte Muster, Format oder den Stil zu verdeutlichen.

PROMPT ENGINEERING
PROMPT-TECHNIKEN

HOCHSCHULE HEILBRONN

Präzisierende Prompt-Techniken

4. Chain-of-Thought Prompting (CoT):
   Modell wird aufgefordert, den Denkprozess Schritt für Schritt zu erklären.

5. Self-Consistency Prompting:
   Modell soll mehrere Lösungswege generieren und konsistenteste Antwort wird gewählt.

6. ReAct Prompting (Reason + Act):
   Kombination aus Schlussfolgern (Reasoning) und Handeln (z. B. Informationen nachschlagen, Tools nutzen).

PROMPT ENGINEERING
PROMPT-TECHNIKEN

HOCHSCHULE HEILBRONN

Formatierungs- und Strukturtechniken

7. Instruction Prompting:
   Sehr klare Anweisungen

8. Role Prompting:
   Du setzt das Modell in eine Rolle

9. Persona / Style Prompting:
   Vorgabe eines Stils oder Charakter

10. Output-Formatting:
    Du definierst das gewünschte Format (z. B. JSON, Tabelle, Markdown).

PROMPT ENGINEERING
PROMPT-TECHNIKEN

HOCHSCHULE HEILBRONN

Fortgeschrittene Techniken

11. Prompt Chaining:
    Komplexe Aufgabe wird in mehrere Prompts zerlegt, die aufeinander aufbauen.

12. Meta-Prompting:
    Du erklärst explizit wie das Modell vorgehen soll.

13. Reflexion Prompting:
    Modell wird aufgefordert, seine Antwort zu überprüfen und ggf. zu verbessern

14. Multimodal Prompting:
    Kombination aus Text und Bild/Audio (falls unterstützt).

PROMPT ENGINEERING
BEWERTUNG VON PROMPTS

HOCHSCHULE HEILBRONN

1. Manuelle Evaluation
2. Definierte Metriken
3. A/B-Tests
4. LLM-as-a-Judge
5. User Feedback
6. Stress-Tests / Robustheitstests

› 3. MASCHINELLES LERNEN I: SUPERVISED LEARNING

SUPERVISED LEARNING: DEFINITION UND AUFGABEN

HOCHSCHULE HEILBRONN

Supervised Learning (überwachtes Lernen) ist ein Teilbereich des maschinellen Lernens,

> bei dem ein Modell anhand von Beispieldaten trainiert wird,

> die sowohl *Vorhersagekriterien (Features)*

> als auch die dazugehörigen *Ausgaben (Labels/Zielwerte*) enthalten.

Typische Aufgaben:

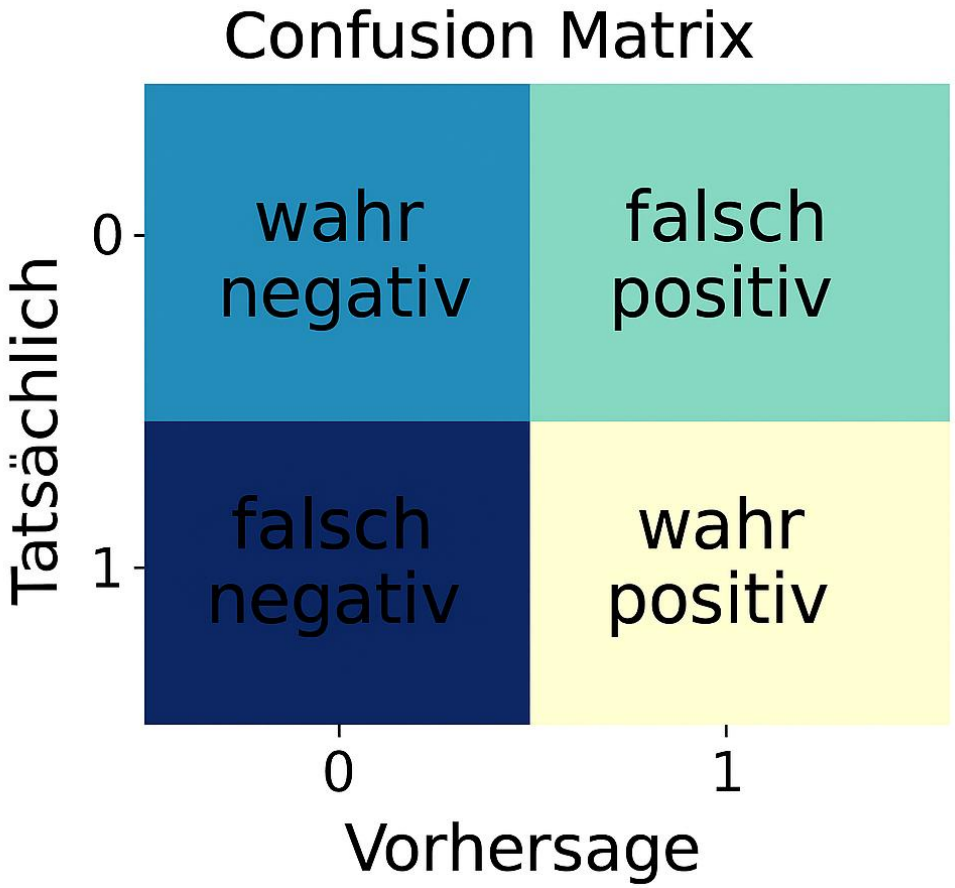
> *Klassifikation* → Vorhersage einer Kategorie (z. B. Spam vs. Nicht-Spam)

> *Regression* → Vorhersage eines Werts (z. B. Temperatur morgen, Aktienkurs)

SUPERVISED LEARNING: EVALUATION VON KLASSIFIKATIONSMODELLEN

HOCHSCHULE HEILBRONN

Basis Wahrheitsmatrix (Confusion Matrix):



$$Accuracy = \frac{Wahr\ Positiv + Wahr\ Negativ}{Gesamtanzahl}$$

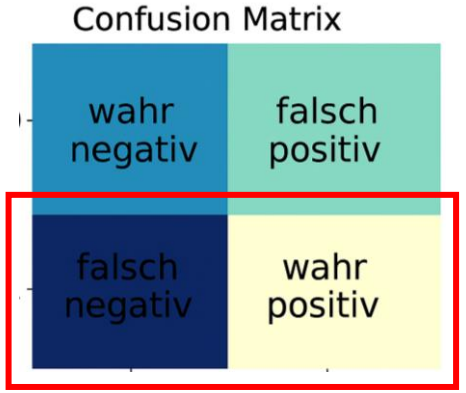
SUPERVISED LEARNING: EVALUATION VON KLASSIFIKATIONSMODELLEN

HOCHSCHULE HEILBRONN

Weitere Metriken:

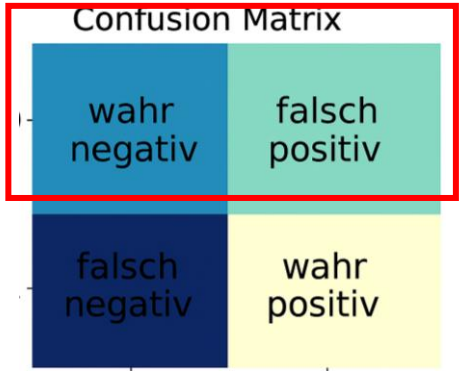
> Sensitivität / True Positive Rate / Recall_1:

$$Sensitivität = \frac{Wahr\ Positiv}{Wahr\ Positiv + Falsch\ Negativ}$$



> Spezifizität/ True Negative Rate / Recall_0:

$$Spezifizität = \frac{Wahr\ Negativ}{Wahr\ Negativ + Falsch\ Positiv}$$



SUPERVISED LEARNING: EVALUATION VON REGRESSIONSMODELLEN

HOCHSCHULE HEILBRONN

Gegeben wahre Werte (y_i), vorhergesagte Werte ($\hat{y}_i$) und Stichprobengröße (n), dann

- ist der Mean Squared Error

$$MSE = \frac{1}{n}[(y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2]$$

- ist der Root Mean Squared Error

$$RMSE = \sqrt{MSE}$$

- ist das Bestimmtheitsmaß R^2

$$R^2 = 1 - \frac{Erklärte\ Varianz}{Totale\ Varianz} = 1 - \frac{(y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2}{\left(y_1 - Mittelwert_y\right)^2 + \cdots + \left(y_n - Mittelwert_y\right)^2}$$

SUPERVISED LEARNING: VORGEHENSWEISE

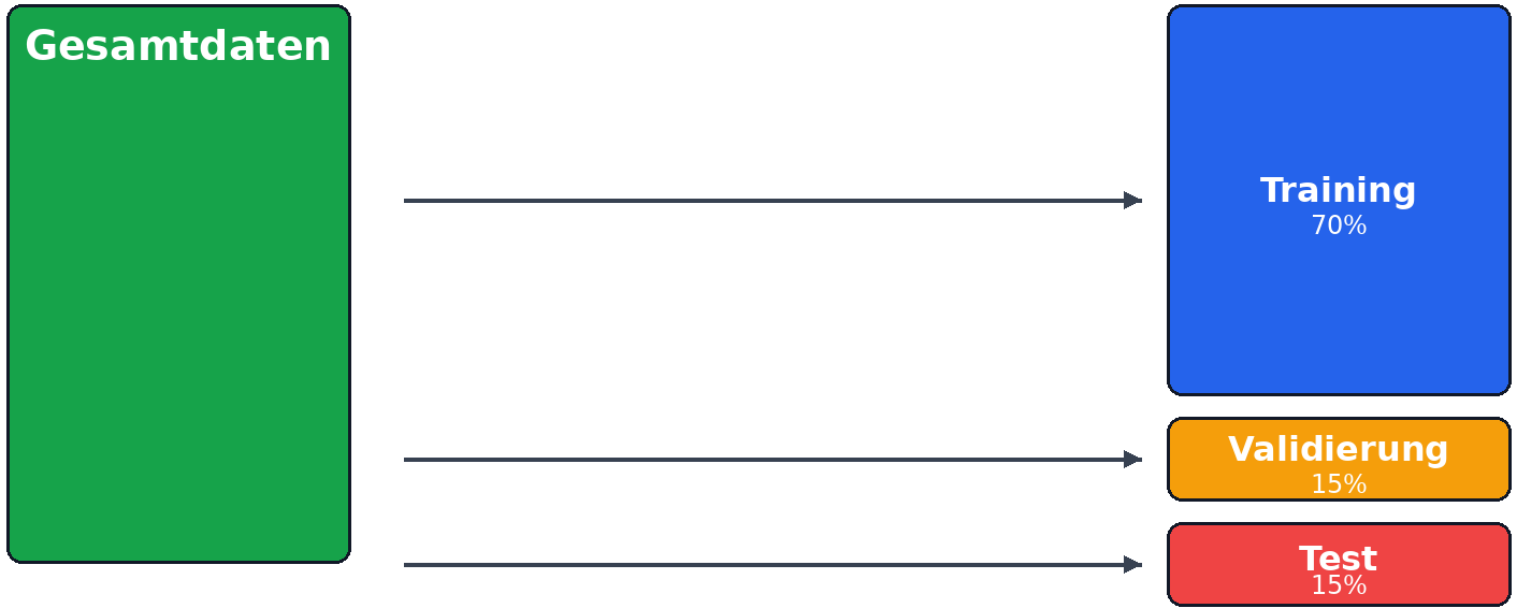
HOCHSCHULE HEILBRONN

Daten → Aufbereiten → Modell wählen → Trainieren → Testen → Optimieren → Anwenden

SUPERVISED LEARNING: VORGEHENSWEISE

HOCHSCHULE HEILBRONN

Train-Test-Split:



Train/Validation/Test Split

SUPERVISED LEARNING: STANDARD ALGORITHMEN

HOCHSCHULE HEILBRONN

Lineare Modelle	Tree-basierte Modelle	Neuronale Netze	Sonstige
Lineare Regression	Decision Trees	Einfache KNN	Support Vector Machines
Logistische Regression	Random Forest	Deep Learning	k-Nearest Neighbors
	Gradient Boosting		Naive Bayes

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Decision Tree Algorithmus:

1. Wurzel: Startet mit allen Trainingsdaten

2. Splits: Wählt das Merkmal aus, das die Daten am besten trennt (gemessen durch Kriterien wie Gini-Impurity oder Entropie)

3. Rekursion: Wiederholt den Prozess für jeden Ast, bis ein Stopp-Kriterium erreicht ist

4. Blätter: Endknoten enthalten die finale Vorhersage (Klasse oder Wert)

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Decision Tree Algorithmus: Gini-Impurity und Entropie

Gegeben Daten mit K Kategorien; p_i sei die relative Häufigkeit der Kategorie $i \in \{1, \dots, K\}$,

dann

> ist die *Gini-Impurity* definiert durch:

$$Gini = p_1(1 - p_1) + \cdots + p_K(1 - p_K) = 1 - (p_1^2 + \cdots + p_K^2)$$

> ist die *Shannon-Entropie* definiert durch:

$$Entropie = -(p_1 \log_2 p_1 + \cdots + p_K \log_2 p_K)$$

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

Decision Tree Algorithmus: Berechnungsbeispiel (Vorlesung)

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Random Forest und Gradient Boosting

> Ensemble-Methoden = Kombination mehrerer Modelle (hier: Trees)

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Hauptarten von Ensemble-Methoden:

1. *Stacking / Voting*
   - Mehrere unterschiedliche Modelle werden trainiert (z. B. SVM, Neuronales Netz).
   - Stacking: Meta-Modell lernt, wie man deren Vorhersagen optimal kombiniert.
   - Voting: Mehrheitsentscheidung (hard voting), Mittelung (soft voting)

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Hauptarten von Ensemble-Methoden:

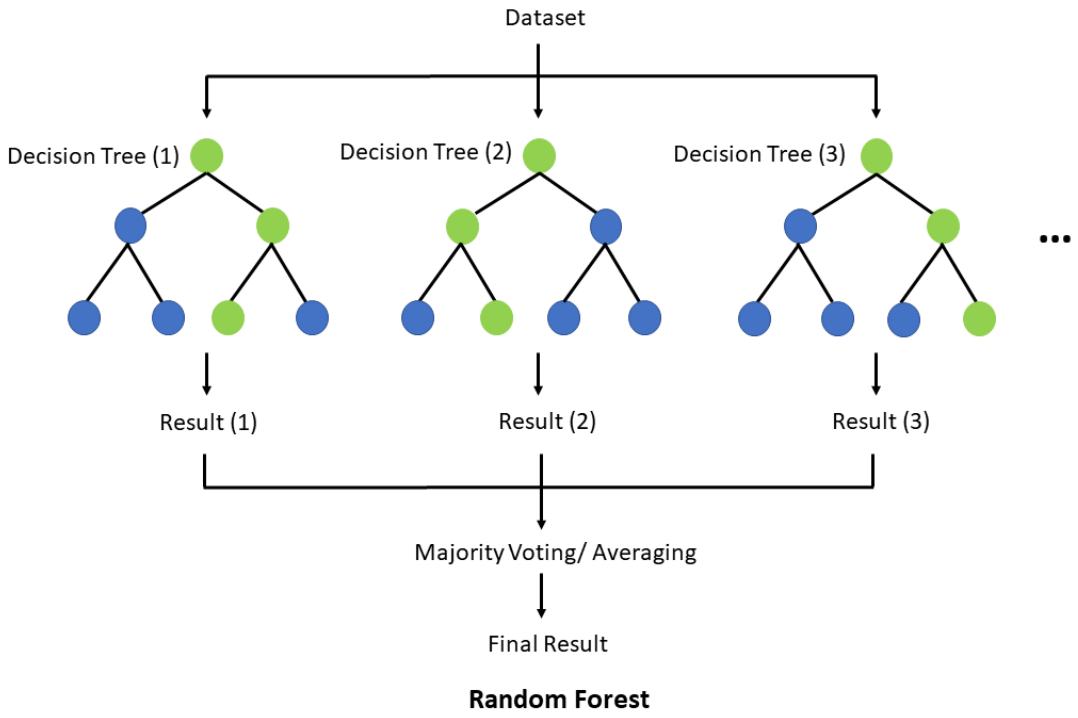
2. *Bagging (Bootstrap Aggregating)*
   - Trainiert viele Modelle parallel auf zufälligen Teilmengen (Ziehen mit Zurücklegen).
   - Ergebnisse werden gemittelt oder abgestimmt.
   - Beispiel: **Random Forest**

3. *Boosting*
   - Trainiert Modelle nacheinander. Neue Modelle reduzieren Fehler der vorherigen.
   - Schwierig vorherzusagende Datenpunkte bekommen mehr Gewicht.
   - Beispiel: **Gradient Boosting**

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

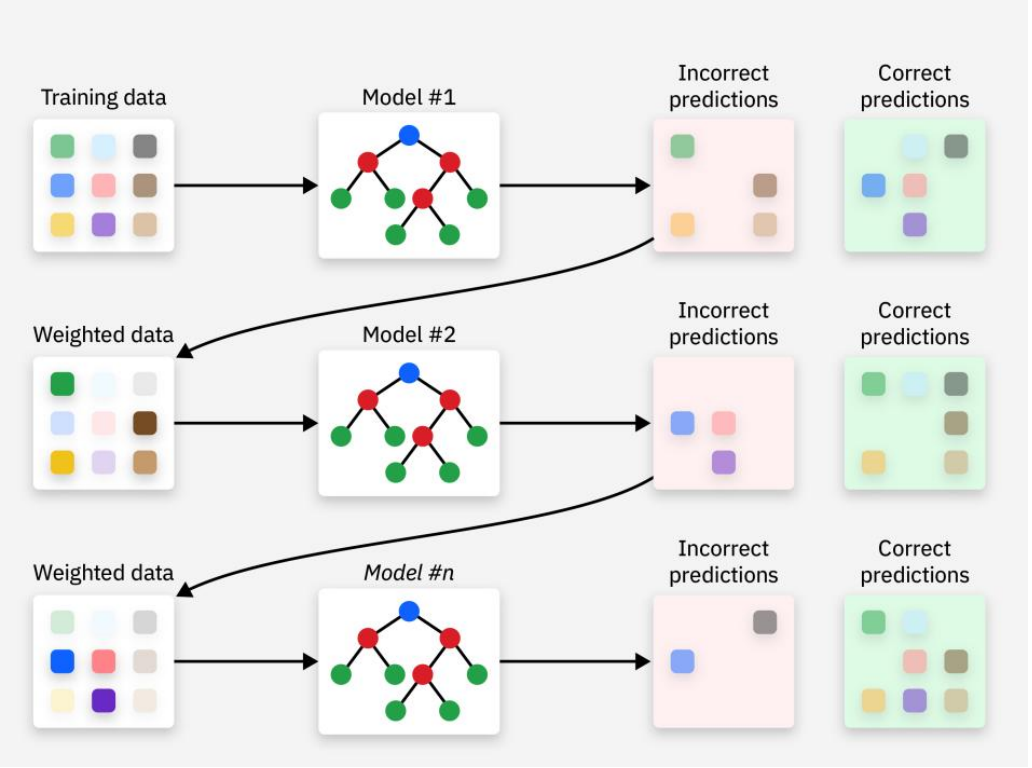
HOCHSCHULE HEILBRONN

Random Forest:



Quelle: https://de.wikipedia.org/wiki/Random_Forest

Gradient Boosting:



Quelle: https://www.ibm.com/think/topics/gradient-boosting

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

Gradient Boosting Algorithmus:

1. Initialisierung: Startet mit einer einfachen Vorhersage (oft der Mittelwert)

2. Iterativer Prozess:
   - Berechne Residuen (Fehler) der aktuellen Vorhersage
   - Trainiere einen neuen schwachen Lerner auf diese Residuen
   - Addiere die neue Vorhersage mit einem Gewichtungsfaktor (Learning Rate)

3. Wichtige Parameter
   - Learning Rate (α): Kontrolliert Schrittgröße (typisch 0.01-0.3)
   - Anzahl Bäume: Mehr Bäume = komplexeres Modell
   - Baumtiefe: Meist shallow trees (3-8 Ebenen)

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

Gradient Boosting Algorithmus: Odds und Log-Odds

Sei p die relative Häufigkeit eines Merkmals, dann

> sind die Odds (o) des Merkmals

$$o = \frac{p}{1-p}$$

> und die Log-Odds sind:

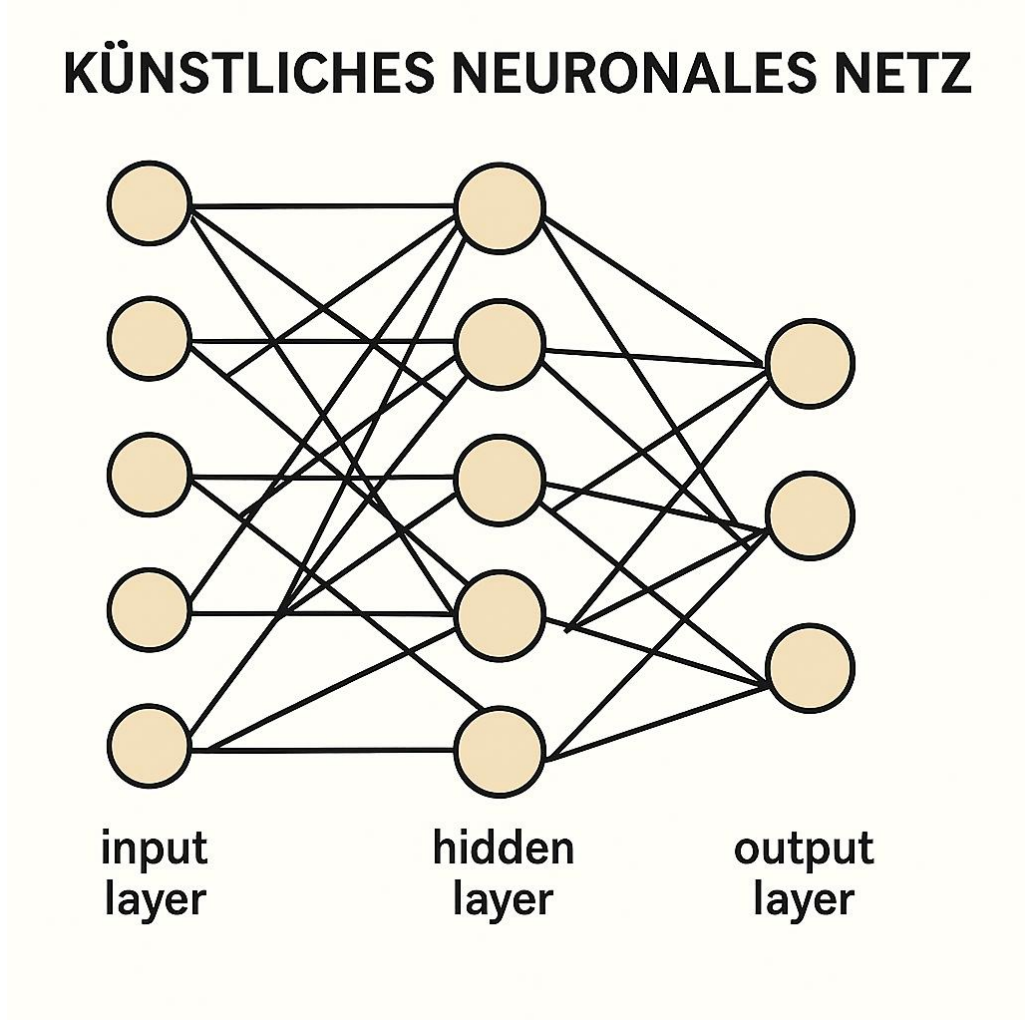
$$\ln(o)$$

SUPERVISED LEARNING: TREE-BASIERTE MODELLE

HOCHSCHULE HEILBRONN

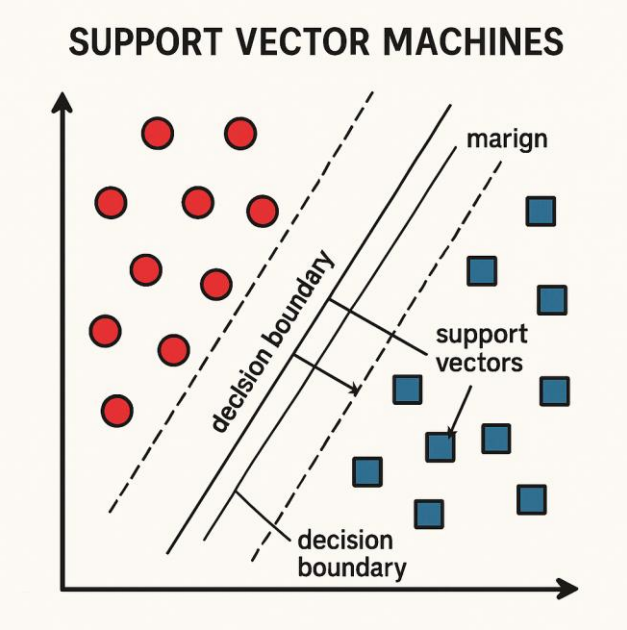
Gradient Boosting Algorithmus: Berechnungsbeispiel (Vorlesung)

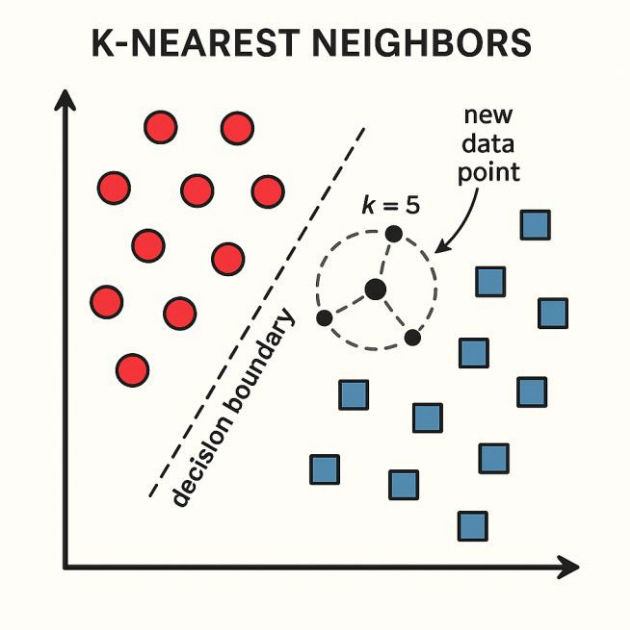
SUPERVISED LEARNING: NEURONALE NETZE

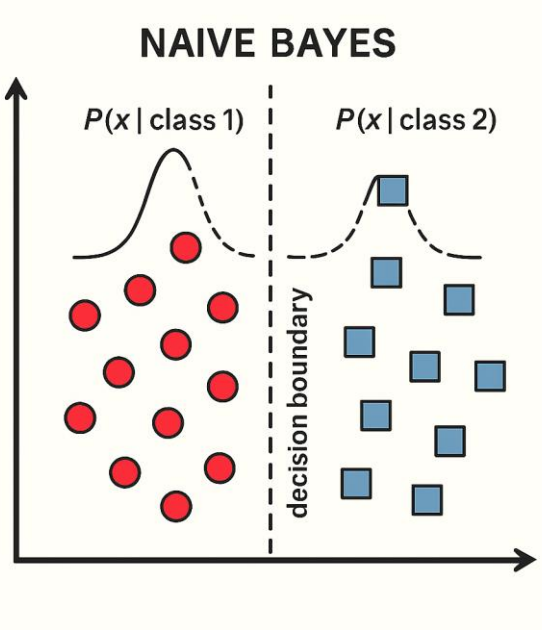


SUPERVISED LEARNING: SONSTIGE MODELLE

HOCHSCHULE HEILBRONN







› 4. MASCHINELLES LERNEN II: UNSUPERVISED LEARNING

UNSUPERVISED LEARNING: DEFINITION UND AUFGABEN

HOCHSCHULE HEILBRONN

Unsupervised Learning (unüberwachtes Lernen) ist ein Bereich des maschinellen Lernens,

> bei dem ein Modell anhand von Beispieldaten trainiert wird,

> *ohne das Ausgaben (Labels/Zielwerte)* vorgegeben sind.

> Aufgabe ist es selbstständig Muster und Strukturen in den Daten zu entdecken.

UNSUPERVISED LEARNING: STANDARD ALGORITHMEN

HOCHSCHULE HEILBRONN

Clustering	Dimensionsreduktion	Anomalieerkennung
K-Means	Principal Component Analysis	Isolation Forest
Hierarchisches Clustering	t-distributed Stochastic Neighbor Embedding	One-Class SVM
DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	Uniform Manifold Approximation and Projection	
Gaussian Mixture Models (GMMs)	Autoencoder	

UNSUPERVISED LEARNING: K-MEANS ALGORITHMUS

HOCHSCHULE HEILBRONN

1. Wähle Anzahl an Clustern k

2. Initialisiere zufällig k Zentren (Centroids)

3. Weise Punkten den Clustern über minimale Distanz zu den Zentren zu

4. Aktualisiere Zentren: Für Cluster C mit Punkten $x_1, \ldots, x_n$:

$$\mu_C = \frac{1}{n}(x_1 + \cdots + x_n)$$

5. Wiederhole Schritt 3 und 4 bis sich die Cluster nicht mehr ändern

UNSUPERVISED LEARNING: K-MEANS ALGORITHMUS

HOCHSCHULE HEILBRONN

Berechnungsbeispiel (Vorlesung)

5. MASCHINELLES LERNEN III: REINFORCEMENT LEARNING

REINFORCEMENT LEARNING: DEFINITION UND AUFGABEN

HOCHSCHULE HEILBRONN

Reinforcement Learning (RL) ist ein Bereich des maschinellen Lernens,

> bei dem ein Agent lernt, optimale Entscheidungen in einer Umgebung zu treffen,

> indem er durch Versuch und Irrtum Erfahrungen sammelt.

REINFORCEMENT LEARNING: KERNKOMPONENTEN

HOCHSCHULE HEILBRONN

- Agent: Das lernende System, das Entscheidungen trifft
- Umgebung: Die Welt, in der der Agent agiert
- Zustand (s): Die aktuelle Situation der Umgebung
- Aktion (a): Was der Agent tun kann
- Belohnung (r): Das Feedback, das der Agent für seine Aktionen erhält
- Policy (π): Die Strategie, die bestimmt, welche Aktion in welchem Zustand gewählt wird

REINFORCEMENT LEARNING: Q-LEARNING

HOCHSCHULE HEILBRONN

1. Initialisierung:
   - Lege Lernrate α, Diskontfaktor γ und Explorationsrate ϵ fest.
   - Erstelle eine Q-Tabelle (Zustände × Aktionen) mit zufälligen oder null Werten.
2. Wähle Aktion
   - Mit Wahrscheinlichkeit ϵ: wähle zufällige Aktion (Exploration).
   - Mit Wahrscheinlichkeit $1 - \epsilon$: wähle Aktion mit dem größten Q-Wert (Exploitation).
3. Führe Aktion aus (Übergang zum neuen Zustand s', erhalte Belohnung r)
4. Aktualisiere Q-Werte:

$$Q(s,a) \leftarrow Q(s,a)+\alpha[r+\gamma \max_{a'} Q(s',a')-Q(s,a)]$$

5. Wiederhole 2.-4. bis Q-Tabelle stabil ist oder maximale Anzahl Episoden erreicht

REINFORCEMENT LEARNING: Q-LEARNING

HOCHSCHULE HEILBRONN

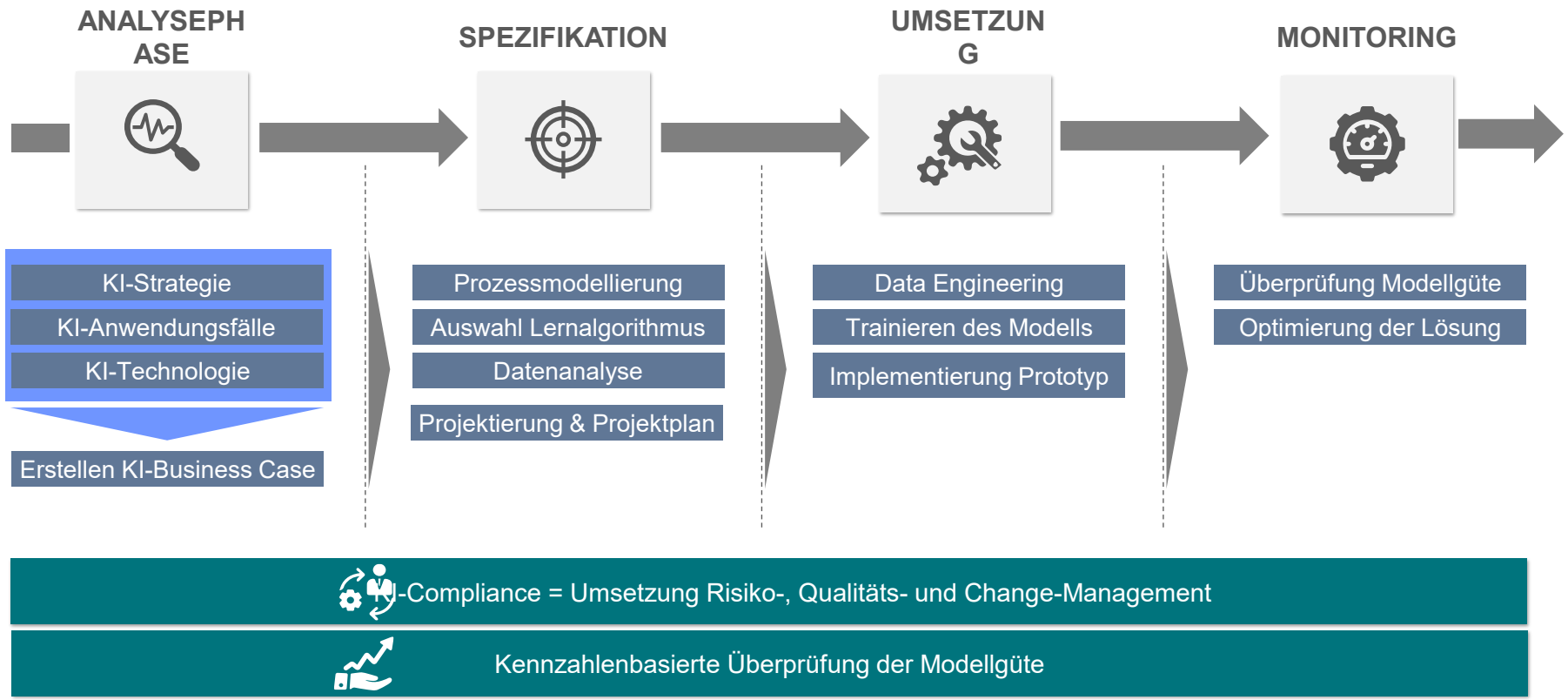
Berechnungsbeispiel (Vorlesung)

6. IMPLEMENTIERUNG VON KI-LÖSUNGEN

IMPLEMENTIERUNG VON KI-LÖSUNGEN

HOCHSCHULE HEILBRONN

Ablauf KI-Projekt:

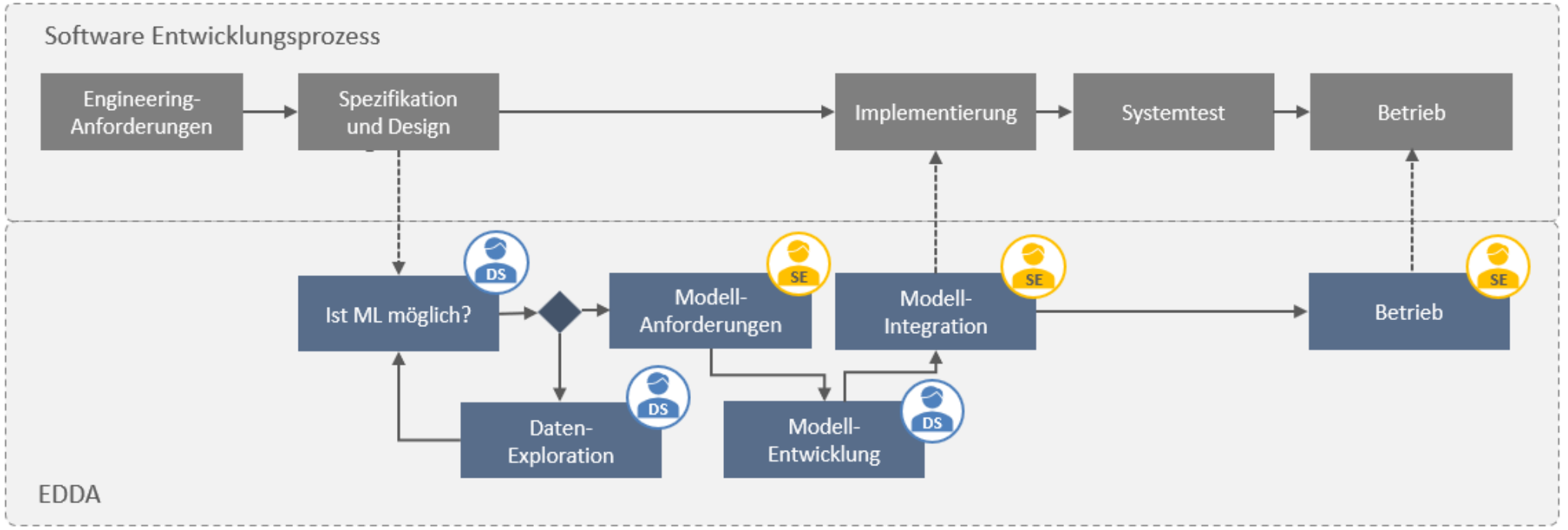


Quelle: DSC GmbH

IMPLEMENTIERUNG VON KI-LÖSUNGEN

HOCHSCHULE HEILBRONN

Der Weg zur KI-Anwendung:



Quelle : EDDA – Engineering Data-Driven Applications, Universität Duisburg-Essen

IMPLEMENTIERUNG VON KI-LÖSUNGEN

HOCHSCHULE HEILBRONN

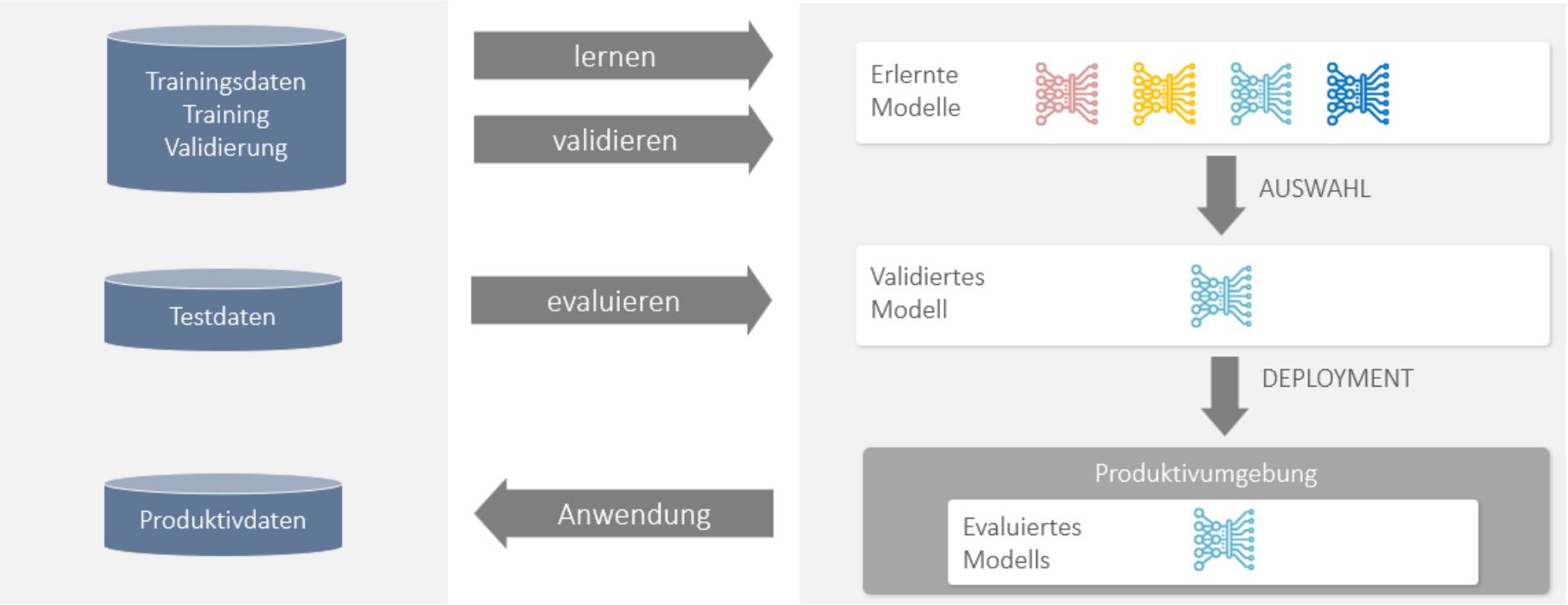
Aufgabenverteilung KI-Projekt:

Quelle: DSC GmbH

IMPLEMENTIERUNG VON KI-LÖSUNGEN

HOCHSCHULE HEILBRONN

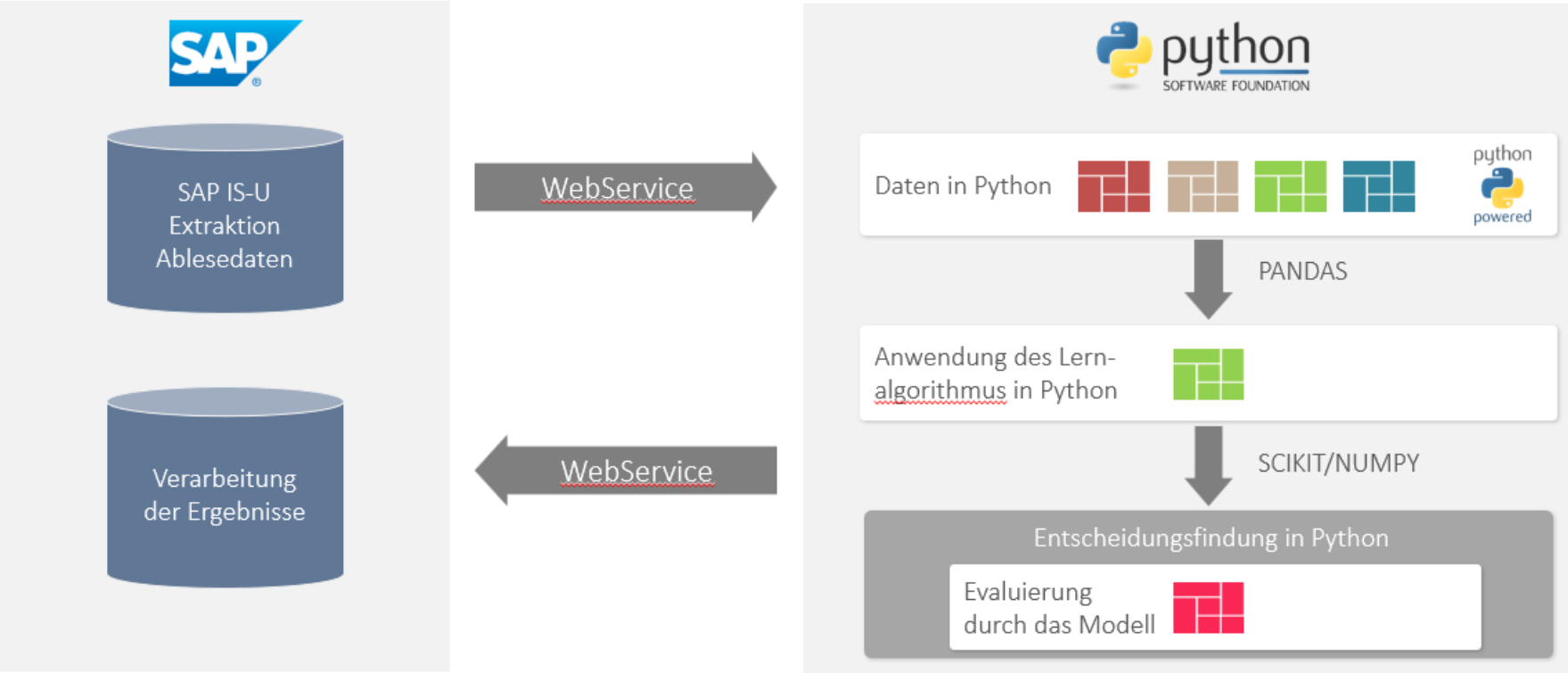
Methodisches Vorgehen:



Quelle: : Introduction to Data Science (Ch04), Prof. Drabant, Hochschule Mannnheim

IMPLEMENTIERUNG VON KI-LÖSUNGEN

Beispiel:



Quelle: DSC GmbH

THE END!

Please refer any questions to:
Prof. Dr. Florian Kauffeldt
Faculty of International Business
florian.kauffeldt@hs-heilbronn.de