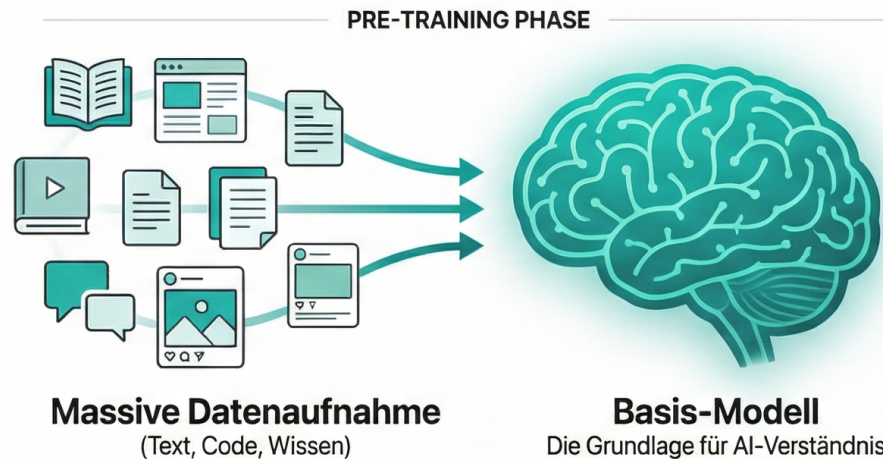
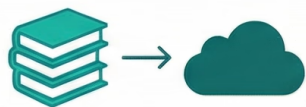


Was ist "Pre-Training"?



Das "Pre-Training" ist die initiale, unüberwachte Lernphase, in der ein großes KI-Modell auf einem riesigen, vielfältigen Datensatz trainiert wird, um allgemeine Sprachmuster, Grammatik, Faktenwissen und Kontext zu verstehen, bevor es für spezifische Aufgaben spezialisiert wird.



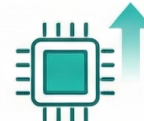
1. Umfangreiche Datenbasis

Nutzt enorme Mengen an Text aus dem Internet, Büchern und mehr, um ein breites Weltwissen aufzubauen.



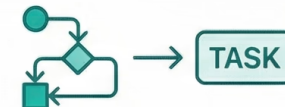
2. Mustererkennung & Struktur

Lernt die statistischen Zusammenhänge der Sprache, Grammatik und logische Strukturen ohne spezifische Anweisungen.



3. Basis für Spezialisierung

Schafft ein fundamentales "Fundament-Modell", das später durch "Fine-Tuning" für bestimmte Aufgaben angepasst wird.

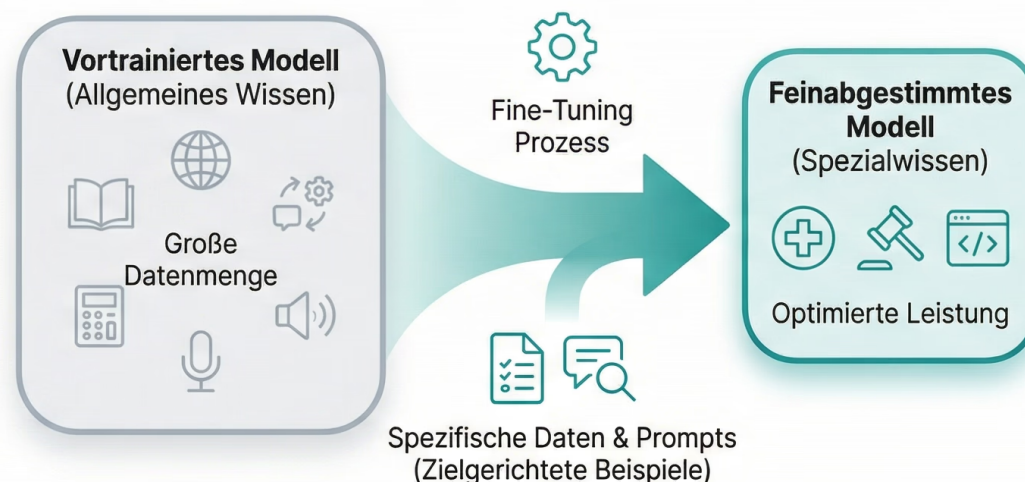


4. Vor dem Prompting

Findet statt, bevor ein Nutzer mit dem Modell durch Prompts interagiert; es bildet die Wissensgrundlage für jede Antwort.

Was ist "Fine-Tuning"?

Die Anpassung eines vortrainierten KI-Modells (Foundation Model) für spezifische Aufgaben oder Fachgebiete durch gezieltes Nachtraining mit einem kuratierten Datensatz, um Leistung und Genauigkeit in bestimmten Kontexten zu verbessern.



Zielgerichtete Optimierung.

Fokussiert das Modell auf präzise Antworten und Aufgaben, minimiert Halluzinationen.



Effizienteres Lernen.

Benötigt weniger Daten und Rechenleistung als ein komplettes Neutrainning, nutzt vorhandene Strukturen.

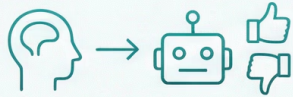


Verbesserte Interaktion.

Ermöglicht konsistentere, kontextbewusste und fachspezifischere Dialoge durch maßgeschneiderte Prompts.

Was ist der Unterschied zwischen PPO und DPO?

Ein Vergleich von Reinforcement Learning Methoden für die Interaktion mit KI-Modellen im Kontext von Prompting und Human Feedback.

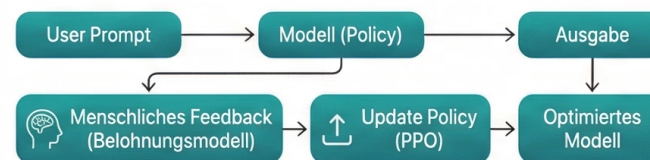


Grundprinzip (RLHF Foundation)

Beide Methoden nutzen Reinforcement Learning from Human Feedback (RLHF), um KI-Modelle basierend auf menschlichen Präferenzen zu verfeinern.

PPO: Proximal Policy Optimization

Kernkonzept:



Iterativer Prozess. Nutzt ein separates Belohnungsmodell, um die Policy schrittweise zu aktualisieren. Stabilität durch Begrenzung der Updates.

Schlüsselpunkte:

- Komplexität & Ressourcen:** Erfordert Training und Wartung eines separaten Belohnungsmodells. Rechenintensiver.
- Stabilität:** PPO begrenzt Policy-Änderungen, um instabiles Lernen zu verhindern, aber das Tuning ist anspruchsvoll.
- Prozess:** Mehrstufig und iterativ, oft mit mehreren Durchläufen für Konvergenz.

DPO: Direct Preference Optimization

Kernkonzept:



Vereinfachter Ansatz. Optimiert die Policy direkt auf Basis von Präferenzdaten, ohne ein explizites Belohnungsmodell zu benötigen.

Schlüsselpunkte:

- Effizienz & Einfachheit:** Kein separates Belohnungsmodell notwendig. Weniger Hyperparameter, schnelleres Training.
- Stabilität:** Mathematisch fundierter und oft stabiler im Training, da es das Optimierungsproblem direkt löst.
- Prozess:** Direkter und stromlinienförmiger, nutzt Präferenzdaten effizienter.

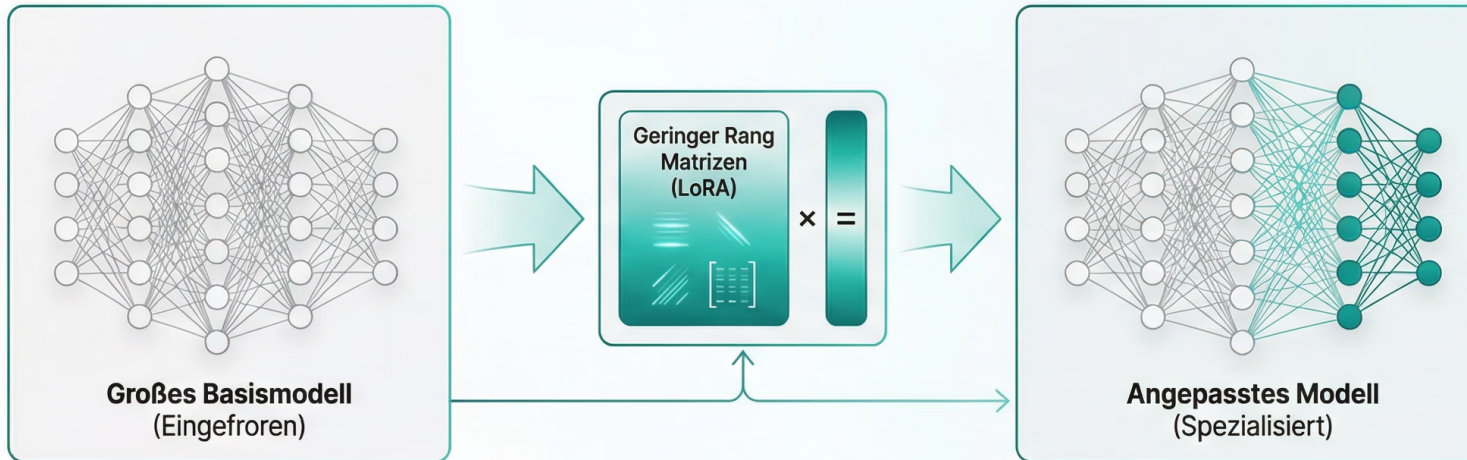


Fazit: DPO bietet oft eine effizientere und stabilere Alternative zu PPO, indem es die Komplexität des RLHF-Prozesses durch direkte Nutzung von Präferenzdaten reduziert, was besonders für die Feinabstimmung von Sprachmodellen vorteilhaft ist.

Was ist LoRA (Low-Rank Adaptation)?

Kapitel 3.6

Effiziente Feinabstimmung für KI-Modelle und Prompting



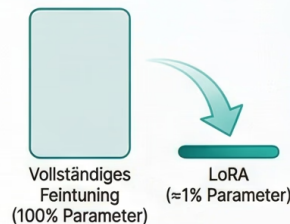
Angepasstes Modell (Spezialisiert)

LoRA ist eine Parameter-effiziente Methode zur Feinabstimmung von großen KI-Modellen. Statt das gesamte Modell neu zu trainieren, werden nur kleine, niedrigrangige Matrizen hinzugefügt, die sich an spezifische Aufgaben oder Daten anpassen. Das ursprüngliche Modell bleibt dabei unverändert.



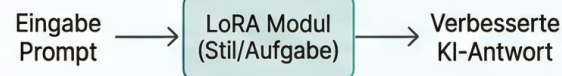
1. Effizienzsteigerung

Drastische Reduzierung der benötigten Trainingsparameter (bis zu 99%). Spart Rechenressourcen und Zeit.



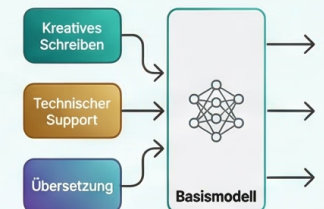
2. Zielgerichtete Anpassung

Ermöglicht schnelle und präzise Adaptation an neue Aufgaben, Stile oder Domänen, ohne das Basiswissen zu verlieren.



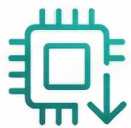
3. Modulare Architektur

Verschiedene LoRA-Modelle können kombiniert und flexibel ausgetauscht werden, um die KI-Fähigkeiten zu erweitern.



Was ist QLoRA?

Quantisierte Low-Rank-Adaption: Eine effiziente Methode zur Feinabstimmung großer Sprachmodelle (LLMs) durch drastische Reduzierung des Speicherbedarfs bei gleichbleibender Leistung, ideal für Prompting und KI-Interaktion auf begrenzter Hardware.



Reduzierter Speicherbedarf

Nutzt 4-bit Quantisierung, um den VRAM-Verbrauch während des Trainings signifikant zu senken, was den Einsatz auf Consumer-GPUs ermöglicht.



Schnelleres Training & Inferenz

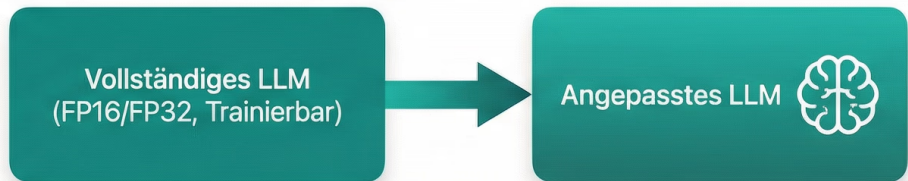
Fokus auf kleine, trainierbare Adapter-Schichten statt des gesamten Modells, was den Rechenaufwand und die Zeit minimiert.



Leistungserhalt

Erzielt nahezu die gleiche Qualität wie eine vollständige Feinabstimmung, indem die Kern-Gewichte des Basismodells eingefroren und nur Adapter angepasst werden.

Traditionelles Fine-Tuning (Vollständig)



Hoher VRAM & Rechenaufwand, erfordert teure Hardware.

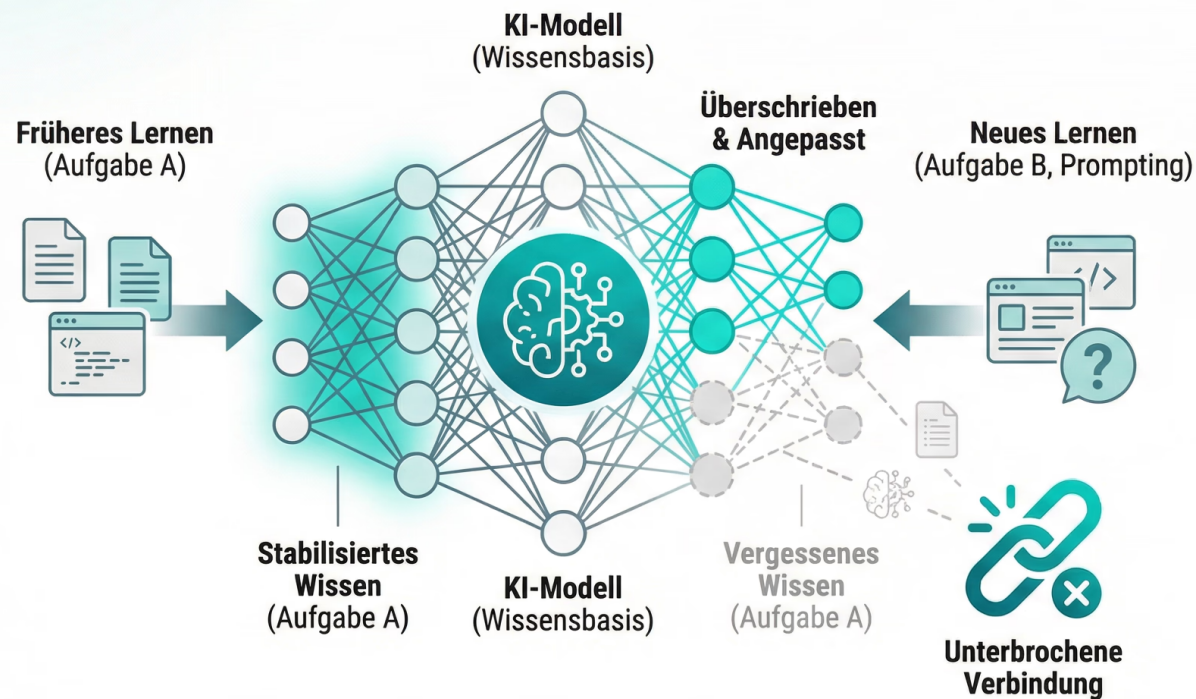
QLoRA Fine-Tuning (Effizient)



Effizient & Zielgerichtet, ermöglicht Anpassung mit minimalen Ressourcen.

Was ist "Catastrophic Forgetting"?

Das Phänomen, bei dem ein KI-Modell beim Erlernen neuer Aufgaben **drastisch das zuvor erworbene Wissen vergisst**, insbesondere wenn es sequenziell und ohne kontinuierliche Wiederholung trainiert wird.



Sequenzielles Lernen (Ursache)

Wenn ein Modell in getrennten Phasen trainiert wird, ohne dass alte Daten erneut präsentiert werden, überschreiben neue Informationen die alten Gewichte.



Überschreiben von Gewichten (Mechanismus)

Die internen Parameter (Gewichte) des Modells werden massiv angepasst, um die neue Aufgabe zu optimieren, wodurch die für die alte Aufgabe kritischen Konfigurationen zerstört werden.



Auswirkung auf die KI-Interaktion (Folge)

In der Praxis führt dies dazu, dass die KI auf ältere Prompts oder bekannte Szenarien plötzlich falsch oder inkonsistent reagiert, nachdem sie für neue Anwendungsfälle optimiert wurde.

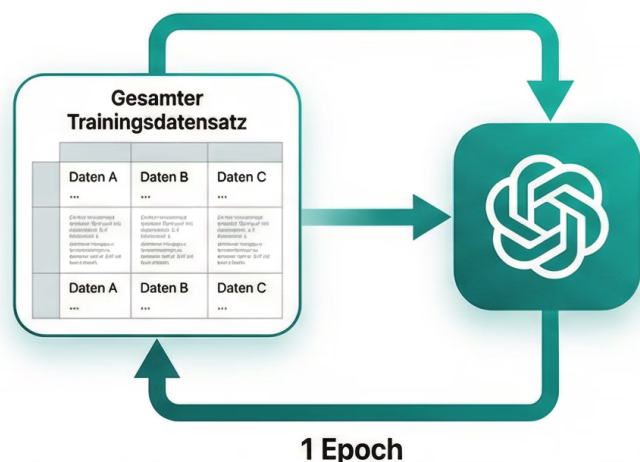
Vorher: Prompt A → ✓ Korrekte Antwort

Nachher: Prompt A → ✗ Falsche/Keine Antwort ?

Was sind "Epochs" beim Training?

Im Kontext von Prompting und AI Interaction.

Ein vollständiger Durchlauf des gesamten Trainingsdatensatzes durch das AI-Modell.

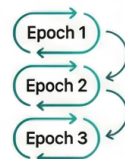


Eine Epoche ist ein vollständiger Zyklus, in dem das Modell jedes Beispiel im Trainingsdatensatz genau einmal 'gesehen' und zur Anpassung seiner internen Parameter verwendet hat.



Mehrere Durchläufe (Iterationen)

Training erfordert typischerweise viele Epochen. Ein Modell durchläuft die Daten mehrfach, um komplexe Muster zu lernen.



Gewichtsanpassung und Lernen

Nach jeder Epoche (oder kleineren Batches) passt das Modell seine internen Gewichte an, um den Fehler zu minimieren.

```

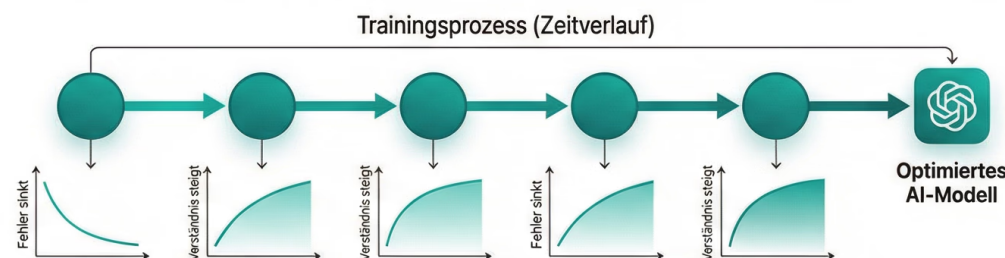
graph TD
    A[Epoche beendet] --> B[Fehler berechnen]
    B --> C[Gewichte anpassen]
    C --> D[Nächste Epoche starten]

```



Balance zwischen Under- und Overfitting

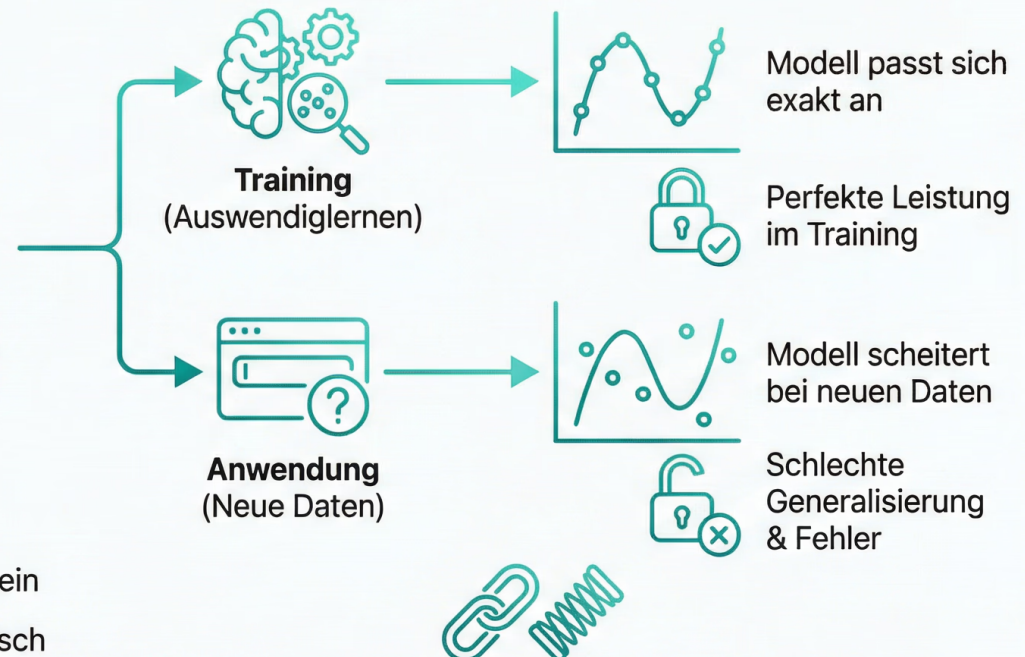
Zu wenige Epochen führen zu "Underfitting" (ungenügendes Lernen), zu viele zu "Overfitting" (Auswendiglernen der Daten). Die richtige Anzahl ist entscheidend.



Das Ziel: Die optimale Anzahl an Epochen finden, um die Leistung des Modells für die AI Interaction zu maximieren, ohne es zu überanpassen.

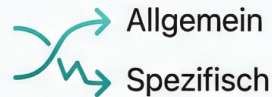
Was ist "Overfitting"?

Overfitting (Überanpassung) beschreibt ein Phänomen in der KI, bei dem ein Modell die Trainingsdaten zu spezifisch und auswendig lernt, anstatt das zugrundeliegende allgemeine Muster zu erfassen. Dies führt zu exzellenten Ergebnissen im Training, aber zu einer schlechten Leistung bei neuen, unbekannten Eingaben (Generalisierung).



Zu spezifisch für Trainingsdaten

Das Modell fixiert sich auf Rauschen und spezifische Details der Trainingsdaten, anstatt das generelle Prinzip zu lernen.



Schlechte Generalisierung

Fähigkeit, das Gelernte auf neue, unbekannte Situationen (Prompts) anzuwenden, ist stark eingeschränkt.



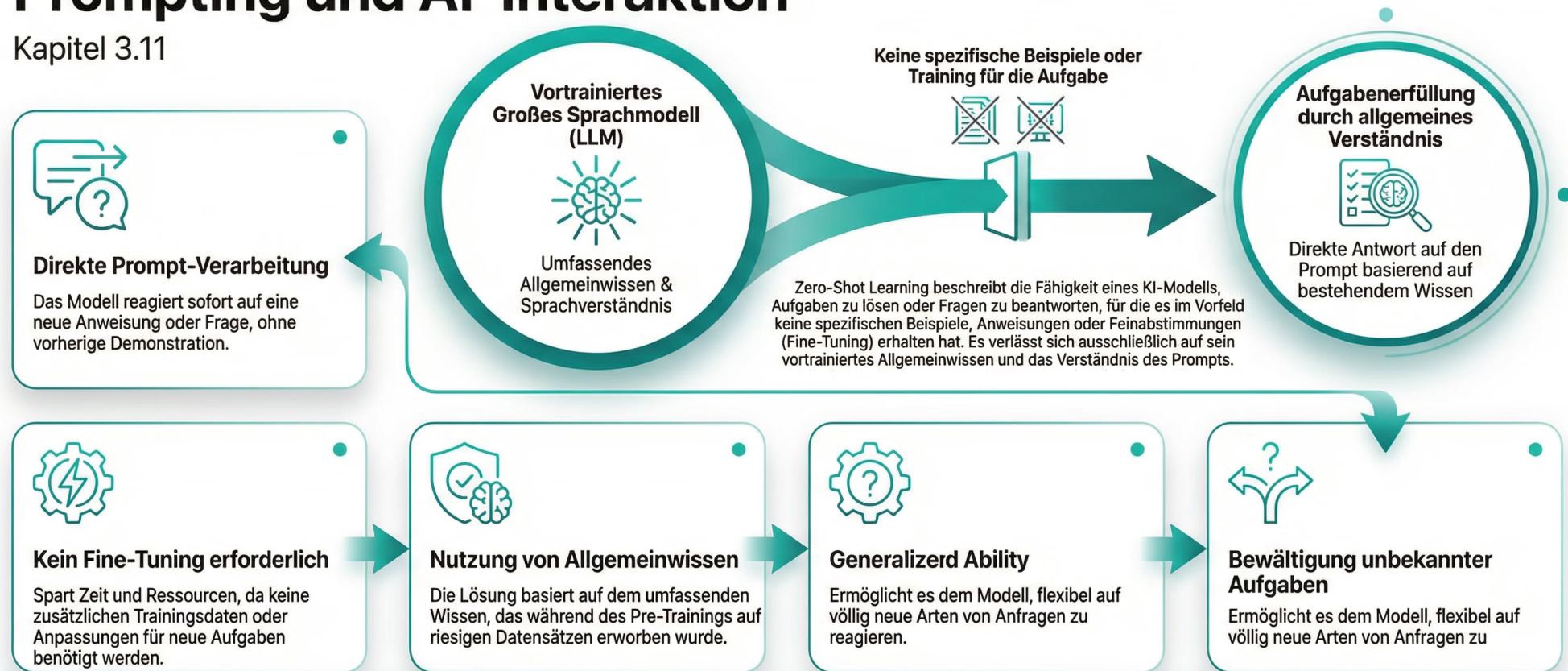
Rigide & Unflexibel

Die KI reagiert starr auf Eingaben, die auch nur minimal von den Trainingsbeispielen abweichen.

➤ Kapitel 3.10 | KI-Prompting Grundlagen

Was ist "Zero-Shot" Learning? – Im Kontext von Prompting und AI-Interaktion

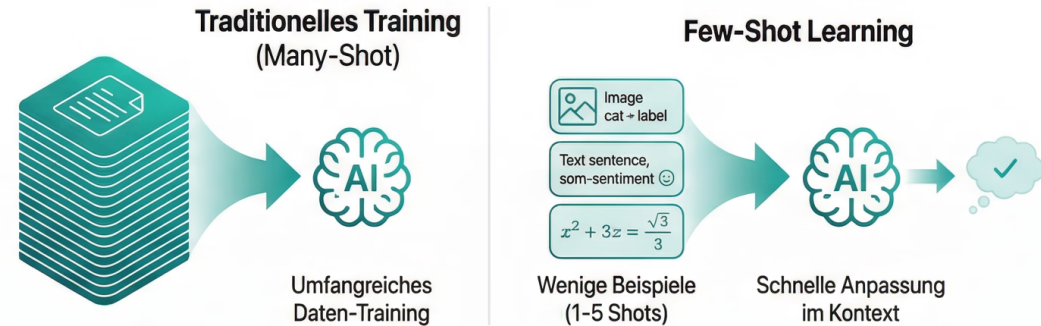
Kapitel 3.11



Was ist "Few-Shot" Learning?

Im Kontext von Prompting und KI-Interaktion.

Few-Shot Learning: Ein maschinelles Lernparadigma, bei dem ein KI-Modell befähigt wird, neue Aufgaben oder Konzepte anhand einer sehr kleinen Anzahl von Beispielen (den "Shots") zu verstehen und zu generalisieren, ohne dass ein umfangreiches Nachtraining erforderlich ist. Dies ermöglicht eine schnelle Anpassung und Interaktion durch spezifische Prompts.



Schlüsselkonzepte & Vorteile



Effizienz & Geschwindigkeit

Ermöglicht schnelle Anpassung an neue Aufgaben ohne ressourcenintensives Neutaining.



Anpassungsfähigkeit durch Prompts

Nutzer können das Modellverhalten durch gezielte Beispiele im Prompt direkt steuern und verfeinern.



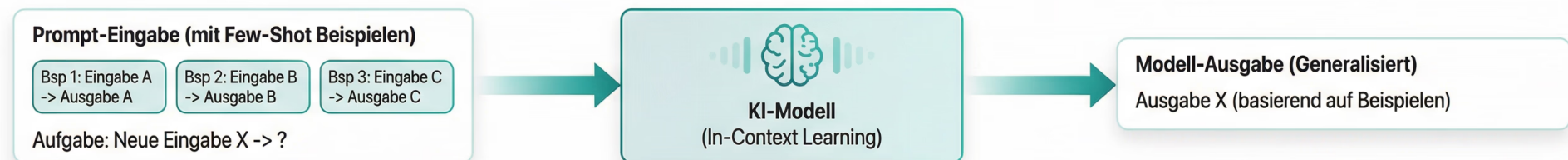
Reduzierter Datenbedarf

Funktioniert effektiv auch in Bereichen mit begrenzten oder teuren Trainingsdaten.



Generalisierung im Kontext

Das Modell nutzt sein Vorwissen, um aus den wenigen Beispielen im aktuellen Kontext zu lernen und auf ähnliche, unbekannte Eingaben zu schließen.



Was ist "Chain-of-Thought" (CoT)?

Ein Ansatz zur Förderung komplexer Denkprozesse bei der KI-Interaktion durch explizite Schlussfolgerungsschritte.

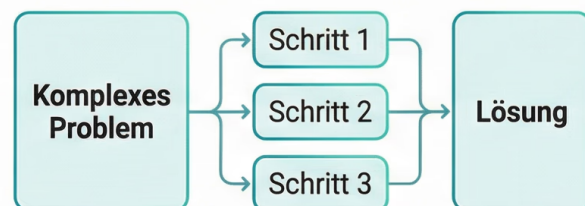


CoT (Chain-of-Thought) ist eine Prompting-Technik, die KI-Modelle dazu anregt, ihre Zwischenschritte beim Problemlösen explizit darzulegen, anstatt direkt eine Antwort zu geben. Dies ahmt menschliches Denken nach und führt zu fundierteren Ergebnissen.



1. Zerlegung in Schritte

Komplexe Aufgaben werden in kleinere, logische Zwischenschritte unterteilt, um den Lösungsweg transparent zu machen.



2. Explizites Denken

Die KI wird aufgefordert, ihren Gedankengang zu artikulieren, was zu einer besseren Nachvollziehbarkeit der Antwort führt.

Prompt:

Wie viele Äpfel habe ich, wenn ich 3 habe, 2 erhalte und 1 esse?

CoT-Antwort:

Denken: Ich starte mit 3. Ich bekomme 2 hinzu, also $3 + 2 = 5$. Ich esse 1, also $5 - 1 = 4$.
Antwort: Du hast 4 Äpfel.



3. Verbesserte Genauigkeit

Durch die strukturierte Herleitung werden Fehlerwahrscheinlichkeiten reduziert, insbesondere bei mathematischen oder logischen Problemen.

Standard KI:

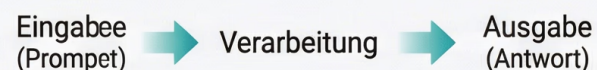


Mögliche Fehlerquelle

CoT KI:



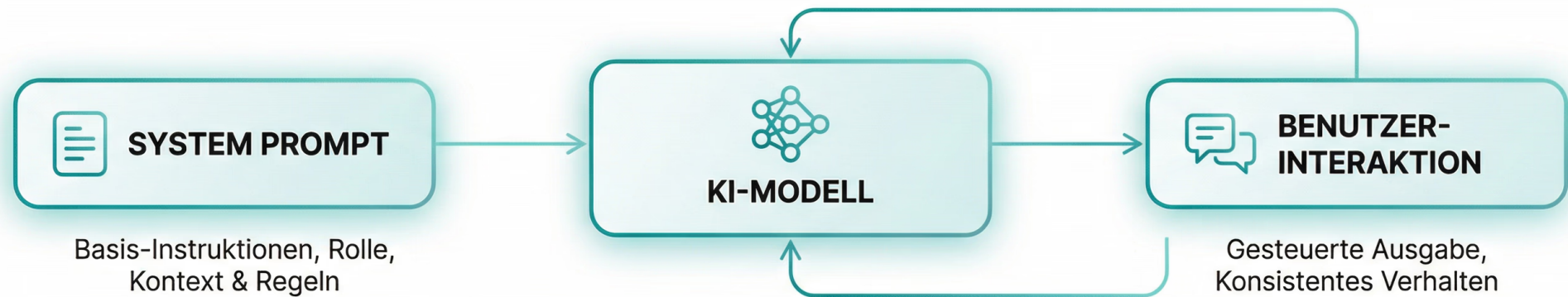
Standard Prompting



Chain-of-Thought (CoT) Prompting



Was ist "System Prompt Engineering"?



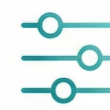
ROLLEN-DEFINITION

Legt die Persönlichkeit, Expertise und den Tonfall der KI fest.



KONTEXT & EINSCHRÄNKUNGEN

Setzt klare Grenzen für Themen, Verhalten und Sicherheitsrichtlinien.



AUSGABE-STEUERUNG

Definiert Format, Struktur und Länge der Antworten.

Was ist "Synthetic Data"?

Künstlich generierte Daten, die statistische Eigenschaften und Muster realer Daten nachahmen, aber keine echten, sensiblen Informationen enthalten. Sie werden häufig verwendet, um KI-Modelle effizient und datenschutzkonform zu trainieren.

