Was KI gut kann — und wo sie danebenliegt

Kapitel online

Stark bei Mustern aus vielen Beispielen, schwach bei allem Neuen — und warum eine Antwort überzeugend klingen und trotzdem falsch sein kann.

Start · Kapitel 3 von 50 · ca. 8 Min. · 50 XP

Einfach gesagt: KI ist stark, wo es viele Beispiele gibt, und schwach bei allem Neuen — und sie klingt auch dann sicher, wenn sie rät.

Du weißt jetzt, wie eine KI lernt: aus Beispielen, mit Raten und Nachbessern. Aus dieser einen Tatsache folgt alles Weitere — vor allem, wo diese Technik großartig ist und wo sie regelmäßig auf die Nase fällt.

Es ist das wichtigste der drei Start-Kapitel — weil es dich vor dem häufigsten Alltagsfehler bewahrt: einer Antwort zu glauben, nur weil sie gut klingt.

Was du hier lernst

- Wo KI wirklich stark ist — und woran man solche Aufgaben erkennt
- Warum sie bei allem **Neuen** schwächelt, für das es keine Beispiele gibt
- Warum eine Antwort **überzeugend klingen** und trotzdem falsch sein kann
- Eine einzige Alltagsregel, die dich davor schützt

Für den Unterricht

Einstieg (2 Min): Kapitel gemeinsam lesen, Alltagsbeispiel aufgreifen.

Erkunden am Gerät (5 Min): die interaktiven Module auf herrki.de/003_was_ki_kann_und_was_nicht ausprobieren.

Sichern mit Quiz (2 Min): „Kurzer Check" besprechen, „Das nimmst du mit" festhalten.

Wo KI stark ist: viele Beispiele, wiederkehrendes Muster

Eine KI ist gut in genau einer Sorte Aufgabe: **etwas wiedererkennen, das sie in ähnlicher Form schon tausendfach gesehen hat.** Je mehr Beispiele es gibt, desto besser wird sie.

Handschrift lesen Millionen Menschen schreiben eine Sieben — alle anders, alle ähnlich. Perfekte Aufgabe.
Gesprochenes in Text Sprache ist überall aufgezeichnet. Aus dem Muster „so klingt dieses Wort“ wird ein sehr zuverlässiges Diktiergerät.
Bilder sortieren Hund, Katze, Strand, Rechnung. Endlos viele Beispiele, und die Aufgabe bleibt dieselbe.
Formulieren und übersetzen Wie Sätze gebaut werden, steht in unfassbar vielen Texten. Deshalb klingen KI-Texte sprachlich so rund.

Wo sie schwach ist: alles Neue

Dreh die Regel einfach um. Wo es **keine Beispiele** gibt, hat eine KI nichts, woraus sie ihre Regel ziehen könnte:

1**Etwas, das noch nie passiert ist.** Eine ganz neue Situation, ein Ereignis von gestern Abend, ein Gerät, das es letztes Jahr noch nicht gab.

2**Seltene Sonderfälle.** Was nur alle paar Jahre vorkommt, taucht in den Beispielen kaum auf — genau dort liegt sie am häufigsten daneben.

3**Dinge, die nur du weißt.** Deine Vertragsnummer, der Beschluss von gestern, der Name deiner Nachbarin — stand in keinem Beispiel, kann sie nicht wissen.

4**Einseitige Beispiele.** Sind die Beispiele schief, wird die Regel schief — die KI übernimmt sie ungeprüft.

Warum sie so überzeugend klingt, wenn sie danebenliegt

Jetzt der Teil, der die meisten überrascht. Ein Sprachmodell — das, was hinter Chat-KIs steckt — macht im Kern etwas Schlichtes: Es setzt **Wort für Wort das dazu, das am besten passt**. Nicht das wahre Wort. Das *passende*.

> **Stell dir jemanden vor, der jedes Gespräch der Welt mitgehört hat und immer weiterreden muss.** Er weiß, wie eine Antwort klingt. Er weiß, welche Wörter in solchen Sätzen vorkommen. Ob es stimmt, ist eine ganz andere Frage — er hat nie gelernt, „ich weiß es nicht“ zu sagen.

Deshalb sieht ein erfundener Satz aus wie ein richtiger: gleiche Sicherheit, gleiche Satzmelodie. Bei einem Menschen hörst du Unsicherheit heraus — hier gibt es nichts, was sich verrät.

Und noch etwas: Bei Zahlen, Namen, Jahreszahlen und Quellenangaben liegt so ein Modell besonders gern daneben. Sie klingen alle gleich plausibel, egal ob sie stimmen.

Die eine Regel für den Alltag

Prüf nach, was dir wichtig ist.

- **Formulieren, zusammenfassen, Ideen sammeln:** lass sie machen. Fehler siehst du sofort.
- **Namen, Zahlen, Daten, Paragrafen, Medizin, Geld:** immer gegenprüfen. Bei einer echten Quelle.
- **Frag nach der Quelle** — und schau sie dir an. Auch eine Quellenangabe kann erfunden sein.
- **Sei misstrauisch, wenn es zu glatt klingt.** Sicherheit im Ton sagt nichts über Richtigkeit.

Das ist keine Warnung vor der Technik. Es ist derselbe Umgang, den du mit jedem Werkzeug hast: Eine Bohrmaschine ist großartig — du hältst trotzdem die Finger weg vom Bohrer.

Kurzer Check – mit Lösung

Lösungen für Lehrkräfte – für Lernende die Online-Version nutzen.

Frage 1: Für welche Aufgabe ist eine KI am besten geeignet?

- Sagen, was morgen in deiner Straße passieren wird.
- ✔ **Aus tausenden Beispielen ein wiederkehrendes Muster erkennen.**
- Entscheiden, was moralisch richtig ist.

Genau. Alles, wovon es viele ähnliche Beispiele gibt, liegt ihr — alles Einmalige und Neue nicht.

Frage 2: Warum klingt auch eine falsche Antwort überzeugend?

- Weil das Modell absichtlich täuschen will.
- ✔ **Weil es Wort für Wort das passende Wort wählt — passend heißt nicht wahr.**
- Weil es die falsche Antwort irgendwo so gelesen hat.

Richtig. Es gibt keine Absicht und kein Nachschlagen — nur „was passt hier als Nächstes“. Deshalb sieht Geraten aus wie Gewusstes.

Frage 3: Was solltest du auf jeden Fall gegenprüfen?

- Vorschläge für ein Geburtstagsgeschenk.
- ✔ **Zahlen, Namen, Jahreszahlen und Quellenangaben.**
- Die Rechtschreibung im Text.

Genau an diesen Stellen kostet ein Fehler etwas — und genau dort liegt so ein Modell besonders gern daneben.

Zurück zur Anwendung

Damit hast du das Wichtigste beisammen: KI erkennt Muster aus Beispielen. Das erklärt Stärken, Schwächen und den zu glatten Tonfall auf einmal — mehr brauchst du für den Alltag nicht.

Wenn du wissen willst, wie das *Lernen* selbst abläuft — mit einem Spiel, bei dem du das Modell bist —, geht es weiter in „Was heißt eigentlich Lernen?“. Und warum selbstsicher erfundene Antworten entstehen und woran man sie erkennt, nimmt sich das Kapitel „Halluzinationen“ später ausführlich vor — dann weißt du auch, was in der Maschine dabei vorgeht.

Das nimmst du mit

- **Stark bei Mustern.** Viele ähnliche Beispiele — das ist das Heimspiel jeder KI.
- **Schwach bei Neuem.** Wofür es keine Beispiele gibt, kann sie keine Regel gefunden haben.
- **Passend ist nicht wahr.** Ein Sprachmodell wählt das passende nächste Wort — nicht das richtige.
- **Sicherer Ton heißt gar nichts.** Es fehlt das Zögern, an dem du bei Menschen Unsicherheit merkst.
- **Prüf nach, was dir wichtig ist.** Zahlen, Namen, Daten, Quellen — immer.