Einstieg: [Künstliche Intelligenz einfach erklärt (explainity® Erklärvideo) (2023) - YouTube](https://www.youtube.com/watch?v=fT69SLK2m9I)

1. **Wer erkennt es besser?**



**Fragen:**

* Woran erkennst du, welches der echte Hund ist?

→ Am Fell, den natürlichen Bewegungen, dem Gesichtsausdruck, den Augen.

* Wie könnte eine KI den Unterschied lernen?

→ Durch viele Bildbeispiele mit echten und Spielzeughunden, aus denen sie typische Merkmale ableitet.

**2. Klassische Programmierung (A) vs. Maschinelles Lernen (B)**

**Frage:** Was ist bei B anders als bei A?

→ Bei A wird alles manuell programmiert. Bei B erkennt die KI selbst Muster in Beispielen und „lernt“ daraus.

**3. Übung: Kann die KI das lernen?**

|  |  |  |
| --- | --- | --- |
| **Beispiel** | **Lernbar für KI?** | **Begründung** |
| Katzen- vs. Hundebilder unterscheiden | ✅ Ja | KI kann Unterschiede in Form, Ohren, Schnauze usw. aus Bildern lernen. |
| Was ist ein fairer Preis für ein Eis? | ❌ Nein | „Fairness“ ist subjektiv, hängt von Meinung, Ort, Gefühl ab. |
| Handschrift erkennen | ✅ Ja | KI kann mit vielen Schriftbeispielen lernen, Buchstaben zu erkennen. |
| Welcher Witz ist lustiger? | ❌ Nein | Humor ist schwer messbar und hängt stark vom Menschen ab. |

**4. Ausblick: Was ist mit schlechten Daten?**

**Analysiere folgende ChatGPT-Aufgabe:**Erstelle mir ein Bild einer Armbanduhr und stelle die Uhrzeit auf halb 10.

* ChatGPT kann leicht eine Armbanduhr zeigen,
* hat aber große Probleme richtige Uhrzeiten anzuzeigen, da die meisten Uhren-Bilder von Uhrenherstellern hochgeladen werden. Diese bevorzugen bestimmte Uhrzeiten, da diese „attraktiver“ wirken.



**Frage:** Was könnte passieren, wenn die KI nicht richtig lernt, weil die Daten falsch sind?

→ Die KI erkennt möglicherweise nur rote Äpfel als Äpfel und hält grüne Äpfel für etwas anderes oder falsch. Sie hat ein „Vorurteil“, weil sie nicht genug verschiedene Daten gesehen hat.

**1. Fehlentscheidungen**Beispiel: In der Medizin könnte eine KI falsche Diagnosen stellen, wenn sie mit unvollständigen oder verzerrten Patientendaten trainiert wurde.  
Folge: Menschen könnten falsch behandelt oder nicht behandelt werden.

**2. Diskriminierung und Vorurteile (Bias)**Beispiel: Wenn eine KI zur Personalauswahl vor allem mit Daten von Männern trainiert wurde, bevorzugt sie Männer gegenüber Frauen.  
Folge: Ungerechte Entscheidungen, soziale Ungleichheit wird verstärkt.

**3. Vertrauensverlust**Wenn Nutzer merken, dass die KI häufig Fehler macht oder unfair handelt, verlieren sie das Vertrauen – auch in korrekte KI-Systeme.

**4. Verbreitung von Fehlinformationen**Beispiel: Eine KI, die Texte generiert (wie Chatbots), kann falsche Fakten verbreiten, wenn sie mit schlechten Quellen trainiert wurde.  
Folge: Fake News, Desinformation, politische Manipulation.

**5. Wirtschaftliche Schäden**Falsche Prognosen in Bereichen wie Börse, Logistik oder Produktion können hohe finanzielle Verluste verursachen.

**6. Gefährdung der Sicherheit**In sicherheitskritischen Bereichen wie autonomem Fahren oder Luftfahrt kann eine KI auf Basis falscher Daten Menschenleben gefährden.

**7. Verstärkung bestehender Probleme**Fehlerhafte Daten können bestehende gesellschaftliche Probleme nicht nur abbilden, sondern sogar verstärken (z. B. Rassismus, Ungleichheit, Armut).

* **Fazit:** KI kann erstaunlich viel – aber sie hängt stark davon ab, was man ihr zeigt!