

Unterschiede und Funktionen einfach erklärt

NZZ [nzz.ch/technologie/kuenstliche-intelligenz-ki-deep-learning-einfach-erklart-id.1632034](https://www.nzz.ch/technologie/kuenstliche-intelligenz-ki-deep-learning-einfach-erklart-id.1632034)

Ruth Fulterer (ful), Eveline Geiser (evg)

Erklärt

Was ist künstliche Intelligenz? Wie funktioniert Deep Learning? Ein Überblick

Was steckt hinter Chat-GPT, und wo in unserem Alltag wird KI noch angewendet? Was können die Maschinen besser als Menschen? Und müssen wir uns Sorgen machen?

Ruth Fulterer, Eveline Geiser 03.03.2023, 13.12 Uhr ⌚ 9 min



Künstliche Intelligenz hat selbstlernende Systeme hervorgebracht. Derweil muss der Mensch lernen, die künstliche Intelligenz richtig zu verwenden. Ein Kind interagiert mit einem Roboter im Hong Kong Science Museum, 2021.

Miguel Candela / Imago

Was ist künstliche Intelligenz, und wie funktioniert sie? →

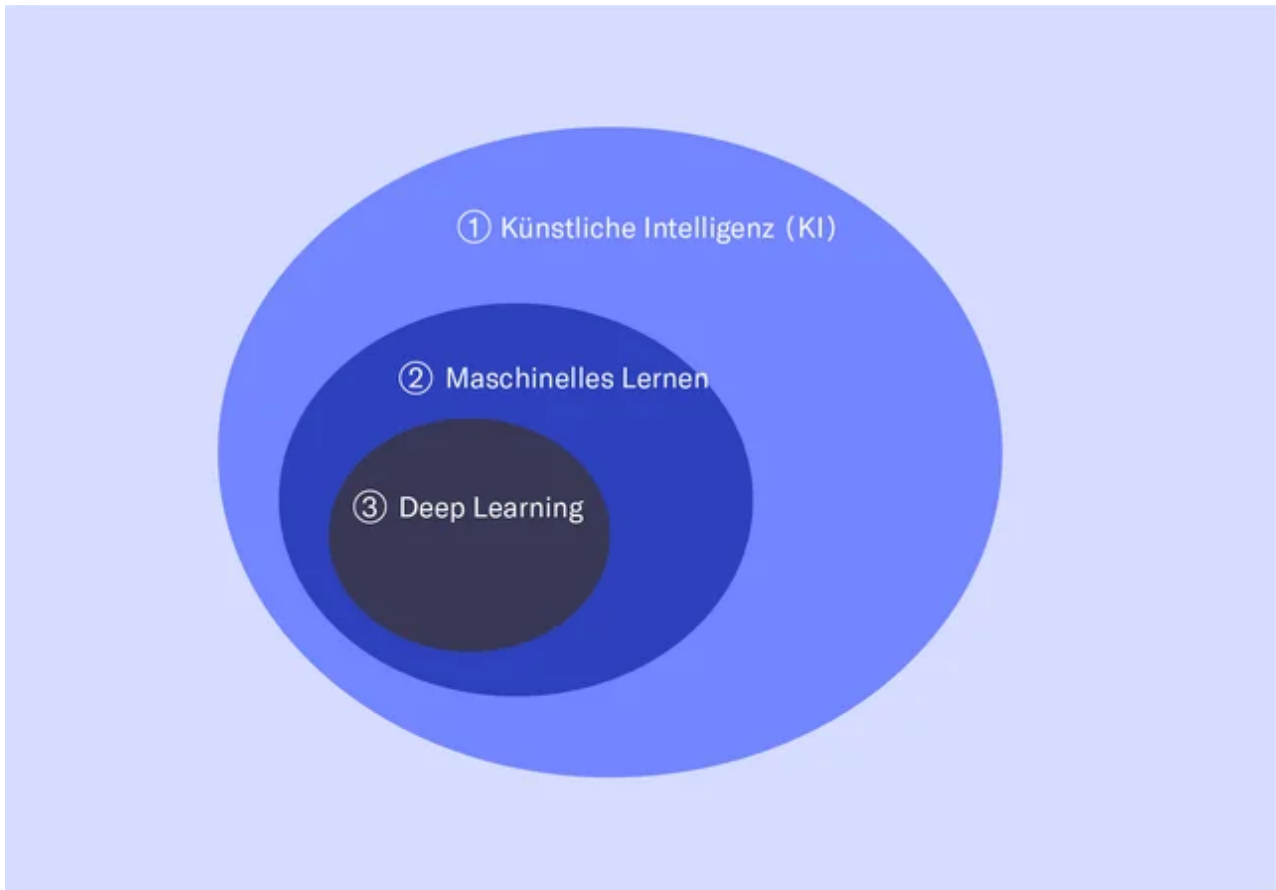
Das deutsche Wort künstliche Intelligenz (KI) übersetzt das englische *artificial intelligence* (AI). Der Begriff wird in mehreren Bedeutungszusammenhängen verwendet.

Erstens: Das **Forschungsfeld**, das Maschinenintelligenz herstellen will: Der Traum, Maschinen intelligent zu machen, ist Hunderte Jahre alt. Mit der Entwicklung von Computern setzte man neue Hoffnung in seine Erfüllung. Man programmierte Computern Expertenwissen ein, das diese mit immer grösserer Geschwindigkeit abrufen und anwenden konnten. So entstanden etwa frühe Formen eines Schachcomputers.

Zweitens wird der Begriff oft verwendet, wenn Computer Aufgaben übernehmen können, von denen man einst dachte, sie seien intelligenten Wesen vorbehalten: Übersetzen, Texte schreiben, Tumoren auf Röntgenbildern erkennen oder Auto fahren. Solche Programme werden im allgemeinen Sprachgebrauch als künstliche Intelligenz oder KI bezeichnet.

Eine engere Definition liefert die Expertengruppe der EU-Kommission. Künstliche Intelligenz sei ein System, das seine Umwelt analysiert und mit einer gewissen Autonomie agieren kann, um ein Ziel zu erreichen. Demnach spricht man immer dann von künstlicher Intelligenz, wenn nicht jede Aktion des Computerprogramms vorher definiert und einprogrammiert wurde.

Bereiche der künstlichen Intelligenz



- ① Computersysteme, die ihre Umwelt analysieren und mit einem gewissen Grad an Eigenständigkeit darauf reagieren.
- ② Teilbereich der KI, bei dem Algorithmen anhand Trainingsdaten lernen, Muster wiederzuerkennen.
- ③ Teilbereich des maschinellen Lernens, der auf sogenannte künstliche neuronale Netze setzt.

NZZ / cia.

Was ist der Unterschied zu maschinellem Lernen? →

Maschinelles Lernen ist eine Form der künstlichen Intelligenz: Programmierer geben in diesem Fall zwar die Strategien vor, nach denen Daten analysiert werden, doch sie legen nicht jedes Detail fest wie in herkömmlichen Programmen. Innerhalb eines vorgegebenen Rahmens verbessert der Algorithmus stattdessen selbständig sein Vorgehen bei der Datenanalyse.

Zum Beispiel beim Erkennen von Gesichtern in Bildern. Frühere Computerprogramme, die nicht selber lernen konnten, gingen nach einprogrammierten Regeln vor, um ein Gesicht zu identifizieren. Es wurde also «erklärt», was ein Gesicht ausmacht: etwa die Form, die Farbe, die Position von Augen oder Mund. Da Bilder aber so vielgestaltig sind, waren solche Programme wenig effektiv.

Besser funktionierte der Ansatz des maschinellen Lernens. Hier verarbeitet der Algorithmus Beispielbilder, um zu erlernen, was ein Gesicht ausmacht. Dabei gibt es mehrere Methoden: Beim sogenannten überwachten Lernen sind Trainingsbilder «beschriftet», der Algorithmus erfährt also von Anfang an, bei welchen Bildern es sich um die Kategorie «Gesicht» handelt. Durch den Vergleich mit Bildern von anderen Objekten lernt er dann, was Gesichter auszeichnet.

Es gibt aber auch Lernmethoden, die ganz auf diese kategorielle Information verzichten. Dann identifiziert der Algorithmus eigene Kategorien in den Trainingsbildern. So kann er zu Lösungen kommen, an die der Mensch nicht gedacht hätte.

Einem selbstlernenden System können verschiedene Arten von Algorithmen zugrunde liegen. Einer davon ist das sogenannte künstliche neuronale Netz.

Was sind künstliche neuronale Netze? →

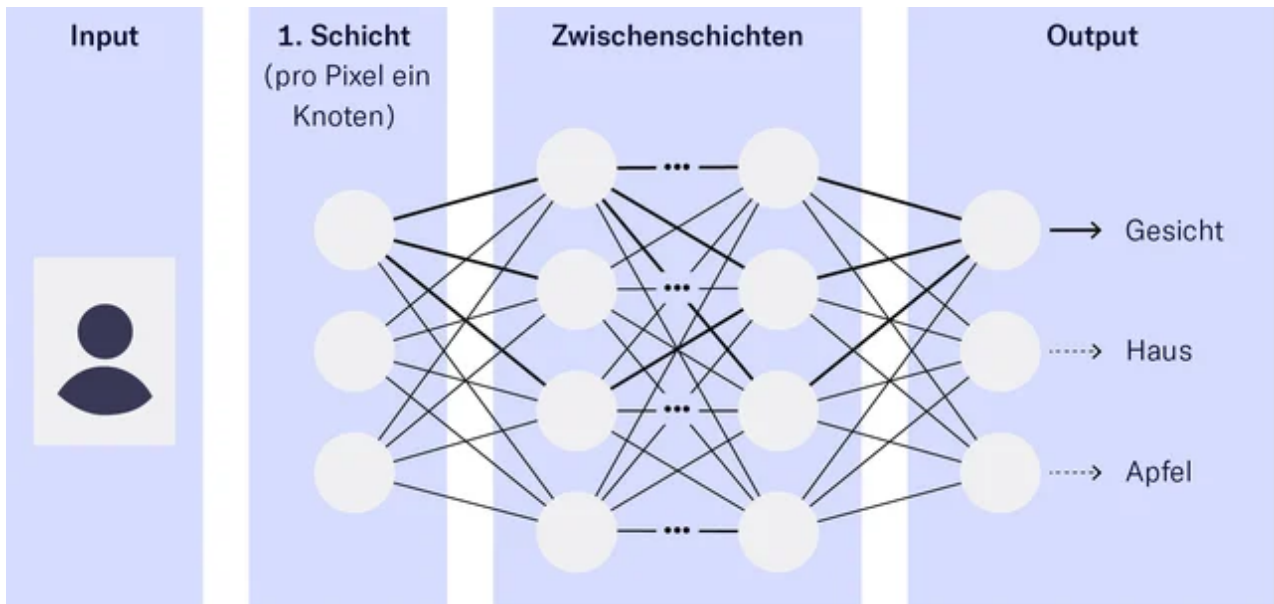
Künstliche neuronale Netze sind Systeme aus Algorithmen, die in ihrer Funktionsweise dem Gehirn nachempfunden sind. Es geht dabei also nicht um eine physische Nachbildung von natürlichen neuronalen Netzen, sondern um eine Nachbildung der Art und Weise, wie das Gehirn arbeitet und lernt.

Das menschliche Gehirn besteht aus einem Netzwerk von unterschiedlichen Nervenzellen, die über eine Vielzahl von Zellfortsätzen miteinander verbunden sind. Lernen in dieser Netzwerkstruktur bedeutet, dass bestimmte Verbindungen schneller, andere langsamer werden.

Ähnlich dem Gehirn können künstliche neuronale Netze Informationen in komplexen Strukturen gewichten und diese Gewichtung laufend neuen Anforderungen anpassen. Dies ist dank mathematischen Optimierungsfunktionen mit einer riesigen Zahl an Variablen möglich. Die Lernmethode dieser Netze nennt man Deep Learning.

Was ist Deep Learning? →

Deep Learning ist eine Methode des maschinellen Lernens, die künstliche neuronale Netze nutzt. Man kann sich diese zum besseren Verständnis als Netz aus Knoten und Verbindungen vorstellen. Die Knoten eines solchen Netzes sind in Schichten organisiert, wie die folgende Abbildung zeigt. Beim Lernen werden die Verknüpfungen zwischen den Knoten stärker oder schwächer. Die Stärke einer Verbindung wird durch einen Wert zwischen 0 und 1 ausgedrückt.



NZZ / cia.

Künstliche neuronale Netze werden in ganz verschiedenen Bereichen der künstlichen Intelligenz angewendet. Hier erklären wir die Funktionsweise anhand der Bilderkennung.

Die erste Schicht eines neuronalen Netzes zur Bilderkennung enthält etwa die Farbe von jedem Pixel auf einem Bild. Die folgenden Schichten verknüpfen die Informationen zu komplexeren Zusammenhängen. Die letzte Schicht enthält das Ergebnis, in diesem Fall etwa, was auf dem Bild zu sehen ist. Durch Übung lernt das System von selbst, auf welche Informationen es ankommt. Es analysiert Unmengen von Bildern und passt dabei seine Verknüpfungen und Knoten an.

Allerdings ist es grossteils unklar, was genau in dem neuronalen Netz passiert. Es ist zu einem gewissen Grad eine Black Box, in die man nicht direkt hineinschauen kann. Dadurch kann es etwa vorkommen, dass unpassende Kriterien für die Unterscheidung von Bildkategorien herangezogen werden.

Zum Beispiel «erkannte» ein Algorithmus den Unterschied zwischen Hunden und Wölfen am Schnee, der im Bildhintergrund zu sehen war, weil in den Trainingsdaten Wölfe fast immer in einer Schneelandschaft abgebildet waren. So ein System kann bei Trainingsdaten gute Resultate liefern, zugleich aber scheitern, wenn man es in der realen Welt anwendet.

Was steckt hinter Chat-GPT? →

Chat-GPT ist ein Chatbot, der auf einem neuronalen Netz basiert. Die neuronalen Netze der GPT-Familie gehören zu den grössten der Welt. Es handelt sich um Sprachmodelle, die dazu gemacht sind, menschliche Sprache zu verarbeiten.

Chat-GPT wurde Ende 2022 gratis der Öffentlichkeit zur Verfügung gestellt. Die erstaunlichen Fähigkeiten des Chatbots, der auf alle möglichen Fragen plausible und scheinbar kreative Antworten gibt, haben ein grosses Publikum begeistert und Sprach-KI

über Expertenkreise hinaus bekannt gemacht.

Grundsätzlich bauen diese Modelle auf einen Algorithmus auf, der vorhersagt, was das nächste Wort ist, wenn man ihm Text vorlegt.

Was steckt hinter Chat-GPT und Co.?

Chatbots wie Chat-GPT können Texte schreiben und Fragen beantworten, als wären sie ein Mensch.

Hinter dem Chatbot steckt ein statistisches Modell. Wie das funktioniert →

Beispiel: Texte vervollständigen

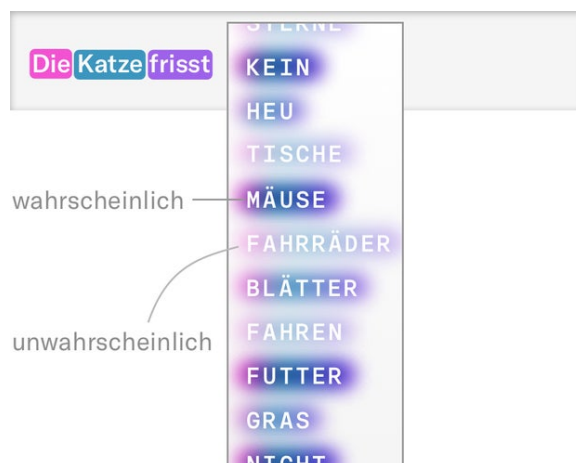
Die Katze frisst MÄUSE

Beispiel: Fragen beantworten



Was steckt hinter Chat-GPT und Co.?

Das Modell berechnet, welches Wort wahrscheinlich als Nächstes kommt. Es berücksichtigt dazu alle Wörter davor als Kontext.



Was steckt hinter Chat-GPT und Co.?

So arbeiten sich Chatbots von Wort zu Wort vor und können plausible Texte erzeugen.

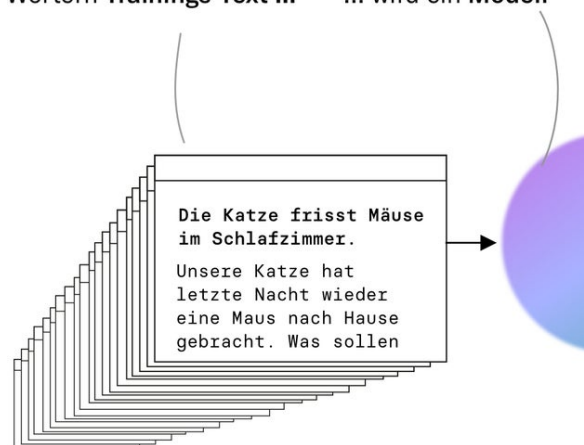
Woher diese Wahrscheinlichkeiten kommen →



Was steckt hinter Chat-GPT und Co.?

Die Wahrscheinlichkeiten errechnen sich aus Trainingstexten. Digitalisierte Bücher sind darin enthalten und ein grosser Teil des Internets.

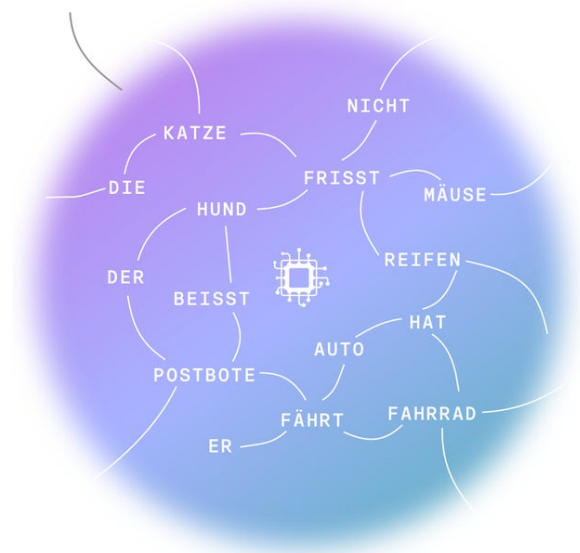
Aus 360 Milliarden Wörtern **Trainings-Text** wird ein **Modell**



Was steckt hinter Chat-GPT und Co.?

Das Modell macht aus den Texten ein Netz von Bedeutungszusammenhängen. So weiss es, welche Wörter tendenziell zueinandergehören. Darauf basierend kann es neue Texte erzeugen. Dann spricht man von generativer künstlicher Intelligenz.

Das Modell

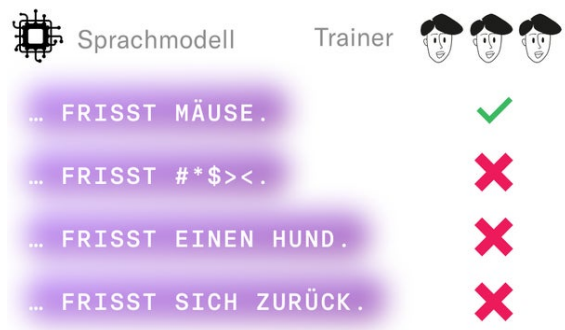


Was steckt hinter Chat-GPT und Co.?

Auch mit vielen Daten sind die Outputs von Sprachmodellen oft seltsam oder moralisch unerwünscht. Mithilfe von Menschen wird dem Modell deshalb beigebracht, was eine gute Antwort ist und was nicht.

Weil das Modell nur Wörter vorhersagt und kein tatsächliches «Wissen» beschreibt, können die Outputs auch falsch sein. Das nennt man auch Halluzination.

NZZ / joe., las.



Wo kann KI Menschen ersetzen? →

Computerisierte Algorithmen verarbeiten heutzutage eine grosse Menge von Daten gleichzeitig. Dadurch übertreffen sie die Rechenleistung eines menschlichen Gehirns um ein Vielfaches. Es ist beispielsweise unmöglich für den Menschen, einen Schachcomputer zu besiegen, der in Sekundenbruchteilen Millionen von Schachzügen und deren Konsequenzen berechnen kann.

Heute kann KI Tumoren auf Röntgenbildern erkennen, Musikhörern das nächste passende Lied vorschlagen oder wissenschaftliche Forschung vorantreiben. Beispielsweise wurde es durch ein neuronales Netz namens AlphaFold erstmals möglich, aufgrund einer gegebenen Folge von Aminosäuren vorherzusagen, wie sich ein Protein falten wird.

Generative künstliche Intelligenz kann zudem Bilder und Texte in ziemlich hoher Qualität erzeugen. Diese beinhalten oft Fehler oder Klischees, aber in Kombination mit menschlicher Kuratierung hat diese Art von künstlicher Intelligenz das Potenzial, viele Berufsfelder zu verändern.

In welchen Bereichen ist der Mensch der KI überlegen? →

Das Gehirn verarbeitet Informationen insgesamt langsamer als ein Computer. Trotzdem ist es dem Computer weit überlegen, wenn es darum geht, in kurzer Zeit ein neuartiges Problem zu lösen.

Wie das Gehirn dabei im Detail vorgeht, ist Gegenstand der neurowissenschaftlichen Forschung. Kognitionsforscher sprechen von sogenannten Heuristiken. Das sind Daumenregeln, die der Mensch anwendet, um Probleme zu lösen. Das Charakteristische daran: Anstatt alle verfügbaren Informationen zu verarbeiten, greift das menschliche Gehirn auf nur wenige, informationsreiche Anhaltspunkte zurück. So erzielt es die optimale Balance zwischen Geschwindigkeit und Perfektion – eine Lösung, die sowohl schnell genug als auch gut genug ist.

Einfache Heuristiken können im Alltag überlebenswichtig sein. Beispielsweise hat unser Gehirn gelernt, dass ein bewegtes Objekt Gefahr bedeuten kann. So können wir einem herannahenden Auto blitzschnell ausweichen, ohne zuerst irrelevante Informationen wie die Farbe des Autos zu verarbeiten. Heuristiken erlauben es uns, mit der limitierten Rechenleistung unseres Gehirns effektiv umzugehen.

Computersysteme sind noch weit davon entfernt, mit der Fülle von Informationen, aus der die echte Welt besteht, ähnlich effizient umzugehen.

Schliesslich ist der Mensch auch einem grossen Sprachmodell überlegen. Denn er versteht den Inhalt des Gesagten und kann diesen auf sein sogenanntes «kommunikatives Ziel» hin beurteilen und anpassen. Das heisst, er weiss, wer zu wem in welchem Zusammenhang spricht und wie der Zuhörer das Gesagte interpretieren könnte.

Gibt es bald echt intelligente KI und Maschinen mit eigenem Bewusstsein? →

Ein System, das nicht nur spezialisierte Aufgaben übernimmt, sondern wie der Mensch ein breites Spektrum an Dingen tun kann, nennt man «generelle» künstliche Intelligenz oder «starke» künstliche Intelligenz. Von der Informatik bis zur Philosophie – es wird aus vielen Perspektiven darüber debattiert, ob solche Systeme prinzipiell möglich sind.

Man fragt sich etwa, ob man nur von echter Intelligenz sprechen kann, wenn ein Bewusstsein vorhanden ist. Dann würden Maschinen vielleicht nie intelligent werden. Damit verknüpft ist die Frage, was überhaupt das menschliche Bewusstsein ausmacht.

Gehört dazu die Fähigkeit, über sich selbst nachzudenken? Oder gehört zu menschenähnlicher Intelligenz auch ein eigener Wille, eine individuelle Handlungsmotivation?

Einige Fachleute sind überzeugt, dass Maschinen menschenähnliche Fähigkeiten erreichen werden. Doch solche Prognosen gibt es schon länger, und bisher verschoben sie sich regelmässig nach hinten. Im Moment stützen sie sich unter anderem auf die Geschwindigkeit, mit der Chips in den letzten Jahren schneller und billiger wurden. Setzt sich diese Entwicklung fort, so das Argument, könnte man in dreissig bis vierzig Jahren die Rechenleistung des menschlichen Gehirns digital nachbauen, also um 2050 oder 2060.

Andere Forscher sind skeptisch, ob es überhaupt je menschenähnliche Maschinen geben könnte. Denn neben der Rechenleistung – die bereits jetzt die des menschlichen Hirns übertrifft – gibt es noch kategorische Herausforderungen, die der Entwicklung menschenähnlicher Maschinen im Weg stehen.

So leisten grosse Sprachmodelle wie Chat-GPT Erstaunliches. Doch die Sprache, die sie verarbeiten, ist eindimensional. Bis jetzt kann kein künstliches neuronales Netz mit der Mehrdimensionalität umgehen, die die Welt um uns herum kennzeichnet.

Wie umweltschädlich ist maschinelles Lernen? →

Das Trainieren von künstlichen neuronalen Netzen benötigt grosse Mengen an Rechenleistung, also auch an Energie. Wie viel Energie genau, darüber geben die Entwicklerfirmen oder Forschergruppen meist keine Auskunft. Eine vielzitierte Studie der Universität von Massachusetts, Amherst, besagt, das Trainieren eines grossen Netzes für Sprachverarbeitung benötige etwa so viel Energie wie ein Amerikaner in siebzehn Jahren Lebenszeit. Einmal trainiert, kann dieses Modell allerdings ganz verschiedene Dinge mit Sprache tun, wie Fragen beantworten oder Texte zusammenfassen. Feinanpassungen für neue Aufgaben sind weniger aufwendig und ohne Neutraining möglich.

Die meisten neuronalen Netze sind noch viel kleiner, doch der Trend geht gen Wachstum. Zudem sind in den Prozessoren knappe Ressourcen wie Lithium verbaut. Es ist also durchaus sinnvoll, dass sich Forschende vermehrt überlegen, welchen Nutzen ein Training mit noch mehr Daten und Rechenleistung bringt und ob sich der Aufwand lohnt.

Welche ethischen Risiken birgt künstliche Intelligenz? →

Künstliche Intelligenz kann Muster in Datensätzen finden. In der Praxis werden diese Muster verwendet, um ihnen neue Daten zuzuordnen. Wird also beispielsweise ein Algorithmus an einem Datensatz trainiert, in dem Frauen in Kaderpositionen unterrepräsentiert sind, so fliesst dieses Muster als «Bias» in die daraus abgeleiteten Prognosen oder Eignungseinschätzungen ein. Mit anderen Worten: Wird künstliche Intelligenz mit Daten einer Welt gefüttert, in der es Diskriminierung gibt, so setzt sich die Diskriminierung möglicherweise fort.

KI-basierte Übersetzungsprogramme illustrieren die Problematik. Beispielsweise kennt die türkische Sprache kein grammatikalisches Geschlecht. «Er» oder «sie» wird gleichermassen mit «o» ausgedrückt. Bei der Übersetzung vom Türkischen ins Deutsche wird aus «doktor» «der Arzt», aus «ist sexy» «sie ist sexy» und aus «ist kompetent» «er ist kompetent». Die häufigere Wortkombination im Trainingsdatensatz ist für die Übersetzung ausschlaggebend.

Diese Diskriminierung zu kontrollieren, ist schwierig. Denn weil künstliche Intelligenz an grossen Datenmengen trainiert werden muss, werden Trainingsdaten oft wenig selektiv direkt im Internet gesammelt. KI-Systeme sind aber in der Anwendung nur so gut wie die verwendeten Trainingsdaten. Viele Gesichtserkennungssysteme etwa funktionieren bei schwarzen Frauen am schlechtesten, weil sie vor allem mit Fotos von Weissen und von Männern trainiert wurden.

Für diese Probleme gibt es keine einfache Lösung. Denn die Kriterien, anhand deren ein Algorithmus beispielsweise zu seiner Kategorisierung kommt, sind nachträglich oft schwer nachvollziehbar. Deswegen arbeiten Forscher daran, Licht in die Black Box der selbstlernenden KI-Algorithmen zu bringen. Auf diese Weise könnten unerwünschte Unterscheidungskriterien, zum Beispiel das Geschlecht von Jobanwärtern, bewusst ausgeschlossen werden. Künstliche Intelligenz hat also das Potenzial, der Diskriminierung gar entgegenzuwirken.

Gibt es Beispiele für KI in unserem Alltag? →

Künstliche Intelligenz steckt in vielen digitalen Produkten: etwa in Spam-Filtern von E-Mail-Programmen, in Suchmaschinen, Sprachassistenten oder in der automatisierten Empfehlung von Musik oder Filmen. Auch bei der Verkehrsplanung und bei medizinischen Abklärungen kommen solche Algorithmen zum Einsatz, aber auch bei Bewerbungsverfahren oder auf Arbeitsämtern. Weil gerade dort eine Fehlentscheidung schlimme Auswirkungen haben kann, wird KI bis jetzt nur unterstützend eingesetzt. Sie gibt Empfehlungen, medizinische Entscheide treffen aber immer noch die Menschen.

Wer ist für Entscheidungen einer lernenden KI verantwortlich?

→

Wie so oft hinkt das Recht bei der KI der technischen Entwicklung hinterher. Weil KI-Systeme sich während des Einsatzes weiterentwickeln können und weil von aussen oft nicht klar ist, wie genau sie zu bestimmten Schlüssen kommen, muss etwa geklärt werden, wer für ihre Entscheidungen haftet.

Neue technische Möglichkeiten werfen auch neue ethisch-rechtliche Fragen auf, etwa, wann es erlaubt sein soll, Menschen anhand von Gesicht und Stimme zu identifizieren.

Produkte wie medizinische Geräte oder Autos müssen bereits jetzt produktspezifische Sicherheitsbestimmungen erfüllen. Die EU ist dabei, darüber hinaus Standards zu beschliessen, die für alle Produkte gelten sollen, in denen KI steckt. Die Kommission hat

dazu einen branchenübergreifenden Regulierungsvorschlag eingebracht. Die Hoffnung ist, dass diese Standards anderswo Anklang finden.

Die EU will nicht nur das Potenzial, sondern auch die Risiken der KI ernst nehmen. Die USA und China sind zurückhaltender mit solchen Vorstößen. Doch Diskussionen um KI-Systeme zur Gesichtserkennung oder solche, die automatisiert Entscheidungen treffen, gibt es auch dort.

Andere Autoren

Eveline Geiser (evg)

Weitere Themen

ChatGPT