Künstliche Intelligenz: Ursprung, Mechanismen und Perspektiven

Entstehung, Hintergrund, Herkunft

Künstliche Intelligenz (KI) hat ihren Ursprung in der Mitte des 20. Jahrhunderts, als Mathematiker wie Alan Turing die Grundlagen für maschinelles Lernen legten. Der Begriff „Künstliche Intelligenz“ wurde 1956 bei der Dartmouth-Konferenz geprägt. Ziel war es, Maschinen zu entwickeln, die menschliches Denken nachahmen können.

Meilensteine: 1996 gewinnt der IBM-Computer Deep Blue gegen den amtierenden Schachweltmeister Garri Kasparow. 2016 gewinnt AlphaGo von Google gegen den damaligen GO-Champion Lee Sedol.

Grundidee, Prinzip, Verfahren

Die Grundidee der KI ist es, Systeme zu schaffen, die Probleme lösen und Entscheidungen treffen können, ähnlich wie der Mensch. Dabei kommen Verfahren wie maschinelles Lernen, neuronale Netze und Algorithmen zum Einsatz. Maschinelles Lernen basiert darauf, dass Systeme aus Daten Muster erkennen und diese nach statistischen Methoden benennen, z.B. Erkennung von Bildern oder Handschriften. Zum Verständnis wichtig ist, dass alle KI-Anwendungen Rechenvorschriften sind und keine sinnentnehmenden oder kreativen Vorgänge.

Mechanismen, Erklärung, angewandte Technik

KI-Technologien nutzen sehr komplexe Algorithmen und Datenstrukturen. Zum Beispiel analysieren neuronale Netze große Datenmengen und lernen daraus, indem sie Gewichte anpassen, um Muster zu identifizieren. Ein neuronales Netz erhält Eingabedaten und berechnet daraus Ausgabedaten. Die Gewichte (Parameter) in einem Netz steuern die Stärke der Verbindung von Neuronen, ob diese aktiviert werden oder unterdrückt. Der Output eines neuronalen Netzes hat zu Anfang des Prozesses eine Abweichung von den erwarteten Ergebnissen. Durch die Fehlerrückführung werden die Gewichte modifiziert. Ziel dieses künstlichen Lernprozesses ist die Minimierung des Fehlers (bzw. der Verlustfunktion). Ein Beispiel ist die mittlere quadratische Abweichung. Diese Verfahren finden Anwendung in Bereichen wie Spracherkennung und Bildverarbeitung. Die Technologie erfährt durch Fortschritte in Rechenleistung und Datenverfügbarkeit eine rasante Entwicklung. Eine Auswahl bekannter Anwendungen: Siri in Apple, Finanzmarktanalysen, Rechtschreibkorrekturen, Spamfilter, Gesichtserkennung, Empfehlungssysteme (z.B. Amazon, Netflix etc.), personalisierte Werbung, Übersetzungsprogramme, Assistenzsysteme im Kfz (autonome Fahrzeuge) u.ä.

Mathematische Grundlagen von ChatGPT (Generative Pretrained Transformer)

ChatGPT basiert auf einem sogenannten "Transformer-Modell", einer speziellen Architektur für neuronale Netze, die 2017 entwickelt wurde. Vereinfacht lässt sich dies wie folgt erklären:

1. **Neuronale Netze:** Die Grundidee des maschinellen Lernens ist ein künstliches neuronales Netz, das dem natürlichen Netz in Gehirnen nachempfunden ist.
2. **Lineare Algebra**: Die Basis vieler Berechnungen sind Matrizen und Vektoren. Beispielsweise wird Sprache als Zahlenvektor dargestellt, der mathematisch bearbeitet werden kann. So kann z.B. die „Verwandtschaft“ von „King“ und „Queen“ als dreidimensionale Nähe von zwei Vektoren dargestellt werden.
3. **Wahrscheinlichkeitsrechnung**: Das Modell sagt voraus, welches Wort mit welcher Wahrscheinlichkeit als nächstes kommt, basierend auf Kontext und gelernten Mustern. Hier ist insbesondere das Bayes-Theorem der bedingten Wahrscheinlichkeit zu nennen. Wie hoch ist die Wahrscheinlichkeit eines Ereignisses A unter der Bedingung von B? Die bedingte Wahrscheinlichkeit wird nach folgender Formel berechnet: $P(A|B) = [P(B|A)\cdot P(A)]/P(B)$
4. **Optimierungsverfahren**: Durch das sogenannte "Gradient Descent" wird das Modell darauf trainiert, Fehler zu minimieren. Das bedeutet, dass es durch viele Wiederholungen lernt, bessere Vorhersagen zu treffen. Die sogenannte Verlustfunktion nimmt immer kleinere Werte an. Das geschieht durch Differenzierung nach der Kettenregel.
5. **Selbstaufmerksamkeit (Self-Attention)**: Diese Methode erlaubt es dem Modell, die wichtigsten Teile eines Textes zu identifizieren. Beispielsweise bezieht sich das Pronomen "er" im Satz "Max ging nach Hause, weil er müde war" auf "Max". Dies wird mathematisch durch Gewichtungen in Matrizen dargestellt.
6. **Trainingsumfang** Die folgenden Zahlen beziehen sich auf ChatGPT 3: Dieses Model wurde mit ca. 500 Milliarden Tokens=Wortbestandteile oder ganze Worte- trainiert. (Der Begriff Token in KI hat nichts mit dem Begriff Token=Anteil in Kryptowährung zu tun.) Der technische Aufwand bestimmt auch die Kosten der KI-Modelle. In das Training sind aber nach wie vor Menschen involviert, die Ergebnisse einer KI nach Wahrheitsgehalt, rechtlichen Verstößen oder weltanschaulichen Prinzipien bewerten. In den physikalischen Grenzen des Trainings liegen auch die Gründe für Falschaussagen (sog. Halluzinieren) oder offensichtlich gesteuerte Ergebnisse. Weniger bekannt ist, dass für solche menschlichen Interventionen zum großen Teil Billig-Arbeitnehmer aus eher armen Ländern rekrutiert werden. (Sog. Click-Worker)

Der Transformer setzt Wort auf Wort nach statistischen Prinzipien. Schreiben ist für die KI ein Lückentext mit der Lücke stets am Ende. Bsp.: „Im Juli ist der Himmel …“

Diese Grundlagen zeigen, wie Mathematik, insbesondere Algebra, Statistik und Optimierung die Basis für die Funktionsweise von ChatGPT bilden. Eine genauere mathematische Darstellung der Architektur von GPT 3 findet sich hier: https://dugas.ch/artificial_curiosity/GPT_architecture.html

Entwicklung, aktueller Stand, Akteure

Heutzutage spielt KI eine zentrale Rolle in der Wissenschaft, Industrie und im Alltag. Unternehmen wie Google, OpenAI und Microsoft gehören zu den Vorreitern. Durch bahnbrechende Modelle wie ChatGPT hat KI ein neues Niveau erreicht und wird kontinuierlich weiterentwickelt. Die technische Dimension von ChatGPT: So hat OpenAI für GPT3 eine Zahl von 175 Milliarden Parametern (Basiswerte und

Gewichte für ein Neuron) angegeben. Für das aktuelle Modell ChatGPT4o wird eine 100mal höhere Anzahl von Parametern geschätzt. Die schiere Größe der Modelle hat zur Folge, dass kein Entwickler sagen kann, was im Inneren eines solchen Modells stattfindet. Man spricht hier auch von der KI als Black Box. Dieser Aspekt wirft eine Reihe von juristischen Fragen auf. Neben des Black-Box Problems wird zunehmend der sogenannte Bias der Modelle diskutiert. Mit Bias wird hier eine inhaltliche „Schlagseite“ verstanden, auch als das Problem des schwarzen George Washington bekannt geworden (von Google KI Gemini erzeugt): Bei der Suche nach einem Bild von George Washington gab die KI das Bild eines offensichtlich schwarzen George Washington aus.

Perspektiven, Erwartungen, Interessenlagen

Die Perspektiven der KI sind ambivalent: Sie könnte die Gesundheitsversorgung revolutionieren, die Wissenschaft voranbringen und neue Geschäftsmodelle ermöglichen. Die Schattenseiten der KI sind schon heute sichtbar: Chat Bots ersetzen menschliche Arbeitskräfte. Mittelfristig werden standardisierte Aufgaben durch KI erledigt: Schon werden geschätzt 30-40 % der journalistischen Texte von KI verfasst. Dieser Anteil dürfte in Schule und Universität deutlich höher liegen, ist aber nicht bekannt. Durch den sog. EU AI Act werden Unternehmen und Organisationen ab 2026 verpflichtet, den Einsatz von KI zu dokumentieren. Zwischen Interessenlagen, etwa von Unternehmen, die KI kommerzialisieren, oder von Regierungen, die ihre ethischen und sicherheitspolitischen Aspekte betrachten, wird sich die Entwicklung abspielen. KI-Unternehmen sind aber noch nicht wirklich profitabel. Google verdient mit Werbung, nicht mit KI sein Geld. Kurios: Die meisten Pro-Anwendungen/Abonnements von KI kosten in 2024 20$ im Monat.

Chancen und Risiken

KI bietet immense Chancen, wie die Automatisierung repetitiver Aufgaben und die Entwicklung neuer Medikamente. Gleichzeitig birgt sie Risiken, darunter Arbeitsplatzverlust, Datenmissbrauch und algorithmische Diskriminierung. Es ist entscheidend, ein Gleichgewicht zwischen Innovation und Regulierung zu finden.

Konsequenzen für die Gesellschaft

Die aktuelle Politik in Deutschland geht mit KI lediglich als Schlagwort um. Während in den letzten Dekaden des 20. Jahrhunderts Technologiefolgenabschätzung ein Großthema war, wird in Deutschland KI nicht prioritär behandelt, außer im Bereich von Bürokratie und Regulierung (siehe auch EU Artificial Intelligence Act). Die offensichtlichen juristischen Probleme, etwa Datenschutz, Urheberrecht, Haftung, die Konsequenzen für den Bildungsbereich u.ä. stehen noch nicht im Fokus des öffentlichen Diskurses. Das kann in Europa auch durch einen eklatanten Mangel an MINT-Kompetenz erklärt werden. Für mittelständische Unternehmen wird die Implementierung einer KI-Anwendung zum Kostenfaktor, der mögliche Produktivitätsgewinne wieder beseitigt. So ist im Unterschied zur öffentlichen Darstellung eine spezifische Anwendung von ChatGPT im Unternehmen (als sogenannter API-Assistent) mit erheblichen Implementierungskosten verbunden. Für den Normalbürger wird die kostenlose Nutzung von KI zunehmend unattraktiver wegen Kapazitätsbeschränkungen. Zudem ist bei den sogenannten Pro-Versionen ab 2025 ein Preissprung zu erwarten.

Einschätzung, Bewertung, Kritik

Die Entwicklung von KI ist ambivalent. Die immer stärkere Produktion von Texten oder Artefakten wird dazu führen, dass Trainingsdaten aus dem World Wide Web, Wikipedia o.ä. auch von KI generiert sein werden, sodass man von einer Art Datenkannibalisierung sprechen kann: Der Anteil von KI generierten Trainingsdaten wird steigen, der Anteil von originär menschlichen Artefakten sinken. Die physikalischen Grenzen sind schon heute durch den enormen Energieverbrauch der verbauten Graphikprozessoren und der Mächtigkeit der Transformer Architektur sichtbar. Auf der anderen Seite kann man die Potentiale von KI nicht hoch genug bewerten, ohne Übertreibung könnte man sicher von einer neuen technologischen Basisinnovation sprechen (etwa einem neuen Kondratieff-Zyklus), die zu erheblichen ökonomischen Strukturveränderungen und Wandlungsprozessen führen und die insbesondere den White-Collar-Arbeitsmarkt treffen wird.

Für die europäische Gemeinschaft ist zu hoffen, dass man neben Regulierungen auch in technische Entwicklungen und Ausbildung investieren wird, um nicht noch mehr in eine Abhängigkeit von US-amerikanischen und asiatischen Unternehmen zu geraten. So hat der neue US-Präsident mit dem Projekt „Stargate“ eine Investition von 500 Mrd. Dollar in Hardware für KI angekündigt.

Weiterführende Literatur und Quellen

- **Bücher**:
  - Rechtsleitfaden KI im Unternehmen Mühleis, Niklas 2024
  - Mathematik und ChatGPT
    Helfrich-Schkarbanenko, Andreas]
    [2023]
  - Künstliche Intelligenz kapieren & programmieren *Michael Weigend 2023*
  - Hacking und Cyber Security mit KI
    Dalwigk, Florian André]
    2023
  - Russell, S., & Norvig, P. *Artificial Intelligence: A Modern Approach* 2020
- **Artikel**:
  - Vaswani, A., et al. (2017). "Attention Is All You Need."
- **Online-Ressourcen**:
  - OpenAI Blog: https://openai.com/blog

Anmerkung der Autoren:

Der Text wurde zu 80% von Werner Michael und Walter Dörhage und zu ca. 20% von ChatGPT 4o verfasst. Der Text wurde mit Zero GPT detektiert.